Atmos. Meas. Tech. Discuss., 6, C2720–C2727, 2013
www.atmos-meas-tech-discuss.net/6/C2720/2013/

Atmospheric Measurement Techniques Discussions

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.

Anonymous Referee #2

Received and published: 26 September 2013

This paper provides a comprehensive comparison of ozone measurements in the UTLS between ozonesondes and Mozaic aircraft measurements. The results are interesting, useful, and have a wide application. The paper is well written, the data mostly presented well, and it should be published, but not before addressing the following major issues.

1) Provide a systematic assessment, during the development of the methods, of the uncertainty introduced through trajectory matching. Three "instruments/tools" are involved in this paper: ozonesondes, Mozaic photometers, and air motion calculations. All three introduce uncertainties. The characteristics and variances of the first two are

mostly covered. The third is not.

2) Reduce the bias displayed here in attributing reasons, for the differences observed, primarily to the ozonesondes, either their manufacturer or their application. That the Mozaic photometers may also have some issues should not wait until the last paragraph of the paper to be acknowledged.

3) Be careful when discussing whether differences in solution strength between buffered 1.0 and 0.5% in the same sonde matter. This is an important issue which has led to some confusion in the sonde community. The confusion is finally nearly resolved with good laboratory and field comparisons. Statements in this paper, in several locations, suggest that the differences are not important; however, the authors have not provided the detailed analysis in those cases to make this statement stand. The authors have to either decide that they will carefully address the issue, or refrain from making such statements as detailed below.

4) Either treat the figures in the appendix as an appendix, and restrict reference in the paper to the minimum, or bring them fully into the paper as full figures. The present manuscript just serves to irritate this reader with the detailed discussion of figures in the main paper, then to the appendix, then to the main figures, then back to the appendix, back to the main paper, . . .

Major issues detailed.

1) Influence and variance of trajectory matching.

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 inter-

preted? Throughout the paper, the authors attribute reasons for the differences to the sondes. Is this correct?

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.

7113.20. 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.

7122.10-13. 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.

2) Bias toward the Mozaic as the standard reference instrument

That the Mozaic photometer may have issues is only acknowledged in the conclusions. This idea must be brought forward and this fact used to temper the discussion of differences. Note my comments in the conclusions on page 7120, before reading to the end on 7122.

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.

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.

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.

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.

4) Paper and figure presentation

Although the paper is well written, it is irritating to read because of the figure organization. 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.

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)

for backward trajectories. Use the same line pattern for sonde (solid) and Mozaic (dashed) in all figures. Then the comparisons can be easily discerned by the reader every time. Since these differences are often discussed the reader needs to be able to easily distinguish the measurements without continually referring to the legend as now, and even with the legend the pairs to compare are hard to distinguish since they are so close.

Fig 2, 3, ... c), d), e). Aadd a label indicating the LS, Trop, UT regions. Thanks for ordering these by altitude with the LS on top.

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.

Insure the figure caption and the figure are on the same page!

Minor comments.

7101.13 – do you mean NO longer used?

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

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?

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

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.

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.

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?

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