

The paper “An Improved, Automated Whole-Air Sampler ...” by Brian M. Lerner et al. is a generally well written and structured paper presenting a new, cryogen-free GC-MS system capable of analyzing a large number of VOC compounds in fairly short analysis time and at low detection limits. It has demonstrated successful operation during the 2 flight campaigns SENEX (2013) and SONGNEX (2015) with impressive number of samples analyzed, and comparison results to other VOC online-techniques are presented.

The sampling and analytical systems are well described. Characterization experiments are presented and the DL and precision of the GC-MS system is determined as well as recovery efficiencies and humidification and storage effects of the canisters. The comparisons to PTRMS, H_3O^+ -CIMS, and optical ethane detector are appreciated yielding an impressive data set for characterization of the various techniques. The quality of the VOC measurements by the new instrument is hard to assess for the reader but should generally be in the 15% + x range for high concentrations, and worse for OVOC, heavier aromatics, and at lower concentrations. This is not the state of the art with respect to high quality VOC measurements (see recent intercomparisons and GAW Data Quality Objectives (Plass-Duelmer et al., 2006; Rappenglueck et al., 2006, Hoerger et al., 2015), however, the here presented technique has the focus on short analysis time, many different VOC, and suitability for continental NH conditions. It provides very experienced and excellent concepts and corresponding compromises with respect to data quality must be accepted.

The paper is highly recommended for publication, however, there are some general and specific comments to be considered for revision.

General comments

1. The authors do not, but should present an uncertainty budget covering all uncertainty contributions. This should separately describe the precision contributions and calibration issues of the GC-MS system, and then should set-up an uncertainty budget for the entire measurements covering inlet, canister issues, calibration and GC-MS for the various compounds. This needs to consider blank and bias issues (absolute contributions) as well as precision (reproducibility, relative contributions) related issues. Table 2 currently only lists partial uncertainty contributions and the unexperienced reader might take those numbers as the overall uncertainty. Quantitative estimates are possible based on the presented material, and in the end they characterize the capability of the new system. Furthermore, these are needed to quantitatively interpret the comparison results presented in section 4.
2. This paper lacks some more thorough discussion versus existing “knowledge”. Comparisons with other canister conditioning procedures (e.g. Irvine group, GAW) and canister issues (e.g. GAW and AMOHA intercomparisons), airborne intake systems (e.g. FAAM), calibration procedures (e.g. GAW), etc. Though generally reliable, the choice of methods and materials often appears somewhat arbitrary and could be better justified by/discussed versus corresponding references.
3. Canister issues with OVOCs (alcohols, ketones...) and light alkenes are well known and should be made clearer in the paper. The assumption that OVOC issues are due to the humidification step is part of the story but also, adsorption/desorption artefacts on stainless steel surfaces and in water traps are known to occur. Thus, compounds showing artifacts should be more rigorously identified.

4. It is argued that the measurements of section 4 were not taken at the same time and differences may be due to the variability of VOC compounds. However, the online instruments operated at higher time resolutions than the sampling of canisters, and the scatter in those data could be used to identify the variability. The authors should discuss the deviations considering this information.

Specific Comments

p. 2, l. 8: Please, consider also citation of the recent GAW paper by Schultz et al., 2016.

p. 5: Sampling is via a stainless steel hose, metal bellows pumps and SS manifold. Though the flows are high and corresponding fill times are short (3-15 s), there is sufficient contact time to non-electropolished and badly flushed (in hose) SS surfaces. Sampling artefacts, especially of OVOCs, are known to occur and the authors should verify how these were tested and under which level they are. (It is well recognized, that in the subsequent analysis Sulfinert-treated SS is used)

p. 6, l. 24-27: The impression is raised that only 6 seconds flushing time are used prior to trap loading to passivate surfaces, which is misleading. Thus, it is recommended to point out that an additional flushing time of 210 s is used to achieve equilibrated surfaces.

p. 7, l. 23: Please provide Valco valve details (Rotor material, channel diameter) and operation temperature of valves and lines.

p. 11, l. 11: Please, provide title and Journal of reference Isaacmann-Van Weertz et al. (2016), if not published yet, please provide the main concepts.

p.11, 2nd §: Could the authors provide information on the quality of automated peak integration relative to hand-integration? Ideally, the results obtained by applying both methods on the same data should be compared with respect to scatter in correlation plots and estimated integration uncertainties. Part of the differences in section 4 could be due to the peak integration, this should be discussed.

p. 11, section 2.23: Could you please refer to other canister conditioning procedures, e.g. GAW Guidelines on Canister sampling (GAW Report 204), and discuss deviations from other procedures.

p. 11, l. 27-29: Which type and purity of water is used, is the VOC content analyzed, and how is the water brought into the canisters?

p. 11, ll. 29...: Please make clear what the argument is: when pressurizing ambient air to 3450 hPa, water vapor will condense at relative humidity above some 30% in the canisters anyways, so why change the humidification process? Which problems are expected when there is a condensed water phase in the canister versus a layer of adsorbed water molecules on the inner canister-surface? Were issues with water soluble VOC observed when a condensed water phase can be expected?

p. 12, section 3.1: There are various possible contributions to drifting sensitivity of VOC analysis, comprising (1) MS sensitivity drifts, (2) trapping efficiency changes, (3) mass discrimination and nonlinearity in MS, and (4) active sites on the Al₂O₃ column to partly destroy compounds, etc. The authors only consider (1) and (2). The chosen calibration method is reasonable as it follows drifts run

by run, though it is not usual. Usual is the use of calibration gases and interpolation of the sensitivity changes. No justification for the choice of normalization and interpolation is provided, based on the data. Thus, it is left open whether this is really the best choice. Furthermore, no thorough uncertainty estimate for the applied calibration factors is provided for the various VOC. Please, provide more information on the quality of thus performed calibrations, e.g. by plotting quality charts with time series of the various normalization factors based on each halocarbon analyzed on each column (e.g. three NF for column 1). Further, the calibrations of hydrocarbons could be displayed the same way by plotting their respective ratios to the initial raw peak area. Thus, similar temporal changes are expected given the cause of sensitivity change is (1) or (2). Deviation and scatter should then be discussed with respect to uncertainty of individual calibrations. This could be provided in the Supplemental material and only the result should go into the paper.

p. 13, section sensitivity: Issues have been identified with the linearity of the GC-MS, potentially the linearity of the water trap, and with commercial standards yielding differences of 3-18% which is partly substantially larger than the stated uncertainty of the standards. I guess the stated uncertainty is at 95% confidence, whereas the differences are as observed. Is the message here that the stated uncertainties are wrong or that the instrument yields up to 18% unexplained variation?

The water trap equilibration should affect water soluble and insoluble compounds very differently, and the authors should provide corresponding information.

The calculation method of the total uncertainty of the calibration method of 12% is not transparent. Furthermore, it is not clear at what confidence level this uncertainty is stated. It is amazing that the authors consider all compounds' calibrations equal uncertain though they should show very different behavior. It is recommended to provide calibration uncertainties for the individual compounds.

P. 14, section 3.3: The parameters described in this section appear to be the repeatability and detection limit for very specific conditions. This should not be mixed up with the precision as a measure of the random uncertainty of the analytical system (see comment to Fig. 5 below).

Please provide the mixing ratios of the secondary standard measurements.

The detection limit corresponds to conditions of standard dilution measurements of the analytical system and should be clearly separated from the detection limit in canister measurements, which is further enhanced by blank issues of the canisters for various compounds.

What is missing here is a precision and detection limit for the entire measurements of canisters (see also general comment).

p. 15, l. 9-11 and Fig. 5: Please provide the range of deviation in the form of $\pm (x \text{ ppt} + y\%(\text{mixing ratio}))$ for 95% of the data, because below 10 ppt the deviation is determined by an absolute deviation rather than a relative deviation. Fig.5 and the suggested evaluation of the scatter should be applied to all compounds analyzed on both channels. The deviations do very well characterize the precision and detection limit of the analytical system for real ambient air without sampling and canister issues. A rough inspection of Fig. 5 by eye makes clear, that detection limit and precision stated in Table 2 (1 ppt DL and 3% precision) appear not realistic. The authors are encouraged to provide a realistic estimate of the random uncertainty contributions and prove the consistency with the data.

P. 16, l. 22: It is hard to understand why contamination of the canisters due to the humidification process should have been different for SONGNEX and SENEX. Both needed water which might have been contaminated. Please, provide proof why SENEX is not affected by blanks.

P. 17, l. 17-22: Earlier studies have seen some problems with light alkenes in canisters, also Table 2 indicates 33% standard deviation in ethene retrieval, propene values are not reported. I'd recommend more caution in the statement with regard to light alkenes (also see Plass-Duelmer et al., JGR, 2006).

The intercepts in Fig. 7 should be given with units.

The oxygenated VOC show substantial deviations. Though it is understandable that the authors claim issues with the humidification system, they do not provide evidence that other reasons for these issues can be excluded. Thus, at the current state, the canister technique appears inappropriate for alcohols and many of the ketones. This should be clearly stated.

P. 17, l. 26: The statement "entire system" should be more careful as it does not comprise the step in-situ to canister.

Section 4 (see also general comment 1): Though there is a generally good impression from the comparisons with fairly good correlations, the authors do not provide quantitative estimates of the consistency of the scatter plots with the uncertainties of the analyses. This would be of high importance in order to prove the understanding of the analytical system and the uncertainty budget. Figure 9 displays considerable scatter, especially at low concentrations, and as mentioned above, there should be a description of the uncertainty and scatter as combination of an absolute and a relative quantity. This is important for judging the VOC system capability especially at lower concentrations.

Furthermore, it is hard to understand comparisons based on different scales, as demonstrated for ethane (p. 20, l. 8). Though it is surprising that nowadays ethane calibration still appears to be an issue, this opens the question of how well defined are the scales for all VOC compounds and accordingly the results of the presented measurements. The scientific community is aware of the importance of good calibration methods since various intercomparisons and there have been standardizations in quality assurance like in Global Atmosphere Watch. The authors are encouraged to comment on their calibration with respect to the standards set by, e.g. GAW.

p. 20, 2nd §: Different ratios may be due to different scales or systematic errors in the analytical instruments. As similar compounds are compared in Fig. 11, no differences in the analytical system are expected and this kind of study is appropriate to characterize the scale ratios and the scatter in the data, but less the capability of the analytical system.

Table 2: As stated above, the table should contain absolute and relative information on the repeatability, specify ideal analytical system conditions (DL), list information on compounds from both columns, and discuss some of the unusual data in the text (ethane and propane show highest scatter of alkanes in Rtv Eff though they are expected to be most stable, several alkanes are significantly changing over time based on Replicate and standard deviations (e.g. propane 0.93 with s.d. of 0.01), ethene indicates large scatter in Rtv eff and what about other light alkenes, aldehydes should be strictly characterized as canister artifacts, how is a negative value for acetone understandable (Rtv Eff)...).

Another Table should sum the overall precision (random uncertainty contributions with specified confidence) and calibration uncertainty (systematic, 95 % confidence) for canister sampling and analysis by the presented system for all VOC compounds.

Figure 10a: Benzene appears to go down to zero for the CIMS instrument whereas for the iWAS not. Benzene at 0 ppt seems unrealistic for continental air over US and points to a bias problem of the CIMS. This also underlines the importance of considering the uncertainties of the compared methods.

Figure 11: the symbol colors are hardly discernable (grey and light blue) and should be changed.

References:

Hoerger, C. C., Claude, A., Plass-Duelmer, C., Reimann, S., Eckart, E., Steinbrecher, R., Aalto, J., Arduini, J., Bonnaire, N., Cape, J. N., Colomb, A., Connolly, R., Diskova, J., Dumitrescu, P., Ehlers, C., Gros, V., Hakola, H., Hill, M., Hopkins, J. R., Jäger, J., Junek, R., Kajos, M. K., Klemp, D., Leuchner, M., Lewis, A. C., Locoge, N., Maione, M., Martin, D., Michl, K., Nemitz, E., O'Doherty, S., Pérez Ballesta, P., Ruuskanen, T. M., Sauvage, S., Schmidbauer, N., Spain, T. G., Straube, E., Vana, M., Vollmer, M. K., Wegener, R., and Wenger, A.: ACTRIS non-methane hydrocarbon intercomparison experiment in Europe to support WMO GAW and EMEP observation networks, *Atmos. Meas. Tech.*, 8, 2715-2736, doi:10.5194/amt-8-2715-2015, 2015.

Schultz MG, Akimoto H, Bottenheim J, Buchmann B, Galbally IE, Gilge S, Helmig D, Koide H, Lewis AC, Novelli PC, Plass-Dülmer C, Ryerson TB, Steinbacher M, Steinbrecher R, Tarasova O, Tørseth K, Thouret V, Sabatier P, Zellweger C : The Global Atmosphere Watch reactive gases measurement network. *Elementa: Science of the Anthropocene* 3: 000067. doi: 10.12952/journal.elementa.000067

C. Plass-Dülmer, N. Schmidbauer, J. Slemr, F. Slemr, and H.D'Souza, European hydrocarbon intercomparison experiment AMOHA part 4: Canister sampling of ambient air, *J. Geophys. Res.*, 111, D04306, doi:10.1029/2005JD006351, 2006.

B. Rappenglück, E. Apel, M. Bauerfeind, J. Bottenheim, P. Brickell, P. Cavolka, J. Cech, L. Gatti, H. Hakola, J. Honzak, R. Junek, D. Martin, C. Noone, C. Plass-Dülmer, D. Travers, and D. Wang, The first VOC intercomparison exercise within the Global Atmosphere Watch (GAW), *Atmos. Environ.* 40, 7508-7527, 2006.

GAW Report 204, WMO, Geneva,

http://www.wmo.int/pages/prog/arep/gaw/documents/FINAL_GAW_204.pdf