

Review of “Organic aerosol composition measurements with advanced offline and in-situ techniques during the CalNex campaign” by Timkovsky et al.

This study presents results from the in situ measurement of aerosol organic compounds by TD-PTR-MS during the CalNex study. The basis for the comparison is GC-GC-MS analysis of hi-
vol filter samples. The new instrument under consideration shows a lot of promise as tool for the measurement of organic aerosol constituents. The manuscript is well written, and the work certainly fits the scope of *AMT*. However, I think that there are some fairly significant issues that need to be addressed before I can recommend the manuscript for publication.

Specific Comments:

1. A significant portion of the manuscript is redundant: the comparison of total OA measured with the TD-PTR-MS and AMS during CalNex was thoroughly discussed in Holzinger et al. (2013). The novel aspect of this study is the comparison of individual compounds measured with the TD-PTR-MS vs. the GC-GC-MS. This should be the focus of the present study.
2. Related to the above comment, there is a lot of discussion about the “123 compounds measured with the GC-GC” and the “64 corresponding masses measured with the TD-PTR-MS.” Because both instruments measure individual compounds, I suggest removing all comparisons where the individual compounds are summed and compared (e.g., Fig. 3; abstract lines 9-11; Section 3.2.1). I think that this presents a misleading representation of the instrument’s capabilities. For some compounds, the agreement is quite good, while for others, the level of agreement seems to be very poor. Presenting these capabilities and limitations is crucial in a methods development paper. For this reason, I strongly suggest expanding the quantitative comparison between the two instruments to many more of the detected compounds beyond the four alkanolic acids that seem to serve as the core of the manuscript (e.g., Fig. 5 and Table 2).
3. The results in Section 3.1 (and Figure 2) demonstrate the *necessity* of characterizing all compounds using the offline TD-PTR-MS technique first, before proceeding with the in situ measurement and comparison. Why is it that acids are the only class of compound analyzed with the offline method? How does this potentially affect the results in Fig. 4? In the off-line TD-PTR-MS, was there any investigation of recovery as a function of compound loading? This also seems critical to demonstrating the method capabilities and determining LODs for all compounds.
4. I think that the idea behind Figure 4 is good, but why is it that ‘average concentrations’ are used for the comparison when individual data points are available? This would be a far more appropriate way to carry out the comparison.
5. Finally, in the comparison of alkanolic acids measured by the two techniques, the differences are almost entirely attributed to ‘filter sampling artifacts’. The TD-PTR-MS measured hexadecanoic acid at ~half the concentration that was measured by the GC-GC, even though hexadecanoic acid is predicted to be entirely in the particle phase. Further, the results of Williams et al. (2010) are also dismissed as a product of sampling artifacts. The feeling one gets from reading the manuscript is that the TD-PTR-MS measurement is the established

'reference' and differences (at least for the alkanolic acids) are explained from the vantage of problems with the other measurements. However, it seems plausible – or even more likely – that the TD-PTR-MS measurements are biased low, due to some combination of sampling/desorption efficiency. In a 'methods/instrumentation' paper, significantly more effort and discussion are needed to address these measurement differences, and I would encourage the authors to start from the position that the more established methods to which they are comparing their new instrument are the 'reference' methods.

Technical Corrections:

1. Pg. 12452, line 11 – CalNex not defined yet.
2. Pg. 12454, line 11-12: Holzinger et al. (2013) describes one of the inlets as having a filter in place? The present results are from the same CalNex deployment?
3. Pg. 12459, line 13: changed 'put' to 'spiked'
4. Table A1 should be in the main paper
5. Pg. 12463, line 22-23: this seems highly speculative (see comment 5 above)
6. Pg. 12467, line 13: I don't understand how the diurnal cycle demonstrates 'consistency of the measurements' (or what this phrase even means)?
7. Pg. 12468, line 16: why is it that thermal decomposition is the explanation here even though other methods use similar thermal desorption temperatures, for example the GC-GC method in this paper desorbs at 320 °C, and do not have such an apparent problem? See Specific Comment 5 above.
8. Pg. 12469, line 7: I disagree with the assessment that "all classes of compounds were detected well"; or perhaps a qualifier should follow that indicates compound classes may have been detected well, but quantification is highly uncertain for some classes.

References:

Holzinger, R., Goldstein, A. H., Hayes, P. L., Jimenez, J. L., and Timkovsky, J.: Chemical evolution of organic aerosol in Los Angeles during the CalNex 2010 study, *Atmos. Chem. Phys.*, 13, 10125–10141, doi:10.5194/acp-13-10125-2013, 2013.

Williams, B. J., Goldstein, A. H., Kreisberg, N. M., and Hering, S. V.: In situ measurements of gas/particle-phase transitions for atmospheric semivolatile organic compounds, *P. Natl. Acad. Sci. USA*, 107, 6676–81, doi:10.1073/pnas.0911858107, 2010.