Supplemental Information for:

On-line derivatization for hourly measurements of gas- and particle-phase semi-volatile oxygenated organic compounds by Thermal desorption Aerosol Gas chromatography (SV-TAG)

Gabriel Isaacman[1], Nathan M. Kreisberg[2], Lindsay D. Yee[1], David R. Worton[1,2], Arthur W. H. Chan[1,*], Joshua A. Moss[1], Susanne V. Hering[2], Allen H. Goldstein[1,3]

[1]{Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA}

[2]{Aerosol Dynamics Inc., Berkeley, CA}

[3]{Department of Civil and Environmental Engineering, University of California, Berkeley, CA}

[*]{now at: Department of Chemical Engineering and Applied Chemistry, University of Toronto, Canada}

Correspondence to: G. Isaacman (gabriel.isaacman@berkeley.edu)

S1. Derivatization efficiency tests

As discussed in the text, derivatization efficiency is tested using repeated injection of a mixture including 43 oxygenated compounds. In Table S1, names of the compounds injected and their molecular formulas and number of OH groups are shown, as well as the retention index of the derivatized compound. For every OH group, the derivatized compound eluted contains a trimethylsilyl group, adding C_3H_8Si to the formula of the observed peak. Observed relative retention times are shown as a Kovats-like (Kovats, 1958) retention index relative to *n*-alkanes, with i.e. *n*-pentacosane (C_{15}) having an index of 1500. Compounds span a retention time of approximately tridecane to tetratriacontane, but most (38 compounds) elute earlier than pentacosane.

Table S1. Compounds injected to assess derivatization efficiency: name, molecular formula, number of OH groups, and retention index of the derivatized peak relative to an *n*-alkane series.

Injected Compound Name	Molecular Formula	Number OH Groups	Derivatized Retention Index
Glyceric acid	$C_3H_6O_4$	3	1323
2,6-Dimethoxyphenol (Syringol)	$C_8H_{10}O_3$	1	1398
3,3-Dimethylglutaric acid	$C_7H_{12}O_4$	2	1433
n-Decanoic acid	$C_{10}H_{20}O_2$	1	1453
Threitol	$C_4H_{10}O_4$	4	1493
Erythritol	$C_4H_{10}O_4$	4	1501
Ketopinic acid	$C_{10}H_{14}O_3$	1	1507
cis-Pinonic acid	$C_{10}H_{16}O_3$	1	1526
3-Methoxy-4-hydroxybenzaldehyde (Vanillin)	$C_8H_8O_3$	1	1537
2-Methoxy-4-propenylphenol (Isoeugenol)	$C_8H_{10}O_3$	1	1568
Diethyltoluamide	$C_{12}H_{17}NO$	0	1582
Benzophenone	$C_{13}H_{10}O$	0	1645
cis-Pinic acid	$C_9H_{14}O_4$	2	1663
γ-Dodecalactone	$C_{12}H_{22}O_2$	0	1688
Levoglucosan	$C_6H_{10}O_5$	3	1698
α-Bisabolol	$C_{15}H_{26}O$	1	1742
n-Tridecanoic acid	$C_{13}H_{26}O_2$	1	1746
1,9-Nonadioic acid	$C_9H_{16}O_4$	2	1791
1,10-Decadioic acid	$C_{10}H_{18}O_4$	2	1887
Methyl palmitate	$C_{17}H_{34}O_2$	0	1924
n-Hexadecanol	$C_{16}H_{34}O$	1	1956
cis-9-Hexadecenoic acid	$C_{16}H_{30}O_2$	1	2022
Homosalate	$C_{16}H_{22}O_3$	1	2025
n-Hexadecanoic acid	$C_{16}H_{32}O_2$	1	2041
1,12-Dodecadioic acid	$C_{12}H_{22}O_4$	2	2082
Methyl stearate	$C_{19}H_{38}O_2$	0	2125
n-Heptanoic acid	$C_{17}H_{34}O_2$	1	2139
n-Octadecanol	$C_{18}H_{38}O$	1	2154
cis,cis-9,12-Octadecadienoic acid	$C_{18}H_{32}O_2$	1	2206
cis-9-Octadecenoic acid	$C_{18}H_{34}O_2$	1	2213
cis-11-Octadecenoic acid	$C_{18}H_{34}O_2$	1	2220
n-Octadecanoic acid	$C_{18}H_{36}O_2$	1	2240
1,14-Tetradecanoic acid	$C_{14}H_{26}O_4$	2	2275
n-Eicosanol	$C_{20}H_{42}O$	1	2349
Isopimaric acid	$C_{20}H_{30}O_2$	1	2353
16-Hydroxyhexadecanoic acid	$C_{16}H_{32}O_3$	2	2385
12-Hydroxyoctadecanoic acid	$C_{18}H_{36}O_3$	2	2423
Abietic acid	$C_{20}H30O_2$	1	2433
Deoxycholic Acid	$C_{24}H_{40}O_4$	3	3065
Cholesterol	$C_{27}H_{46}O$	1	3166
β-Stigmasterol	$C_{29}H48O$	1	3297
β-Sitosterol	$C_{29}H_{50}O$	1	3360
Lupeol	$C_{30}H_{50}O$	1	3441

S2. Derivatization reproducibility tests

When correcting analytes for run-to-run variability using internal standards, an internal standard must be selected to use for the correction. Several possible selection criteria are available for correction of oxygenated compounds. Table 1 in the main text lists the schemes tested for correction of all oxygenates with hydroxyl groups and the error of each scenario, measured as the relative standard deviation from the average ratio of one internal standard to another one selected based on the criteria of the scenario. Figure S1 shows the cumulative error distribution for a subset of the test scenarios selected to apply to most, if not all, operating conditions. Correcting for compounds using an internal oxygenated standard of only similar volatility (i.e. in the case of analytes of unknown structure or a functionally similar internal standard is unavailable) is modeled by correcting all oxygenates to the nearest *n*-acid in volatility, of which there are 4 of various volatility in the standard used (Fig. S1a). The error present in correcting for only general changes in derivatization efficiency is quantified by correcting all oxygenates using a single, relatively stable oxygenate, *n*-octadecanoic acid-d_{35} (Fig. S1b). Under operating conditions requiring minimal internal standards or maximizing standard stability by not including oxygenates, oxygenates can be corrected for variability in transfer efficiency and detector sensitivity using only alkanes of similar volatility, but this results in large errors with a relatively non-Gaussian distribution (Fig. S1c). Due in part to operational improvements after the SOAS field campaign, Study 2 has lower error in all cases shown in Fig. S1 owing to more reproducible measurements of multi-functional acids.

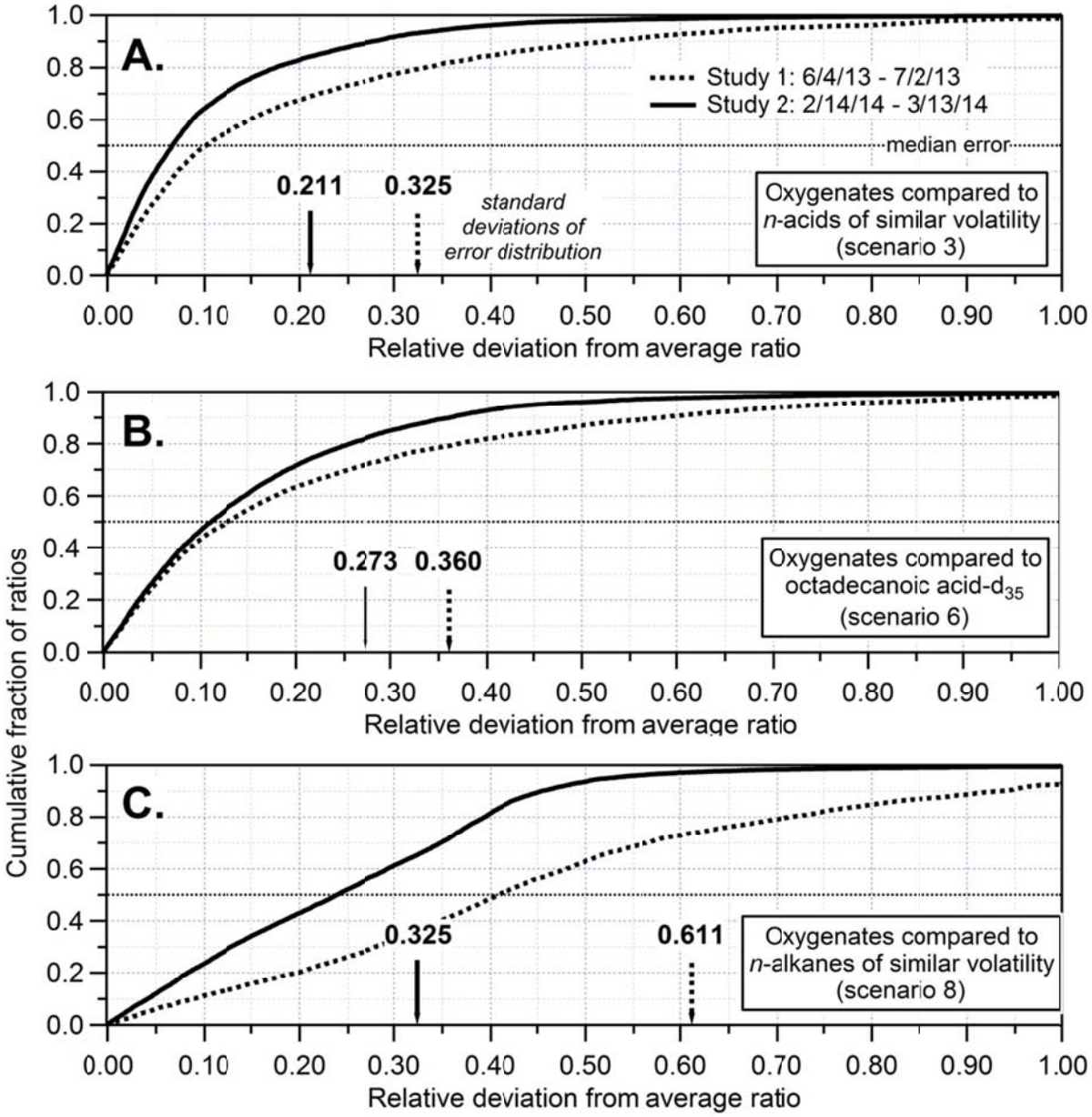

Figure S1. Cumulative distribution of error in correction of oxygenates for run-to-run variability in derivatization efficiency and instrument response measured as relative deviation of two internal standards from their average ratio. Boxes and Sect. S2 describe each correction scenario, which are numbered corresponding to Table 1 in the main text. Arrows show relative standard deviation for each study. Dotted line is median error (50% of points halve less than this error).

S3. Total uncertainty

The main text of this manuscript addresses reproducibility and precision in derivatization. To explicitly calculate total uncertainties in reported masses and fractions in the particle, the two independent cells of the instrument also have to be calibrated and then compared. Uncertainty in mass calibration using a linear calibration curve is described by NIST (2014a) and is applied in Sect. S3.1 without significant modifications to recommended practices. Differences between parallel sampling cells are removed through normalization to the mean response of the two cells to identical samples, with the magnitude and uncertainty in this normalization included in estimation of mass uncertainty as described in Sect. S3.2. Finally, in Sect. S3.3, uncertainty in particle fraction is found to be a function of this normalization, as well as derivatization precision as discussed in the main text, but not mass calibration.

Known compounds are found below to be quantified with 20-25% accuracy, though detailed error analysis is actually compound-, study-, and even point-specific. By calculating fraction in particle before performing mass calibration, partitioning can be measured with less uncertainty than simply compounding the error in two compared mass measurements. Fraction in particle is thus found to also be measured with approximately 15-25% uncertainty; formal estimation of this error is confirmed through empirical estimates. Calculating F_P from signals also allows measurement of particle fraction even for compounds which are not unambiguously identified or for which no authentic standard exists. Data from this instrumented is found to be best reported as total mass and fraction in the particle, with particle mass calculated from these two value, because particle mass is typically lower and thus more uncertain than total mass.

S3.1. Uncertainty in mass calibration of a single cell

Raw signal, R_A^i, of an analyte in either cell, *i*, is ratioed to the raw signal of an internal standard, R_{IS}^i, to generate a corrected signal, S_A^i, that accounts for any systematic variability in instrument response. The uncertainty in this correction for run-to-run variability is detailed in the main text and is defined here as σ_S^i, a relative uncertainty in precision. It depends on the similarity between the analyte and the internal standard, but in typical operation is usually 10-15% depending on the compound (and includes uncertainty in the AutoInject injections).

Calibration curves as shown in Fig. 5 are calculated in terms of calibrant corrected signal, S_c^i, (raw calibrant signal, R_c^i, ratioed by internal standard, R_{IS}^i), against mass injected (using the AutoInject) for all injections within a given operational period. A linear fit of calibrants ($S_C^i = a^i + b^i M_C^i$) is used, such that the mass of an analyte M_A^i is:

$$M_A^i = \frac{S_A^i - a^i}{b^i} \quad \text{(Eq. S1)}$$

Absolute error in mass for any point, j, $\Delta_{M\,j}^i$, is derived from the linear fit (NIST, 2014a):

$$\Delta_{Mj}^i = \sqrt{\left(\frac{1}{b^i}\right)^2 {\Delta_S^i}^2 + \left(\frac{-1}{b^i}\right)^2 {\Delta_a^i}^2 + \left(\frac{-\left(S_{Aj}^i - a\right)}{{b^i}^2}\right)^2 {\Delta_b^i}^2 + 2\left(\frac{-1}{b^i}\right)\left(\frac{-\left(S_{Aj}^i - a\right)}{{b^i}^2}\right)\sigma_{ab}^i} \quad \text{(Eq. S2)}$$

Where Δ_S^i is absolute uncertainty in the corrected signal (S_A^i multipled by σ_S^i), Δ_a^i and Δ_b^i are the absolute uncertainty in the intercept and slope, respectively, and σ_{ab}^i is their covariance. Injections of only internal standard (lacking any calibrant) are used as zeroes to constrain the intercept. Inclusion of an intercept term therefore amounts to background subtraction, which is observed in most cases to be within uncertainty of the origin, suggesting background subtraction is minor and any error introduced is included in the error of the intercept term, Δ_a^i.

Error in mass can also be calculated empirically from the fit as the difference between measured and injected mass (NIST, 2014b)

$$\Delta_{Mj}^i = M_{Aj}^i - (\text{mass injected}) \quad \text{(Eq. S3)}$$

As signal approaches the intercept, the relative calibration error increases in importance due to the uncertainty in the intercept term, even when constrained by injections of "zeros". This is demonstrated in Fig. S2, in which the uncertainty for calibration of pinic acid is calculated both formally from Eq. S2 (blue squares) and empirically from Eq. S3 (red circles) and normalized to the observed mass. The formal calculation is in good agreement with the empirical measurements.

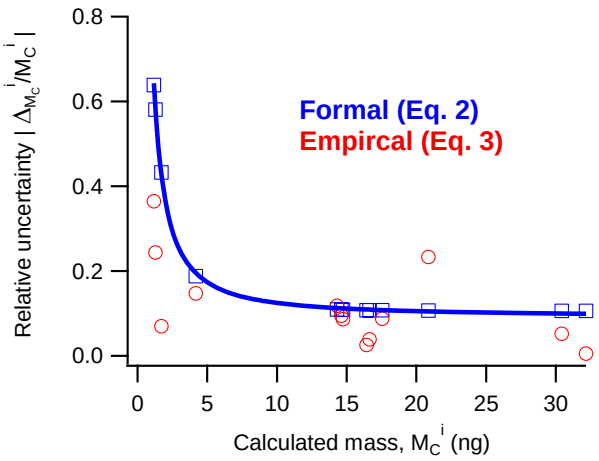

Figure S2. Relative error in mass calibration of pinic acid during one operational period, calculated with the formal equation (blue) and as the residual of the linear fit for each point (red).

While Δ_M^i is point-specific, uncertainty in the intercept, Δ_a^i, is negligible for large samples so uncertainty for most points is dominated by error in the corrected signal and the slope, which are a constant fraction independent of measured mass. However, uncertainty can be very large at measured masses near the lower limit of the calibration range (Fig. S2). The detection limit for this compound (3 times the background chromatographic signal) is approximately 0.5 ng on column, in good agreement with Fig. S2 as uncertainty in the measurement approaches 100% near the detection limit. For compounds with concentrations typically higher than the detection limit, relative mass uncertainty, σ_M^i, can be considered to be 15% (assuming σ_S^i = 10%), though it should be explicitly calculated for each compound. For cases in which most points fall well above detection limit, an average estimate of error for each compound is sufficient, though it is in all cases more desirable to report absolute uncertainty for each measurement when possible.

The uncertainty introduced by the intercept term is expected to be more important in denuded samples, where signals are typically lower, though including an adequate zero or blank measurement can help reduce this source of uncertainty. Consequently, it can reasonably be expected that $M_A{}^{den}$ typically has a higher uncertainty than $M_A{}^{byp}$. The degree to which this is the case depends on uncertainty in the intercept term and the size of denuded signals relative to bypass signals, but suggests error will typically be lower in bypass samples.

S3.2. Normalizing between cells

In an ideal world, the parallel sampling cells are exactly the same, but this is not the case due to small but consistent differences in derivatization efficiency and/or transfer losses not accounted for by the calibration, so there is a need to correct for these differences. Generating continuous timelines of mass concentrations or to calculate fraction in the particle relies on this correction because during typical operation a sample is denuded on one cell and bypass on the other, then the cells are switched to avoid bias. Variations on this sampling scheme can be employed, but in all schemes there is a fundamental need to intercompare cells. To correct for systematic differences between the cells, bypass samples are periodically collected simultaneously in both cells, providing a direct comparison using real air samples. This comparison is shown in Fig. S3 using both fully calibrated masses (Fig. S3a) and uncalibrated signals (Fig. S3b). Note that calibrating the signal reduces the difference between the cells, but also marginally increases the relative uncertainty in the slope.

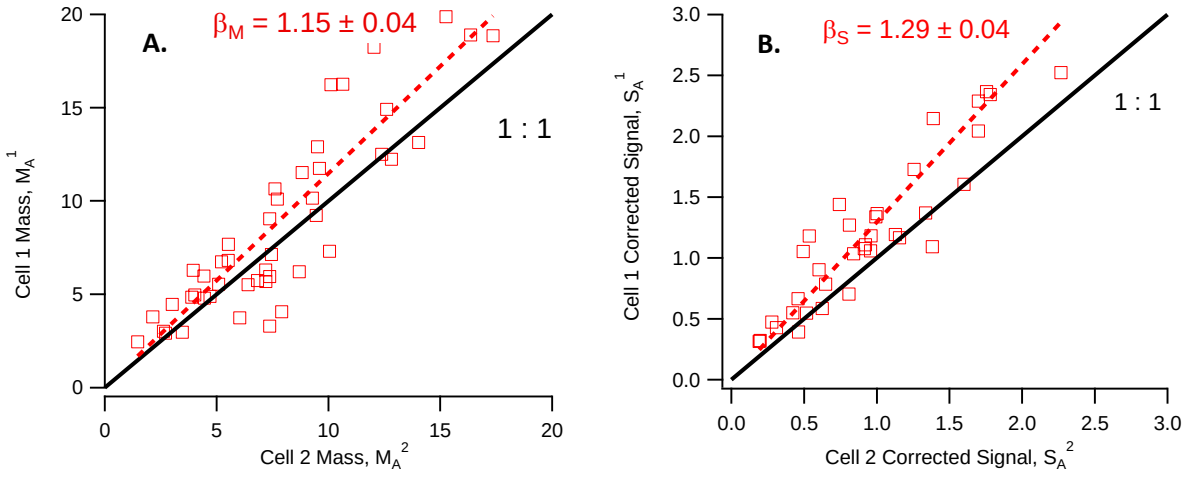

Figure S3. "Bypass-bypass" comparison of pinic between two cells, Cell 1 and Cell 2, in (a) mass terms, M_A^1 and M_A^2 and (b) signal terms, S_A^1 and S_A^2.

The bypass-bypass comparison of identical samples on the two cells prior to normalization provides an equalization factor, E_M^i, to adjust the cells to their mean value. This equalization can be applied across an entire measurement period, or to subsets thereof. It has been observed that there can be some temporal variability in cell-to-cell differences, so when possible, equalization is performed in small subsets, comparing only the bypass-bypass points nearest in time to each

point. This is unlikely to affect average calibrated values, but is expected to yield more accurate temporal and diurnal variability.

The equalization factor is calculated from the best-fit slope, β, (Fig. S3) which is forced through zero because sample cannot exist on one cell and not the other, so an intercept has no physical meaning. In most cases the intercept is within uncertainty of the fit so forcing through zero simplifies calculations with no detriment to the fit. The equalization factor is therefore:

$$E_M^2 = 0.5(\beta + 1) \quad \text{and} \quad E_M^1 = 0.5(\frac{1}{\beta} + 1) \qquad \text{(Eq. S4)}$$

and, in relative terms, $\sigma_{E_M} = \sigma_{\beta_M}$ (in the shown example, approximately 4%) (Eq. S5)

To equalize the cells, the original mass calculated is multiplied by this factor:

$$\text{Adjusted mass, } M_A^{i\,*} = E_M^i \times M_A^i \qquad \text{(Eq. S6)}$$

$$\text{with error: } \sigma_{M_p}^{i\,*} = \sqrt{\left(\sigma_M^i\right)^2 + \left(\sigma_{E_M}\right)^2} \qquad \text{(Eq. S7)}$$

This error accounts only for uncertainty in precision, designated with subscript *p*, which incorporates uncertainty in the equalization factor, σ_{E_M}, but does not account for systematic errors or instrument biases. The equalization factor is itself a systematic error in accuracy, while other additional instrument uncertainties, σ_I, may also contribute to accuracy error, to yield:

$$\text{total error in mass measurement, } \sigma_M^{i\,*} = \sqrt{\left(\sigma_{M_p}^{i\,*}\right)^2 + (\sigma_I)^2} \qquad \text{(Eq. S8)}$$

Because equalization modifies the absolute calibrated mass in each cell, systematic biases must be as great as or greater than the magnitude of the equalization factor. Additional known sources of error, i.e. uncertainty in sample volume, liquid injection, etc., may also cause systematic uncertainties and instrument errors. However, because cells are largely independent – i.e. separate sample volume control and liquid injection volumes – many potential significant sources of instrument error would negatively impact E_M. For example, if uncertainty in sample flow were 10%, it is unlikely the bypass-bypass comparison for any compounds would ever be in good agreement. Therefore, E_M incorporates in large part most other large uncertainties so is expected to be dominate instrument error:

$$\text{total instrument error, } \sigma_I = |E_M - 1| + \sigma_1 + \sigma_2 + \ldots \approx |E_M - 1| \qquad \text{(Eq. S9)}$$

Final precision uncertainty for the example compound (pinic acid) is therefore approximately 15% (Eq. S7: $\sqrt{(15\%)^2 + (4\%)^2}$), with an additional bias error greater than 8% for a total error of approximately 20%. It should be noted that any compound for which an authentic standard error is unavailable has an additional bias error, which is in most cases difficult to constrain.

S3.3. Uncertainty in Partitioning Fraction, F_P

Fraction in the particle, $F_{P,}$ is calculated by comparing a "bypass" sample in one cell to a simultaneously collected "denuded" sample in the other cell, which requires equalization between the cells. Though this can most intuitively be considered in mass terms, $F_{P,M}$, calculation of F_P relies solely on the ability to intercompare samples on different cells, not necessarily quantitative mass measurements. F_P can therefore also be considered in signal terms, using a signal-based equalization factor, which will be shown here to result in reduced uncertainty:

$$F_{P,M} = \frac{M_A^{den}}{M_A^{byp}} \times \beta_M \quad \equiv \quad F_{P,S} = \frac{S_A^{den}}{S_A^{byp}} \times \beta_S \quad \text{(in the case where Cell 2 denuded)} \quad \text{(Eq. S10)}$$

As a comparison between the cells, systematic biases in accuracy do not add uncertainty to this calculation. Instead, only uncertainty in precision and the equalization factor are relevant:

$$\sigma_{F_P,M/S} = \sqrt{\left(\sigma_{M/S}^{den}\right)^2 + \left(\sigma_{M/S}^{byp}\right)^2 + \left(\sigma_{E_{M/S}}\right)^2} \quad \text{(Eq. S11)}$$

In signal terms, the bypass-bypass ratio, β, is expected to be higher than in mass terms, as is observed in Fig. S3. However, the magnitude of the equalization does adversely affect uncertainty in F_P because uncertainty in intercomparison between cells does not depend on the size of this equalization, only on the uncertainty in the ratio, which is similar or slightly lower in signal terms. Given that σ_M^i is a function of both signal uncertainty, σ_S^i, and calibration error, formulating Eq. S11 in signal terms using only σ_S^i necessarily yields lower uncertainty.

The formal calculation of error in Eq. S11 can be tested against an empirical error estimate by investigating scatter around the cell-to-cell equalization line in an ideal case. When the internal standard is very similar to the analyte, as in the example compound (pinic acid – $C_9H_{14}O_4$, using as an internal standard a deuterated adipic acid – $C_6H_6D_4O_4$), the scatter around the equalization

line, β, is a result of the uncertainties in corrected signals (σ_S^i) as well as any uncertainties in cell-to-cell equalization and is therefore a good estimate of the error in F_P, similar to Eq. S3 (NIST, 2014b). The standard deviations of the residual from the equalization line for pinic acid in signal and mass terms (Fig. S3) are 14% and 18% respectively, very similar to the calculated errors of $\sigma_{F_P,S} \approx 15\%$ and $\sigma_{F_P,M} \approx 20\%$, so formal calculation of error is found to be reasonable.

If no internal standard is available that is similar to the analyte of interest, equalization using bypass-bypass analyses may result in a bias when comparing bypass to denuded samples due to compound differences in sensitivity to sample concentration. Instead, a formal estimation of error from Eq. S11 is the best estimate because calculation of σ_S^i using the scenarios shown in the main text does includes error caused by differences between analyte and internal standard. From the estimates in Table 1, an internal standard containing approximately the same number of OH groups is sufficient to greatly reduce this error. However, a large suite of internal standards is recommended and is typical in SV-TAG operation, allowing relatively unbiased measurement of F_P for all compounds with a robust formally estimated error of 15-25%

It should be noted that particle fraction is calculated as the ratio of one cell to another, so a value of greater than 1 is possible due to measurement uncertainties. Uncertainties are reported as standard deviations, so greater than approximately 70% of points with a particle fraction greater than 1 should be within uncertainty of 1.00, and compounds entirely in the particular phase are expected to be measured as an approximately normal distribution around $F_P = 1$. An example of such a compound is hydroxy glutaric acid (Fig. S4), which is calculated from Eq. S11 to have an unceratinty, $\sigma_{F_P,S}$, of approximately 22% using deuterated hydroxy glutaric acid as an internal standard. The distribution of points is centered on $F_P = 1.02 \pm 0.02$ (standard error) with a distribution well-described by $\sigma_{F_P,S}$ (97% of points with $F_P > 1$ are within 2 $\sigma_{F_P,S}$). Uncertainty in the particle fraction is only a weak function of signal size as demonstrated by Fig. S4, in which larger analyte signals (green line) have a slightly narrower distribution around $F_P = 1$ than signals smaller than the internal standard (red line), which is approximately 4 times the level of quantification. Therefore, due to errors in chromatographic integration and mass spectrometric background signals, it is possible that the average estimation of error underpredicts uncertainty at signals very close to the limit of detection.

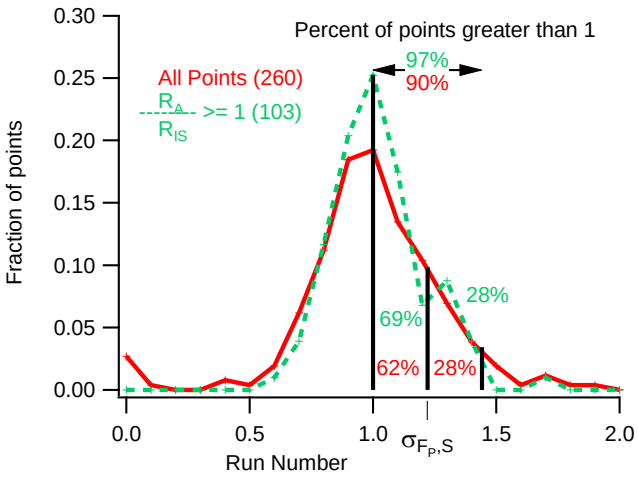

Figure S4. Histogram of fraction in particle, $F_{P,S}$, for hydroxy glutaric acid for all points (red), and for points greater than the internal standard (green). $\sigma_{F_{P,M/S}}$ = 22% from Eq. S11.

While the formal error estimate of partitioning relies on the precision of the instrument and the ability to compare the cells, systematic biases can exist that must be explored. Both penetration of gases through the denuder and loss of gases to the inlet, for instance, will bias the instrument toward a higher F_P. Using various forms of zeroes, no evidence is found in SOAS data for a non-negligible influence of either of these potential biases. However, such biases need to be considered and generally added into F_P error if appropriate.

References

Kovats, E: Gas-chromatographische Charakterisierung organischer Verbindungen. Teil 1: Retentionsindices aliphatischer Halogenide, Alkohole, Aldehyde und Ketone, Helv. Chim. Acta, 41(7), 1915-1932, doi:10.1002/hlca.19580410703, 1958

NIST/SEMATECH e-Handbook of Statistical Methods: Section 2.3.6.7.3. Comparison of check standard analysis and propagation of error. Available online at http://www.itl.nist.gov/div898/handbook/mpc/section3/mpc3673.htm, Accessed June 12, 2014a.

NIST/SEMATECH e-Handbook of Statistical Methods: Section 2.3.6.7.2. Uncertainty for linear calibration using check standards. Available online at http://www.itl.nist.gov/div898/handbook/mpc/section3/mpc3672.htm, Accessed June 12, 2014b.