Atmos. Meas. Tech. Discuss., 6, C3613–C3623, 2013
www.atmos-meas-tech-discuss.net/6/C3613/2013/

Atmospheric Measurement Techniques Discussions
Open Access

AMTD
6, C3613–C3623, 2013

Interactive Comment

Interactive comment on "Trajectory matching of ozonesondes and MOZAIC measurements in the UTLS – Part 2: Application to the global ozonesonde network" *by* J. Staufer et al.

J. Staufer et al.

johannes.staufer@lsce.ipsl.fr

Received and published: 6 December 2013

We thank you for your review. Your comments and recommendations have certainly helped to improve the manuscript. Please find our responses to your comments below.

Major issue (1) Influence and variance of trajectory matching:

Comment: Since the a/c photometers are regularly calibrated and checked in flight they are assumed, by the authors, to be the standard as the paper is presented. The ozonesondes have also been well characterized in terms of precision and accuracy, by laboratory and field work, and differences resulting from differences in solution strength and ECC sonde manufacturer are also well characterized. Thus when there are differences, outside the known precisions between the sondes and Mozaic, how should it be interpreted? Throughout the paper, the authors attribute reasons for the differences to the sondes. Is this correct?

Response: We modified the revised manuscript by confirming the possible calibration drift as indicated by Logan et al. (2012) and Staufer et al. (2013) as the motivation of the study. It is now stated that, based on the findings of Logan et al. (2012) and Staufer et al. (2013), who both raised some concern about the long-term stability of MOZAIC, one of the purposes of the revised paper is to check whether MOZAIC or the ozonesondes, or both, have some problems in the period 1994-1998.

Comment: To achieve the comparisons, trajectory calculations are used to find matches with air parcels sampled by both instruments. How accurate are these trajectories? How different would the comparisons be if an ensemble of trajectories, or variable trajectory matching calculations, isentropic, 3-D, were used in each case, or in example cases? Some variance due to small differences in trajectories for the same air parcel matches must be provided. How important is time? How do 3-day trajectories compare with 6-day trajectories in terms of the comparisons? A more systematic analysis of the differences, which may be attributed to uncertainties of the trajectories alone, should be provided, and should be organized into a section presented early in the paper. Presently this discussion is done to some extent with the forward / backward trajectory separation, but this discussion is scattered throughout the text, never satisfactorily carried through to explain why such differences should be observed, other than in the case of Izana when completely different air masses are sampled, and in several places loosely attributed to differences in chemistry between the forward and backward trajectory, which, without more detail, makes no sense to this reader.

Response: We agree that some more explanation is appropriate since the present paper, although a companion paper, should be self-contained. However, most of the questions raised are already explained in the companion paper (Staufer et al, 2013,

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

accepted for publication in AMT) to which we refer. We decided to extend the model description of the revised manuscript. The accuracy of the trajectories was estimated to be +/- 2%, using MOZAIC-MOZAIC “self-matches”. To derive this uncertainty we assumed that the uncertainty of the MOZAIC observations is negligibly small. Stohl and Seibert (1998) showed that fully three-dimensional trajectories are more accurate than kinematic isentropic or isobaric trajectories in the troposphere, while in the stratosphere, isentropic trajectories are of similar accuracy to fully three-dimensional trajectories. So, in both cases fully 3D trajectories are amongst the most accurate trajectories. We also calculated clusters of trajectories to see if the errors can be reduced but found that too tight criteria reduce the sample size too much while too soft criteria did not affect the results. The effect of the length of the trajectories will be discussed more thoroughly in the revised paper, since the time difference of the matches needs to be considered because trajectory errors typically accumulate with time. When we derived our uncertainties (Staufer et al., 2013) for the ozonsonde station Payerne we got a large number of matches and the majority of them where found within the first three days. However, we found that when the matches in the first 24 h were excluded, the uncertainty increased by 1-2%, particularly for tropospheric matches. A quantitative assessment of the increasing uncertainty with time is difficult to ascertain given the uneven temporal distribution of the matches. The results, for example for Sapporo, Tsukuba, and other stations mostly outside of Europe, however, clearly show the limitation of the matching technique when most matches are found after the trajectories have traveled for more than 50 h.

Comment P7113.20. Here and elsewhere. Chemical processing along the trajectories is not a very satisfactory explanation for the forward/backward trajectory differences. Why would this cause a difference? Even though it is explained elsewhere, it is important enough in this work to repeat the essence of the explanation.

Response: We argue that ozone production along the trajectory path could cause a potential problem with this method. However, we agree that chemical processing

AMTD
6, C3613–C3623, 2013

Interactive Comment

along the trajectories cannot be solely responsible for backward/forward trajectory differences. This has been also discussed by Staufer et al. 2013, who showed that the backward/forward trajectory differences for Payerne is particularly pronounced in summer, which would, on one hand, support the importance of chemical processes, but on the other hand, the differences are too large to be plausibly explained by photochemistry. Therefore, it is rather a combination of different factors (photochemistry, unresolved vertical motions, when/where does the measurement take place (behind a trough/high), and uneven distributions (temporal and spatial) of the matches). It is planned to explain the differences between forward- and backward-only trajectories more thoroughly in the methodology section (Section 2.3). We have modified the revised version of the manuscript accordingly.

Comment P7122.10-13. The second to last paragraph of the paper is too late to be making the statement. "Overall, results are more uncertain and less consistent when the majority of the aircraft measurements match with trajectories after they have traveled more than three days, which is the case, for example, for the Japanese and most of the tropical stations." What is meant by, "more uncertain"? Do the authors even know? They should. It is a tractable problem. At some point, if the uncertainty is too large, there is little point in publishing the comparisons.

Response: We plan to discuss the influence of the trajectory length already in the methodology section and at the beginning of Section 3 (Results) in the revised manuscript. Additionally, we separated the results for stations with a large number of matches (Uccle, MOHP, Payerne, DeBilt, Legionowo) from those with a low number of matches (in the revised manuscript there is a paragraph after the discussion of the station Legionowo). We do not know exactly (in terms of absolute numbers) how much the uncertainty increases with increasing trajectory travel time due to the limited number of matches in this time frame (the uncertainty was derived using MOZAIC-MOZAIC self-matches). However, as already pointed out above, the analysis of Staufer et al. (2013) showed a 1-2% increase of uncertainty when the matches are not allowed in the first

AMTD
6, C3613–C3623, 2013

24 h of the trajectories' travel. Additionally, there are several studies regarding errors in trajectory analyzes (e.g., Stohl, 1998), so we know that trajectory differences increase with travel distance, but due to the uneven temporal distribution of the MOZAIC self-matches we cannot give an exact number. However, the statement “”Overall, results are more uncertain . . . “ is removed from the conclusions.

References: Stohl, A.: Computation, accuracy and applications of trajectories – a review and bibliography, Atmos. Environ., 32, 947–966, 1998.

Major Issue (2) Bias toward the Mozaic as the standard reference instrument:

7120 last paragraph. The discrepancy between Mozaic and sonde from 1994 – 1998 and the subsequent improvement is, in the authors' statements here, attributable only to changes/improvements in the ozonesondes. This conclusion is in spite of the fact that the same discrepancy is observed with both Brewer Mast and electrochemical sondes. This conclusion is too narrow. Throughout the paper the authors appear convinced that Mozaic coupled with the trajectory matching technique is the reference to which the ozonesondes are to be compared, and either found wanting, 1994-1998, or blessed, after 1998. Is it not possible that, in the period 1994-1998, there was a systematic bias in the Mozaic measurements, or in the application of the trajectory matching? What assurances have been presented to rule this out? The authors should at least consider the possibility that their own measurements and techniques may have some systematic problems, particularly in view of the cross ozonesonde agreement in the differences observed with Mozaic 1994-1998.

7122.14-end. Now finally in the last paragraph of the paper, the authors acknowledge what I just said. This is much too late to acknowledge there may also be some issues with the Mozaic instruments, after the reader is led the whole way through attributing all differences to changes in sonde manufacturer or application. I whole heartedly agree with the last sentence of the paper, but object that this view is not brought much further forward, so that it is considered by the reader as the data are presented.

Response to both comments above: In the revised manuscript a calibration drift in MOZAIC as a possible cause for the larger difference in the first years is discussed right from the beginning and gives the motivation of this work. It is now stated that (based on earlier findings of Logan et al. (2012) and Staufer et al. (2013), who both raised some concern on the long-term stability of MOZAIC) one of the purposes of the revised paper is to check whether MOZAIC or the ozonesondes have some problems in the period 1994-1998. We intend to change the structure of the paper to bring more attention to the MOZAIC instrument issue. After analyzing all European BM stations, DeBilt and Legionowo, which are the stations with the largest numbers of matches, we discuss the possibility that there was a systematic bias in the MOZAIC measurements. Furthermore, this issue will be mentioned in both the Conclusions and Abstract.

Major Issue (3) Superficial description of differences, or the lack of differences, when sonde solution strengths are changed.

The description of ozonesondes in the introduction provides the standard detail, but does not quantify the well-determined differences of the different solution strengths with respect to a standard photometer or in sonde to sonde comparisons, both in the laboratory and in the field. Such information is brought up later in the discussion of several stations, but treated in a rather off-hand, cavalier way. Because of confusions in the community about solution strengths, this loose use of the language is to be avoided. Specific examples are the following.

7117.5-18. Please be careful with statements like SP and ES sondes were flown with 1.0% and no differences were observed. This conclusion is not supported by extensive laboratory and dual sonde field work to test for such a difference. Here the statement is true because the ES sondes with 1.0% made up a small fraction of the data, so they won't change the average behavior of the sonde-Mozaic comparison during this period. When the switch is really made to 0.5% ES sondes, the results reproduce the SP 1.0% just as they should, illustrating clearly the point above concerning the solution concentration in ES sondes. It has taken a long time for the community to establish and quantify these differences. It is not helpful when publications like this are not careful with their analysis, and imply that 1.0% in ES sondes does not lead to any differences. The differences seen at Scoresbysund are exactly as expected, but then without qualification in the next paragraph, 7117.26-27, the authors state, “Lerwick frequently changed between SP and ES sondes, but it seems this does not influence the agreement with MOZAIC data.” What was the frequency of the shift? How are the data weighted, more SP or more ES sondes? The Mozaic trajectories could be separated into those against ES and those against SP sondes. Then the authors may have a different result. While the statements made here may be true, they may not really reflect the differences, which could be masked by large weighting to one sonde or another. So if further investigation into the frequencies of the two types of sondes and their relative matches is not made, such statements should be removed.

Response: We agree that it is appropriate to avoid loose language, as this topic is of particular importance for the ozonesonde community. We have changed these paragraphs in the revised manuscript. We have made a clear distinction between Scoresbysund, where the ECC sondes have changed but a 1.0% KI was retained, and Sodankylä, which changed ECC sondes as well, but which used a sonde configuration following the recommendations of the scientific community and manufactures. At Scoresbysund, we will explain that differences of 5% are similar to those found by extensive laboratory and field studies such as JOSIE and BESOS. Lerwick is a different case. According to the WOUDC both SP and ES sondes were used but it was not exactly clear which sensor was used when. We therefore contacted the principle investigators but they could not report the exact switch dates either. In this sense, we cannot perform the analysis suggested. In the manuscript, the statements regarding no influence of the sensor change have been avoided.

7118.12-14. Why all of a sudden is the following statement included, “which is 2–3 times larger than what was shown in the JOSIE experiments for SP sondes operated with a 1.0% KI cathode solution (Smit et al., 2007).” That was in a laboratory with a photometer. Here as the authors state, the differences are similar to several European stations, so this reference now to differences in the laboratory for a particular sonde type, when no such discussion was provided for the earlier stations which showed a similar difference, is odd. If that information is to be included then it should be in the introduction when the different sonde types are discussed, and then differences with laboratory standards could be discussed. Here it is much too late.

Response: The reference to JOSIE have been removed in the revised manuscript at this point. However, later in the conclusion, we have made references to JOSIE and BESOS for all stations considered in this work.

7121.12-16. The concern with changes in solution strength is not with the NOAA measurements. The solutions used by NOAA are unique to NOAA, changes are carefully monitored to assure continuity of the record, and the data analysis method may account for changes through the application of an internal transfer function. The concern is with those stations using one of the standard buffered strengths, 1.0% or 0.5%, but not with the right manufactured ECC sonde. Thus this paragraph should not lump "all other stations" with Boulder. The NOAA stations can be compared amongst themselves, but comparing them with the other stations does not make sense. The sonde solution chemistry is quite different between the NOAA measurements and all other stations. It is comparing apples and oranges, never helpful, when the implication is comparing apples and apples.

Response: In the revised manuscript, we have clearly stated that the sonde techniques and solutions used at the NOAA sites are unique. First, this issue has been brought to more attention in the ozonesonde description section (Section 2.1). Second, the sentence "...but no obvious change in the sonde-MOZAIC agreement is found in the UTLS" in 7121.12-16. will be removed as the concern is not the NOAA sites.

Major Issue (4): Paper and figure presentation

Although the paper is well written, it is irritating to read because of the figure organiza-

AMTD
6, C3613–C3623, 2013

tion. Why are 7 figures in an appendix, but all are called out and discussed in detail in the paper? If the figures in the appendix really have a secondary role, then move the discussion of them into the appendix as well. Right now the reader has to continually jump back and forth between the main figures and figures in the appendix.

Response: We corrected this. All Figures now have a number and are arranged in order according to their appearance in the text.

Figure suggestions: Fig 2,3, a) It is very difficult to separate the forward and backward trajectories with monochrome symbols and lines. Use one color for forward (black) and another (red).

Response: The climatologies from both ozonesondes and MOZAIC (forward-only and backward-only) are very close to each other for most European and Canadian sites. Even colored lines will make a distinction difficult (particularly with the error bars also being displayed). For those stations with large differences, for example those in the tropics and those with a small number of matches, the different lines and symbols, in our opinion, are discernible. Therefore, it was felt unnecessary to re-do all plots.

Fig 2, 3 b), Include in the figure caption the reasons for the date breaks for each comparison. This information is most useful when viewing the figure rather than requiring the interested reader to leave the figure to find Table 1, where the information is still hard to extract.

Response: We mostly broke the comparison into three periods. The first covers the 1994-1998 period to better illustrate the MOZAIC UV-photometer issue, except when changes in sensor type or solution strength occurred. The other two periods typically comprise 4 years but for certain stations changing operating procedures we tried to analyze the periods before and after the change. We added in the Figure caption if and when operating procedures changed.

Minor comments: 7101.13 – do you mean NO longer used?

AMTD

6, C3613–C3623, 2013

Interactive Comment

Response: Yes.

7105.1-22. This is a lot of detail in a paragraph. Could it be added to Table 1?

Response: We tried but the table looked overloaded and unclear.

7112.14. There is something wrong with this sentence. “In terms of data treatment, but an altitude- declining background current with altitude is applied.” Is it even a sentence?

Response: It is corrected.

7113.5. I think Komhyr et al., (1995) recommended buffered solutions, not unbuffered.

Response: Yes, Komhyr et al., (1995) recommended buffered solutions. This is corrected in the revised manuscript.

7115.26. Here and elsewhere. The phrases, “typically above 250 hPa” are ambiguous and all should be changed to make them clear. To me above 250 hPa means, e.g. 300 hPa, not 200 hPa, which is what I think the authors mean here, and evidently earlier, but earlier I thought they were using it in the sense that I understand it so at altitudes below 250 hPa. This could be the reason I could not see in the figure the point of the discussion. Change all of these to use something like pressures < 250 hPa, or at altitudes above 250 hPa. Both of these are clear.

Response: To avoid any ambiguity, we use in the revised manuscript always phrases like “pressures < 250 hPa”.

7115.25. Why are the results encouraging? Not very descriptive. They are either in good or bad agreement. Encouraging suggests that these data were already suspect before being tested.

Response: We agree. We replaced this sentence with “The 16 yr mean O3 concentrations obtained from sondes flown at Tsukuba and Sapporo, and from MOZAIC show an agreement better than 5% in the troposphere (Fig. 10a).”

7118.25. “Differences between the two data sets are less than 5% in the stratosphere” This statement is supported by Fig. 7b but not by 7c. Why the difference?

Response: It’s due to the different calculations used in Fig. 7b and 7c (Figure 13b and 13c in the revised manuscript). The calculation of O3 in Fig. 13b is based on monthly mean differences. Some months contain more ascents than others. Ascents in months with few matches are therefore weighted stronger than ascents in months with many matches which is not the case for the calculation used in Fig. 13b, where all matched sonde ascents contribute with the same weight to O3.

Interactive comment on Atmos. Meas. Tech. Discuss., 6, 7099, 2013.