Supplemental information for "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], Stéphanie 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)

S1. ISO method

A standard approach to interpreting interlaboratory comparison data is outlined in ISO 5725-2 – Basic method for the determination of repeatability and reproducibility of a standard measurement method (ISO, 2019), as was applied to TOA data by Panteliadis et al. (2015). As ISO 5725-2 was used in a previous ILC for TOA, and may be more familiar to some researchers, the analysis in this study was repeated using the ISO 5725-2 method and compared below to the MCMC method used in the text. There are two significant caveats to the use of the ISO method here.

(1) Given the destructive nature of the test, there is only a single measurement for each laboratory-filter combination, such that $n_{ij} = 1$. Note that this n_{ij} applies to all of the measurements, such that the nomenclature can be simplified to $n_{ij} = n = 1$. This requires some adaptation to apply the ISO method.

(2) As noted in the main text, the laboratory-reported $\tilde{s}_{ij}$ represents a broader set of uncertainties for the method derived by the instrument manufacturer, rather than a proper measure of repeatability within a laboratory. However, in the absence of a recommendation to calculate the repeatability, the laboratory-reported uncertainties are used instead.

Then, each filter is considered as a separate *level*, consistent with the treatment of Panteliadis et al. (2015). The applicable random effects model when $n_{ij} = 1$ is given by Eq. (2) in the main text.

$$Y_{ij} = F_j + L_{ij} + E_{ij} \tag{1}$$

where F_j is equivalent to the general mean, m, in the ISO standard and L_{ij} is equivalent to B.

Proceeding with the ISO method, the filter effect is considered as fixed for each filter and is the expected value for a given filter,

$$\hat{m}_j = \frac{1}{p}\sum_{i=1}^{p} y_{ij} \quad , \tag{2}$$

where, as noted above, $n_{ij} = n = 1$ disappears given that it is a constant for all of the measurements and $p = 6$ is the number of laboratories. The within-laboratory variance is then taken as the arithmetic mean of the laboratory-reported uncertainties,

$$\tilde{s}^2_{\mathrm{r},j} = \frac{1}{p}\sum_{i=1}^{p} \tilde{s}^2_{ij} \quad , \tag{3}$$

which represents a simplification of Eq. (25) in ISO 5725-2 under the assumption of the same number of measurements by each laboratory, avoiding the singularity caused by $n - 1 = 0$. We also make use of the symbol $\tilde{s}_{\mathrm{r},j}$ to denote that this is not a true measure of repeatability. The remaining between-laboratory variance is given by Eq. (26) in ISO 5725-2. Given that $n =$ 1,

$$\tilde{s}^2_{\mathrm{L},j} = s^2_{\mathrm{d},j} - \tilde{s}^2_{\mathrm{r},j} \quad , \tag{4}$$

where

$$s^2_{\mathrm{d},j} = \frac{1}{p-1}\sum_{i=1}^{p} \left(y_{ij} - \hat{m}_j \right)^2 \quad . \tag{5}$$

Similar to $\tilde{s}_{\mathrm{r},j}$, this between-laboratory variance is not a true measure of the differences between laboratories. Rather, it is a measure of the additional uncertainties that are not already accounted for in the laboratory-reported $\tilde{s}_{ij}$. As these uncertainties would be hidden outside of an interlaboratory study, we rather use the term *dark* to describe them here (Thompson and Ellison, 2011). To differentiate the two, we use the term dark here instead of between-laboratory uncertainties. Note that dark uncertainties cannot be negative. As such, individual filters where the laboratory-reported uncertainties exceed the observed variance result in zero dark uncertainties. When averaging over all of the filters, this tends to increase the overall uncertainties relative to the MCMC method. Finally, the *reproducibility* variance is taken as

$$s^2_{\mathrm{R},j} = \tilde{s}^2_{\mathrm{L},j} + \tilde{s}^2_{\mathrm{r},j} \quad , \tag{6}$$

representing the combined within- and between-laboratory uncertainties, which should be similar to the reproducibility estimated using the method in the text (it is only the decomposition of the uncertainties that should be substantially different).

For EC and TC, Table S1 indicates that reproducibility is very similar between the ISO and MCMC methods. For OC, the reproducibility is substantially worse for the ISO method, likely due to a combination of high laboratory-reported uncertainties (which exceed the reproducibility from the MCMC method) and the constraint for non-negative dark uncertainties (which happens more often for OC than for EC and TC). Laboratory-reported uncertainties are more significant than the within-laboratory reported uncertainties determined by the MCMC method, in particular for EC and TC. This is due to the consistent bias observed in the data, which causes the measurements to be more consistent for a single laboratory than the laboratory-reported uncertainties would indicate. For EC and TC, laboratory-reported uncertainties remain insufficient to accommodate the overall variance, resulting in significant dark uncertainties.

Table S1. Breakdown of uncertainties in the TOA measurements, stated as expanded (*k* = 2) coefficients of variation based on the ISO method. The reproducibility using the MCMC method in the text are repeated here for convenience; to compare other uncertainties, refer to Table 3 in the text. All quantities correspond to an average of the corresponding variance across the filters.

		Expanded coefficient of variation (k = 2) [%]*		
Uncertainty component	Symbol	EC	OC	TC
Laboratory-reported†	$\tilde{s}_r$	12.4	14.3	13.1
Dark uncertainties‡	$\tilde{s}_L$	12.3	9.2	3.9
Reproducibility	s_R	17.5	17.1	13.7
Reproducibility (MCMC)	s_R	16.5	8.0	12.1

† Within-laboratory uncertainties are approximated by the laboratory-reported values for the purposes of the ISO method. *Coefficient of variation are stated using mean EC, OC, and TC measurements of 8.1, 4.7, and 12.9 μg/cm^2 and EC/OC and EC/TC ratios of 1.74 and 0.63. The term *expanded* refers to the use of a coverage factor of k = 2, roughly equivalent to 95% intervals. ‡While computed analogous to the between-laboratory uncertainties in ISO 5725-2, the use of the laboratory-reported uncertainties rather than a true measure of repeatability means that this is not a true between-laboratory contribution. Rather, we use the term *dark*, denoting that these are additional, otherwise hidden uncertainties but not true between-laboratory contributions.

References

ISO: ISO 5725-2:2019: Accuracy (trueness and precision) of measurement methods and results — Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method, International Standards Organization, 2019.

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., 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.

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