Reviewer 1

The article presents new statistical methods for determination of uncertainties connected to inter-instrument differences and measurement noise and applies them to micro-aethalometers MA300 and MA200. The experiment was performed by comparing tested instruments with Centrifugal Particle Mass Analyzer-Electrometer Reference Mass Standard (CERMS). The article is well written with appropriate introduction and methodology and description of the results. There are two issues which need to be addressed.

Filter photometers suffer from particle size dependent response, which makes it important to use realistic particle size for testing. The soot particles selected by CERMS with a mobility diameter of 235 nm are larger compared to the diesel engine and wood stove emissions (Laborde et al., 2012) and are more comparable to wildfire emissions. Measurement of larger particles is expected to result in higher "quantum" noise (from discrete arrival of particles on the filter), as already observed by the authors. Particle size is expected to influence the instrument response to BC mass.

There are two components to this comment.

The first pertains to Poisson noise. There is some subtilty here. Yes, if the particles are small, the same Δm would require more particles, leading to the intuitive suggestion that the Poisson noise would be higher. However, the number of particles would also be multiplied by a smaller number to obtain a mass, canceling out this difference. Mathematically, we start form the fact that, roughly,

$\Delta m = \bar{m_p}\,\Delta N,$

where ΔN is the number of particles collected in a given interval. The underlying variances can thus be stated as

$var(\Delta m) = \bar{m_p}\,var(\Delta N).$

Given pure Poisson noise due to the arrival of particles on the filter, $var(\Delta N) = \Delta N$. We also note that $\Delta N = \Delta m / \bar{m_p}$. Substitution gives,

$var(\Delta m) = \bar{m_p}\,var(\Delta N) = \bar{m_p}\,\Delta N = \Delta m.$

Note, then, that Poisson noise can be stated directly in terms of Δm without requiring any accounting for the particle mass. The extra p accounts for the fact that measurements involve absorption, not the mass directly, which can act to amplify the Poisson noise. We have clarified this in the manuscript:

> *"One might expect the single particle mass to affect the Poisson noise, as smaller particles require a higher number of particles to give the same change in mass. However, the number of particles in that instance would also be multiplied by a smaller particle mass to yield Δm, cancelling out this effect."*

The second pertains to the average instrument response. It is well established that the device will have a different response to particles of different sizes, compositions, etc. It is noted, however, that this is not a noise or inter-device reproducibility issue (except for the Poisson

noise contribution noted earlier). It will impact the overall slope of the line in the parity plots, but this should be accounted for using calibration factors that are designed to accommodate some of the particle size and composition effects. Quantifying the inter-device variability requires that these contributions be minimized, motivating the current experimental design.

However, additional notes explicitly stating this are due.

In the introduction:

> *"Since our aerosol is a simple source of eBC, with negligible content of non-absorbing PM and stable gas-phase composition, our results provide a lower limit on between-instrument reproducibility. In other words, artifacts stemming from non-ideal aerosols are expected to affect the devices similarly, such that these artifacts could be targeted by calibration against real, location-specific aerosol sources. The model derived in this work would then broadly describe the remaining uncertainties that cannot be removed by such a calibration procedure."*

We have explicitly acknowledged this following the list of observations from the parity / error plot (Sec. 3.1):

> *"As noted in the introduction, this model represents a minimum uncertainty that does not take into account various causes of artifacts that can occur in aethalometer measurements (such as particle size, composition, and rapid changes in gas composition). Such additional artifacts will add further uncertainty."*

Finally, we have added a note in the conclusions:

> *"It should be noted that these expressions only capture repeatability and inter-device variability. They do not intend to account for other systematic artifacts, including cross-sensitivity to scattering, humidity, and temperature. These effects are often location-specific and should be addressed by using appropriate calibration factors. Rather, these expressions give the minimum uncertainty that can be achieved after the aethalometers have been calibrated for the properties of the specific particles being measured. We also do not assert that this model will account for all uncertainties in the measurements, as, while systematic artifacts in the data can largely be removed by way of calibration (similar to calibration to remove inter-device variability), the calibration factors used to perform this correction will themselves have uncertainties (whether due to physical fluctuations during the measurements or incomplete knowledge of the artifacts) that must be considered alongside the uncertainties here."*

The second issue is the influence of the filter loading effect. Figure 2 show jumps in instrument response during tape advance which indicate the presence of filter loading effect. Same can be concluded by analyzing relative instrument response (eBC/ref_mass) as a function of attenuation

using data in the Supplement (please see the graph in the attached pdf). Authors should discuss the implications of filter loading effect, for example the increased variation of the instrument response.

First, it is worth emphasizing that regions where the dual spot correction was not applied to correct for the filter loading effect were not included in the analysis, and this includes the regions that the reviewer points to in Figure 2. This was already explicitly stated in the manuscript in several spots, including the first paragraph of Sec. 3 and the caption of Figure 3:

"Measurements taken when the attenuation was below 3 are excluded."

The dual spot correction was otherwise found to be sufficient to correct for the filter loading effect. Further, after the dual spot correction was applied, no trend in the residual or structure in the errors was observed as a function of attenuation. This has now been made explicit, with a paragraph and figure dedicated to showing this very feature of the data:

"Figure 5 shows the standardized residual, akin to Figure 4, but as a function of attenuation. Two observations are of note. First, the central tendency of the data does not exhibit a trend and is centred about a standardized residual of zero. This indicates that the applied attenuation correction is effective in generating a linear response from the aethalometer relative to the underlying mass concentration, even if the correction adds a substantial amount of noise to the measurements. Second, the overall spread in the data is constant across the domain, an indication that the attenuation had no impact on the retrieved error model.

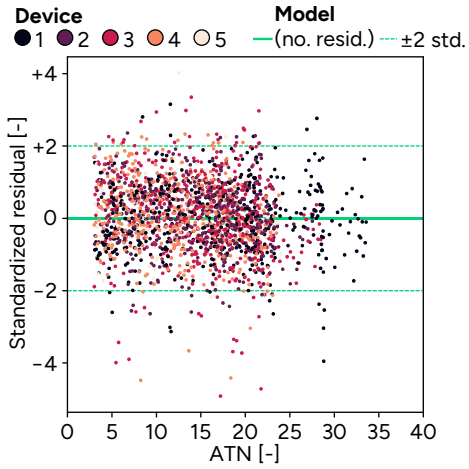

Figure 5. Residual between the model and aethalometer measurements as a function of attenuation (ATN), normalized by the estimated error for each point (i.e., standardized). Data shown is for a single averaging interval of 10 s."

Line by line comments

Page 1, Line 17: “A quantitative expression is provided for the uncertainty in the aethalometer measurements as a function of mass concentration, sampling interval, and flow rate.”

It should be noted that these uncertainties present one part of the uncertainties connected to the filter photometers. Uncertainties connected to filter loading effect, particles size, cross-sensitivity to scattering and mass absorption cross-section can be higher than inter-instrument variability and noise.

We acknowledge that this is true and had noted these effects in the introduction. However, as stated above, it also worth noting that many of these uncertainties are expected to be systemic and are instead taken into account using location-specific calibration schemes.

As per our response to the first questions, we have added several clarifying notes to the manuscript, including in the introduction. We have also added “minimum uncertainties” to the abstract.

Page 2., Line 42:

There have been more measurements of the influence of particle size on instrument response. It seems to be a general feature of filter photometers: it was observed for AE33, CLAP and MAAP (Ramshoo et al, 2022; Drinovec et al., 2022; Yus-Dies et al., 2025).

Ramshoo et al, 2022; https://doi.org/10.5194/amt-15-6965-2022-supplement

Drinovec et al., 2022; https://doi.org/10.5194/amt-15-6965-2022

Yus-Diez et al., 2025; https://doi.org/10.5194/amt-18-3073-2025

As per the comment above, other than the impact on the Poisson noise, these effects are expected to impact the slope of the line, not so much the noise in the measurements.

These references do provide a sense of the inaccuracies associate with particles size and morphology, which, while stated in the introduction, was not adequately tied to the corresponding literature. The suggested citations have been added to the second paragraph of the introduction where these artifacts are acknowledged:

“Accuracy may also be affected by particle size and morphology (Romshoo et al., 2022; Drinovec et al., 2022; Yus-Díez et al., 2025).”

Page 4, Line 114: “size distribution had a geometric mean mobility diameter of 235 nm”

The selected particle size depends strongly on charge distribution from the unipolar charger. How stable was the charge distribution during the experiment?

We have explicitly added an uncertainty to the stated GMD, corresponding to the variability in the SMPS-measured GMD over the duration of the experiment:

"235 nm ± 18 nm"

The particle size was generally stable. It is worth noting that the experimental procedure of filling the chamber with an aerosol and making measurements on that aerosol over time averages over some variability in the particle charging and largely ensures a stable particle size for a suitably long period after the box has been filled. We have also noted this:

"The large volume of the experimental chamber, combined with the fact that we periodically rather than continuously refilled it, resulted in stable concentrations during the experiments that were insensitive to fluctuations in the charger."

Page 7. Fig 3.

In Figure 3. eBC measured by aethalometers is compared to particle mass concentration of CERMS. What is the contribution of organics to the sample mass? What would be the effect of thermodenuder or catalytic stripper?

We have added a note that states that the EC/TC ratio exceeded 0.9, so there was very little organics:

"For testing, the aethalometer chamber was periodically filled with soot, after which the inlet to the chamber was closed and the concentration of particles in the chamber was allowed to decay slowly over time, as shown in Figure 1. BC particles were generated using a Mini Inverted Soot Generator operated on ethylene (MISG; Argonaut Scientific; (Kazemimanesh et al., 2019)) operating with 0.1 SLPM of ethylene and 10 SLPM of air. This is expected to produce particles with EC/TC values in excess of 0.9 (Kazemimanesh et al., 2019), with the balance of the particulate mass made up of hydrogen and oxygen at the soot surface (Corbin et al., 2020)."

Page 7. Line 158. "with structured artifacts as a function of attenuation appearing in the measurements when not correcting the data (cf. Figure 2b)"

This is in contrast with the caption of Figure 2 which indicates that the data is corrected for the filter loading effect (except for data where ATN<3). Please clarify which data is corrected for the filter loading effect.

All of the data used in the analysis is attenuation corrected, and none of the data below an ATN of 3 was used in the analysis, for the very reason that it was not corrected. The cited text

provides more confusion than clarification and was removed. Rather we reference forward to the new Figure 5 and the surrounding discussion that demonstrates that the corrected data did not exhibit any artifacts:

"We note that the correction effectively removed biases due to attenuation (cf. Figure 5 and surrounding discussion later in this work)."

Page 9, Line 202: "While not shown in Figure 3, it is noted that the attenuation coefficient has minimal effect above an attenuation of 3,".

Data on Figure 3 and in the supplement suggest that the filter loading effect is still present.

Figure 3 does not show the effect of filter loading. If anything, the figure shows the opposite, where filter loading would cause a structured pattern lines of data would stray away from the main cloud.

To resolve this, as stated above, we have added a paragraph and figure explicitly showing that the standardized residuals do not show an attenuation / filter loading effect.

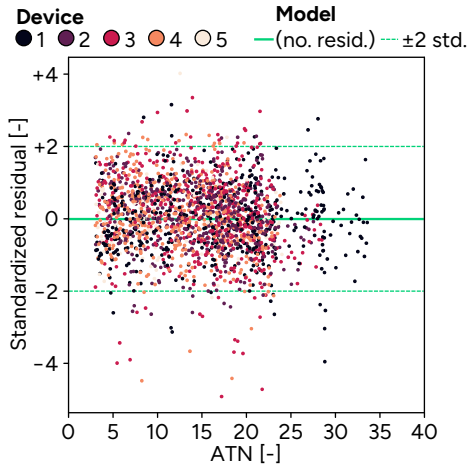

Figure 5. Residual between the model and aethalometer measurements as a function of attenuation (ATN), normalized by the estimated error for each point (i.e., standardized). Data shown is for a single averaging interval of 10 s.

Page 18, Line 376: "While the dual spot correction algorithm was found to be effective in correcting biases in the measurements"

Please see the comments on filter loading effect above.

Also see previous author response. Given that the standardized residuals do not have any structure as a function of attenuation, the statement in the manuscript stands.

Page 19, Line 394. “Partially processed data is included alongside the manuscript.”

Please provide both raw and processed data.

We changed the manuscript text to “Raw data (after time synchronization) are included alongside the manuscript.”