https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

Atmospheric Measurement Techniques
Discussions

Quantifying the uncertainties in thermal-optical analysis of carbonaceous aircraft engine emissions: An interlaboratory study

Timothy A. Sipkens[1], Joel C. Corbin[1], Brett Smith[1], Stephanie Gagné[1], Prem Lobo[1,#], Benjamin T. Brem[2,3], Mark Johnson[4], Gregory J. Smallwood[1]

[1]Metrology Research Centre, National Research Council Canada, Canada
[2]Empa, Swiss Federal Laboratories for Materials Science and Technology, Switzerland
[3]Laboratory of Atmospheric Chemistry, Paul Scherrer Institut, Switzerland
[4]Rolls-Royce plc, Derby, United Kingdom
[#]Current address: Office of Environment and Energy, Federal Aviation Administration, Washington, D.C. 20591, USA

Correspondence to: Timothy A. Sipkens (timothy.sipkens@nrc-cnrc.gc.ca)

Abstract. Carbonaceous particles, such as soot, make up a notable fraction of atmospheric particulate matter and contribute substantially to anthropogenic climate forcing, air pollution, and human health. Thermal-optical analysis (TOA) is one of the most widespread methods used to speciate carbonaceous particles and divides total carbon (TC) into the operationally defined quantities of organic carbon (OC; carbon evolved during slow heating in an inert atmosphere) and elemental carbon (EC). While multiple studies have identified fundamental scientific reasons for uncertainty in distinguishing OC and EC, far fewer studies have reported on interlaboratory reproducibility. Moreover, existing reproducibility studies have focused on complex atmospheric samples. The real-time instruments used for regulatory measurements of aircraft engine non-volatile particulate matter (nvPM) mass emissions are required to be calibrated to the mass of EC determined by TOA of the filter-sampled emissions of a diffusion flame source. However, significant differences have been observed in the calibration factor for the same instrument based on EC content determined by different calibration laboratories. Here, we report on the reproducibility of TC, EC, and OC quantified using the same TOA protocol, instrument model (Sunset 5L), and software settings (auto split-point: Calc405) across five different laboratories and instrument operators. Six unique data sets were obtained, with one laboratory operating two instruments. Samples were collected downstream of an aircraft engine after treatment with a catalytic stripper to remove volatiles. We compared laboratory-reported uncertainties with actual variability in the data set, the difference of which (dark uncertainty) was substantial. Interlaboratory (dark) contributions increase uncertainties by a factor of 1.2 – 1.6 relative to the laboratory-reported uncertainties, even for these relatively simple samples (combustion particles downstream of a stripper), resulting in uncertainties of 26% (k = 2) for EC. Uncertainties were a little larger for EC than for OC. These results indicate that current TOA uncertainties are underestimated and should be adjusted upwards to reflect these interlaboratory differences.

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

Atmospheric Measurement Techniques Discussions

1 Introduction

Carbonaceous particles contribute to both natural and anthropogenic climate forcing, air pollution, and human health impacts. The aviation industry remains a notable source of these particles, and air transportation continues to expand. Unlike CO_2, particulate matter (PM) emissions from the aviation industry contain larger uncertainties, as is their effect on contrails and cloud formation (Righi et al., 2021). For aircraft engine emissions, thermal-optical analysis (TOA) is currently the reference standard for measuring the mass concentration of non-volatile particulate matter (nvPM) emitted by aircraft engines (SAE, 2018; Lobo et al., 2015b; Lobo et al., 2020).

In TOA, the total carbon (TC) mass on collected on a quartz filter is measured in two parts. First, the total carbon mass evolved from a sample during controlled heating in an inert environment is considered organic carbon (OC), while the remainder, heated in an oxidizing environment, is considered EC, after correction for pyrolysis (Birch and Cary, 1996). If the mass fraction of carbon in OC (40–80%; (Turpin and Lim, 2001; Bae et al., 2006)) or in EC (90–98%; (Figueiredo et al., 1999; Singh and Vander Wal, 2020; Corbin et al., 2020)) is known, these quantities can then be used to estimate the total mass of carbonaceous particles on the filter.

It is well known that the widely variable properties of carbonaceous materials leads to significant uncertainties in the separation of TC into OC and EC using TOA (Watson et al., 2005; Lack et al., 2014). In particular, inorganic carbonates may generate spurious signals; soot may partly vaporize at the OC stage; materials such as tarballs or highly-oxidized organics may generate EC signals; and inorganic compounds may catalyze the formation of EC or confound the optical quantification of pyrolysis (Corbin et al., 2020). It is also well known that different temperature ramp protocols lead to differences in the OC/EC ratio reported by TOA (e.g. Bautista et al., 2015; Schauer et al., 2003; Cavalli et al., 2010; Brown et al., 2017; Cheng et al., 2010; Giannoni et al., 2016; Wu et al., 2016; Cheng et al., 2012).

Less well studied are the uncertainties in TOA across multiple laboratories. Interlaboratory studies allow for an assessment of measurement reproducibility (changing laboratories, instruments, and operators), rather than simply repeatability (e.g., replicate measurements performed by the same operator). Here, the few reproducibility studies that exist have often focussed on atmospheric aerosols or surrogates thereof. Schmid et al. (2001) analyzed urban air pollution samples from Berlin, Germany, using 9 different protocols obtained in 17 different laboratories. They reported 7%, 9%, and 11% standard deviation between their TC measurements. Schauer et al. (2003) evaluated EC and OC reproducibility for filter samples of Asian and North American air pollution, as well as secondary organic aerosol, reporting 4–13% reproducibility for OC and 6–21% for EC. They additionally evaluated the reproducibility of the EC/OC division (split point) for various other samples, focussing on this ratio after identifying it as a major source of uncertainty. Ten Brink et al. (2004) sampled rural air pollution in Germany and analyzed the filters in four different laboratories, reporting less than 10% variability in TC and EC. Finally, in a pan-European study, Panteliadis et al. (2015) gathered results from 17 different laboratories to determine a reproducibility of 12–15% for TC, and 20–26% for EC, while Brown et al. (2017) reported a combined uncertainty of < 13 % for a reproducibility study between four laboratories. The known technical shortcomings of TOA instruments cannot explain the magnitude of these uncertainties (Boparai et al., 2008).

We note that neither Schmid et al. (2001), Ten Brink et al. (2004), nor Panteliadis et al. (2015) presented a detailed statistical analysis of OC concentrations, and reported up to a factor of two difference between OC measured by different protocols. This is related to the fact that the accurate quantification of OC in atmospheric samples is extremely difficult, due to the potential vaporization and/or adsorption of volatile organic compounds during and after sampling, especially for low filter loadings, and even when attempting to measure these artifacts (discussed below). This difficulty is one of the reasons that emissions testing protocols typically specify the removal of volatile OC by devices such as catalytic strippers, which remove all volatiles (typically at 350 °C) prior to filter collection. Consequently, any carbon measured as OC on downstream filters must either represent pyrolysis products or contamination. Importantly, Corbin et al. (2020) showed that once gas-phase contamination is accounted for, the remaining OC is also measured by in-situ (filter-free) techniques, and is therefore not a sampling or TOA artifact.

Overall, despite a very significant body of work on the fundamentals and statistical uncertainties behind TOA measurements, there have been few studies where the sample was (*i*) non-volatile, (*ii*) taken from the same or identical filter, and (*iii*) of known composition. Here, we present an intercomparison study where the same filters were punched six times for analysis by five different laboratories, after sampling aircraft engine exhaust denuded at 350 °C. Identical instruments and protocols were used. Our study provides a general estimate of the dark uncertainty (differences between laboratories that would be hidden, or dark, in measurements by a single laboratory) of TOA analyses from similar emissions tests, and a lower limit for the TOA reproducibility in atmospheric studies where additional uncertainties are introduced.

2 Methods

2.1 Experimental protocol

Sampling was performed in accordance with SAE ARP6320A (SAE, 2018), with the experimental setup shown schematically in Figure 1. Emissions were collected from the exhaust of a helicopter turboshaft engine (Olfert et al., 2017) using a single point sample probe. The sample stream was mixed with heated dilution air before passing through a catalytic stripper (Catalytic Instruments CS15). Flow was split to pass through a pair of Dekati diluters (DI-1000, operated with HEPA filtered compressed air) and a pair of cyclones, each with a 1.0 µm cutoff at 50 LPM, before being directed through a sampling manifold. Samples were distributed from the manifold to a suite of instruments, including other instruments for online mass quantification (e.g., as in Corbin et al. (2020)) and TOA. Particles for TOA were collected on quartz filters in stainless steel filter holders. The quartz filters were then sealed in Analyslide Petri dishes (28145-473) and kept at room temperature until analysis.

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

Atmospheric Measurement Techniques
Discussions

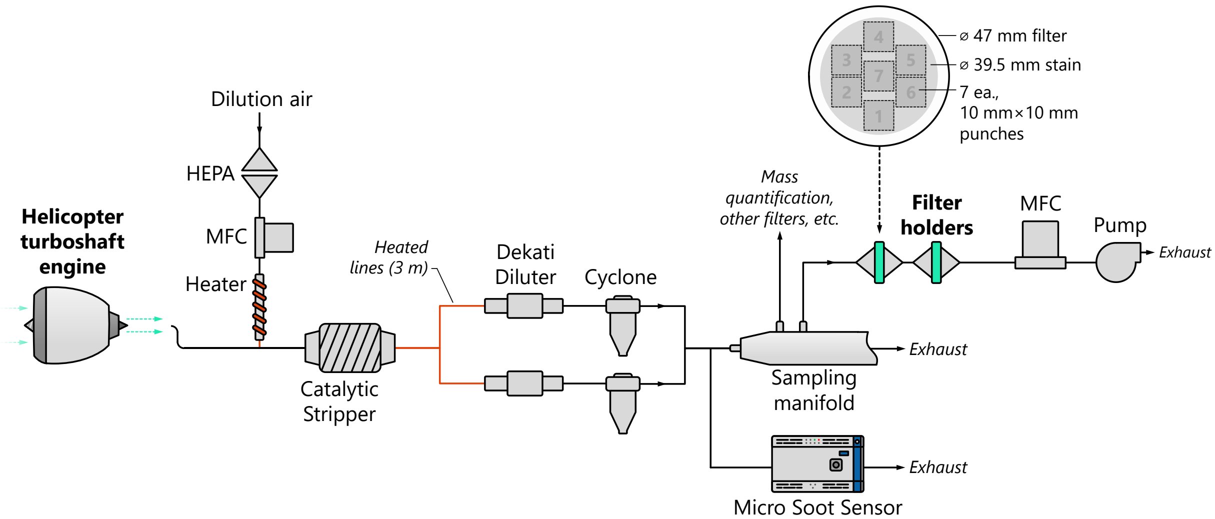

Figure 1. Schematic of the experimental setup where emissions from a helicopter turboshaft engine transit to filter holders used for thermal-optical analysis (TOA). Cyclones had a 1.0 µm cutoff at 50 LPM, with the actual sample flow rate for each cyclone being 56 LPM. MFC stands for mass flow controller, while HEPA refers to a HEPA filter. Dual diluter/cyclone system is consistent with (SAE, 2018). Inset at the top, right depicts the punch positions on the filter. Note that the angular position of the punches on the filter was not constrained.

Samples were composed of 20 filters, with five each sampling at approximately 50, 100, 250, and 500 µg/m^3 (based on measurements made by a AVL Micro Soot Sensor on a separate parallel line connected to the sampling manifold). Sampling times were adjusted to keep filter loadings approximately constant. The engine was operated at the same condition in all of the cases, except the 500 µg/m^3. For that highest sample loading, a higher engine RPM was required to generate sufficient nvPM mass concentration. Five laboratories which were compliant with ISO 17025 (demonstrating competence) for TOA were selected for this intercomparison. Each of the laboratories was instructed to take one (or two, in the case of one laboratory) punches from each of the twenty filters. Seven punches were possible on each filter with an allowance of one spare punch per filter in addition to one (or two) per laboratory, arranged in a ring of six with one central punch as shown in the inset to Figure 1. Punch positions on each sample were implicitly randomized by not otherwise providing further instruction to the laboratories. While this introduces a slight risk in the case of uneven filter loading, symmetry in the sampler and random filter orientations would minimize such risks in all but the center punch. Moreover, the darkness of most filters was visually homogeneous. We therefore treat inter-filter variability as negligible in our analysis below, which is supported by observations.

The protocol for aircraft engine emissions, a refinement based upon NIOSH 5040 (SAE, 2018; Lobo et al., 2015a) with the final oxidizing temperature step at a higher temperature of 930 C and a longer duration to ensure complete oxidation of the particles was used to perform the analysis, with the EC/OC split determined automatically by the instrument software (Sunset, Calc405). The protocol and sample data are shown in Figure 2.

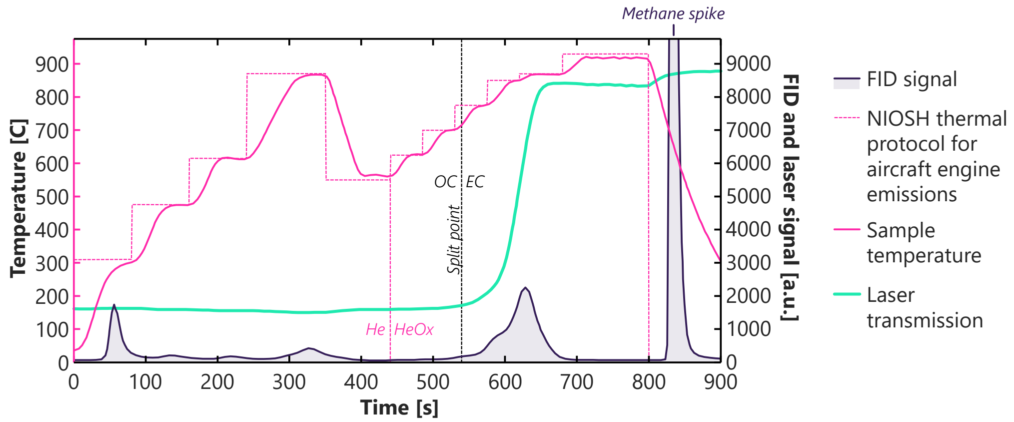

Figure 2. Sample TOA thermogram alongside the thermal protocol for aircraft engine emissions (SAE, 2018; Lobo et al., 2015a), sample temperature, and laser transmission measurement. HeOx corresponds to 2% oxygen in He. FID is a flame ionization detector. The methane spike corresponds to the introduction of methane that is used for calibration after analysis.

Of the six sets of measurements considered, two belonged to a single laboratory and analyst, and are denoted in subsequent figures and discussion as Laboratory 1A and 1B. The remaining laboratories contributed a single set of data and are numbered in ascending order in terms of the average EC measurement across all the filters. (Two of the laboratories were commercial service providers and did not contribute scientifically to the work.)

2.2 Statistical treatment

Results are analyzed using a hierarchical random effects model (e.g., as in Melanson et al. (2018)). In this framework, measurements, y_{ij}, are modeled as a combination of effects:

$$y_{ij} = c_i + l_{ij} + \varepsilon_{ij} \tag{1}$$

where i and j denote the ith filter and jth laboratory. Thus, c_i is a filter-specific consensus value, that represents the expected value of a quantity. The l_{ij} is the effect for the jth laboratory, which represents a systematic shift or mean bias in the measurements measured by that laboratory for the ith filter. Thus, l_{ij} will be realized as a positive value if a laboratory has a bias above the consensus value and vice versa. If there is no such bias in a laboratory, l_{ij} will be realized as zero. The remaining term, ε_{ij}, represents the additional random error in the individual measurements reported by each laboratory, i.e., the mismatch between the expected laboratory bias and the actual measurement. A given laboratory reported their measurements, denoted as y_{ij}, and the their uncertainty, denoted as $u_{\varepsilon,ij}$. The uncertainty values reported by the different laboratories are automatically generated by Sunset's analysis software, which is a combination of the limit of detection of the instrument (0.2 μg/cm^2) and a percentage based upon statistical analysis of duplicate filter punches (±5%). This model is shown schematically in Figure 3.

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

Atmospheric Measurement Techniques
Discussions
Open Access

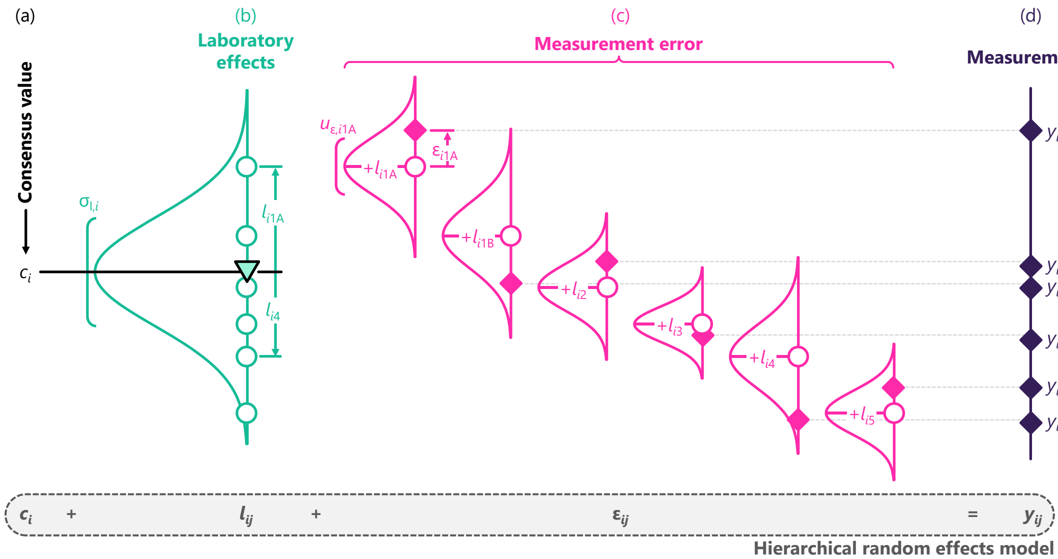

Figure 3. Schematic demonstrating the hierarchical random effects model used in the present analysis. Data is fictional, intended for demonstration only.

Here, these effects (bias and random errors) are perceived as random variables, holding some distribution (i.e., a *Bayesian* representation), with the terms above representing realizations of these random variables for a specific set of measurements. Thus, the set of l_{ij} for the ith filter represents the biases across the laboratories, such that the corresponding standard deviation of this set, $\sigma_{l,i}$, represents *dark* (Thompson and Ellison, 2011), interlaboratory uncertainties and in our study is an assessment of *reproducibility* for a given filter. If the laboratories measured identical values, $\sigma_{l,i} = 0$ and the measurements are considered entirely reproducible. Note that the standard deviation of the set of l_{ij} for the jth laboratory across the filters, i.e., $\{\sigma_{l,j}\}$, represents the equivalent experiment if performed by a single laboratory. This quantity would be a measure of variations in loading combined with variability within a given laboratory (without including interlaboratory variability).

The distribution of these various quantities was determined using a Markov Chain Monte Carlo (MCMC) approach, similar to the method presented by Melanson et al. (2018). MCMC seeks to find the range of inputs, in this case the magnitude of various effects and dark uncertainties, that would cause the distribution of the outputs, namely the observed measurements. Here, we assume that the laboratory effects are normally distributed,

$$l_{ij} \sim \mathcal{N}\left(0, \sigma_{l,i}^2\right), \tag{2}$$

where "~" denotes distributed as, $\mathcal{N}(\mu, \sigma^2)$ denotes a normal distribution with a mean μ and a variance σ^2, and $u_{l,i}^2$ is the variance of the set of measurements, y_{ij}, across the laboratories for the ith filter. We also assume that the measurement errors are unbiased (i.e., have a mean of zero) and are normally distributed, with a standard deviation corresponding to that reported by each laboratory:

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

Atmospheric Measurement Techniques
Discussions

$$\varepsilon_{ij} \sim \mathcal{N}\left(0, u^2_{\varepsilon,ij}\right). \tag{3}$$

MCMC uses this as an input and estimates the consensus value, the value of l_{ij} for each laboratory, and the value of $\sigma_{l,i}$ for each filter. To restrict the solution space and improve convergence with this Bayesian approach, we also apply priors (encoding approximate information known before the statistical analysis) to these quantities, which are summarized in Table 1. To minimize the impact of the burn-in period of the MCMC, the set of l_{ij} were initiated about y_{ij}. A total of 25,000 samples were generated, after thinning the MCMC data by a factor of 20 (to avoid short range correlation in the samples) spread across four independent chains. MCMC samples were realized using the Just Another Gibbs Sampler (JAGS) code (Hornik et al., 2003). Visual inspection of the samples indicated that the chains had converged. Further increasing the number of samples did not have an impact on the statistical outcomes.

Table 1. Table of quantities related to the statistical treatment, including those on which likelihood and priors are directly stated. Note that measurements, y_{ij}, are the input to the MCMC procedure and are thus not sampled.

Quantity	Effect symbol	Variance	Quantities sampled	Likelihood	Prior
Filter consensus value	c_i	$\sigma_{c,i}^2$	Effect	-	$c_i \sim \mathcal{N}(\overline{c}_i, \overline{c}_i^2)$
Laboratory effect (bias)	l_{ij}	$\sigma_{l,ij}^2$	Effect	$l_{ij} \sim \mathcal{N}(0, \sigma_{l,i}^2)$	$c_i + l_{ij} \sim \mathcal{N}(y_{ij}, u_{\varepsilon,ij}^2)$
			Variance	-	$\sigma_{l,i} \sim t(1, 0, \sigma_{l,i}^2)^\dagger$
Measurement error	ε_{ij}	$u_{\varepsilon,ij}^2$	-	$\varepsilon_{ij} \sim \mathcal{N}(0, u_{\varepsilon,ij}^2)$	-
Measurement	y_{ij}	- (fixed)	-	-	-

† $t(\mathrm{df}, \mu, \sigma^2)$ denotes a non-standardized Student's t distribution where df, μ, and σ^2 are the degrees of freedom, mean, and variance, respectively. The variant here corresponds to the half Cauchy distribution, where df = 1 and $\mu = 0$.

Outlier results were identified at the outset using a generalized extreme Studentized deviate (GESD) test with a threshold factor of 0.15 on a per filter basis.

This model is applied for each of the twenty filters separately, allowing for a filter-specific consensus value, and is then repeated for each of EC, OC, and TC. Inter-filter (or between-group) uncertainties are post-processed from the MCMC results.

3 Results

3.1 EC, OC, and TC

Figure 4 shows a sample of the results for EC for 5 of the 20 filters, corresponding to the 100 μg/m^3 mass concentration case. In Figure 4, laboratories are ordered according to the median EC measured over all 20 filters. This order is roughly respected across all of the filters. Results are generally consistent with the remaining 15 filters, not shown, though filter-to-filter differences in some cases exceeded the range shown in this subset. Uncertainties did not show any structure with mass

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

concentration or the measured value for EC, OC, or TC. Uncertainties, alongside their decomposition into their respective components, are reported in Table 2.

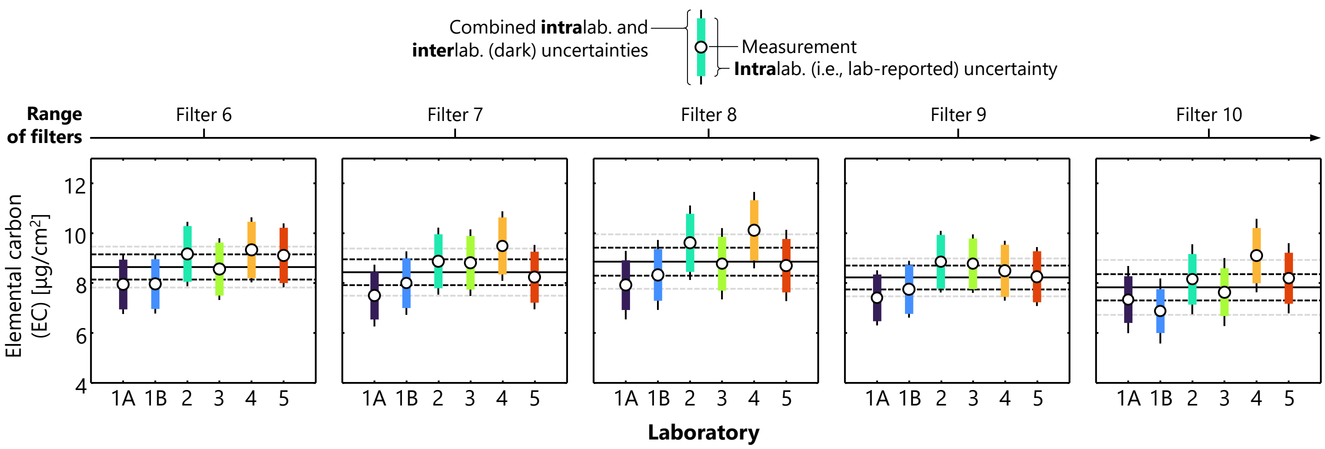

Figure 4. Sample results across a range of filters. Circles correspond to laboratory-reported data. Wide, coloured bars correspond to laboratory-reported uncertainties, $u_{\varepsilon,ij}$, while whiskers correspond to both laboratory-reported and interlaboratory (dark) uncertainties, combined in quadrature as $(u_{\varepsilon,ij}^2 + \sigma_{l,i}^2)^{1/2}$. Horizontal, solid lines correspond to the consensus value for a given filter, while black dashed lines correspond to uncertainty intervals (k = 2) in this consensus value. Dotted grey lines add the dark, interlaboratory uncertainties for a given filter to these uncertainties (reproducibility). Results shown are for elemental carbon and the 100 μg/m³ case. Vertical axes are identical across all of the panels.

Figure 5 shows a decomposition of the different kinds of intra-filter uncertainties present in the measurements, averaged over all of the filters and laboratories and presented as a proportion of the observed variance. For each filter, uncertainties are typically dominated by the uncertainties reported by the laboratory, $u_{\varepsilon,ij}$, which are relatively consistent across all of the measurements. Some exceptions existed for individual filters, namely filters 3 and 14 for EC and filter 16 for OC. These filters coincide with cases where the overall variance is larger and represent a minority of cases. Despite this observation, dark, interlaboratory uncertainties are still significant, expanding reported intralaboratory uncertainties by factors of 1.59, 1.24, and 1.20 for EC, OC, and TC, respectively (shown visually in Figure 5). Table 2 indicates that inter-filter variability adds to these intra-filter uncertainties, where OC is dominated by intra-filter uncertainties, while EC and TC are dominated by inter-filter variability.

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

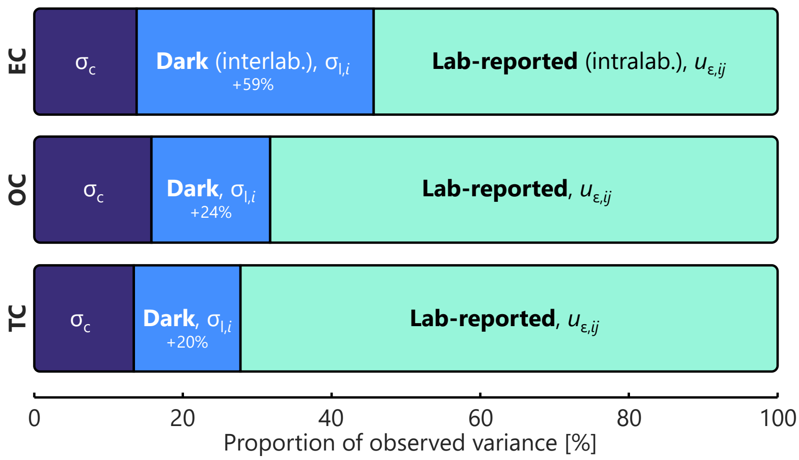

Figure 5. Breakdown of the intra-filter variance in the TOA measurements for the uncertainties in the consensus value, σ_c; dark (interlaboratory) uncertainties, $\sigma_{l,i}$; and uncertainties reported by the laboratory, $u_{\varepsilon,ij}$, stated as a proportion of the overall intra-filter variance, such that all span 0 – 100 %. See Table 2 for the numerical values of the uncertainties. Percentages in the dark bar correspond to the required increase in uncertainties over the lab-reported values to account for dark uncertainties.

Table 2. Breakdown of uncertainties in the TOA measurements, stated as uncertainties relative to the nominal value of EC, OC, and TC (k = 2). Total uncertainties, σ_{tot}, are a combination of the intra- ($u_{\varepsilon,ij}$) and inter-filter (σ_f) uncertainties (in quadrature: $\sigma_{tot}^2 = u_{\varepsilon,ij}^2 + \sigma_f^2$). Intra-filter uncertainties correspond to the sum of the corresponding rows (also in quadrature). The bottom row, $\sigma_{l,j}$, corresponds to the uncertainties that would be reported by a single lab based on hypothetical replicate measurements of multiple similar filters, and is estimated from the MCMC calculations. This quantity is not included in the overall total, as this would double count uncertainties present on other rows (specifically, σ_f).

		Coefficient of variation* [%]		
Uncertainty component	Symbol	EC	OC	TC
Intra-lab. (lab-reported)	$u_{\varepsilon,ij}$	12.6	14.2	13.2
Inter-lab. (dark)	$\sigma_{l,i}$	9.6	7.0	5.8
Consensus	$\sigma_{c,i}$	6.4	6.8	5.6
Intra-filter† (reproducibility)	σ_{intra}	17.0	17.4	15.4
Inter-filter	σ_f	20.8	9.4	14.8
Total	σ_{tot}	26.8	19.6	21.4
Inter-filter, in-lab. (repeatability)	$\sigma_{l,j}$	22.8	16.0	15.6

† Intra-filter uncertainties are taken as average values of the variance over all of the filters. This same averaging is then propagated to the total uncertainties. *Coefficient of variation are stated using mean EC, OC, and TC measurements of 8.1, 4.7, and 12.9 μg/cm^2.

Figure 6 complements Figure 5 with a plot of the measurements for each laboratory across the filters, with the filters sorted in ascending order for each of EC, OC, and TC, such that the filter order differs between the panels. Results for EC and TC

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

Atmospheric Measurement Techniques Discussions

exhibit substantial structure across the different filters, where a laboratory that measured a value below average generally did so for all of the filters. Trends for TC were similar, which is not surprising since EC concentrations were typically double OC concentrations. In other words, the laboratories showed structured biases for these two quantities. Structural trends were less evident for OC, which is more scattered about the central trend.

This structure becomes particularly relevant when considering repeats within a given laboratory (intralaboratory) for uncertainty analysis. Repeatability in TC is similar to the combined uncertainties reported by Conrad and Johnson (2019), in that work across a range of conditions. Ideally, if a measurement approach is entirely reproducible, variability in repeat measurements, $\sigma_{l,j}$, will capture all of the variability. However, since repeats cannot capture interlaboratory reproducibility, repeat measurements performed by the same laboratory result in underestimated intra-filter uncertainties, as shown in the final row of Table 2.

These structural trends may give insight into the physical causes of these uncertainties. For example, minor biases in calibration would lead to the observed structural errors, while random operator error would not. Other potential sources of error (e.g., in terms of FID response) have been discussed in detail elsewhere (Boparai et al., 2008). In this data, Laboratory 5 produced EC and TC values consistently above the other laboratories, and Laboratory 1 (in both the 1A and 1B samples) produced EC and TC values consistently below the other laboratories.

For these aviation particulate emissions samples, overall uncertainties in EC, OC, and TC are around 20% (26.8%, 19.6%, and 21.4%, respectively, $k = 2$) of the nominal values for the full set of measurements. The overall EC uncertainty is consistent with the results of Panteliadis et al. (2015) which were evaluated on ambient atmospheric PM samples.

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

Atmospheric Measurement Techniques
Discussions

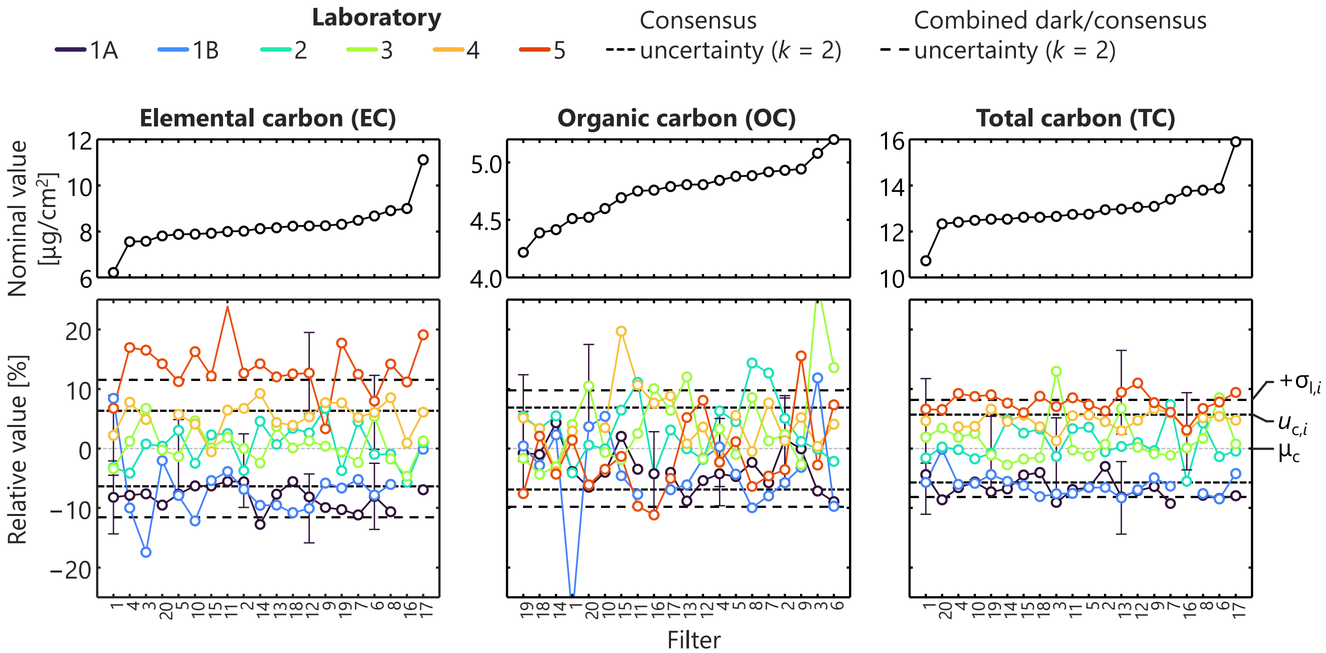

Figure 6. Laboratory measurements across the different filters, showing the structured uncertainties. In each case, filters are re-sorted such that the consensus values for each filter are monotonically increasing for each of the quantities. As such, the order of filters is not the same across the panels. Upper panels show the consensus values for the different filters. Bottom panels show measurements from each laboratory normalized by those consensus values. Breaks in lines correspond to results that were not available. Vertical scales are the same in the lower panels. Error bars are excluded for clarity but would offer some overlap between the laboratories.

3.2 Analysis across EC, OC, and TC values

Little to no correlation was observed between EC and OC measured by the different laboratories (R = 0.11), while TC was dominated by, and thus highly correlated with, the EC contributions (R = 0.94). Combining this with the fact that the measured EC showed structural bias across the laboratories, it is logical that this is equally reflected in TC results. OC and TC were poorly correlated (R = 0.43), given that the TC incorporates but is not dominated by OC. The low level of correlation between EC and OC indicates that the split point is unlikely to be the leading driver of variabilities in the results, as this would result in a negative correlation between EC and OC, where more of the total carbon is attributed to one of the two components at the cost of the other.

Unlike the absolute values for EC, OC, and TC, the EC/OC ratio is expected to be similar across all of the filters, regardless of loading and is a widely used quantity for characterizing the particles emitted. For the EC/OC ratio, simple propagation of errors yields (Sipkens et al., 2023; Jcgm, 2008):

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

Atmospheric Measurement Techniques Discussions
Open Access

$$\mathrm{var}(\mathrm{EC/OC}) = \left(\frac{\mathrm{EC}}{\mathrm{OC}}\right)^2 \left[\frac{1}{(\mathrm{EC})^2}\mathrm{var}(\mathrm{EC}) + \frac{1}{(\mathrm{OC})^2}\mathrm{var}(\mathrm{OC}) - \frac{2}{(\mathrm{EC})(\mathrm{OC})}\mathrm{cov}(\mathrm{EC},\mathrm{OC})\right]. \quad (4)$$

As noted above, EC and OC are not significantly correlated for these measurements, such that the covariance term can be neglected. This produces a relatively uniform estimate of ± 0.43 (k = 2) for the variance in the EC/OC ratio across all of the measurements, around 25% of the nominal value. If only laboratory-reported uncertainties are considered, this reduces to ± 0.33 (k = 2), resulting in an underestimation in the variance of the EC/OC ratio by a factor of 0.65. Inter-filter variability was significant, adding ± 0.39 (k = 2) in addition to the laboratory-stated uncertainty. Overall, uncertainties in the EC/OC ratio are roughly evenly distributed in their source between contributions from the EC measurements, the OC measurements, and filter-to-filter variability, resulting in and EC/OC ratio of 1.74 ± 0.58 (k = 2) for the full set of measurements, such that the uncertainty in the EC/OC ratio is ~ 33 % of the nominal value. Note that this is larger than the uncertainties in the individual EC and OC measurements, as it incorporates uncertainties in both EC and OC at the same time. Again, as noted above, repeat measurements over different filters taken within the same laboratory, $\sigma_{l,j}$, would result in an underestimation of the overall uncertainties, as it is impossible for a single laboratory to determine its own reproducibility.

In this data, Laboratory 5 produced an EC/OC value consistently above the other laboratories, a consequence of measuring higher than average EC in combination with a generally lower than average OC. There was also some trend in EC/OC with mass concentration and sampling time, for all laboratories, indicated by the generally increasing filter number on the x-axis of Figure 7. This results from a similar slight increase in EC and a slight decrease in OC as the sampling period decreases. Since this effect was minor, our preceding discussion summarized the data using the means of EC, OC, TC, EC/OC ratio.

Atmospheric Measurement Techniques
Discussions

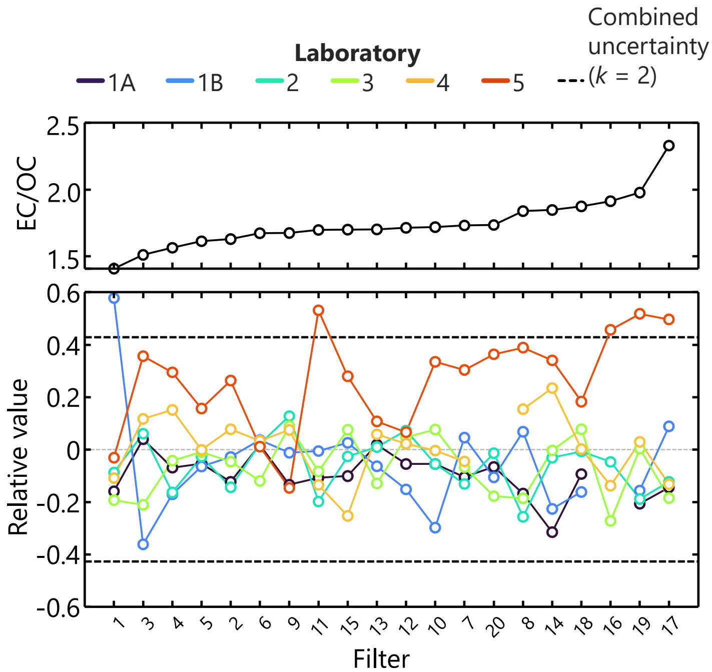

Figure 7. Variability in the EC/OC ratio over the measurements. Uncertainty interval corresponds to the average propagated intra-filter uncertainties at a level of *k* = 2. Upper panel shows consensus EC/OC values against which the filters are sorted. Bottom panel shows measurements from each laboratory normalized by those consensus values.

4 Conclusions

This work investigated the dark, interlaboratory uncertainties associated with thermal-optical analysis (TOA) applied to aircraft engine particulate emissions. These conditions represent optimal samples for TOA, in that they are primarily composed of combustion particles that are stripped of their volatile components. Nevertheless, uncertainties are poorly captured by existing estimates for these measurements, in all cases except perhaps for measurements of OC. Generally, uncertainties need to be expanded by a factor of 1.2 for OC and TC and a factor of 1.6 for EC to account for dark, interlaboratory uncertainties. Uncertainties are not expected to be related to the split point, due to a lack of correlation between EC and OC (where a reduction in OC results in an increase in EC).

EC and TC measurements are also highly correlated with the laboratory (i.e., reflected by a fixed bias), with some laboratories consistently measuring results above and some below the average. These structured laboratory biases suggest a potential link to laboratory-specific calibration that affects the EC (and, by extension, the TC) measurement. Replicates, that is repeat measurements by a single laboratory, are unlikely to properly capture these uncertainties, given structure between laboratories. Fortunately, this could be accounted for by combining uncertainty procedures implemented by laboratories with expanded uncertainties to include an interlaboratory contributions, as noted in this work. For data sets comparable to ours (i.e., PM dominated by soot, treated to remove volatile organic carbon, and containing negligible elemental impurities), net uncertainties of 26% (k = 2) for EC, 20% for OC and TC (k = 2), and 33% (k = 2) for the EC/OC ratio are expected, which accounts for an expansion of the uncertainty bounds to account for reproducibility. This expanded uncertainty should be used in future measurements with this test method.

The treatment in this work does not directly question the interpretation of OC and EC concentrations reported by TOA, nor does this work evaluate the accuracy of the TOA TC concentration (e.g., by indicating traceability to an SI unit). Rather, this work addresses metrological reproducibility of the TOA method by comparing results from the same sample, measured by different laboratories and analysts.

Data availability

Data will be made available on request.

Author contribution

Timothy A. Sipkens: Data Curation, Formal analysis, Methodology, Software, Writing - Original draft, Writing - Review & Editing, Visualization. **Joel C. Corbin**: Conceptualization, Investigation, Writing - Review & Editing. **Brett Smith**: Investigation, Writing - Review & Editing. **Stephanie Gagné**: Investigation, Writing - Review & Editing. **Prem Lobo**: Supervision, Project administration, Writing - Review & Editing. **B.T. Brem**: Investigation, Writing - Review & Editing. **A. Fischer**: Investigation, Writing - Review & Editing. **M. Johnson**: Investigation, Writing - Review & Editing. **Gregory J. Smallwood**: Conceptualization, Investigation, Methodology, Project administration, Supervision, Writing - Review & Editing.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We are grateful to the Rolls Royce Team for their support in obtaining these samples. This work was supported by Transport Canada Aviation (A1-023090). Filter analysis was supported in part by the Swiss Federal Office of Civil Aviation Project "EMPAIREX" SFLV- 2015-113.

References

Bae, M.-S., Schauer, J. J., and Turner, J. R.: Estimation of the monthly average ratios of organic mass to organic carbon for fine particulate matter at an urban site, Aerosol Science and Technology, 40, 1123-1139, 2006.

Bautista, A. T., Pabroa, P. C. B., Santos, F. L., Quirit, L. L., Asis, J. L. B., Dy, M. A. K., and Martinez, J. P. G.: Intercomparison between NIOSH, IMPROVE_A, and EUSAAR_2 protocols: Finding an optimal thermal–optical protocol for Philippines OC/EC samples, Atmos. Pollut. Res., 6, 334–342, 10.5094/apr.2015.037, 2015.

Birch, M. and Cary, R.: Elemental carbon-based method for monitoring occupational exposures to particulate diesel exhaust, Aerosol Science and Technology, 25, 221-241, 1996.

Boparai, P., Lee, J., and Bond, T. C.: Revisiting thermal-optical analyses of carbonaceous aerosol using a physical model, Aerosol Science and Technology, 42, 930-948, 2008.

Brown, R. J. C., Beccaceci, S., Butterfield, D. M., Quincey, P. G., Harris, P. M., Maggos, T., Panteliadis, P., John, A., Jedynska, A., Kuhlbusch, T. A. J., Putaud, J.-P., and Karanasiou, A.: Standardisation of a European measurement method for organic carbon and elemental carbon in ambient air: results of the field trial campaign and the determination of a measurement uncertainty and working range, Environ. Sci. Processes Impacts, 19, 1249–1259, 10.1039/c7em00261k, 2017.

Cavalli, F., Viana, M., Yttri, K. E., Genberg, J., and Putaud, J. P.: Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the {EUSAAR} protocol, Atmos. Meas. Tech., 3, 79-89, 10.5194/amt-3-79-2010, 2010.

Cheng, Y., Duan, F.-k., He, K.-b., Du, Z.-y., Zheng, M., and Ma, Y.-l.: Intercomparison of thermal-optical method with different temperature protocols: Implications from source samples and solvent extraction, Atmos. Environ., 61, 453–462, 10.1016/j.atmosenv.2012.07.066, 2012.

Cheng, Y., He, K. B., Duan, F. K., Zheng, M., Ma, Y. L., Tan, J. H., and Du, Z. Y.: Improved measurement of carbonaceous aerosol: evaluation of the sampling artifacts and inter-comparison of the thermal-optical analysis methods, Atmos. Chem. Phys., 10, 8533–8548, 10.5194/acp-10-8533-2010, 2010.

Conrad, B. M. and Johnson, M. R.: Split point analysis and uncertainty quantification of thermal-optical organic/elemental carbon measurements, JoVE (Journal of Visualized Experiments), e59742, 2019.

Corbin, J. C., Moallemi, A., Liu, F., Gagné, S., Olfert, J. S., Smallwood, G. J., and Lobo, P.: Closure between particulate matter concentrations measured ex situ by thermal–optical analysis and in situ by the CPMA–electrometer reference mass system, Aerosol Science and Technology, 54, 1293-1309, 2020.

Figueiredo, J. L., Pereira, M., Freitas, M., and Orfao, J.: Modification of the surface chemistry of activated carbons, carbon, 37, 1379-1389, 1999.

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

Giannoni, M., Calzolai, G., Chiari, M., Cincinelli, A., Lucarelli, F., Martellini, T., and Nava, S.: A comparison between thermal-optical transmittance elemental carbon measured by different protocols in PM2.5 samples, Sci. Total Environ., 571, 195–205, 10.1016/j.scitotenv.2016.07.128, 2016.

Hornik, K., Leisch, F., Zeileis, A., and Plummer, M.: JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling, Proceedings of DSC,

JCGM: Evaluation of measurement data — Guide to the expression of uncertainty in measurement, BIPMJCGM 100:2008, 2008.

Lack, D. A., Moosmüller, H., McMeeking, G. R., Chakrabarty, R. K., and Baumgardner, D.: Characterizing elemental, equivalent black, and refractory black carbon aerosol particles: a review of techniques, their limitations and uncertainties, Anal Bioanal Chem, 406, 99-122, 2014.

Lobo, P., Hagen, D. E., Whitefield, P. D., and Raper, D.: PM emissions measurements of in-service commercial aircraft engines during the Delta-Atlanta Hartsfield Study, Atmospheric Environment, 104, 237-245, 10.1016/j.atmosenv.2015.01.020, 2015a.

Lobo, P., Durdina, L., Brem, B. T., Crayford, A. P., Johnson, M. P., Smallwood, G. J., Siegerist, F., Williams, P. I., Black, E. A., and Llamedo, A.: Comparison of standardized sampling and measurement reference systems for aircraft engine non-volatile particulate matter emissions, Journal of Aerosol Science, 145, 105557, 2020.

Lobo, P., Durdina, L., Smallwood, G. J., Rindlisbacher, T., Siegerist, F., Black, E. A., Yu, Z., Mensah, A. A., Hagen, D. E., and Miake-Lye, R. C.: Measurement of aircraft engine non-volatile PM emissions: Results of the aviation-particle regulatory instrumentation demonstration experiment (A-PRIDE) 4 campaign, Aerosol Science and Technology, 49, 472-484, 2015b.

Melanson, J. E., Thibeault, M.-P., Stocks, B. B., Leek, D. M., McRae, G., and Meija, J.: Purity assignment for peptide certified reference materials by combining qNMR and LC-MS/MS amino acid analysis results: application to angiotensin II, Anal Bioanal Chem, 410, 6719-6731, 2018.

Olfert, J. S., Dickau, M., Momenimovahed, A., Saffaripour, M., Thomson, K., Smallwood, G., Stettler, M. E. J., Boies, A., Sevcenco, Y., Crayford, A., and Johnson, M.: Effective density and volatility of particles sampled from a helicopter gas turbine engine, Aerosol Science and Technology, 51, 704-714, 10.1080/02786826.2017.1292346, 2017.

Panteliadis, P., Hafkenscheid, T., Cary, B., Diapouli, E., Fischer, A., Favez, O., Quincey, P., Viana, M., Hitzenberger, R., Vecchi, R., Saraga, D., Sciare, J., Jaffrezo, J. L., John, A., Schwarz, J., Giannoni, M., Novak, J., Karanasiou, A., Fermo, P.,

https://doi.org/10.5194/amt-2024-1
Preprint. Discussion started: 27 February 2024

and Maenhaut, W.: ECOC comparison exercise with identical thermal protocols after temperature offset correction – instrument diagnostics by in-depth evaluation of operational parameters, Atmos. Meas. Tech., 8, 779–792, 2015.

Righi, M., Hendricks, J., and Beer, C. G.: Exploring the uncertainties in the aviation soot–cirrus effect, Atmospheric Chemistry and Physics, 21, 17267-17289, 2021.

SAE: Procedure for the Continuous Sampling and Measurement of Non-Volatile Particulate Matter Emissions from Aircraft Turbine Engines (SAE ARP6320), Warrendale, PA, USA, 10.4271/ARP6320., 2018.

Schauer, J. J., Mader, B., Deminter, J., Heidemann, G., Bae, M., Seinfeld, J. H., Flagan, R., Cary, R., Smith, D., and Huebert, B.: ACE-Asia intercomparison of a thermal-optical method for the determination of particle-phase organic and elemental carbon, Environmental science & technology, 37, 993-1001, 2003.

Schmid, H., Laskus, L., Abraham, H. J., Baltensperger, U., Lavanchy, V., Bizjak, M., Burba, P., Cachier, H., Crow, D., and Chow, J.: Results of the "carbon conference" international aerosol carbon round robin test stage I, Atmospheric environment, 35, 2111-2121, 2001.

Singh, M. and Vander Wal, R. L.: The role of fuel chemistry in dictating nanostructure evolution of soot toward source identification, Aerosol Science and Technology, 54, 66-78, 2020.

Sipkens, T. A., Corbin, J. C., Grauer, S. J., and Smallwood, G. J.: Tutorial: Guide to error propagation for particle counting measurements, Journal of Aerosol Science, 106091, 2023.

Ten Brink, H., Maenhaut, W., Hitzenberger, R., Gnauk, T., Spindler, G., Even, A., Chi, X., Bauer, H., Puxbaum, H., and Putaud, J.-P.: INTERCOMP2000: the comparability of methods in use in Europe for measuring the carbon content of aerosol, Atmospheric Environment, 38, 6507-6519, 2004.

Thompson, M. and Ellison, S. L.: Dark uncertainty, Accreditation and Quality Assurance, 16, 483-487, 2011.

Turpin, B. J. and Lim, H.-J.: Species contributions to PM2. 5 mass concentrations: Revisiting common assumptions for estimating organic mass, Aerosol Science & Technology, 35, 602-610, 2001.

Watson, J. G., Chow, J. C., and Chen, L.-W. A.: Summary of organic and elemental carbon/black carbon analysis methods and intercomparisons, Aerosol and Air Quality Research, 5, 65-102, 2005.

Wu, C., Huang, X. H. H., Ng, W. M., Griffith, S. M., and Yu, J. Z.: Inter-comparison of NIOSH and IMPROVE protocols for OC and EC determination, Atmos. Meas. Tech., 9, 4547–4560, 10.5194/amt-9-4547-2016, 2016.