Interactive comment on “Application of GC/Time-of-Flight-MS for halocarbon trace gas analysis and comparison with GC/Quadrupole-MS” by J. Hoker et al.

E. Atlas (Referee)
eatlas@rsmas.miami.edu

Received and published: 7 January 2015

The manuscript by Hoker et al. presents a useful comparison of two mass spectrometer detectors for atmospheric halocarbon analysis. One detector is a conventional quadrupole mass spectrometer (QP-MS) which has an established performance record for trace gas analysis. The second detector is a time-of-flight mass spectrometer (TOF-MS), which is an emerging technology for environmental analysis, but with little application thus far for atmospheric halocarbon measurement. Thus the detailed comparison will be useful for someone evaluating analytical methods and instrumentation.

The manuscript is well organized and the material is presented clearly, though I have several comments and questions which are listed below.

1) My first suggestion is to change the title to “Comparison of GC/Time-of-Flight-MS with GC/Quadrupole-MS for halocarbon trace gas analysis.” This seems to be the essence of the manuscript. Since only 4 gases are evaluated and discussed, there is really very little in the manuscript that deals with the actual use and application of the TOF-MS. Some discussion of linearity issues are mentioned, but this is mainly in the context of comparison to the QP-MS.

2) 12325, l. 16. One main advantage of the TOF is the full-time, high-sensitivity mass spectral acquisition. A second advantage should be the reduction in background noise by selecting the exact mass (+ 50 ppm) of a target ion, which should eliminate interferences from those ions with the same unit mass. I think this aspect, esp. in comparison to the QP-MS, deserves more discussion.

3) 12426, l. 8. I would suggest that instrumental drift is a 6th “key parameter”. This can refer to drift in sensitivity or to drift in mass accuracy.

4) 12326, l. 22. Suggest removing “cryofocu-” and change to “sample enrichment on cooled adsorptive material...”

5) 12326, l. 25. Please provide length and i.d. of 1/16” tube.

6) 12327, l. 1. Please provide mg of Hayesep used in trap. Also, change “cryofocu-” to “enrichment”. What was the sample introduction flow rate, and how was it controlled?

7) 12328, l. 3. How was this split determined? Calculators available to me suggest that the split ratio is theoretically 55% TOF/45% QP. It seems to me that the uneven split between TOF and QP is a major problem in the experimental design. Since some of the properties depend on the absolute amount of mass reaching the detector, comparison of different mass amounts (which may be off by my calculation) biases the results. At a
minimum, some of the tests should be repeated with the splitters reversed. This would add confidence to the split ratio calculation and to any potential biases due to mass differences reaching the detector.

8) 12328, l. 26. Why set the MS scan to 500? For this test, maximum mass is 142. Presumably some improvement in S/N could be obtained by the QP by scanning a smaller mass range. A scan range of 100 amu could increase S/N by a factor of 2.

9) 12329, l. 11. Change to “regular”. How often is “regular”? How often were mass calibrations done on the TOF? Were any manual adjustments done to alter the autotune voltages? In my experience the autotune voltages can be manually adjusted to improve S/N. Though clearly good performance characteristics were found in these experiments, I wonder if additional improvement could come from manual adjustments of source or detector voltages.

10) 12330, l. 11. Define “u”.

11) 12331, l. 6. Change “the according” to “each” (if I understand the meaning).

12) 12331, l. 11. Please elaborate about “well-equilibrated” conditions. What are the matrix effects that are seen? How do you know if the system is well-equilibrated? Plus, I did not see a discussion later that described the mass accuracy over multiple runs, and the mean mass accuracy of the target masses.

13) 12331, Limits of Detection. Calculation and definition of instrumental limits of detection are often confusing and sometimes arbitrary. One definition is that used by the authors (e.g LOD = 3 x noise). However, it is critical how and where the noise is calculated, and how the baseline is drawn. I would suggest showing actual examples of the noise calculation for TOF and QP, esp. for the lowest level analyte, CH3I. In fact, the LOD will change over time due to various instrumental factors, so the determination and comparison is only marginally useful for “optimum” conditions. Furthermore, this determination might be significantly affected by the sample split sent to the different MS. Another approach that is being used to characterize instrument performance is the Instrument Detection Limit (IDL). Noise measurement can be quite variable, depending on the location and the width of the calculated noise segment, especially in low noise detectors. The IDL is based on ~ 3x the standard deviation of a compound measured at approximately 5 x the estimated LOD. This statistical approach seems to give a more meaningful evaluation of instrument performance. It would be helpful to have this information to compare for at least one of the analytes.

14) 12332, l. 7. Could you please explain the choice of 0.28 L sample size, when apparently 1.0 L is the normal sample volume (I assume then that the QP run alone would normally get 0.33 L and the TOF run alone would get 0.66 L?)

15) 12332, l. 22-25. This repeats an earlier description and can be deleted.

16) 12333, Mass Resolution. I found this whole section not very relevant. If the mass resolution of the TOF somehow improves performance, this should be shown with practical examples rather than some theoretical discussion. Can you demonstrate that the ability to separate the halocarbon fragments from a hydrocarbon fragment improves the halocarbon analysis? Is there any co-elution of fragments in the sample? Is there improvement in the baseline noise due to exact mass selectivity? Please provide some specific applications of the impact of mass resolution on the actual analysis of a real sample.

17) 12335, Mass Accuracy. I have the same comments about mass accuracy. A 50 – 100 ppm mass error may exclude some hydrocarbon fragments (though there is no indication that these fragments cause a problem), but this mass error does not exclude potentially other mass fragments of different elemental composition. Mass accuracy of <5 ppm or so would be a more relevant standard for discriminating ion fragments. Also, it would be helpful to understand how the mass accuracy drifted or varied over the different runs.

18) 12335, Limits of detection. I believe there are some mistakes in the description of
dwell time effects. The number of ions that reach the detector are not influenced by the dwell time. Lower dwell times do not reduce signal intensity (at least on the MS systems I use), but rather increase the noise level. Longer dwell times improve S/N by averaging the signal.

19) 12336, LOD. The increased LOD of the QP-MS in operational mode can be mitigated somewhat to use short dwell times for more abundant ions and longer dwell times for low concentration species. Further, one can often do without qualifier ions, and thus allow more dwell time for small peaks.

20) 12336. Reproducibility. Please include some detailed discussion of the method of peak integration. To get sub % precision requires some very reproducible methods for peak integration (esp. for the low abundance peaks). It would be valuable to know specifically how the different data were processed.

21) 12337, l. 13. I don't think the linearity issue is specific to the model of TOF used. All TOF suffer from less linear range compared to QP MS.

22) 12337, l. 24. I don't see how it is possible to quantitatively measure any substance retrospectively if a standard reference was not run at the same time. At best, retrospective analysis is only semi-quantitative.

23) 12338, l. 3. Delete "exemplary"

24) 12338. Linearity. Though it may be the subject for another manuscript, it would be interesting to know if the authors have examined the non-linear response as a function of ion-intensity. While the author’s discuss the impact of detector saturation on linearity (but also on mass accuracy), it is unclear why the low level samples demonstrate lower responses. This could be related to lower detector efficiency at low ion counts, or might it be related to some other baseline attribution issue (i.e. choosing a baseline that eliminates some response due to compound). Also, do the authors note the same functional behavior for all “non-linear” compounds in the group of 35, or is the shape of the response curve different?

25) 12339, l. 6-7. Change deferring to different, and suggest including parentheses around ... (for example ... 0.136).

26) 12340, conclusions. A brief mention of current relative costs of a QP vs TOF might be another useful comparison.

27) General. It would have been interesting to see a demonstration of the ability of a TOF to identify new compounds. For example, could the authors show the signal for one of the emerging or low concentration perfluorocarbons with TOF? This would be unlikely to be detected with a QP-MS in scan mode.

28) Table 3. Typo? I calculate 96.961 for the C2H3Cl2 fragment. The listed fragment must be with one Cl-35 and one Cl-37.

29) Table 4. Caption. Delete “used”. Define the error (2 sd?, 3 sd?). Also for Table 5.

30) Figure 1. This is not a particularly revealing diagram. Could more detail be provided about the preconcentration unit, and also show the splitter inside the oven (as described in the text). In the caption change “splitted by a 3-way split” to just “split”.

31) Figure 4. As noted above, I was not crazy about the discussion of mass resolution, and this theoretical plot does little to help. If something like this is to be shown, please use data from your instrument.