Supplemental Information for Development of a Multichannel Organics *In situ* enviRonmental Analyzer (MOIRA) for mobile measurements of volatile organic compounds

Audrey J. Dang[1], Nathan M. Kreisberg[2], Tyler L. Cargill[1], Jhao-Hong Chen[1], Sydney Hornitschek[1], Remy Hutheesing[1], Jay R. Turner[1], Brent J. Williams[1]

[1]Energy, Environmental, and Chemical Engineering, Washington University in St. Louis, St. Louis, 63130, United States
[2]Aerosol Dynamics, Inc., Berkeley, 94710, United States.

Correspondence to: Brent J. Williams (brentw@wustl.edu)

S1. Instrument design

S1.1 Vacuum chamber and ceramic interface design

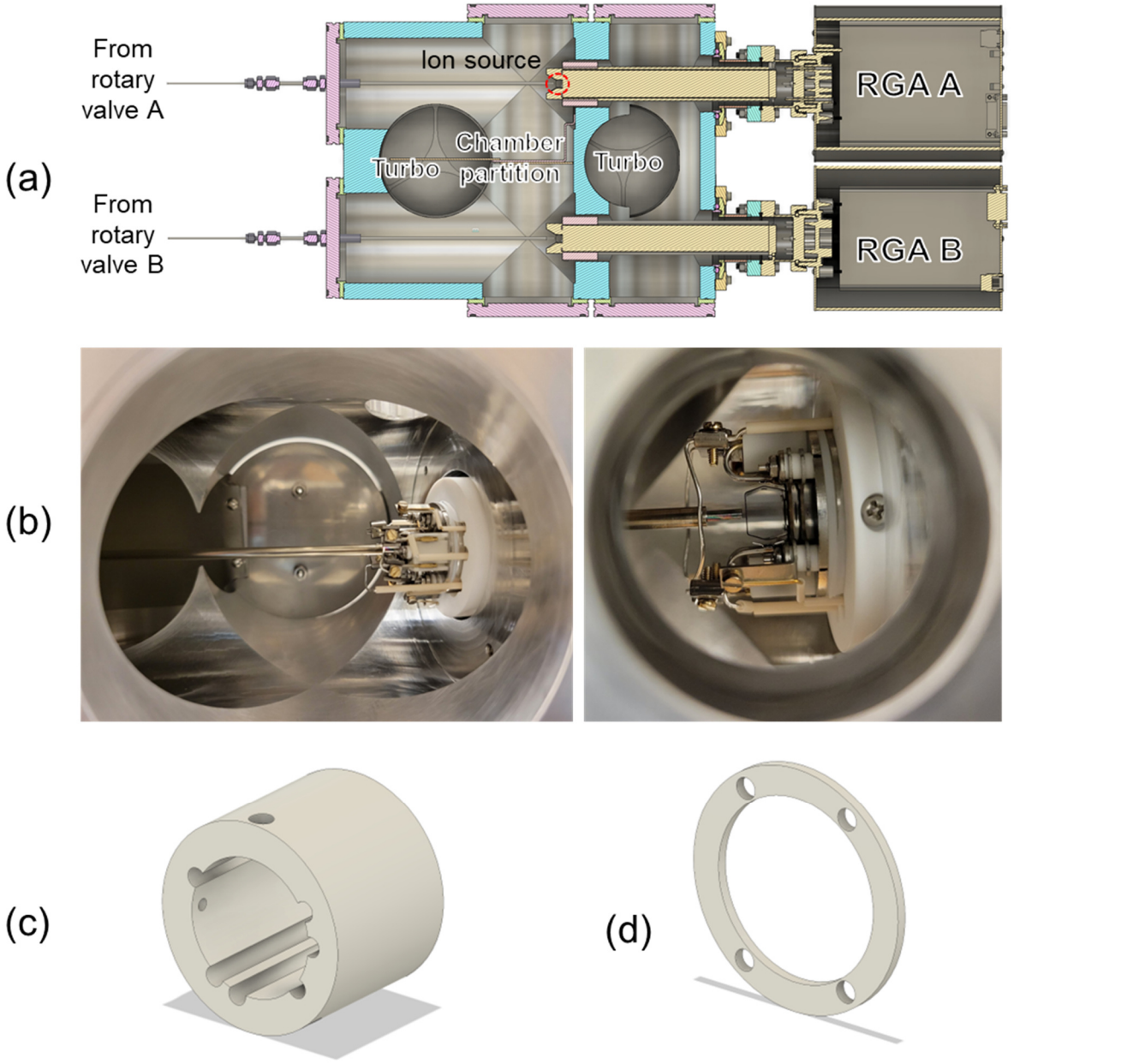


Figure S1. (a) Cross-section of vacuum chamber, RGAs, and transfer line from above. (b) Photographs of the interior of the vacuum chamber, RGA source, and ceramic interface from the side. (c) CAD rendering of the ceramic sheath. (d) CAD rendering of the ceramic washer, which is inserted into the RGA source.

S1.2 Heater Control

The custom LabVIEW program staggers the powering of the 26 heaters, whose duty cycles are proportional to the number of time steps (50 time steps of 20 ms each) for which the heater is powered on during each 1 s control period. In order of greatest to least power (as is typical in bin packing algorithms), the heaters are assigned time steps during which they will be powered on such that the maximum instantaneous total power for all heaters across the 50 time steps is minimized. The algorithm accounts for delays in turning off the heater due to solid state relay delay times.

If the user specifies a maximum instantaneous total heater power, then the LabVIEW program will limit the heating of low priority heated zones if necessary to achieve that threshold. For example, when the instrument is cold after maintenance, power requirements to heat the large number of temperature zones at once to their setpoints can be very large. The maximum power setting is a safeguard to controlling this initial burst of power consumption, which might otherwise exceed the maximum of a typical circuit. In addition, during measurements, the greatest power consumption occurs during the initial 45 seconds of thermal desorption, and heating of less critical zones can be limited in this short period. After the LabVIEW PID controllers calculate the heater duty cycles for the next 1 s control period, the algorithm calculates total heater power for each time step, fractionally decreasing the calculated duty cycles according to user-specified priority until the calculated total heater power is below the specified threshold during all time steps. For example, the rotary valves and GCs have high priority, as operating them below their temperature setpoints could cause component damage or data quality issues, respectively. In contrast, small, temporary decreases in transfer line temperatures are unlikely to cause instrument damage or data quality issues, though very large decreases could degrade sample transfer. To avoid integral wind-up effects and maintain stable control, these decreases in the PID output must be accompanied by a correction to the LabVIEW PID block's "integral error buffer," which is used in the discretized form of the PID algorithm to calculate the integral contribution to the output. In cases where the PID output is modified for power consumption, the integral term is approximated as the difference between the actual, modified output and the proportional term.

To detect anomalies which could indicate an instrument fault, the measured and predicted temperatures are compared for each zone, and a rule set determines whether a heater should be automatically shut-off. The proportionality and time constants of each heated zone are used to predict the temperature response to the duty cycle time series. The temperature prediction is calculated with an exact recursive solution (Eq. (S1)) to Eq. (S2) (Seborg et al., 2010):

$$T(t) = e^{\frac{\Delta t}{\tau_C}} T(k-1) + K\left(1 - e^{\frac{\Delta t}{\tau_H}}\right) u(k-1) \tag{S1}$$

$$\frac{dT(t)}{dt} + \frac{T(t)}{\tau_C} = \frac{Ku(t)}{\tau_H} \tag{S2}$$

where T is temperature, t is time, τ_C is the cooling time constant, τ_H is the heating time constant, u is the duty cycle, and K is the proportional constant. The constants K, τ_C, and τ_H are estimated from the temperature response to a step change in duty cycle. If the predicted temperature is above a certain threshold, or if the predicted increase in temperature deviates from the measured increase in temperature, the heater is shut-off. This prevents heater runaway in cases of inaccurate temperature

measurement. For example, this automated fault detection has prevented instrument damage when a thermocouple fell out of its slot in a heater block and when an analog voltage amplifier malfunctioned.

S2 Shadow peaks

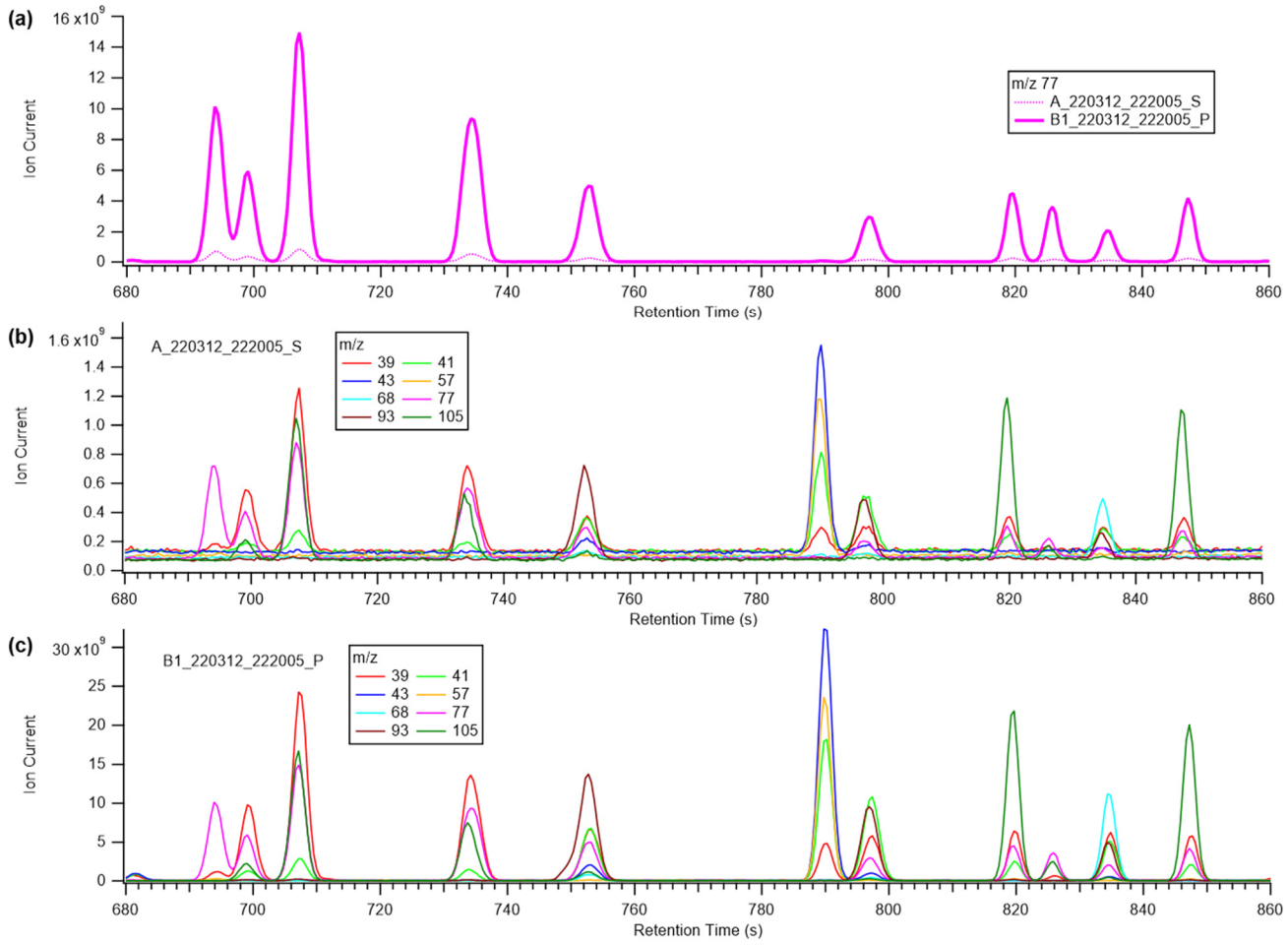


Figure S2. When a calibration standard was analyzed on the prime channel B1, a shadow signal of lesser abundance was simultaneously detected by RGA A: (a) Overlay of m/z 77 for both RGAs (b) Shadow ion chromatograms for multiple m/z values (c) Prime ion chromatograms for multiple m/z values

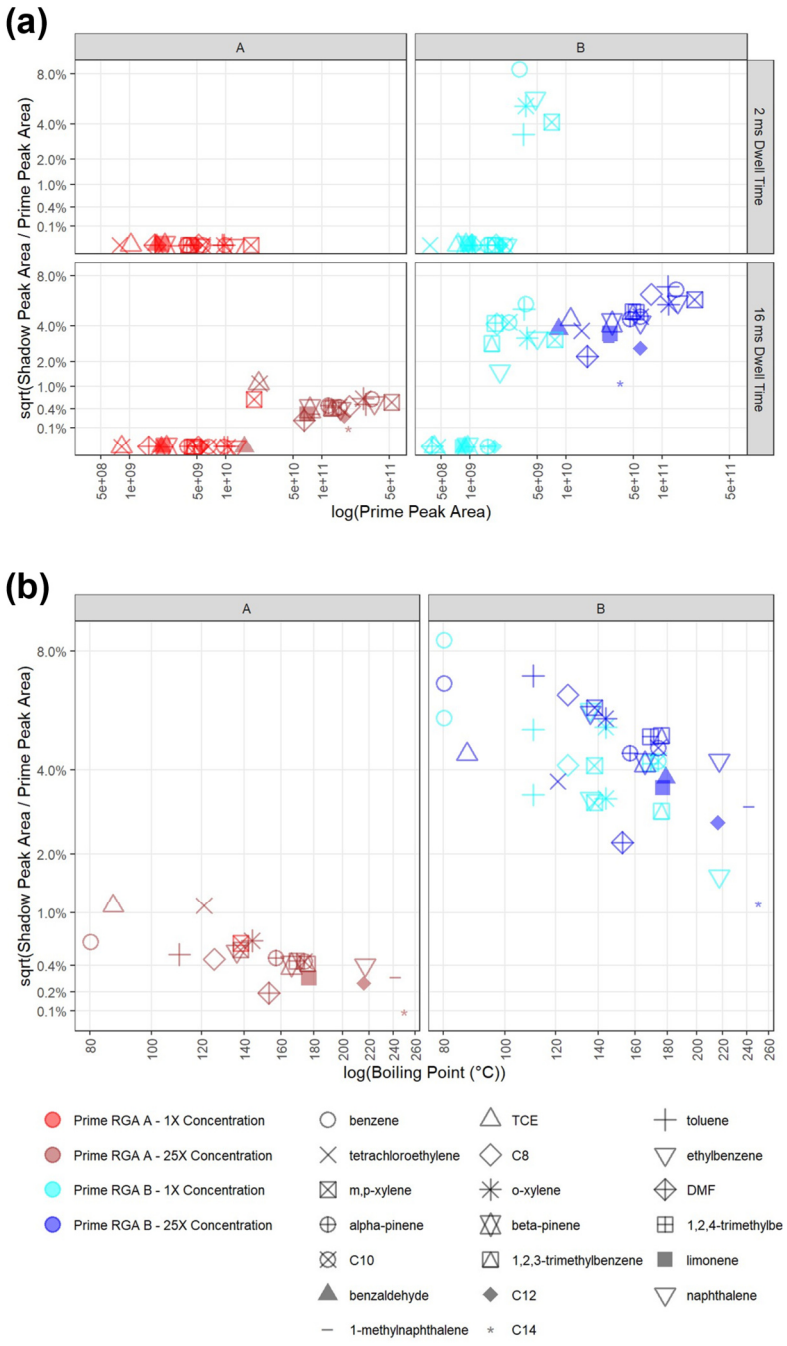


Figure S3. (a) When the signal of an eluting peak at the prime RGA is sufficiently high, a shadow peak is simultaneously detected at the other RGA. (b) Compounds of higher volatility have higher shadow peak areas as a proportion of the prime peak area.

60

S3 Relative Quadrupole Transmission Efficiency

The ion transmission efficiency of the Pfeiffer PrismaPro's quadrupole mass filter is dependent on mass-to-charge ratio, with lower efficiency at higher mass-to-charge ratios. Ng (2011) characterized this dependence by placing an effusive source of naphthalene within the vacuum chamber of the Aerosol Chemical Speciation Monitor (ACSM), which used an earlier model of the RGA used by MOIRA.(Ng et al., 2011) The mass spectrum of naphthalene measured by the ACSM's RGA was compared with the mass spectrum from the NIST library (as well as with measurements by higher performance instruments) to determine a correction for the relative ion transmission of the quadrupole.

In this study, a similar approach is applied to the measured mass spectra of multiple compounds from chromatograms of calibration standards. Representing different functionalities, the chosen compounds have mass spectra with contributions from a wide range of mass-to-charge values: decane (0.69 ppbv), dodecane (0.57 ppbv), tetradecane (0.49 ppbv), m,p-xylene (3.68 ppbv), naphthalene (0.76 ppbv), 1-methylnaphthalene (0.69 ppbv), alpha-pinene (0.72 ppbv), beta-pinene (0.72 ppbv), limonene (0.72 ppbv), and trichloroethylene (TCE, 0.74 ppb). In these calibration chromatograms, these compounds do not have interference from co-eluting species.

To obtain the mass spectrum of each peak, the average of background scans at the base of each side of the peak is subtracted from the average of the three scans closest to the peak maximum:

$$A_i = \frac{1}{3}\left(A_i|_{t_r-\Delta t} + A_i|_{t_r} + A_i|_{t_r+\Delta t}\right) - \frac{1}{2}\left(A_i|_{t_r-2\sigma} + A_i|_{t_r+2\sigma}\right) \quad \text{(S3)}$$

where A_i is the abundance of mass-to-charge i in the mass spectrum, t_r is the retention time of the scan closest to the fitted peak maximum, Δt is the time resolution of the scan, and σ is the standard deviation of the fitted peak. Both t_r and σ are determined by fitting each individual peak with the TERN software package (version 2.2.18, Igor Pro 8) (Isaacman-Vanwertz et al., 2017).

The library comparison ratio R_i for a given ion i is defined as the ratio of the normalized measured abundance and library abundance:

$$R_i = \frac{\left(\frac{A_i}{A_b \cdot 10^{C_j}}\right)}{A_{i,library}} \quad \text{(S4)}$$

where $A_{i,library}$ is the abundance of mass-to-charge i in the library mass spectrum (which is itself normalized to the most abundant ion b in the library mass spectrum), A_b is the abundance of ion b in the MOIRA mass spectrum, and C_j is the compound factor for compound j. The compound factor C_j represents a compound-specific adjustment in the normalization of the mass spectrum and results in a vertical offset in the log plot of the library comparison ratio with respect to mass-to-charge ratio (Fig. S4). For example, in the case where $C_j = 0$, the measured mass spectrum is normalized to ion b, which is most abundant in the library spectrum for that compound.

If there is a perfect match between the library and measured mass spectra, then $\log(R_i) = 0$ for all ions. While other phenomena dependent on mass-to-charge could also contribute to deviations from a perfect match, a single correction is

applied here and named the relative quadrupole transmission efficiency correction. First, a transmission efficiency function $\log(T(i))$ for each RGA is fitted to the $\log(R_i)$ values for the calibration data with respect to mass-to-charge i, iterating to determine C_j for each compound which minimizes the sum of the residuals for that compound (Fig. S4):

$$\log_{10}(\mathrm{T_i}) = \mathrm{m} * \mathrm{i} + \mathrm{d} \quad \text{(S5)}$$

where m is the slope of the regression line, and d is the intercept. Data points are included in the fit if there is high abundance in both the measured mass spectrum ($\mathrm{A_i}$ exceeding the absolute abundance of a blank signal by approximately 25-fold) and the library mass spectrum ($A_{i,library} > 0.05$)). The included mass-to-charge ratios range is 20 to 129, above which there were few abundant ions in the mass spectra of the selected compounds. At mass-to-charge values greater than 129, $T(i)$ is defined as a constant value:

$$T(i) = 10^{m*i+d} \; for \; 20 \leq i \leq 129 \quad \text{(S6)}$$

$$T(i) = 10^{m*129+d} \; for \; i > 129 \quad \text{(S7)}$$

Finally, the mass spectrum can be corrected by dividing by T_i normalized to the transmission efficiency of the most abundant ion b (T_b), for which $A_{b,corrected} = A_b$.

$$A_{i,corrected} = \frac{A_i}{\left(\frac{T(i)}{T(b)}\right)} \quad \text{(S8)}$$

The correction was quite stable over a four month period, which included a retuning of both RGAs (Fig. S5). Compared to the correction used for a similar RGA in Ng (2011), there is more correction for smaller m/z fragments and less correction for higher m/z fragments.

The correction was evaluated by determining its impact on several metrics (Fig. S6, Fig. S7). The cosine similarity, Pearson correlation coefficient, and Spearman coefficient generally improved or remained constant for the selected compounds, especially those that were poorer before the correction. A smaller impact was observed for the match score used by the NIST library search, which is itself a weighted sum of two factors (Stein, 1994). (Given the thresholds for both library and measured abundance for inclusion in these calculations, the NIST score calculated here is a modified version of the "reverse match score," which would include only ions in the library mass spectrum.) The first contributing factor is similar to cosine similarity and has similar trends as that metric. The second factor describes the similarity between library and measured ratios of ion pairs which are nearest neighbors in mass-to-charge (among those with $A_{i,library} > 0.05$ included in this analysis). Since the magnitude of the correction would be similar for ions of similar mass-to-charge value, this second factor (and the overall Reverse Match Score) were less impacted than other metrics. Thus, while this correction improves the similarity of measured and library mass spectrum, the uncorrected mass spectrum would likely yield similar match scores and library compound matches as the corrected mass spectrum.

It is interesting that compounds of the same class have similar trends for the residuals of the relative quadrupole transmission efficiency with respect to mass-to-charge (Fig. S4). For example, ions of lower mass-to-charge are more abundant in MOIRA's measurements of alkanes and less abundant among the monoterpenes (Fig. S4). Ionization energy also affects the library comparison ratio as well as the magnitude of the ion current. In the indoor measurements in this

study, the ionization energy for RGA A was mistakenly set to 80 eV. While library mass spectra are typically acquired at 70 eV, the library comparison ratios are similar from 60 to 80 eV (Fig. S8). Thus, while calibration data used for quantification would need to be acquired at the same ionization energy used for a measurement, the same relative transmission efficiency correction would be appropriate for both 70 and 80 eV.

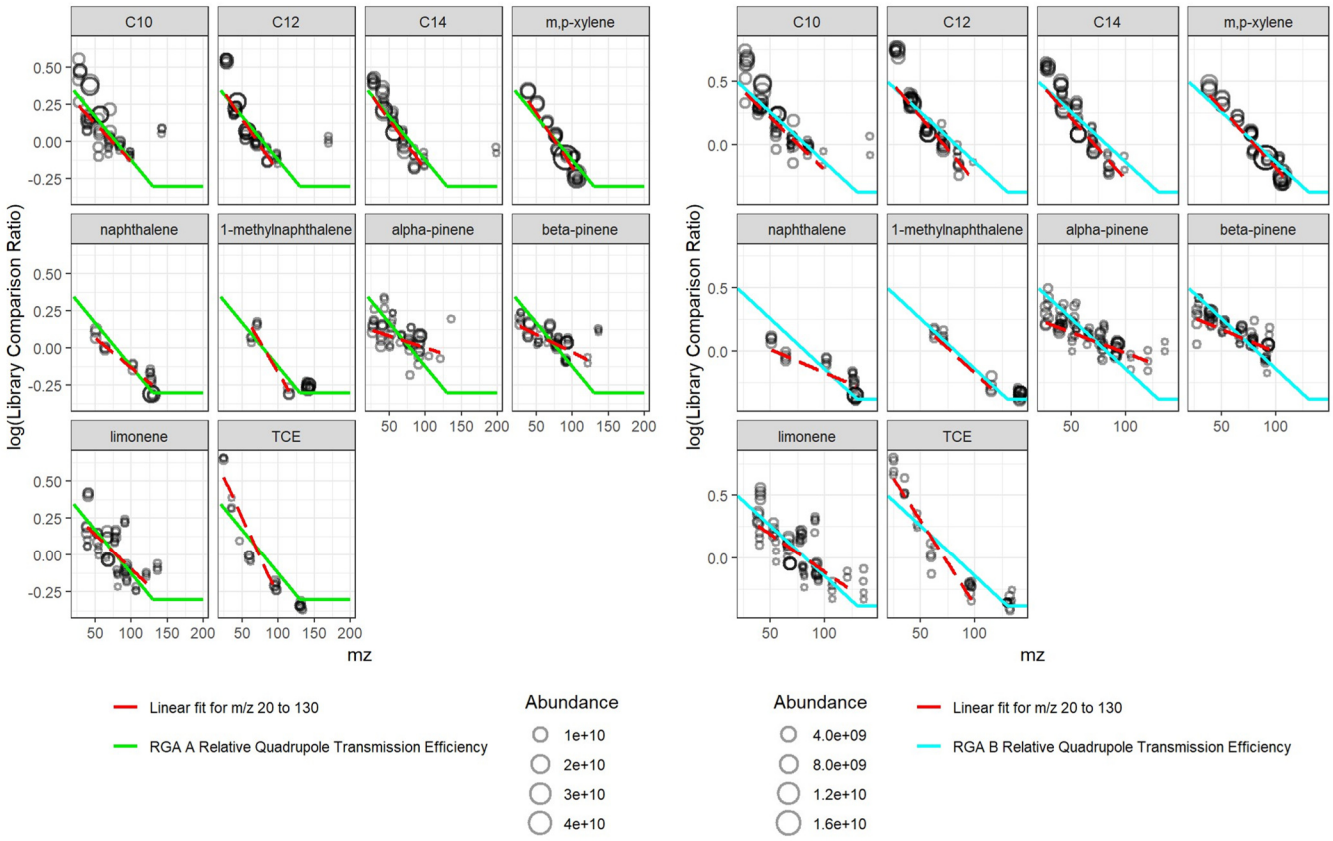
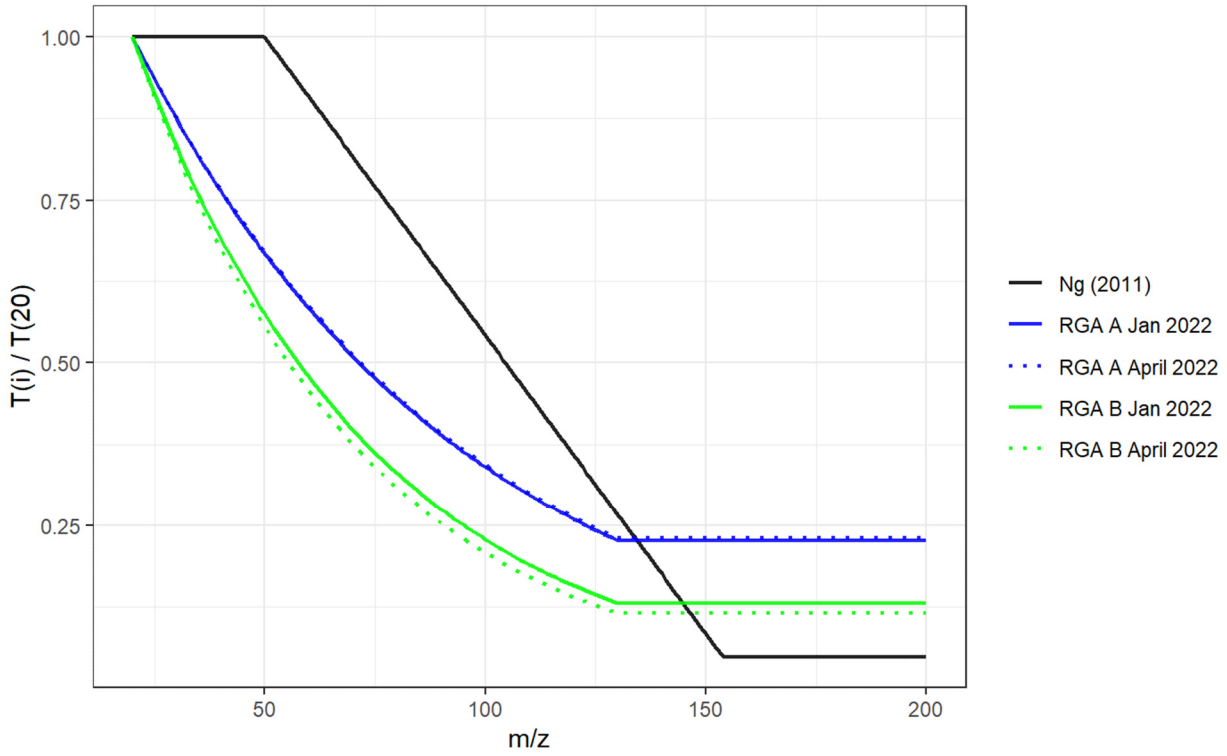


Figure S4. The relative quadrupole transmission efficiency was fitted separately for each RGA's quadrupole. The compound factor represents a vertical shift for the data for each compound and is iteratively determined to minimize each compound's sum of residuals for the RGA-specific fit (green for RGA A, blue for RGA B).



135

Figure S5. Comparison of m/z-dependent transmission efficiency correction for RGA A and RGA B during January and April 2022 with that for a similar RGA used in Ng (2011).

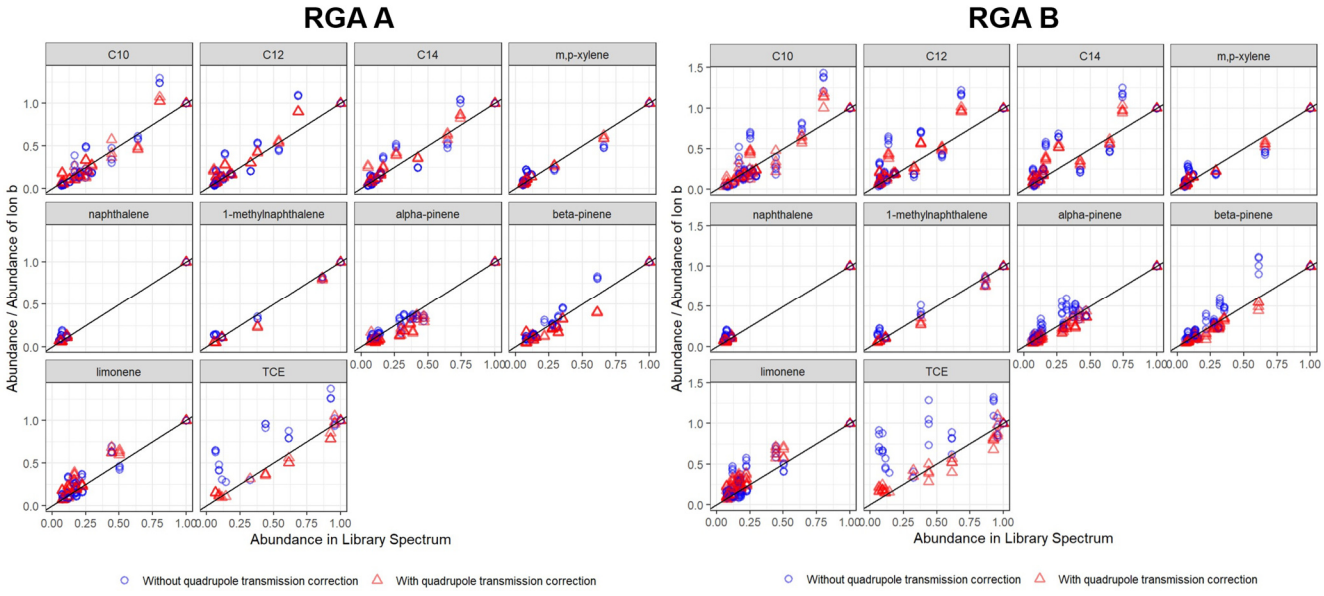


Figure S6. Correcting for quadrupole transmission efficiency increases similarity to the library spectrum.

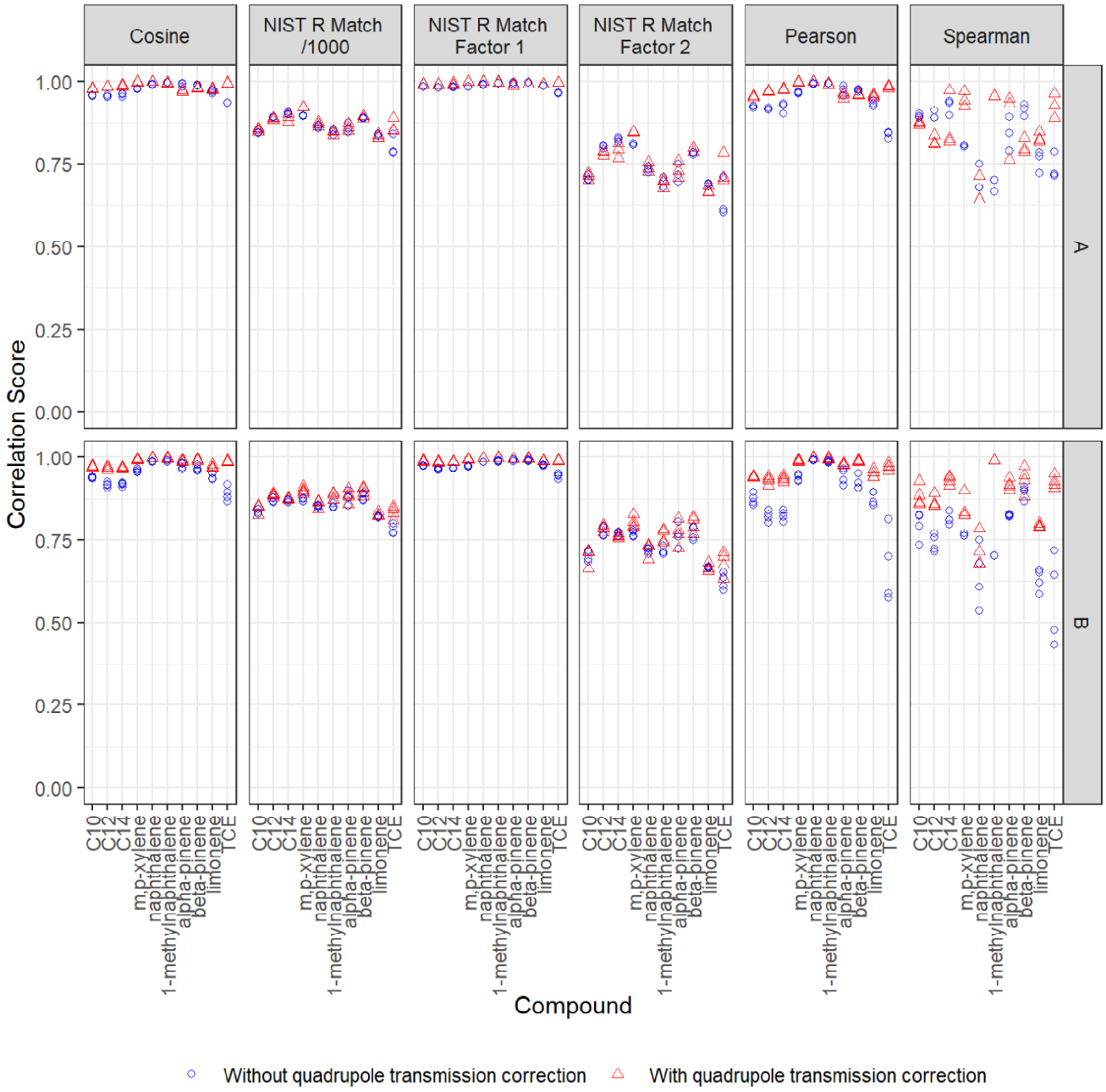


Figure S7. Correcting for quadrupole transmission efficiency increases similarity to the library spectrum for most compounds and metrics.

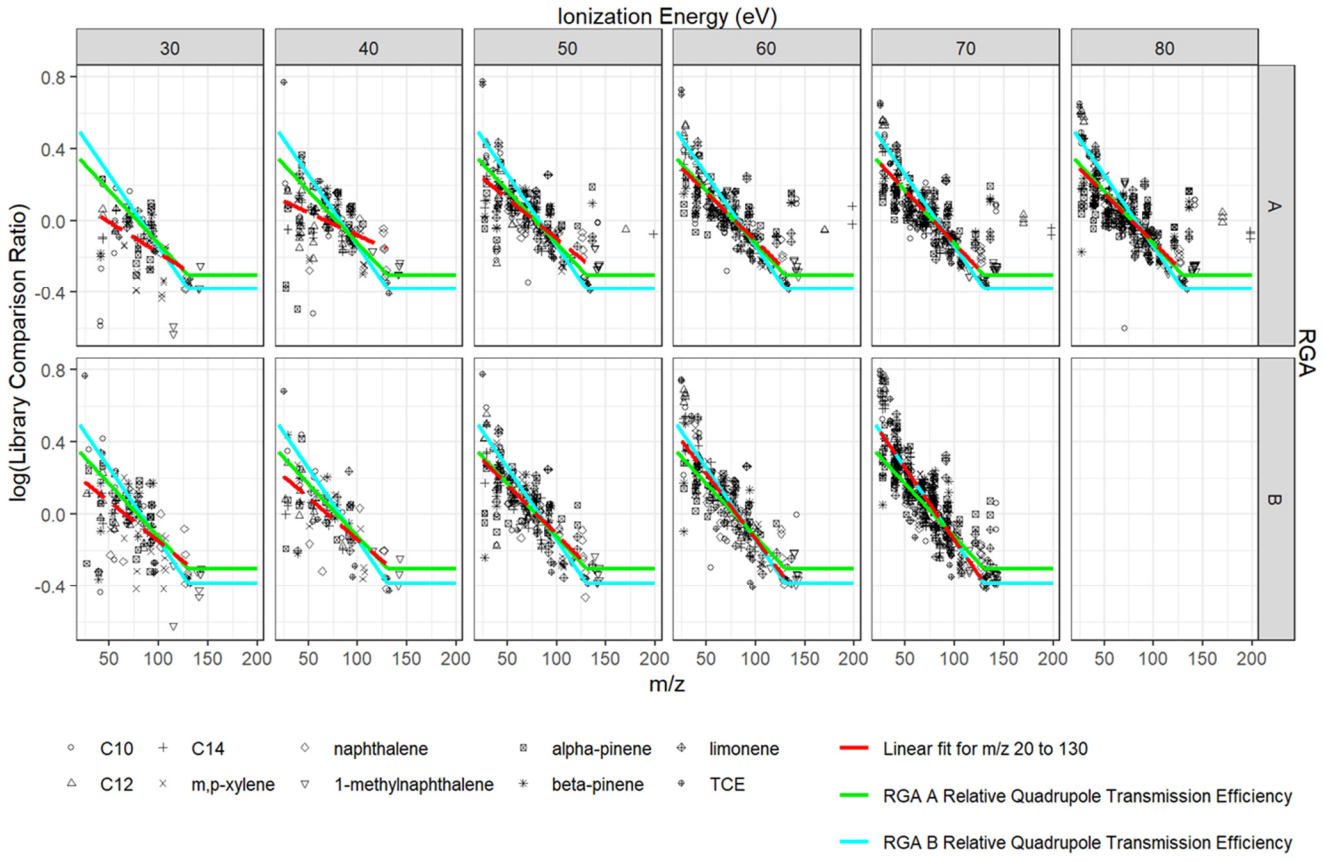


Figure S8. In addition to quadrupole transmission efficiency, ionization energy also affects the similarity of the library and measured mass spectrum, though a single correction addresses both.

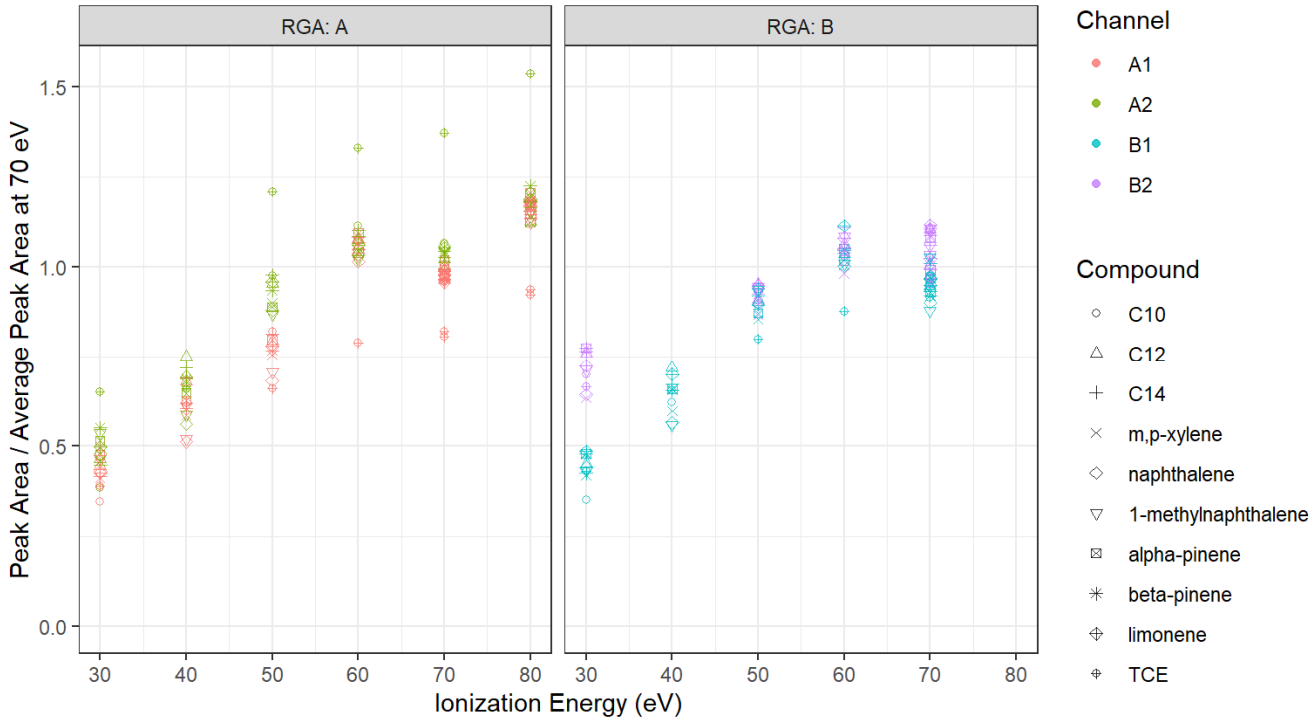


Figure S9. Ion current increases with ionization energy.

S4 Selective ion monitoring (SIM) methods

Table S1. SIM method used for calibration curve experiment and for mobile measurements by RGA B

m/z	Dwell Time (ms)	
	Calibration Curve	RGA B Mobile Measurements
27	64	16
28	2	2
31	n/a	16
32	2	2
40	2	2
41	n/a	16
42	n/a	16
43	16	16
44	16	16
45	16	16
55	16	16
56	n/a	16
57	16	16
58	16	16
68	64	64
69	16	16
70	16	16
78	64	64
83	n/a	4
91	64	64
93	16	16
105	16	16
128	16	16
142	16	16
Total Scan Time (ms)	518	552

S5 Chromatogram characteristics and quantification

Table S2. Lifetimes for reaction with ozone

	τ_{ozone} (hr)		τ_{ozone} (hr)
C5	$1 \times 10^{+7}$	styrene	7
C6	$1 \times 10^{+7}$	alpha - pinene	1
C8	$1 \times 10^{+7}$	beta - pinene	8
C10	$1 \times 10^{+7}$	limonene	1
C12	$1 \times 10^{+7}$	butanal	$1 \times 10^{+4}$
C14	$1 \times 10^{+7}$	pentanal	$1 \times 10^{+4}$
C15	$1 \times 10^{+7}$	hexanal	$1 \times 10^{+4}$
benzene	$1 \times 10^{+4}$	heptanal	$1 \times 10^{+4}$
toluene	$1 \times 10^{+4}$	octanal	$1 \times 10^{+4}$
ethylbenzene	$1 \times 10^{+4}$	nonanal	$1 \times 10^{+4}$
m,p-xylene	$1 \times 10^{+4}$	decanal	$1 \times 10^{+4}$
o-xylene	$1 \times 10^{+4}$	2-butenal	$7 \times 10^{+1}$
1,2,4-trimethylbenzene	$1 \times 10^{+4}$	MEK	$1 \times 10^{+4}$
1,2,3-trimethylbenzene	$1 \times 10^{+4}$	4-methyl-2-pentanone	$1 \times 10^{+4}$
naphthalene	$5 \times 10^{+2}$	2-hexanone	$1 \times 10^{+4}$
1-methylnaphthalene	$8 \times 10^{+2}$	cyclohexanone	$1 \times 10^{+4}$
isoprene	9	MTBE	$1 \times 10^{+4}$

Table S3. List of detected compounds with molecular weight (MW), retention time (RT), CAS or NIST registry number, boiling point (BP), quantification ion (QI), most abundant ions (AI), identification confidence category (C)[a], NIST database search scores (match score (M), reverse match score (R), probability score (P), rank score (N)), and datasets (D) in which the compound has been identified[b]

		MW (g)	RT (s)	CAS or NIST #	BP (°C)	QI	AI	C [a]	M	R	P (%)	N	D [b]
(1) Alkanes													
2-methylbutane	C_5H_{12}	72	192	78-78-4	28*	43	43;42;41;29;57	B	885	935	71.5	1	I
C5	C_5H_{12}	72	209	109-66-0	36*	43	43;42;41;27;39	A	493	803	16.1	1	C I M
2-methylpentane	C_6H_{14}	86	273	107-83-5	61*	42	43;42;41;27;71	B	907	933	69.4	1	I
3-methylpentane	C_6H_{14}	86	290	96-14-0	63†	57	57;56;41;29;27	C	764	897	39.7	1	I M
C6	C_6H_{14}	86	306	110-54-3	69*	57	41;43;57;29;27	A	521	887	15.6	1	C I M
2-methylhexane	C_7H_{16}	100	405	591-76-4	90†	85	43;41;42;57;85	B	828	926	66.2	1	I
3-methylhexane	C_7H_{16}	100	422	589-34-4	92‡	70	43;41;29;71;70	C	784	861	29.1	1	I
C8 - unknown branched	C_8H_{18}	114	446	540-84-1	99†	56	57;56;41;43;29	E	864	874	21	1	I
C7	C_7H_{16}	100	465	142-82-5	98†	71	43;29;41;57;71	C	761	902	60.2	1	I M
2,3,4-trimethylpentane	C_8H_{18}	114	546	565-75-3	113†	70	43;71;70;41;27	C	783	896	22.6	1	I
2,3,3-trimethylpentane	C_8H1_8	114	556	560-21-4	115†	70	43;70;71;41;55	C	733	833	19.8	1	I
3-methylheptane	C_8H_{18}	114	571	589-81-1	119†	56	56;43;57;41;29	D	636	833	25	1	I
C8	C_8H_{18}	114	598	111-65-9	126*	43	43;29;41;57;27	A	605	892	31.8	1	C I M
C9	C_9H_{20}	128	704	111-84-2	151†	43	43;57;29;41;27	D	620	822	14.5	1	I M
C10	$C_{10}H_{22}$	142	787	124-18-5	174*	43	43;57;41;71;29	A	833	902	24.5	1	C I M
C11	$C_{11}H_{24}$	156	865	1120-21-4	195*	43	43;57;41;29;71	D	624	833	17.3	1	I
C12	$C_{12}H_{26}$	170	939	112-40-3	216*	57	57;43;41;71;29	A	851	906	23.6	1	C I M
unknown branched alkane			952			57	57;41;43;71;29	E					I
6-ethylundecane	$C_{13}H_{28}$	184	996	17312-60-6	222§	57	57;43;71;41;56	C	795	837	11.1	2	I
C13	$C_{13}H_{28}$	184	1017	629-50-5	235†	57	43;57;41;71;29	B	835	868	6.34	1	I M
unknown branched alkane			1078			71	57;71;43;41;85	E					I

		MW (g)	RT (s)	CAS or NIST #	BP (°C)	QI	AI	C [a]	M	R	P (%)	N	D [b]
C14	$C_{14}H_{30}$	198	1093	629-59-4	250*	57	43;57;41;71;29	A	851	906	23.6	1	C I M
C15	$C_{15}H_{32}$	212	1178	629-62-9	267*	57	43;57;41;71;29	A	888	899	8.43	2	C I
methylcyclopentane	C_6H_{12}	84	354	96-37-7	72†	56	56;41;69;42;39	B	819	923	44.2	1	I
(2) Aromatics													
2-methylfuran	C_5H_6O	82	344	534-22-5	63‡	53	53;27;82;39;43	D	541	816	22.2	1	I
benzene	C_6H_6	78	451	71-43-2	80*	78	78;52;51;50; 77	A	887	938	67	1	C I M
toluene	C_7H_8	92	606	108-88-3	111*	91	91;92;39;65;63	A	920	922	55.6	1	C I M
ethylbenzene	C_8H_{10}	106	699	100-41-4	136*	91	91;106;51;39; 65	A	946	954	60	1	C I M
m-xylene and p-xylene (co-eluting)	C_8H_{10}	106	708	108-38-3 / 106-42-3	139 / 138	91	91;106;105; 39;51	A	960 / 921	965 / 925	59.1 / 12.4	1 / 3	C I M
o-xylene	C_8H_{10}	106	734	95-47-6	144*	91	91;106;51;105; 39	A	855	883	19.6	2	C I M
styrene	C8H8	104	739	100-42-5	146*	104	104;51;78;103; 77	A	929	932	59.8	1	C
1-ethyl-3-methylbenzene	C_9H_{12}	120	790	620-14-4	161‡	105	105;120;77;39; 57	C	805	882	28.8	1	I M
1,3,5-trimethylbenzene	C_9H_{12}	120	801	108-67-8	165†	105	105;120;119; 32;77	D	659	832	37.1	1	M
ethyltoluene	C_9H_{12}	120	815	108-67-8	165†	105	105;120;77; 91;51	E	598	825	14.3	1	M
1,2,4-trimethylbenzene	C_9H_{12}	120	818	95-63-6	169*	105	105;120;39; 77;27	A	871	914	16.3	2	C I M

		MW (g)	RT (s)	CAS or NIST #	BP (°C)	QI	AI	C [a]	M	R	P (%)	N	D [b]
1,2,3-trimethylbenzene	C_9H_{12}	120	846	526-73-8	176*	105	105;120;39; 77;27	A	853	902	13.8	2	C I M
acetophenone	C_8H_8O	120	914	98-86-2	202†	105	77;105;51; 120;43	D	599	829	7.51	2	
unknown substituted naphthalene	$C_{12}H_{22}$	166	964	1008-80-6	224‡	81	41;81;55;95;67	C	757	809	16.4	1	I
naphthalene	$C_{10}H_8$	128	1000	91-20-3	217*	128	128;51;64; 127;63	A	955	957	56.3	1	C I M
1-methylnaphthalene	$C_{11}H_{10}$	142	1107	90-12-0	242*	142	142;141;115; 71;63	A	935	941	39.3	1	C I
unknown aromatic	$C_{12}H_{10}$	154	1185	92-52-4	255‡	154	154;153;152; 76;155	E	638	813	21.4	1	M
(3) Selected biogenic markers (including terpenes)													
isoprene	C_5H_8	68	232	78-79-5	34*	68	67;39;53;68;27	A	846	920	10	3	C I M
alpha-pinene	$C_{10}H_{16}$	136	754	80-56-8	157*	93	93;92;41;91;39	A	897	909	6.39	5	C I
camphene	$C_{10}H_{16}$	136	772	79-92-5	160‡	93	93;121;39; 41;79	C	777	905	10.5	1	I
myrcene	$C_{10}H_{16}$	136	792	123-35-3	167‡	93	41;69;93;39;27	B	886	901	47.5	1	I
beta-pinene	$C_{10}H_{16}$	136	797	127-91-3	166	93	93;41;69;27;28	A	813	907	27.3	1	C I
limonene	$C_{10}H_{16}$	136	830	138-86-3	177*	68	68;67;93;39;41	A	907	916	46.2	1	C I
eucalyptol	$C_{10}H_{18}O$	154	848	470-82-6	176‡	43	43;71;105;81; 108	D	539	697	13.2	1	I
(4) Aldehydes													
acrolein	C_3H_4O	56	234	107-02-8	53‡	27	27;26;28;56;29	A	720	898	89.5	1	C I
isobutyraldehyde	C_4H_8O	72	313	78-84-2	63‡	72	43;41;27;29;72	D	650	810	54.1	1	I

		MW (g)	RT (s)	CAS or NIST #	BP (°C)	QI	AI	C [a]	M	R	P (%)	N	D [b]
butanal	C_4H_8O	72	361	123-72-8	75*	72	27;29;44;43;41	A	858	931	89.5	1	C I
pentanal	$C_5H_{10}O$	86	536	110-62-3	103*	44	44;29;27;41;58	A	873	910	83.8	1	C I M
hexanal	$C_6H_{12}O$	100	657	66-25-1	129*	44	44;41;29;56;27	A	914	927	83.2	1	C I M
furfural	$C_5H_4O_2$	96	718	98-01-1	162†	39	39;96;95;29;38	B	943	947	76.7	2	I
heptanal	$C_7H_{14}O$	114	752	111-71-7	154*	70	44;29;43;41;27	A	906	916	86	1	C I M
benzaldehyde	C_7H_6O	106	828	100-52-7	179*	77	77;106;105;51; 50	A	892	902	61.1	1	C I M
octanal	$C_8H_{16}O$	128	834	124-13-0	171*	110	29;43;41;44;27	A	910	917	79	1	C I M
nonanal	$C_9H_{18}O$	142	912	124-19-6	185*	57	41;29;57;43;44	A	913	919	77.4	1	C I M
decanal	$C_{10}H_{20}O$	156	990	112-31-2	209*	82	41;29;43;57;27	A	892	896	49.1	1	C I M
cinnemaldehyde	C_9H_8O	132	1096	14371-10-9	254‡	77	131;51;77;132; 103	B	857	894	35	1	I
(5) Unsaturated aldehydes													
2-butenal	C_4H_6O	70	488	4170-30-3	102*	70	41;39;70;29;69	C	773	886	36.4	1	I
2-pentenal	C_5H_8O	84	624	1576-87-0	126‡	83	55;29;27;39;41	C	717	847	10.8	1	I
2-hexenal	$C_6H_{10}O$	98	723	6728-26-3	145‡	83	41;39;27;55;29	C	782	845	29.9	1	I
2-heptenal	$C_7H_{12}O$	112	812	57266-86-1	166§	41	41;27;55;29;39	B	880	942	68.3	1	I
2-octenal	$C_8H_{14}O$	126	894	2548-87-0	190§	70	29;41;55;70;27	C	804	849	19.5	1	I
(6) Alcohols													
methanol	CH_4O	32	165	67-56-1	65†	29	31;32;29;30;28	A	927	964	92.2	1	C
ethanol	C_2H_6O	46	220	64-17-5	78†	31	31;45;29;46;27	D	601	864	83.5	1	I
isopropyl alcohol	C_3H_8O	60	265	67-63-0	82†	45	45;28;29;43;27	D	492	848	30	1	I M
tert-butyl alcohol	$C_4H_{10}O$	74	292	75-65-0	82†	59	59;31;41;43;29	C	752	876	30.3	1	I
1-propanol	C_3H_8O	60	337	71-23-8	97†	31	31;27;59;29;41	D	458	800	23.8	1	I
1-butanol	$C_4H_{10}O$	74	502	71-36-3	117*	31	31;56;41;27;43	B	898	905	71.8	1	I

		MW (g)	RT (s)	CAS or NIST #	BP (°C)	QI	AI	C [a]	M	R	P (%)	N	D [b]
1-pentanol	$C_5H_{12}O$	88	630	71-41-0	138†	42	42;31;29;41;55	B	875	897	63.7	1	I
1-hexanol	$C_6H_{14}O$	102	726	111-27-3	157†	56	56;43;29;41;31	D	488	796	4.64	2	I
furfuryl alcohol	$C_5H_6O_2$	98	739	98-00-0	168†	98	41;39;98;42;29	B	914	927	75.6	1	I
unknown C8 alcohol	$C_8H_{18}O$	130	855	104-76-7	186†	57	57;41;43;29;56	C	695	859	17.8	1	I
2,6-dimethyloct-7-en-2-ol	$C_{10}H_{20}O$	156	884	18479-58-8	188‡	59	59;43;41;55;29	B	847	909	57.3	1	I
benzyl alcohol	C_7H_8O	108	896	100-51-6	205†	79	79;77;108;51;107	D	539	844	7.72	1	I
unknown alcohol			899			59	59;43;41;69;29	E					I
unknown alcohol	$C_{10}H_{18}O$	154	981	464-45-9	210‡	95	95;41;43;27;55	C	699	826	11.9	1	I
(7) Carboxylic acids													
formic acid	CH_2O_2	46	396	64-18-6	101†	45	29;46;45;44;28	C	702	899	83.3	1	M
acetic acid	$C_2H_4O_2$	60	462	64-19-7	118†	45	45;43;60;42;29	C	729	881	52.9	1	I M
Propanoic acid	$C_3H_6O_2$	74	586	79-09-4	142†	74	28;29;27;45;74	D	518	831	28.1	1	I M
butanoic acid	$C_4H_8O_2$	88	675	107-92-6	164†	60	44;60;73;27;42	D	534	661	2.06	4	M
pentanoic acid	$C_5H_{10}O_2$	102	760	109-52-4	186†	60	60;73;45;41;32	D	588	804	49.9	1	M
hexanoic acid	$C_6H_{12}O_2$	116	837	142-62-1	202‡	60	60;73;41;29;43	D	667	772	9.33	1	M
benzoic acid	$C_7H_6O_2$	122	1027	65-85-0	249‡	105	105;122;77;51;28	C	781	879	20.3	1	M
(8) Esters													
methyl methacrylate	$C_5H_8O_2$	100	528	80-62-6	101†	69	41;69;39;100;40	B	899	920	88.8	1	I
ethyl-2-methylbutanoate	$C_7H_{14}O_2$	130	684	7452-79-1	123*	41	29;57;41;102;27	B	832	921	91.2	1	I
ethyl-2-methylpentanoate	$C_8H_{16}O_2$	144	761	39255-32-8	153‡	102	43;102;29;74;71	B	841	873	77.6	1	I
hexyl acetate	$C_8H_{16}O_2$	144	823	142-92-7	170*	43	43;56;41;61;55	C	794	869	71.9	1	I
heptanoic acid ethyl ester	$C_9H_{18}O_2$	158	885	106-30-9	188‡	88	88;29;43;73;60	C	706	801	74.6	1	I

		MW (g)	RT (s)	CAS or NIST #	BP (°C)	QI	AI	C [a]	M	R	P (%)	N	D [b]
unknown acetate	$C_{10}H_{20}O_2$	172	931	72218-58-7	196§	43	43;70;41;55;69	D	607	878	3.85	1	I
1-(1-methoxypropan-2-yloxy) propan-2-yl acetate	$C_9H_{18}O_4$	190	961	NIST 367086		59	59;43;28;29;73	C	782	862	81	1	I
benzyl acetate	$C_9H_{10}O_2$	150	969	140-11-4	215†	91	108;43;91;90;79	B	910	933	89.1	1	I
unknown acetate	$C_{12}H_{22}O_2$	198	1068	32210-23-4	222‡	57	57;43;82;67;41	E	825	840	45.9	1	I
unknown acetate	$C_{12}H_{22}O_2$	198	1084	32210-23-4	222‡	57	57;43;82;67;41	C	799	849	44.1	1	I
dibutyl ethanedioate	$C_{10}H_{18}O_4$	202	1115	2050-60-4	246‡	57	57;29;41;56;27	A	691	849	17.3	1	C
unknown acetate	$C_{12}H_{22}O_2$	198	1129	32210-23-4	222‡	57	57;43;41;67;82	E	771	868	57.9	1	I
(9) Ketones													
acetone	C_3H_6O	58	248	67-64-1	56†	58	43;58;27;42;28	A	627	934	69.1	1	C I M
methacrolein	C_4H_6O	70	328	78-85-3	69‡	70	41;39;70;30;29	D	512	799	17.4	1	M
methyl vinyl ketone	C_4H_6O	70	358	78-94-4	80‡	70	55;43;27;70;28	C	681	836	53.3	1	M
methyl ethyl ketone	C_4H_8O	72	379	78-93-3	70*	43	43;29;27;72;57	A	855	941	83.5	1	C I M
4-methyl-2-pentanone	$C_6H_{12}O$	100	594	108-10-1	116*	43	43;58;41;29;57	A	867	886	55.6	1	C I
2-hexanone	$C_6H_{12}O$	100	650	591-78-6	127*	43	43;58;29;41;27	A	910	928	72.5	1	C
cyclohexanone	$C_6H_{10}O$	98	772	108-94-1	155*	55	55;42;41;27;39	A	890	903	29.9	2	C
4-cyclopentene-1,3-dione	$C_5H_4O_2$	96	785	930-60-9	226§	96	42;96;26;68;40	B	814	906	92.7	1	I
(10) Ether													
methyl tert-butyl ether	$C_5H_{12}O$	88	291	1634-04-4	55*	73	73;41;29;43;57	C	763	906	68.1	1	
(11) Sulfur-containing													
diallyl disulfide	$C_6H_{10}S_2$	146	901	2179-57-9	173‡	41	41;39;45;28;81	C	765	831	91.1	1	I
unknown sulfur-containing	$C_6H_{14}S_2$	150	915	629-19-6	193‡	108	43;41;108;27;39	C	737	829	68	1	I

		MW (g)	RT (s)	CAS or NIST #	BP (°C)	QI	AI	C [a]	M	R	P (%)	N	D [b]
(12) Nitrogen-containing													
acetonitrile	C_2H_3N	41	273	75-05-8	82†	41	41;40;39;38;28	A	850	979	80.7	1	C
n,n-dimethylformamide	C_3H_7NO	73	686	68-12-2	153*	73	44;73;42;28;30	A	940	959	97.8	1	C
(13) Chlorine-containing													
1,1-dichloroethane	$C_2H_4Cl_2$	99	329	75-34-3	57*	27	27;63;65;28;26	A	604	814	84.9	1	C
1,1,1-trichloroethane	$C_2H_3Cl_3$	133	414	71-55-6	74*	97	97;61;99;28;26	A	686	847	85.5	1	C
1,2-dichloroethane	$C_2H_4Cl_2$	99	460	107-06-2	84*	62	27;62;49;64;26	A	647	844	87.6	1	C
trichloroethylene	C_2HCl_3	131	506	79-01-6	87*	130	95;60;130;97; 132	A	708	852	95.4	1	C
1,2-dichloropropane	$C_3H_6Cl_2$	113	529	78-87-5	97*	27	27;63;62;41;39	A	716	857	94.1	1	C
tetrachloroethylene	C_2Cl_4	166	637	127-18-4	121*	47	47;35;28;129;94	A	695	765	87.6	1	C
(14) Unknown category													
unknown			260			29	43;29;42;72;28	E					I
unknown			482			81	81;41;55;27;39	E					I
unknown			517			83	41;43;55;83;42	E					I
unknown			646			41	41;39;43;45;27	E					I
unknown			721			82	39;82;27;54;53	E					I
unknown			790			91	91;120;92;65;51	E					M
unknown			856			70	43;70;71;57;28	E					I
unknown	$C_7H_{10}O$	110	865	4313-03-5	177§	81	81;39;41;27;53	E	861	914	57.9	1	I
unknown	C_6H_6O	94	881	108-95-2	182†	94	94;39;66;65;40	E	636	824	13.9	1	I M
unknown			922			148	41;45;43;106;39	E					I
unknown			943			81	81;43;41;57;80	E					I
unknown	$C_{11}H_{22}O_4$	218	948	NIST 378331		101	43;101;73;45;41	E	715	784	55.3	1	I

		MW (g)	RT (s)	CAS or NIST #	BP (°C)	QI	AI	C [a]	M	R	P (%)	N	D [b]
unknown			976			45	45;73;39;88;47	E					I
unknown			1033			69	41;69;39;55;94	E					I
unknown			1104			81	81;41;29;39;27	E					I
unknown	$C_{15}H_{24}$	204	1178	87-44-5	253‡	41	41;69;93;79;39	E	908	912	17.4	1	I
unknown	$C_{11}H_{12}O_3$	192	1193	NIST 132074		66	43;66;67;39;83	E	790	835	61.4	1	I

[a] Identification confidence category: Each compound identification is categorized as follows: A = positive identification in calibration standard, B = average of match and reverse match scores in NIST database search is greater than 850, C = average of match and reverse match scores in NIST database search is between 750 and 850, D = average of match and reverse match scores is less than 750, E = User determines that none of the proposed matches from the database search is superior to others (ex. stereoisomers with very similar MS spectra, matches with unsuitable retention index). In some cases, a CAS or NIST number (and compound properties) of the top match are still provided for unidentified compounds in the E category.

[b] Datasets: The letters indicate the datasets (C = calibration standards, I = indoor measurements, M = mobile measurements) in which the compound has been quantified. The absence of quantification does not necessarily indicate that the compound was absent in that dataset.

* Experimental boiling point (NIST Chemistry Webbook, 2023)

† Experimental boiling point (CRC Handbook of Chemistry and Physics Online)

‡ Experimental boiling point (Reaxys, 2023)

§ Modeled boiling point (ChemSpider, 2023)

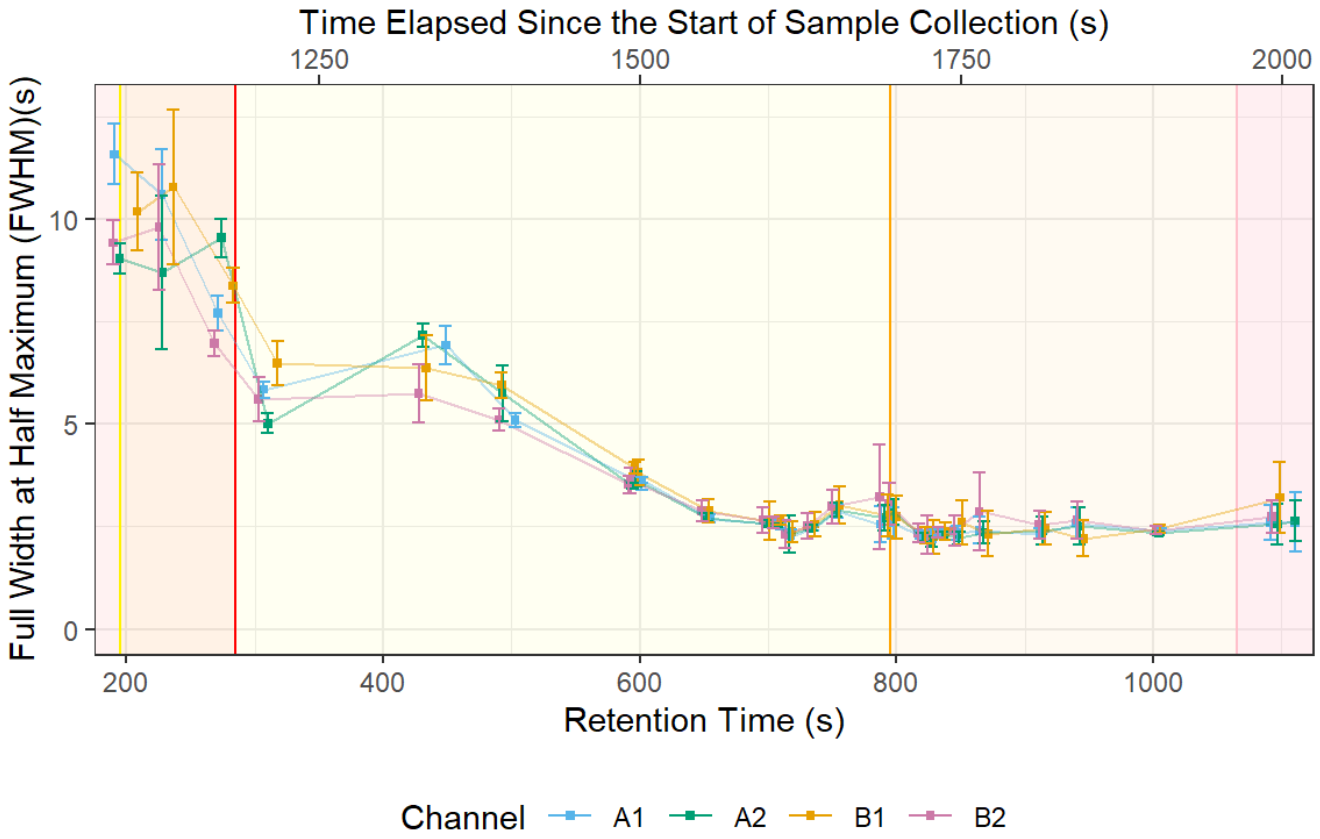


Figure S10. Full width at half maximum (FWHM) of chromatography peaks from the indoor measurements as a function of retention time (error bar is ± 1 standard deviation). The shaded regions indicate thermal desorption (red), the first GC temperature ramp (yellow, overlaps with thermal desorption), the second GC temperature ramp (orange), and the GC hold (pink). Error bars indicate one standard deviation.

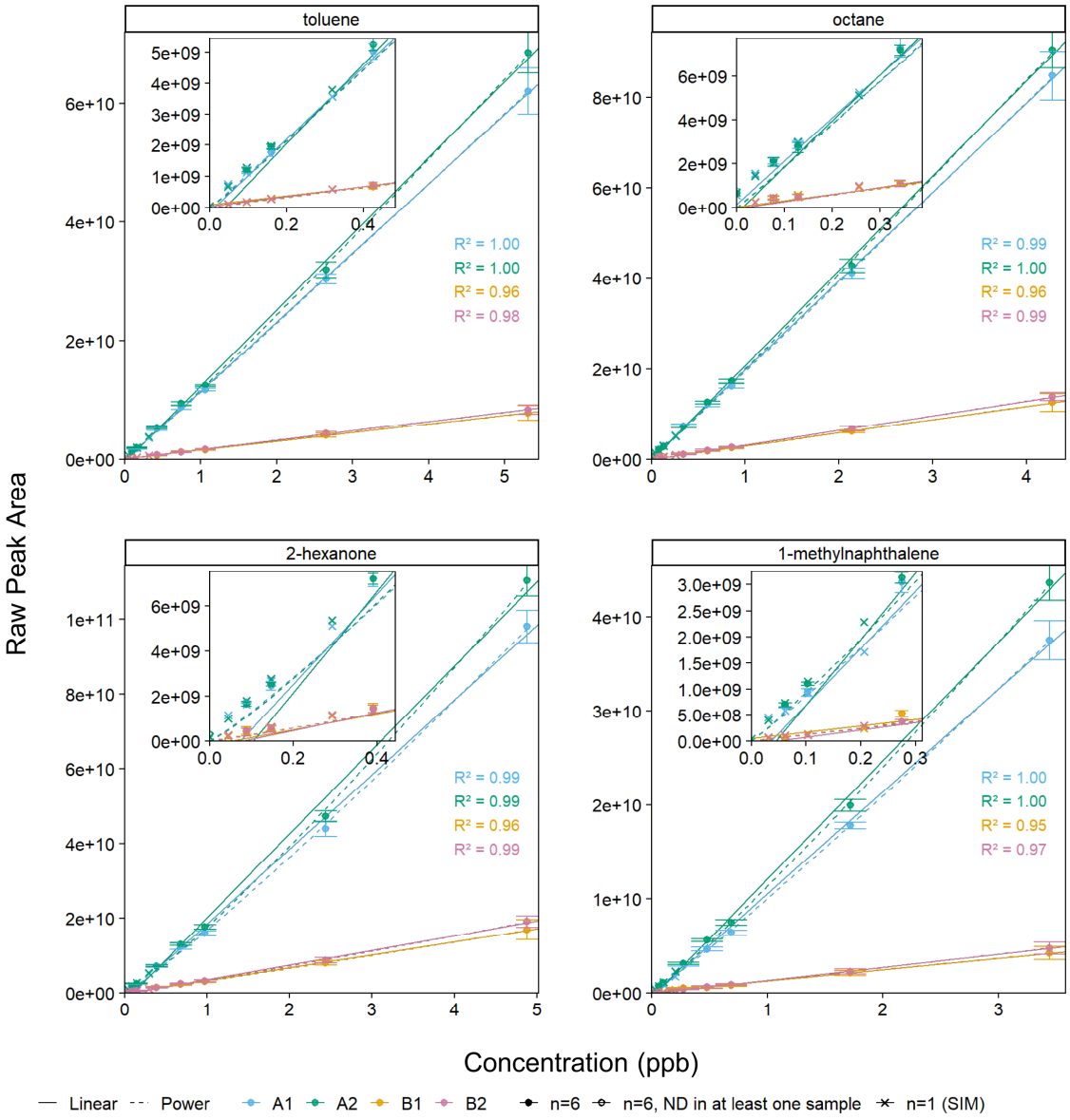


Figure S11. Linear and power calibration curves (error bars ± 1 SD) for raw peak areas of four compounds. The inset plot shows the lower range of data. Open circles indicate concentration levels for which no peak was detected on among at least one of the 2 ms scan replicates (n=6). Pearson R^2 values are for the replicates for which n=6 (filled circles).

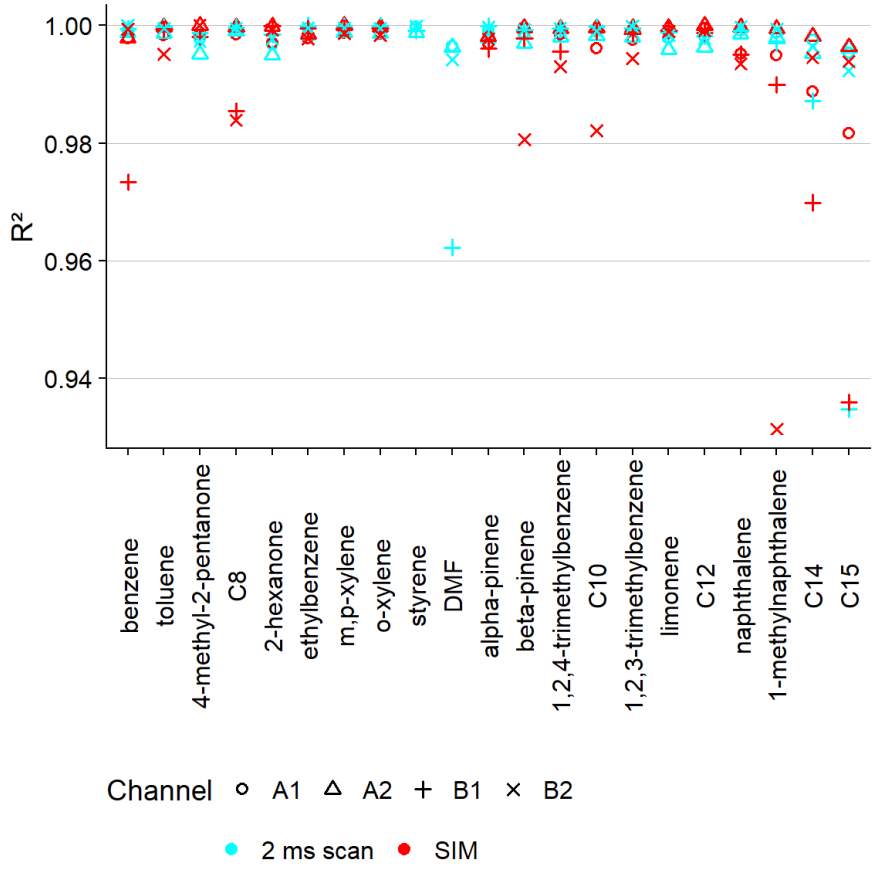


Figure S12. Pearson R^2 values for linear calibration curves. For the seven concentration levels tested with the 2 ms scan method, only levels with six detectable replicates are included. The SIM method data includes one replicate of four concentration levels. Alkanes are indicated with their carbon number.

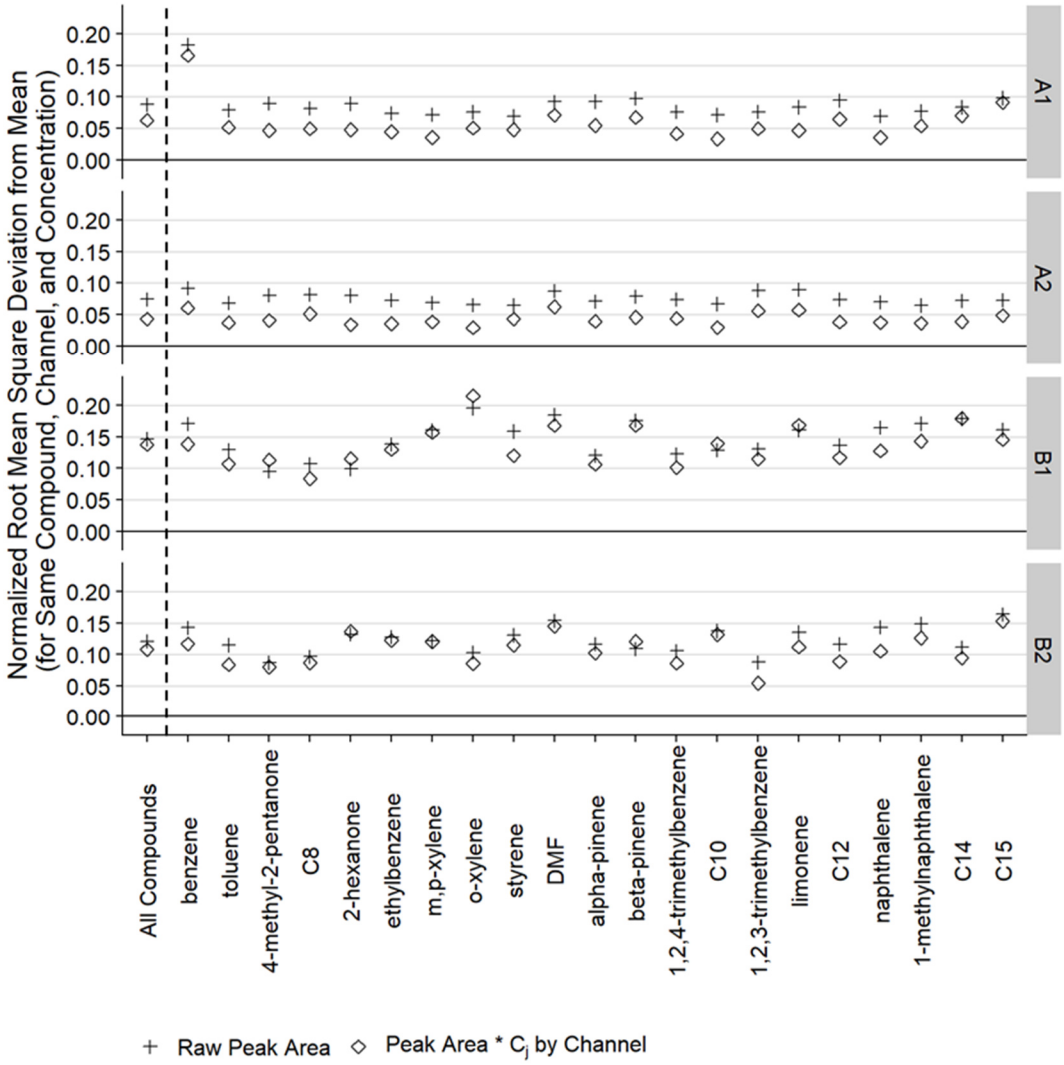


Figure S13. The precision within a single channel is described by the normalized root mean square deviation (NRMSD) from the mean response for the same compound, channel, and concentration as calculated with raw peak area and with the daily correction factor C_j which accounts for MS drift (2 ms scan method, including concentration, compound, and channel sets with six replicates, excluding most volatile compounds which evaporated from liquid calibration standard). Single channel precision is in general better for the more sensitive RGA A and with the daily drift correction. Alkanes are indicated by their carbon number.

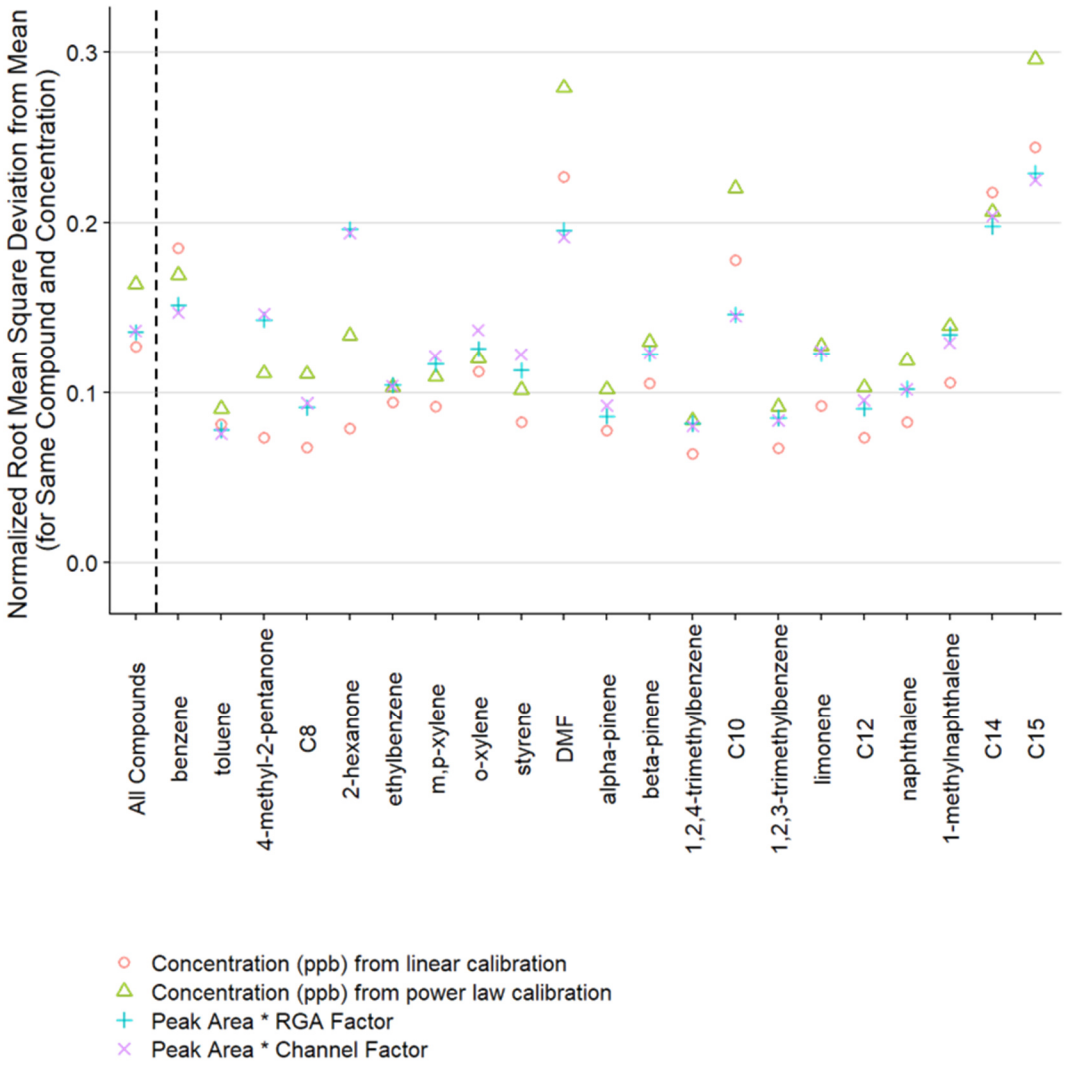


Figure S14. The 4-channel measurement precision is calculated by the normalized root mean square deviation (NRMSD) from the mean response across all channels for the same compound and concentration (2 ms scan method, including concentration, compound, and channel sets with six replicates, excluding most volatile compounds which evaporated from liquid calibration standard, accounting for the daily MS drift correction factor C_j). Alkanes are indicated by their carbon number.

S6 Indoor measurements

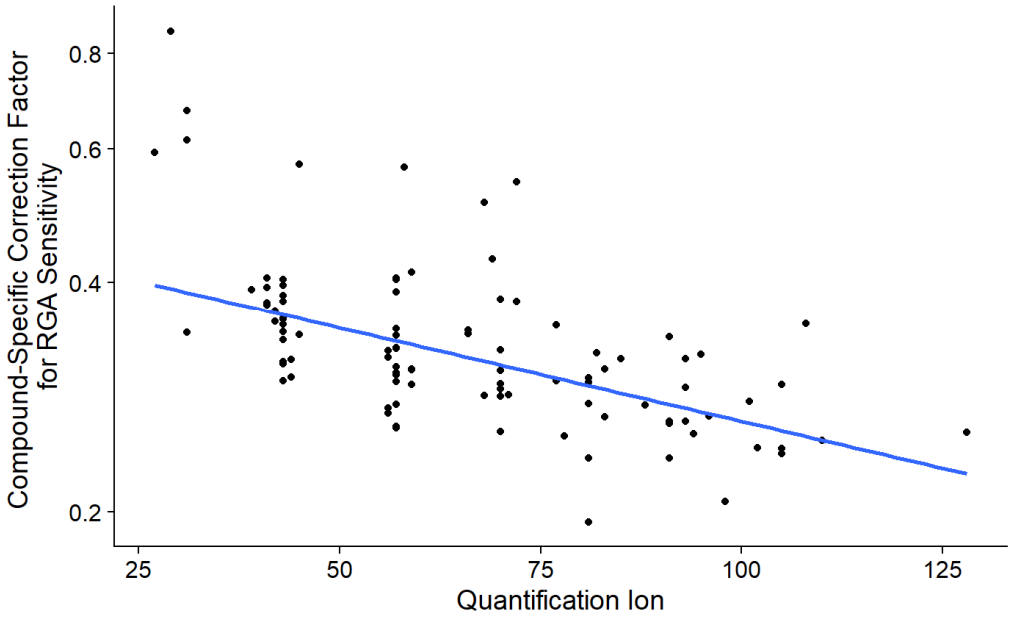


Figure S15. The log of the compound-specific correction factor for RGA sensitivity for the indoor measurements versus the quantification ion. The best fit line is used if a compound has less than six measurements remaining after Cook's distance outliers are removed.

Table S4. Nominal times, start times, and end times of events during December 9, 2021 indoor measurements

Event	Nominal Time	Start Time	End Time	Description
1	12/9/2021 14:20	12/9/2021 14:20	12/9/2021 14:30	Air freshener plugged in
2	12/9/2021 15:10	12/9/2021 15:10	12/9/2021 15:20	Chopped vegetables
3	12/9/2021 15:20	12/9/2021 15:27	12/9/2021 15:30	Heated vegetable oil
4	12/9/2021 15:30	12/9/2021 15:30	12/9/2021 15:40	Stir-fried onion, garlic, and ginger
5	12/9/2021 15:40	12/9/2021 15:40	12/9/2021 15:50	Stir-fried mushroom with black pepper
6	12/9/2021 16:50	12/9/2021 16:56	12/9/2021 17:15	Washed dishes
7	12/9/2021 17:20	12/9/2021 17:20	12/9/2021 17:32	Cleaned floors with pine cleaner, opened kitchen door
8	12/9/2021 17:30	12/9/2021 17:32		Closed kitchen door
9	12/9/2021 18:00	12/9/2021 18:00		Opened kitchen door
10	12/9/2021 18:10	12/9/2021 18:10		Closed kitchen door
11	12/9/2021 18:30	12/9/2021 18:30		Opened windows
12	12/9/2021 19:20	12/9/2021 19:20	12/9/2021 19:27	Closed windows, heated oil and black pepper
13	12/9/2021 19:30	12/9/2021 19:30	12/9/2021 19:40	Stir-fried onion
14	12/9/2021 19:40	12/9/2021 19:42	12/9/2021 19:50	Stir-fried mushrooms
15	12/9/2021 20:10	12/9/2021 20:10		Opened windows, turned on vent fan
16	12/9/2021 21:00	12/9/2021 21:06		Closed windows, turned off vent fan
17	12/9/2021 21:40	12/9/2021 21:42	12/9/2021 21:54	Cleaned floors with pine cleaner, opened kitchen door

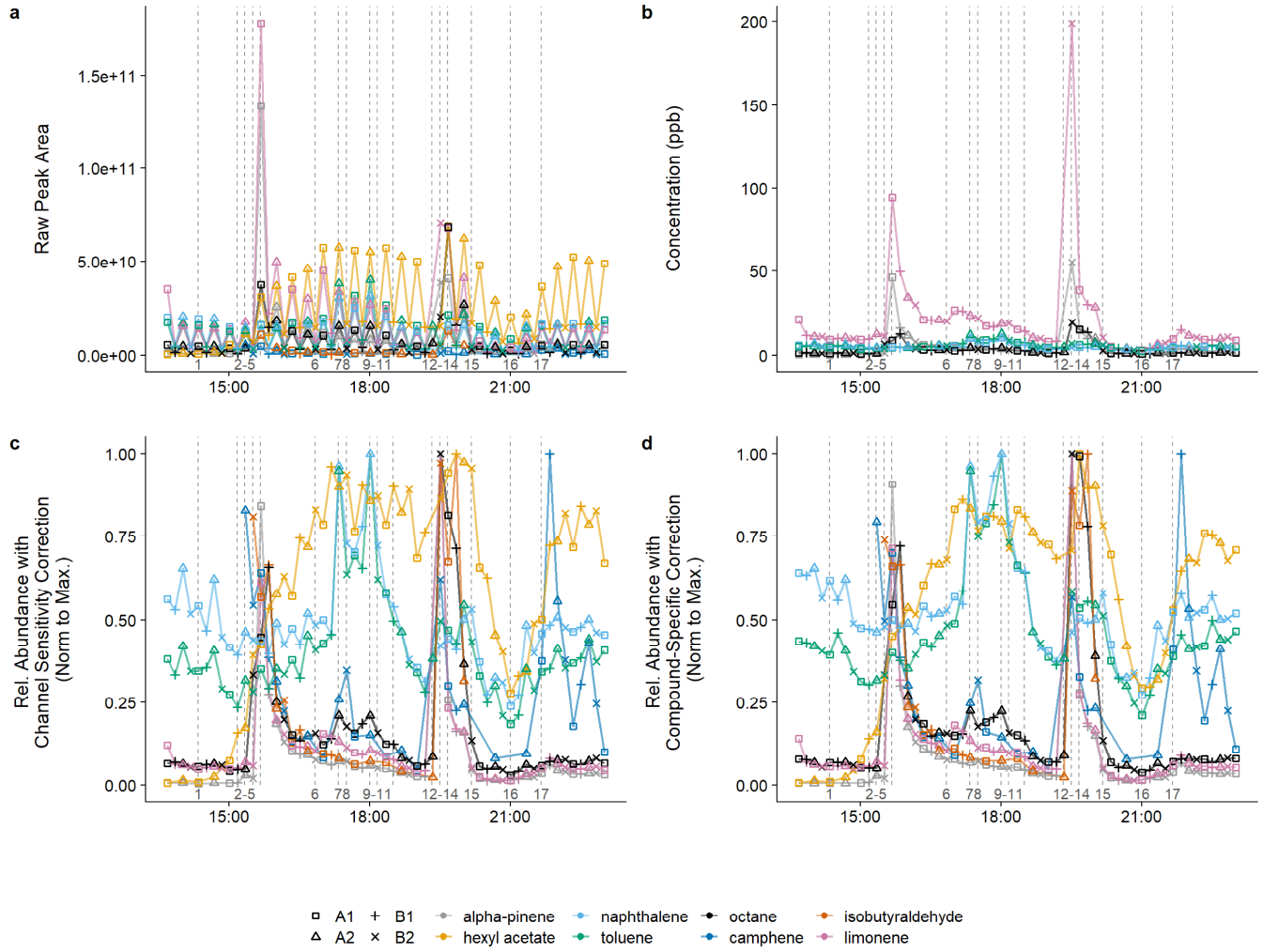


Figure S16. Timeseries of (a) raw peak area, (b) mixing ratios (ppb) calculated with power law calibration curves, (c) relative abundance with channel sensitivity factor correction and normalization to the maximum value, and (d) relative abundance with compound-specific correction and normalization to the maximum value. The vertical lines correspond to the following activities (with times rounded down to the nearest sampling time): 1 - Air freshener plugged in, 2 - Chopped vegetables, 3 - Heated vegetable oil, 4 – Stir-fried onion, garlic, and ginger, 5 – Stir-fried mushroom with black pepper, 6 - Washed dishes, 7 - Cleaned floors with pine cleaner and opened kitchen door, 8 – Closed kitchen door, 9 - Opened kitchen door, 10 - Closed kitchen door, 11 - Opened windows, 12 - Closed windows, heated oil and black pepper, 13 - Stir-fried onion, 14 - Stir-fried mushrooms, 15 - Opened windows, turned on vent fan, 16 - Closed windows, turned off vent fan, 17 - Cleaned floors with pine cleaner, opened kitchen door

S7 Mobile measurements

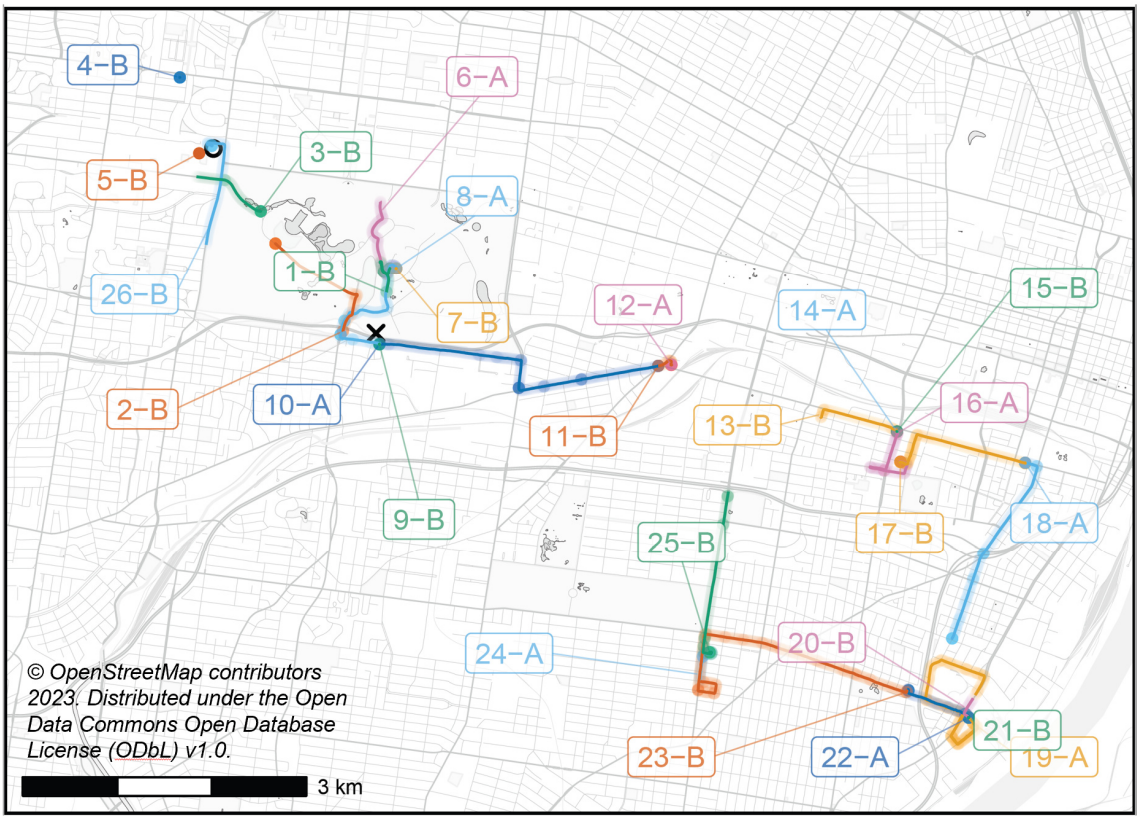


Figure S17. Map of locations of pilot mobile measurements in St. Louis, Missouri. The labels indicate the order of measurement and the RGA which acquired the MS data. Measurements 1 through 5 were from July 20, 2022, and measurements 6 through 26 were from July 21, 2022. The thin line indicates the path of the vehicle during the measurement, and the darkness of the thicker line corresponds to time spent in a given location. If sample flowrate was < 20 ccm (for example, during intermittent pump failures on July 20), the data are omitted, and the locations are not shown. The black open circle indicates the location of supporting air quality instrumentation (including ozone). The black × indicates the position of additional supporting meteorology and air quality instrumentation.

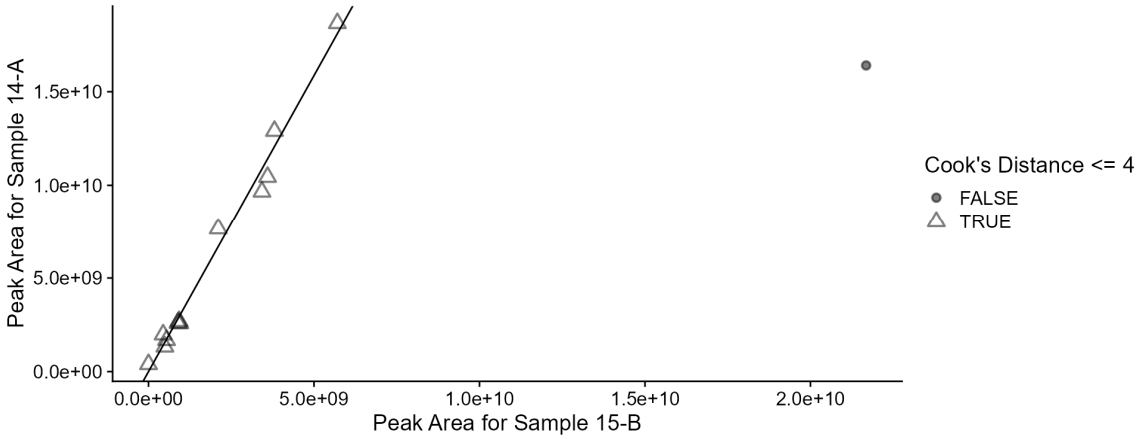


Figure S18. The correction factor for the sensitivity difference in the two RGAs was calculated as the slope of the linear fit through the origin of the peak areas for sample 14-A versus the peak areas for the same compounds for sample 15-B (slope = 3.17, R^2 = 0.99). One point with Cook's distance > 4 was excluded.

References

Reaxys: www.reaxys.com, last access: 26 January 2023.

Isaacman-Vanwertz, G., Sueper, D. T., Aikin, K. C., Lerner, B. M., Gilman, J. B., De Gouw, J. A., Worsnop, D. R., and Goldstein, A. H.: Automated single-ion peak fitting as an efficient approach for analyzing complex chromatographic data, J. Chromatogr. A, 1529, 81–92, https://doi.org/10.1016/j.chroma.2017.11.005, 2017.

NIST Chemistry Webbook: webbook.nist.gov/chemistry, last access: 26 January 2023.

Ng, N. L., Herndon, S. C., Trimborn, A., Canagaratna, M. R., Croteau, P. L., Onasch, T. B., Sueper, D., Worsnop, D. R., Zhang, Q., Sun, Y. L., and Jayne, J. T.: An Aerosol Chemical Speciation Monitor (ACSM) for Routine Monitoring of the Composition and Mass Concentrations of Ambient Aerosol, Aerosol Sci. Technol., 45, 780–794, https://doi.org/10.1080/02786826.2011.560211, 2011.

ChemSpider: www.chemspider.com, last access: 26 January 2023.

CRC Handbook of Chemistry and Physics Online: hbcp.chemnetbase.com.

Seborg, D. E., Mellichamp, D. A., Edgar, T. F., and Doyle, F. J.: Process Dynamics and Control, 3rd ed., Wiley, Hoboken, 528 pp., 2010.

Stein, S. E.: Estimating probabilities of correct identification from results of mass spectral library searches, J. Am. Soc. Mass Spectrom., 5, 316–323, https://doi.org/10.1016/1044-0305(94)85022-4, 1994.