

Interactive comment on “ECOC comparison exercise with identical thermal protocols after temperature offsets correction – instrument diagnostics by in-depth evaluation of operational parameters” by P. Panteliadis et al.

Anonymous Referee #1

Received and published: 27 October 2014

Panteliadis et al describes a round robin effort to compare the performance of ECOC measurements, similar to what has been done previously, but with the addition of a new temperature calibration procedure. The results provide some clarity to the somewhat erratic behavior of ECOC results and the dependence on instrument and operational error. My recommendation is that this version is accepted for publication with the addition of some much needed clarification for the reader's sake as well as some technical corrections.

C3351

Needing to be addressed: The authors claim that all participants made use of an identical instrument setup, but then go on to describe the use of two types, the Sunet Lab and Field instruments. The semi-continuous instrument is not the same as the lab instrument, and it should not be equally compared to the lab instrument. There are different nuances at play for the semi-continuous that can cause additional errors that aren't clear when comparing it to the lab instrument. As such, these results should really be removed from the comparison as they aren't relevant. The authors even show this in their results as outliers were reported for the semi instrument. If the purpose of using the semi-continuous is just to show the unreliability of this instrument compared to the lab one, then that needs to be stated more clearly.

What measures were taken to (mostly) eliminate operational error? Were the instruments “corrected” after the sucrose calibrations? Or did the users just continue on and run the PM samples?

During the introduction the authors state that there is not reference material for EC, so no calibration was done. This is not true, I can think of two: NIST SRM 1648a (Urban particulate matter) and NIST SRM 8785 (Air particulate matter, QFF filter). Both pricey, but perfectly acceptable EC standards. If the authors meant no European standard is available then they need to state that more clearly.

Did all groups use a 98% He/2% O₂ mixture for the oxidation phase? 90/10 is also a common mixture. This will affect the results slightly. Even if all 17 labs used the same mixture, please include a short discussion of how different mixtures will change split points, etc.

The differences in sucrose values for Figure 1 are astoundingly different with a wide range. Over the course of many years of using the Sunset ECOC to run sucrose as an external calibration, I've rarely seen values that low/high above expected. Ours has ranged between 26–36 (for an expected 32.6), at its very worst. Can the authors please provide more insight into the differences and wide range? Why were the instruments

C3352

even being used if sucrose values weren't within 5% of the desired value? This should have signified that something needed to be changed on the instrument (e.g., new back oven, front oven, thermocouple). And, because it was not stated in the methods, what did the groups use to "warm up" the instrument? Were these sucrose values the first three samples ran on an instrument coming out of sleep mode? Typically, and especially for older instruments, two to three samples minimum need to be ran in order for the instrument to reach steady state.

For the thermograms overall. Is there not another way to present these data? The figures are very busy and impossible to follow in black/white. And I'm not sure seeing all the graphs together is useful. What are the authors trying to show is important – split points? What not just have a table of split points and put the figures in the SI?

For the Method Performance section. Multiple paragraphs of statistical analysis results is very hard to follow for the reader, and I don't think having it in this format provides the reader with an answer to what is the significance of these observed values. Can you not put these data into a table and then use the section to discuss the significance of the results?

For the Comparison section: Why are the authors comparing split points between two different temperature protocols using time? When comparing the same program across multiple labs, time is okay, but not for different programs. Because these two programs use different temperatures and time steps, time is irrelevant. What is important for this type of comparison is temperature. If you look at the two figures, the split points are both in the 550 C range, so really there is not that much difference in how these two programs determine EC/OC.

Authors state that the initial transmission reading is not really an issue. I would argue that observing a difference in split points of 100 s is an issue and that it is possibly due to the initial transmission readings. For those that have a low initial reading there is likely some contamination on the laser window or the oven/spoons are old. When the

C3353

signal is <1500, this is almost always the case in our experience. The high temperature offsets are very large (compared to what ours was when we did the calibration). Can the authors please provide some insight as to why the offsets were high? Were new thermocouples also installed to ensure accurate readings? Was it because the thermocouples weren't initially installed correctly on the front and back ovens?

Regarding the interpretation of the POC formation between the two programs, NIOSH and EUSAAR2. Can the authors please provide more insight on this? This includes more references and more scenarios rather than the one that is outline, which is speculative. As stated, the two programs use a different maximum temperature and difference residence times. The max temperature can affect the amount and type of POC formed and some POC is only formed at lower temperatures. Since the residence times are much greater for the EUSAAR2 program, this would give these "lower temperature POC" more time to form, thus increasing the amount as was observed by the authors. For example, the combination of the last three temperature steps for NIOSH is 270 s, compared to 360 s for EUSAAR2 comparable temperatures (last two steps).

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 8697, 2014.

C3354