Supplement of Atmos. Meas. Tech., 10, 291–313, 2017
http://www.atmos-meas-tech.net/10/291/2017/
doi:10.5194/amt-10-291-2017-supplement

Atmospheric Measurement Techniques
Open Access
EGU

Supplement of

An improved, automated whole air sampler and gas chromatography mass spectrometry analysis system for volatile organic compounds in the atmosphere

Brian M. Lerner et al.

Correspondence to: Brian M. Lerner (blerner@aerodyne.com)

Ambient sampling humidity during the SENEX and SONGNEX campaigns.

For canister samples collected from the NOAA WP-3D, samples collected in the planetary boundary layer [PBL] are of primary importance to the scientific goals of the SENEX and SONGNEX campaigns. The range of water mixing ratios encountered in the PBL during the SENEX and SONGNEX field campaigns were significantly different (Figure S1). For simplicity, we conservatively define the mixing height of the PBL as 1500m for all samples (Peischl et al., 2015). During SENEX, the majority (≈85%) of canister samples were collected in the summertime PBL, primarily in the southeastern United States. The ambient water mixing ratio of >99% of PBL samples was 9 ppth water or higher, so that pressurizing canisters to 3500 hPa would produce condensed water in the canister at room temperature. For the SONGNEX campaign, ≈92% of samples were collected in the PBL throughout the United States in springtime, with the majority in the western U.S. The ambient water mixing ratio of 68% of PBL samples was 5 ppth or less, so that condensed water would not be expected in these canisters at room temperature without the additional water vapor added during canister cleaning.

Description of TERN peak integration software

With the increasing size of datasets collected by our GC-MS instruments, both in terms of number of chromatograms and number of species reported, we recently developed a new software tool named TERN for the handling of chromatogram data files and peak integration. TERN extracts single ion chromatogram for peak fitting of individual compound peaks, although peak full mass spectra are retained for assessment of the fits. Chromatographic data is reduced to peak areas by fitting a segment of detector signal with multiple Gaussian and/or exponentially-modified Gaussian [EMG] functions. The use of EMG fits for chromatographic data has been well-described in the literature and is the preferred function used by TERN (Anderson et al., 1970; Di Marco and Bombi, 2001; Phillips and White, 1997). Gaussian fits are used only in cases where satisfactory EMG fits are not possible, typically in cases of multiple overlapping peaks of similar size as the tailing or fronting of the peaks cannot be well-constrained. Further discussion of the multi-peak fitting functions are available at WaveMetrics website (https://www.wavemetrics.com/products/igorpro/dataanalysis/peakanalysis/multipeakfitting.htm).

The peak fitting method follows this work-flow:

1. Extract a 40-second window of single ion chromatogram data to analyze, centered on expected retention time (RT) of target compound
2. Trim the chromatogram data
    a. Automatically reduce the size of the window due to:
        i. missing signal if the ion was not scanned at the edges of the window
        ii. peaks at the edges of the window that do not have a peak maximum inside the window
    b. Manually allow the user to specify a smaller data window size if desired

3. (*Optional*) Identify constant baseline value via smoothing and low-pass filter, and subtract from data
4. Search for peaks (local maxima) that can be fit with a Gaussian function
5. Eliminate peaks too far from expected RT or too small, relative to largest peak. Note that the criteria here can be specified by the user; default values are:
    a. time from expected RT = 5 sec
    b. minimum relative peak area = 5%
6. Find a mathematical solution for the chromatogram data
    a. Solve for linear or constant baseline if step 3 (above) was skipped
    b. Run preliminary fits using Gaussian function for all peaks
    c. If Gaussian fit fails, repeat steps a. and b. with increased boxcar smoothing of the data signal. If Gaussian fit succeeds, continue to step d.
    d. Attempt an improved EMG fit based upon Gaussian parameters from b.
7. Identify peak closest to expected RT ("winner").
8. Confirm "winner" mass spectrum fragmentation pattern matches expected, based upon linear fit of relative mass responses

Peak fitting using TERN requires a tight match between the expected and actual retention time of the analyte peak. This can be difficult to achieve in the field environment, where instrument stability can be challenged by environmental conditions, e.g. diurnal changes of room temperature of up to 10°C. To reduce this potential timing mismatch, we fit peaks in an iterative manner, where a subset of peaks (typically 5 to 10 per channel) easily identified by retention time and ion mass are automatically fit. A retention time shift algorithm is then performed for each channel, in which a polynomial fit of expected versus observed retention time for the subset of peaks is used to calculate new expected retention times for all compounds to be analyzed. These new expected retention times are used for a second pass of the data where the remaining peaks of interest are fit. The time-shift algorithm typically results in a match between expected and observed retention times within one second.

To demonstrate the quality of the returned peak areas from TERN, we have reprocessed a data set from the UBWOS 2012 (section 4.2) field campaign that had been previously analyzed by manual integration (Warneke et al., 2014). The returned raw peak areas from TERN are approximately one-tenth of manually integrated peaks using Agilent Chemstation (Version E.02.02); since all raw data are in arbitrary units and are subsequently normalized before calculating mixing ratios, this offset is irrelevant. For all compounds for which we have produced intercomparisons (Table S1). Example scatterplots for four of these intercomparisons are shown in Figure S2. The goodness of the orthogonal linear fit is excellent ($r^2 > 0.998$) between the manual and automated peak area integrations; for compounds with larger dynamic ranges during the field campaign the fits are better, typically with $r^2 > 0.999$ for compounds spanning more than three decades.

References

Anderson, A. H., Gibb, T. C., and Littlewood, A. B.: Computer Resolution of Unresolved Convoluted Gas-Chromatographic Peaks, Journal of Chromatographic Science, 8, 640-646, doi: 10.1093/chromsci/8.11.640, 1970.

Di Marco, V. B., and Bombi, G. G.: Mathematical functions for the representation of chromatographic peaks, Journal of Chromatography A, 931, 1-30, doi: http://dx.doi.org/10.1016/S0021-9673(01)01136-0, 2001.

Peischl, J., Ryerson, T. B., Aikin, K. C., de Gouw, J. A., Gilman, J. B., Holloway, J. S., Lerner, B. M., Nadkarni, R., Neuman, J. A., Nowak, J. B., Trainer, M., Warneke, C., and Parrish, D. D.: Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions, Journal of Geophysical Research: Atmospheres, 120, 2119-2139, doi: 10.1002/2014JD022697, 2015.

Phillips, M. L., and White, R. L.: Dependence of Chromatogram Peak Areas Obtained by Curve-Fitting on the Choice of Peak Shape Function, Journal of Chromatographic Science, 35, 75-81, doi: 10.1093/chromsci/35.2.75, 1997.

Warneke, C., Geiger, F., Edwards, P. M., Dube, W., Pétron, G., Kofler, J., Zahn, A., Brown, S. S., Graus, M., Gilman, J. B., Lerner, B. M., Peischl, J., Ryerson, T. B., de Gouw, J. A., and Roberts, J. M.: Volatile organic compound emissions from the oil and natural gas industry in the Uintah Basin, Utah: oil and gas well pad emissions compared to ambient air composition, Atmos. Chem. Phys., 14, 10977-10988, doi: 10.5194/acp-14-10977-2014, 2014.

Table S1. A comparison of peak areas integrated via Agilent Chemstation and TERN for select compounds measured during UBWOS 2012. The dynamic range of peaks areas found by TERN are represented by minimum and maximum values. The fit coefficients (1σ uncertainties) are based upon orthogonal least squares linear fits, while the coefficients of determination (r^2) are based upon a linear regression.

5

	kcounts						
Compound	**Min**	**Max**	**Slope**		**Intercept**		**r^2**
Ethyne	3.9	66.3	0.0977	(0.00007)	-0.05	(0.01)	0.9999
Propane	11.9	7410	0.1016	(0.00005)	-8.6	(0.5)	0.9998
n-Butane	14.4	14996	0.1018	(0.00009)	-22.2	(1.5)	0.9992
n-Hexane	4.0	8985	0.0988	(0.00005)	-0.30	(0.32)	0.9997
n-Heptane	1.5	6626	0.0969	(0.00003)	1.1	(0.13)	0.9999
n-Octane	0.7	4939	0.0950	(0.00002)	-0.54	(0.05)	1.0000
Benzene	14.5	1840	0.0951	(0.00003)	-1.1	(0.1)	0.9998
Toluene	2.9	4777	0.0946	(0.00005)	-5.7	(0.3)	0.9997
0-Xylene	0.2	613	0.0931	(0.00004)	-0.3	(0.02)	0.9998
Methylcyclohexane	2.5	7757	0.9960	(0.00002)	-0.14	(0.15)	0.9999
Methanol	11.5	8496	0.0972	(0.00005)	-6.8	(0.3)	0.9997
Acetone	9.4	87.2	0.0959	(0.00008)	-0.6	(0.04)	0.999
Methyl ethyl ketone	0.8	21.8	0.0927	(0.00016)	-0.05	(0.01)	0.9981

Table S2. Individual results for comparison of simultaneously filled canister samples of ambient air. The canister samples were aged 1, 2 or 4 days before analysis. Bolded values are significantly different (at 95% confidence) than at least one other test result for that compound.

		2 day v 1 day		4 day v 1 day		4 day v 2 day	
Compound	**Channel**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
Ethane	1	0.94	0.09	1.14	0.10	1.06	0.01
Propane	1	0.99	0.22	1.30	0.30	1.08	0.01
i-Butane	1	n/a	n/a	n/a	n/a	n/a	n/a
n-Butane	1	1.19	0.13	1.21	0.10	1.04	0.02
i-Pentane	1	0.92	0.06	1.00	0.04	0.97	0.03
n-Pentane	1	0.91	0.06	0.96	0.05	1.03	0.02
n-Hexane	1	0.93	0.08	0.94	0.07	0.93	0.05
n-Hexane	2	0.84	0.09	0.90	0.11	1.08	0.03
n-Heptane	2	0.89	0.11	0.90	0.14	1.04	0.10
n-Octane	2	0.80	0.09	0.75	0.12	0.92	0.08
n-Nonane	2	0.73	0.08	0.60	0.15	0.93	0.06
Ethene	1	0.98	0.09	1.09	0.10	1.06	0.06
Isoprene	1	0.82	0.14	0.77	0.09	0.92	0.05
α-Pinene	2	0.74	0.06	0.72	0.07	0.91	0.06
β-Pinene	2	0.72	0.13	0.78	0.20	0.93	0.10
Ethyne	1	1.07	0.11	1.12	0.07	0.98	0.02
Methylcyclopentane	2	0.82	0.12	0.87	0.11	1.11	0.07
Cyclohexane	2	0.85	0.05	0.93	0.07	**1.09**	**0.06**
Methylcyclohexane	2	0.86	0.09	0.88	0.08	1.04	0.09
Benzene	2	0.89	0.03	0.88	0.03	0.99	0.03
Toluene	2	0.80	0.06	0.82	0.07	0.94	0.03
Ethylbenzene	2	0.69	0.09	0.59	0.12	**0.92**	**0.03**
m,p-Xylenes	2	0.70	0.06	0.63	0.10	0.89	0.04
o-Xylene	2	0.65	0.10	0.66	0.11	0.89	0.05
Nitrate, ethyl	2	1.4	2.1	-0.3	1.6	1.03	0.38
Nitrate, i-propyl	2	1.1	1.2	0.2	1.3	1.01	0.09
Nitrate, n-propyl	2	0.65	0.92	0.3	1.2	1.14	0.42
Methanol	2	-1.0	1.8	16	482	0.98	0.31
Ethanol	2	2.7	5.9	1.4	2.0	1.02	0.06
Acetone	2	1.3	1.7	1.1	1.1	0.84	0.82
Methyl ethyl ketone	2	2.3	3.6	1.4	2.6	1.29	0.49
Methyl vinyl ketone	2	1.4	1.8	1.1	1.6	0.54	0.50
Acetaldehyde	2	0.73	0.74	2.1	2.8	2.5	2.0
Propanal	2	2.7	4.3	1.5	3.1	0.56	0.64
Methacrolein	2	1.30	0.85	1.00	0.72	0.79	0.26

Table S3. Individual results for comparison of canister samples vs. co-sampled ambient air. The canister samples were aged 1, 2 or 4 days before analysis. Bolded values are significantly different (at 95% confidence) than at least one other test result for that compound.

		1 day v amb.		2 day v amb.		4 day v amb.	
Compound	**Channel**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
Ethane	1	1.08	0.16	1.08	0.05	1.12	0.05
Propane	1	1.09	0.20	1.04	0.05	1.09	0.05
i-Butane	1	n/a	n/a	n/a	n/a	n/a	n/a
n-Butane	1	**0.77**	**0.04**	1.12	0.02	1.13	0.02
i-Pentane	1	1.02	0.06	0.94	0.06	0.87	0.05
n-Pentane	1	0.97	0.05	1.02	0.03	1.04	0.03
n-Hexane	1	0.96	0.07	0.99	0.05	0.86	0.06
n-Hexane	2	0.97	0.14	0.88	0.08	0.95	0.09
n-Heptane	2	1.19	0.57	1.03	0.38	1.00	0.39
n-Octane	2	1.7	1.8	1.00	0.47	0.92	0.43
n-Nonane	2	1.2	2.3	1.1	1.4	0.9	1.1
Ethene	1	1.22	0.60	1.06	0.31	1.14	0.34
Isoprene	1	1.12	0.33	1.01	0.23	0.85	0.19
α-Pinene	2	1.5	2.5	0.88	0.83	0.78	0.75
β-Pinene	2	1.5	10.6	0.9	4.6	0.8	4.1
Ethyne	1	1.11	0.40	0.93	0.18	1.05	0.21
Methylcyclopentane	2	1.15	0.40	1.01	0.27	1.10	0.29
Cyclohexane	2	1.15	0.66	1.07	0.52	1.16	0.57
Methylcyclohexane	2	1.17	0.72	0.98	0.54	0.95	0.49
Benzene	2	1.10	0.09	0.94	0.07	0.91	0.07
Toluene	2	0.89	0.03	0.91	0.02	**0.81**	**0.02**
Ethylbenzene	2	1.20	0.38	0.84	0.13	0.73	0.11
m,p-Xylenes	2	**0.91**	**0.08**	0.76	0.05	0.61	0.04
o-Xylene	2	1.02	0.24	0.69	0.12	0.62	0.10
Nitrate, ethyl	2	2	106	0.5	2.0	0.6	2.0
Nitrate, i-propyl	2	2	26	0.86	0.83	0.86	0.84
Nitrate, n-propyl	2	2	140	0.9	7.4	0.9	6.7
Methanol	2	1.37	0.44	1.83	0.52	1.55	0.65
Ethanol	2	1.85	0.29	1.19	0.07	1.26	0.09
Acetone	2	-0.9	1.4	0.6	2.1	-0.7	2.1
Methyl ethyl ketone	2	1.8	2.1	0.91	0.31	1.19	0.38
Methyl vinyl ketone	2	5	24	4	11	-10	43
Acetaldehyde	2	3.0	4.0	22	81	20	52
Propanal	2	10	40	-58	580	19	83
Methacrolein	2	2	11	1.3	2.3	1.2	2.7

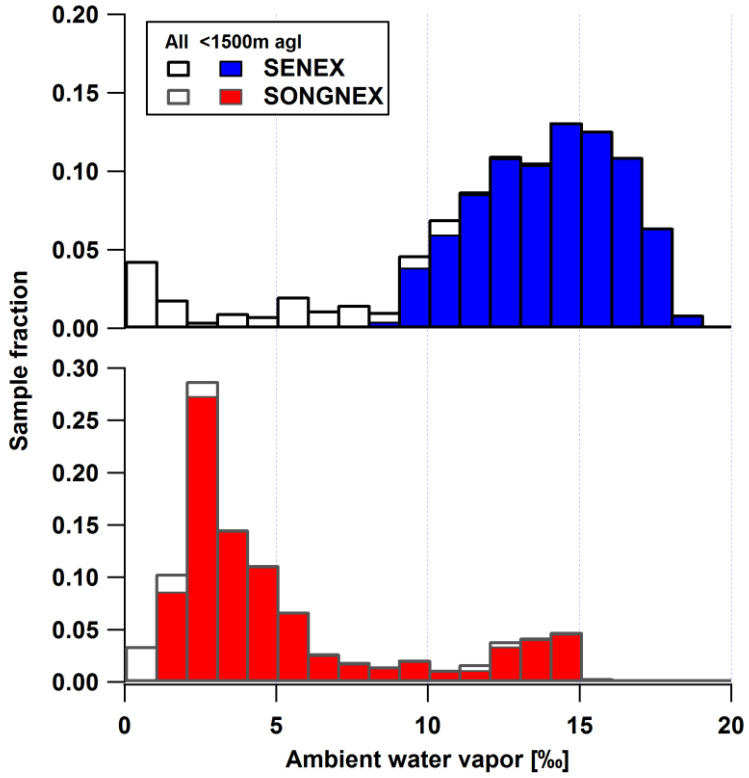

Figure S1. Histograms of ambient water vapor for all WAS canisters filled during the SENEX and SONGNEX field campaigns. The open bars represent all samples, while the colored bars are the subset of samples collected below 1500m above ground level (agl).

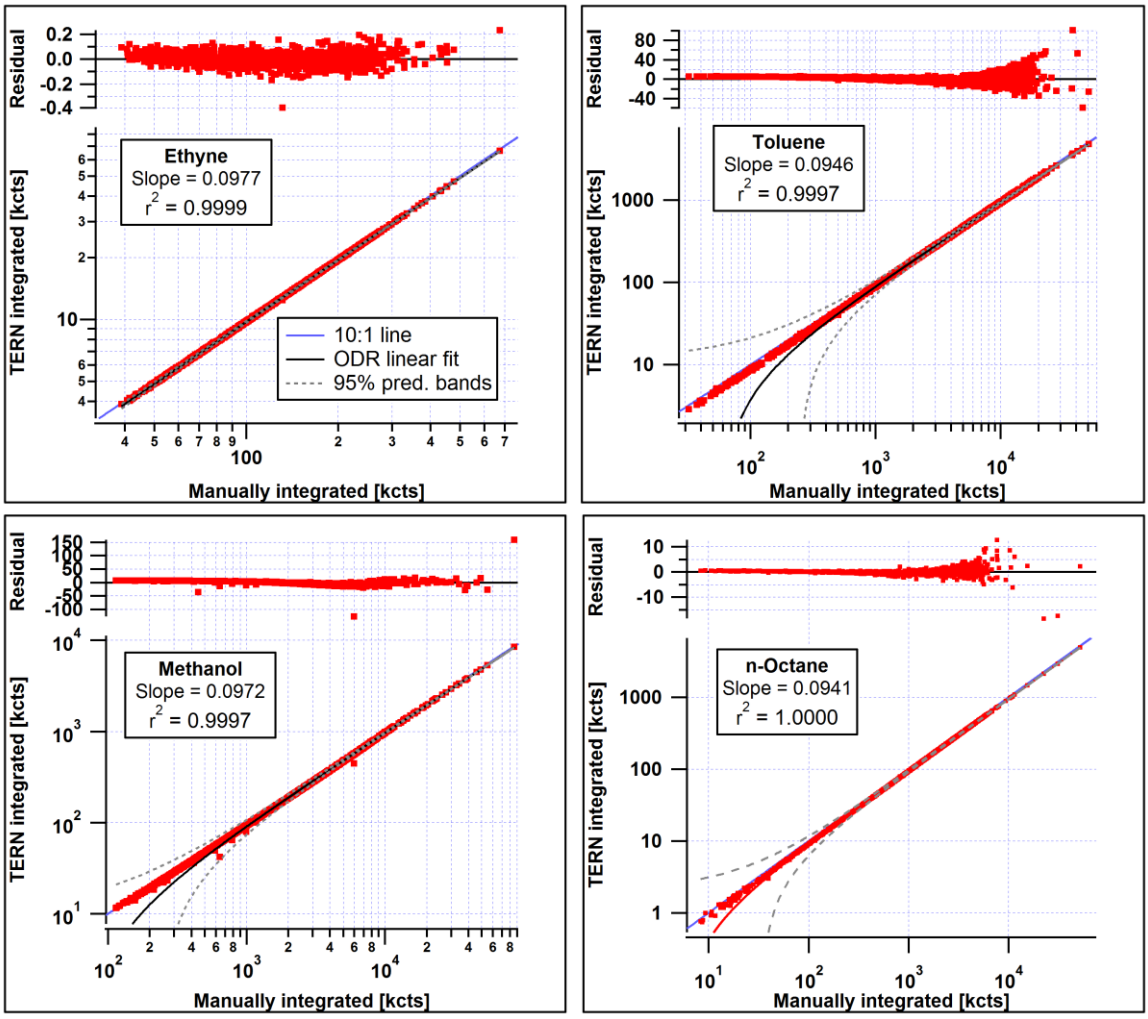

Figure S2. Example scatter plots showing comparison of single-ion peak fit areas to traditional manual baseline integrated peak areas for select compounds measured during UBWOS 2012.

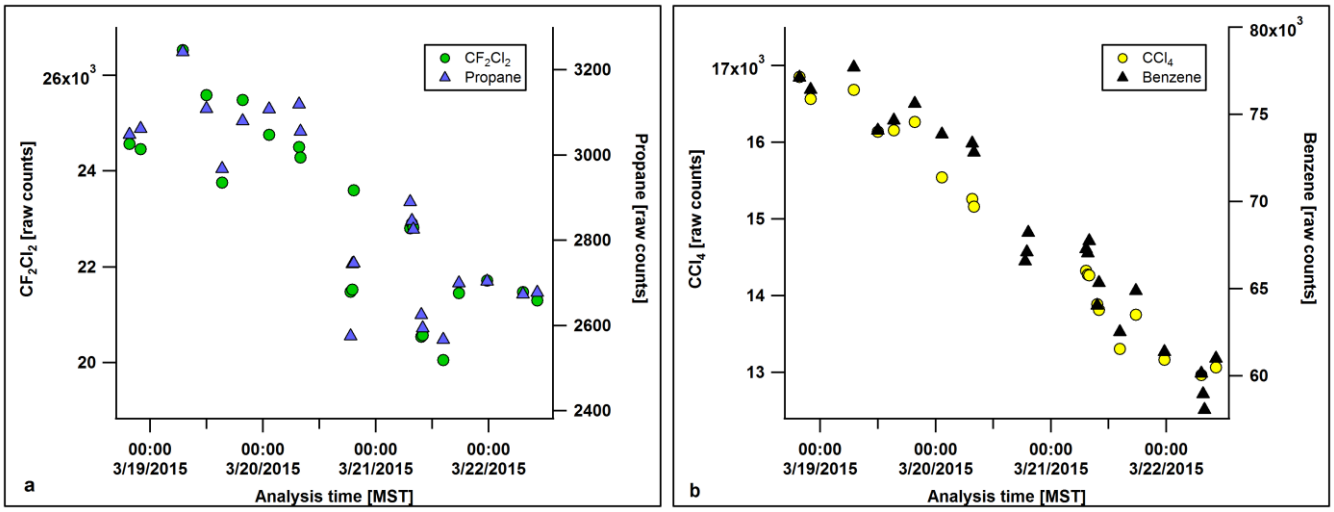

Figure S3. Time series showing change in measured raw counts for multiple analysis of a secondary standard over 3.5 days. Panel a) shows changes in response for two species from channel #1, CF_2Cl_2 and propane; panel b) shows changes in response for two species from channel #2, CCl_4 and benzene.

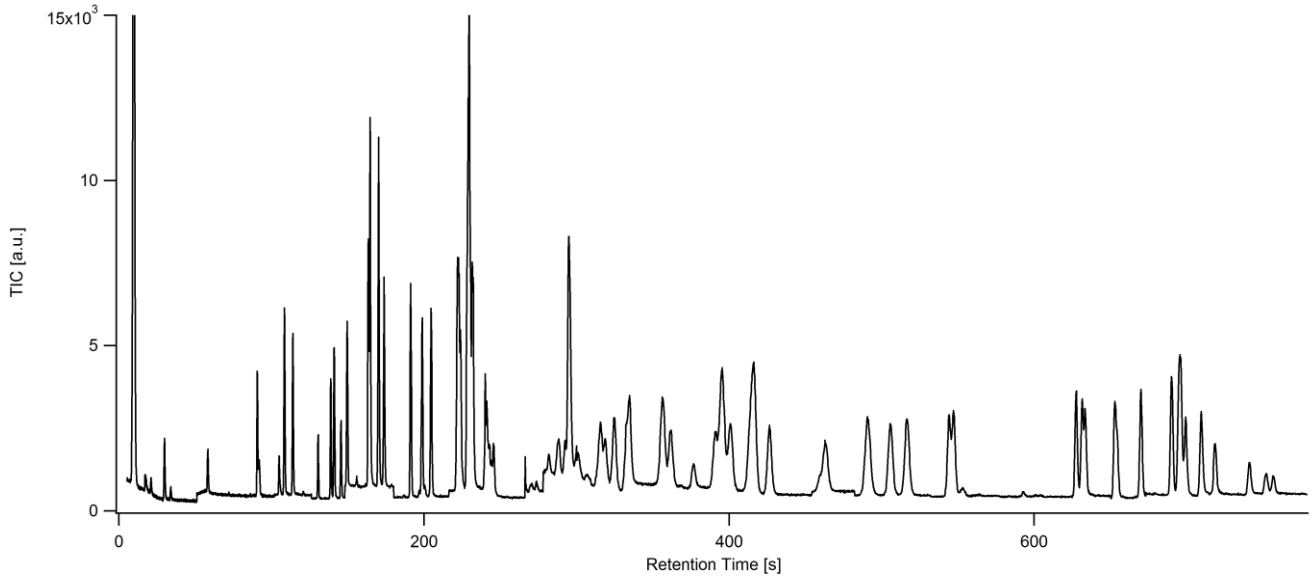

Figure S4. Example TIC chromatogram of 57-component PAMS standard diluted to 26 pptv in humidified UHP N2, the lowest of seven calibration levels tested prior to the SONGNEX field campaign. Replicates of measurements of the lowest calibration point are used as a basis for calculating limits of detection. Note that the chromatogram scaling is 1/10 of the chromatogram shown in Figure 4.