Response to Reviewer #1: AMT-2022-278

The authors would first of all like to thank the reviewer for the insightful comments on the work we have submitted for publication, and the editor for the opportunity to improve the manuscript. Under each reviewer comment (in red) there is a summary of the response (black text), in addition to the text from the manuscript (in blue) that was modified, if applicable.

(1) Lensing is important, it **should at least be mentioned.** Lensing can lead to inflated eBC values using manufacturer's default MAC values especially during wildfire periods; impact on source apportionment is uncertain (whether lensing has a wavelength dependence). E.g. Saleh et al. https://doi.org/10.1002/2015JD023697 and Bond et al. https://doi.org/10.1029/2006JD007315

Author's Response (MC): We thank the reviewer for highlighting the importance of the lensing effect on an aethalometer's measurement. Even though mixing state plays a critical role, the fact that IR specific light absorption is only due to pure BC has been well accepted in the literature. Since, source apportionment calculations utilize the two-wavelength measurement, where light absorption measurement in lower wavelengths can be significantly impacted by the lensing effect due to the presence of BrC compounds, we agree that this effect may cause uncertainty in eBC apportioned components. We included this discussion in the introduction section. (**L62-L67**)

In addition, the mixing state of aerosol can influence the light absorption measurements of bulk aerosol (Bond and Bergstrom, 2006; Saleh et al., 2015; Healy et al., 2015). At lower wavelengths (near-UV), light absorption measurements were found to be enhanced (by a factor >2) due to the presence of brown carbon (BrC) during wildfire-affected aerosols in Canada (Healy et al., 2015). The same study identified minimal light absorption enhancements due to the lensing effect at higher wavelengths (near-IR) regions. As such, light apportionment-based BC measurements and their source apportionment can be further complicated by the bulk aerosol's source and mixing state.

Also in Section 4, we mention that our study could not explicitly estimate the impact of lensing in light absorption estimations, particularly during the Wildfire period. We identify this as a source of uncertainty in our work and highlight as a potential important area for future work.

Section 4: This study did not take into account the lensing effect of BC, which has been identified as being particularly relevant during wildfire periods (Healy et al., 2015), and can impact the light absorption coefficient measurement. Evaluating how lensing impacts the b_{abs} measurements of MA300 is an important area for future work.

(2) Effect of filter loading (**BC mass per unit filter area**) should be considered, which can explain key results - e.g. noisier MA300 (0.15 lpm) data at lower BC values, better performance (lower NRMSE) during wildfire periods, and greater precision of AE31 in a previous study (as sample flow rate is 2-5 lpm).

Author's Response (MC): Thank you for this helpful suggestion. We have adopted the filter loading calculations you suggest. The results of the filter loadings were estimated for all the devices for individual periods (**Table S6**). For individual devices, filter loadings were estimated from real-time flow measurements and then averaged for the regular and wildfire periods. We then use instrumental noise estimates (as suggested by Reviewer #2) and MAE to identify instrumental variability during the wildfire episode. We verified that larger errors in Wildfire periods are likely related to higher filter loading and updated the text accordingly. The discussion on filter loading was included in **Section 3.1.2 (L359-L368).**

Text from Section 3.1.2

As mentioned in Section 3.1.1, filter loading during a high BC event can lead to measurement errors that vary by each instrument's sensitivity. Mean filter loading (eBC mass per unit filter area, see Table S6) during the Regular period were 1.63 μg cm^{-2}, 1.46 μg cm^{-2}, 1.96 μg cm^{-2} and 5.8 μg cm^{-2} for devices MA300A, MA300B, MA300C and AE33, respectively. For the Wildfire period, the filter loading increased across all the devices (to 5.49, 4.85, 5.79 and 19.0 μg cm^{-2}, respectively) as expected. During the Wildfire period, filter loading per unit volume of air sampled increased by a factor of 2 in AE33 (see Table S6). However, for MA300 units, this factor varied (2.4 for MA300A, 2.27 for MA300B and 1.9 for MA300C). Even though, all these devices measure the same environment, we find differences in the aerosol loading on MA300 filters. We hypothesize this variability might be occurring due to the variability in sampled airflow and instrumental noise. By studying the instrumental noise and filter loading estimates, we find that the error contribution in MA300's eBC measurement can be sensitive to their exposed concentration range. This observation aligns with a previous study on MA-series aethalometers, where the impact of high eBC concentration has been found to impose large errors from more pronounced filter loading corrections (Alas et al., 2020).

Table S6: Filter Loadings experienced by aethalometers during the measurement periods.

Device	Filter Loading (μg/cm^2)		Filter Loading per unit volume of air sampled (μg/cm^2/mL)	
	Reg	WF	Reg	WF
AE33	5.8	19.0	0.002	0.006
MA300A	1.63	5.49	0.016	0.055
MA300B	1.46	4.85	0.015	0.049
MA300C	1.96	5.79	0.020	0.058

(3) Drinovec algorithm considers the effect of variable flow rates, and flow rate variability should impact all wavelengths equally. So it is unclear why flow rate variability should result in the Drinovec algorithm not working for MA300 (unless they did not include that factor) nor why it performs quite well for UV absorption but not for IR absorption. How did the authors verify flow rates in the MA300?

Author's Response (MC): This is a great point raised. The Drinovec algorithm considers non-linearities in the flow measurements in estimating the loading correction factors (Equation 4 of this manuscript).

$$\frac{F_2}{F_1} = \frac{\ln(1 - k \times ATN_2)}{\ln(1 - k \times ATN_1)}$$

The reason for the impact of variability in flow rate across different wavelengths of light absorption measurement has been explained by Drinovec et al., (2015). Briefly, filter spot advancement is driven by the lowest wavelength, i.e. UV (until a set attenuation value is reached). Corresponding measurement signals from other channels tend to be lower as wavelength increases. The differences in wavelength-specific attenuation values determine the face velocity ratio factors (referred as FVRF by Drinovec.) which was found to significantly impact the k-values. AE33's onboard correction mechanism uses flow dependent-corrections using Drinovec et al., (2015), and any loading correction estimates will be sensitive to small errors in face velocities. As compared to AE33, the MA300 has significantly lower face velocities, which can increase the overall sensitivity of the correction method (modified Drinovec method, Section 2.4) tested on MA300s. MA300's on-board correction algorithm (Section B of Supplementary Information) is independent of the flow measurements, however, comparatively higher flow variability has been observed in MA300, as compared to AE33. Raw data from both MA300 and AE33 includes real-time flow

measurements from spot 1 and spot 2, which have been used to calculate the percentage deviation in MA300's flow compared to the setpoint. We have added discussions on the flow measurements and their relationship to the MA300-adopted modified Drinovec correction algorithm in **Section 3.4**. (**L469-L475**).

From the onboard mass flowmeter readings, we find that filter spots 1 in all the MA300 units were drawing comparatively lower airflow with wider variability. MA300 uses a sampling flow rate of 150 mL min^{-1}, and ideally, 2/3rd (100 mLmin^{-1}) is split to filter spot 1. MA300's flow setpoint deviation ranged from -9.7% to 2.4% for MA300A, -8.9% to 5.8% for MA300B, and -14.2% to 1.4% for MA300C. In contrast, AE33 was run at 5 L min^{-1} (with set airflow of 3333.33 mL min^{-1} on filter spot 1) and had smaller deviations from the setpoint (-0.44 % to 3.2%). A high range of flow variability can lead to additional noise in corrected eBC measurements when a flow-based correction technique, such as the Drinovec et al. (2015) algorithm, is adopted in MA300 devices.

Also, in the conclusion (**L605-L606**) we now discuss the regular flow audits we have performed in this work.

Since MA300 operates at very low-flow conditions, we regularly audited and calibrated the flow (twice a month) and we recommend that MA300 users conduct routine flow audits while doing continuous sampling, particularly in a highly polluted environment.

(4) Was Fig 1b for SD of MA300 averaged across all three units and Fig S8 for unaveraged data? The latter seems more representative - though perhaps you should (a) normalize response of each MA300 (as that is a known bias, not noise) and then calculate SD for normalized response.

Author's Response (MC): In Fig 1b, our goal was to look at the between-unit SD of MA300, which could be considered a measure of unit-to-unit variability. We created a pool of MA300's measurements (for absolute and normalized data separately) for each bin (of width 1 μg/m^3) of AE33 reported eBC data. We then plot the standard deviation of the pooled measurement against the AE33's measurement. The slope of the linear fit corresponds to the linear response of MA300's variability across the concentration range.
In our earlier submission, Fig 1b was for SD of MA300 averaged across all three units and Fig S8 for unaveraged (pooled) data. As per the suggestion, we replaced Fig 1b with the normalized responses of MA300's pooled data. Additionally, we updated supplementary information (Fig S7), to show the pooled MA300 variation (non-normalized). Normalized eBC concentrations, when fitted against normalized binned AE33 concentrations showed a 5% unit-to-unit variability. The normalization approach we used is summarized in the equation below, and is explained in **Section 2.7** (**L270-L274**) and supplementary **Section E**. In addition to the normalized response, in the updated discussion, we have also reported the calculated variability of 21% from the non-normalized measurements of eBC. The updated discussion can be found in **Section 3.1.2 (L337-L345)**

Normalization process discussed in Section 2.7

Relative accuracy assessments were performed using the slope of the linear fits. The measured and derived parameters had different scales or ranges over wavelength channels, MA300 units, and periods. Hence, to remove measurement bias and to focus on instrument's variability, we choose to normalize the data for individual groups by scaling with respect to their range of measurement, also known as min-max normalization (Géron, 2022).

From Section E of Supplementary Information:

$$N_{normalized} = \frac{N - N_{min}}{N_{max} - N_{min}}$$

For any parameter N, normalized quantity ($N_{normalized}$) was derived by the respective range, i.e. the difference in the maximum (N_{max}) and minimum (N_{min}) values.

From the Section 3.1.2:

First, we normalize the absolute measurements from MA300 units by the range of concentration (max – min criteria) to only consider the measurement bias and eliminate any device-specific noise. Next, we created a pool of MA300's measurements (for absolute and normalized data separately) for each bin (of width 1 μg m^{-3}) based on AE33 reported eBC data. The standard deviation of the pooled measurement from MA300 were fitted against the AE33's measurement. The slope of the linear fit corresponds to the linear response of MA300's variability across the concentration range (Figure 1(b)). The slope of this linear fit line is 0.049, and can be interpreted as an approximately 5% variability across MA300 units of hourly eBC mass concentration. When compared against non-normalized measurements, MA300 exhibited 21% unit-to-unit variability (see Figure S7). In Figure 1(b), we also see that a linear fit cannot totally explain MA300's unit-to-unit variability (R^2 = 0.6) and depend on the observed concentration range, which suggests estimates of unit-to-unit variability can change based on the range of eBC concentration.

Updated Figure 1 (b):

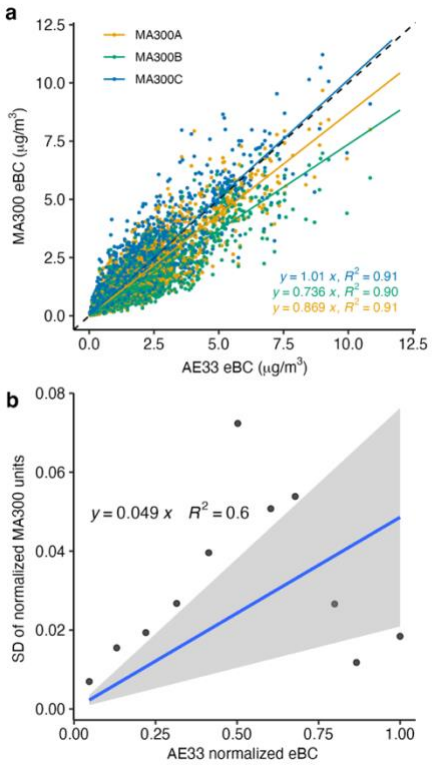

Figure 1 Scatter plot of eBC mass concentration for individual MA300 units A,B and C vs AE33. The dashed line represents the 1:1 line, and solid colors are the regression fit lines for the individual MA300 units; (b) Linear relationship of multi-unit pooled Standard Deviation (SD) from normalized MA300 measurements for each μg m^{-3} of normalized AE33 eBC concentration. The fit line (in blue) represents the linear response of MA300's variability across the concentration range. The shaded region represents the 95% CI of the fit.

Updated Figure S7

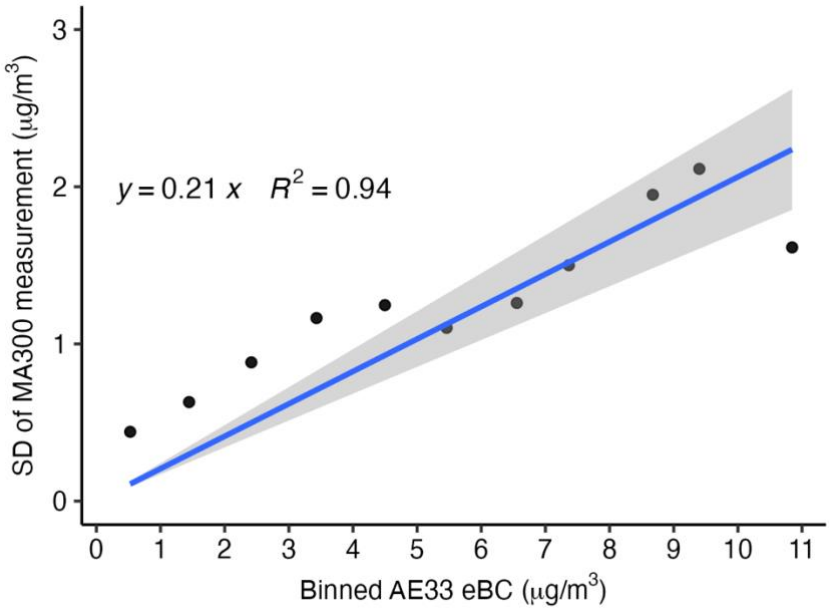

Figure S 1Multi unit pooled standard deviation from MA300 measurements for each one µg. m^{-3} of eBC concentration measured by AE33. The fit line (in blue) represents the linear response of MA300's variability across the concentration range. The shaded region represents the 95% CI of the fit

5. Table 1 shows individual MA300 data, which are useful (given N=3), but Sec 3.2.2 discusses apparently the average of the three units - which masks the significant variability in device performance. The latter may be more useful to the reader as people may buy just one unit ($10k is a lot of money!), and whether they get unit B or unit C makes a huge difference.

Author's Response (MC): This is a good point and is aligned with the objectives of this work. Table 1 serves the purpose of providing readers with a statistical summary of the measurements made in this study. Based on the individual unit's response, we want to highlight the inconsistency observed in MA300 units. A handful of studies have focused on the performance evaluation of multi-wavelength black carbon (BC) measurements from Dual-Spot-based micro-aethalometers (e.g. MA300). However, a one-to-one comparison of individual MA300 units and the reference method may not clarify the overarching performance of such micro-aethalometers. As we understand, the growing focus on BC emission and its source estimates across the globe might lead to multi-unit aethalometer measurements. Therefore, both overall and individual metrics are important for understanding the extent of measurement uncertainty and exploring the reliability for future users of micro-aethalometers.

(6) Lines 364-367 - results for filter loading effect not shown elsewhere; OA hygroscopicity speculative. Suggest deletion or clarification.

Author's Response (MC): We thank the reviewer for pointing this out. As mentioned in the response to Major Comment #2, a discussion on filter loading is now included in **Section 3.1.2 (L357-L366).** In the updated discussion in Section 3.5, we establish the impact of strong loading effects and RH on the MA300's measurements. Additionally, a more detailed explanation of aerosol aging and its relationship with hygroscopicity has been included. Section 5 (**L485-L495**)

In our study, all three MA300s were influenced by strong filter loading (Table S6) in addition to RH changes (45% to 95%) during the sampling periods. Being a near-road emission measurement site, our measurements captured complex aerosol mixtures of various mixing states. During the regular period, local traffic during the daytime contributes to fresh BC-enriched aerosols, which can be hydrophobic in nature (Sarangi et al., 2019; Wang et al., 2020) and by night-time, these fresh BC-enriched aerosols can evolve by aging and change their morphological and optical properties. In contrast, during the wildfire smoke-affected days, the measurement site experienced increased quantities of aged aerosols through long-range transport from the Pacific Northwest. These claims align with our calculated α values, as shown in Fig. 3. With the abundance of organic aerosol components during wildfire days, coated BC particles have been found to dominate and often enhance light absorption in lower wavelengths due to the presence of BrC (Healy et al., 2015). This wildfire smoke-affected BC particles can be mixed with a significant fraction of secondary organics, which can be hygroscopic in nature (Wang et al., 2020).

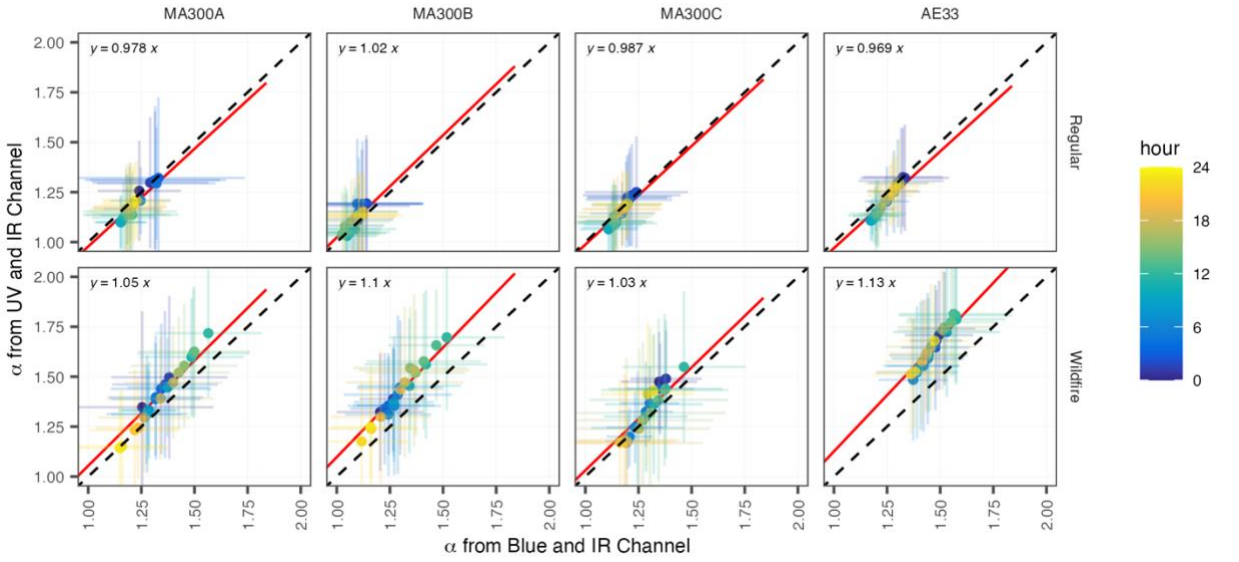

Figure 3: AAE (by the hour of the day) measured by different aethalometers during the Regular and Wildfire period. Average hourly AAE values derived from the Blue-IR wavelength pair (on x-axis) and the UV-IR wavelength pair (on y-axis) with error bars representing respective standard deviation. Red lines represent the linear relationship (forced through the origin), and the dashed line is 1:1.

(7) Suggest running source apportionment for a week before and after the wildfire period to minimize the effect of seasonality on fossil fuel BC. Also, GDI vehicles could also contribute BC especially in urban areas, not just diesel vehicles, as shown by this excellent paper: https://doi.org/10.1021/acs.est.5b04444

Author's Response (MC): This is a great suggestion. We have performed Source Apportionment on a week of data before the wildfire period (August 24 to August 30) and on a week of data after the wildfire period (October 14 to October 20) and included these new results in the discussion (**Section 3.5.1, L527-L532**). In this study, we have also used traffic data measured by MetroVancouver. Our apportioned eBC_{ff} results overlapped with peak traffic periods, however the fuel type of vehicles could not be specified. We used general fossil-fuel based traffic sources in the updated discussion.

To verify the impact of seasonality on eBC_{ff} during the wildfire period source apportionment results, we considered two additional week-long periods (August 24 – August 30 and October 14 – October 20) before

and after the wildfire smoke-affected period (Figure S12). We found that the fossil-fuel component dominated the eBC mass throughout the day and eBC_{ff} range was similar to the Regular period. We find slight difference (increase in pre-wildfire period) in the eBC_{ff} concentrations in the late night hours. Both pre- and post-wildfire weeks followed almost similar diurnal eBC_{ff} profile as the Regular period. Hence, no seasonality in eBC_{ff} were identified with this analysis.

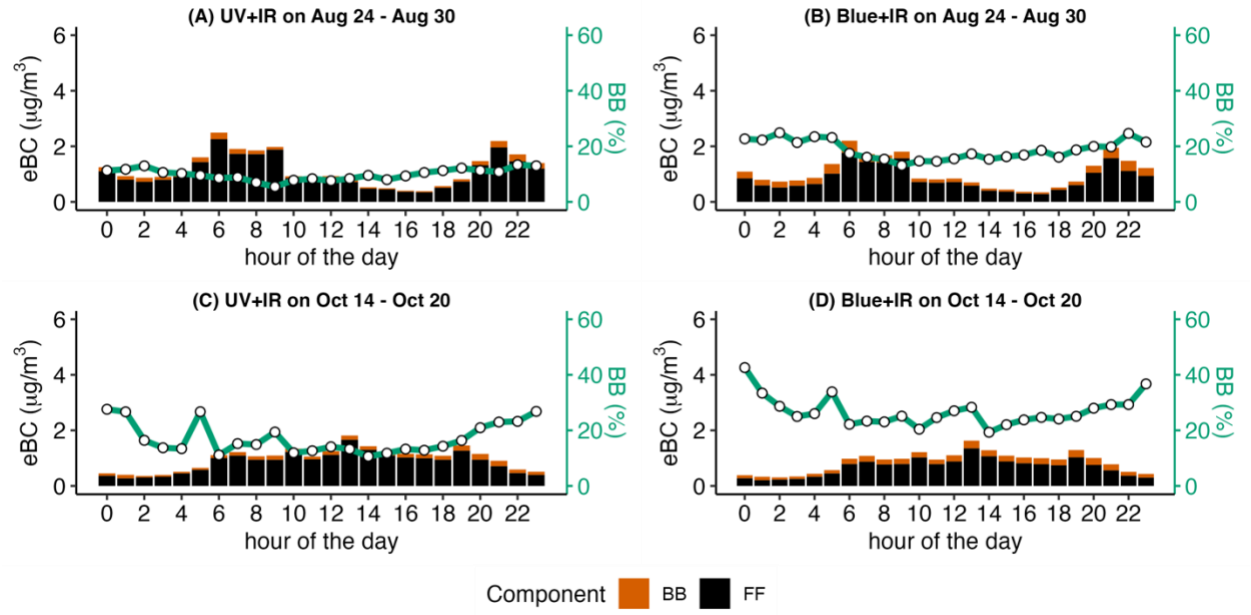

Figure S 12: Source apportionment results (similar to Figure 4 in the main text) but for two specific periods. Panel A and B are pre-wildfire period: August 24 to August 30. Panel C and D are for post-wildfire period: October 14 to October 20.

(8) Lines 395-396 - is that for the regular, non-wildfire period? (Also, just type out "wildfire" and "regular" or "non-wildfire", using acronyms/abbreviations is annoying and this does not reduce the word count.) What is the MDL below which source apportionment is not robust?

Author's Response (MC): We thank the reviewer pointing this out. As recommended, we have updated the manuscript "Reg" to "Regular", and "WF" to "Wildfire". Based on particle free measurements, we calculated MDL of MA300B and MA300C as 0.21 and 0.43 µg m^{-3} (MA300A had measurement error, hence excluded). Clearly instrumental sensitivity played a big role in determining the MDL of MA300. Considering the lowest identified MDL value of eBC measurement, we argue that any source apportionment should not be conducted below the MDL of the black carbon concentration, which is 0.21 µg m^{-3}. However, results from sensitivity analysis (**Figure S10** and **S11**), reflected that, different combinations of α pair can provide negative $b_{abs,bb}$ and $b_{abs,ff}$ compounds which can again be treated as artifacts of source apportionment method.

(9) If the Drinovec algorithm is not appropriate for the MA300, why are source apportionment results with this algorithm discussed in the main text? Delete or move to SI.

Author's Response (MC): We thank the reviewer for this suggestion. In the revised manuscript, Figure 6 have been updated to present the results from MA300's onboard correction only. The results from the responses of modified Drinovec correction were moved to SI. (Figure S13).

Minor Comments

Line 21: improves by 10% or comes to within 10% of reference monitor?

Author's Response (MC): Here we refer that the results of source apportioned components improved by 10% when we choose Blue-IR method as compared to UV-IR method. This has been clarified in the Abstract (**L24-L25**).

When the Blue-IR based source apportionment technique is adopted instead of the UV-IR, there is a 10% (on average) decrease in the percentage difference of the apportioned components from the reference monitor.

Line 53: coating by non-BC material produces a lensing effect that complicates these measurements. Not considered in this manuscript, must be at least mentioned. e.g. Saleh et al. https://doi.org/10.1002/2015JD023697 and Bond et al. https://doi.org/10.1029/2006JD007315

Author's Response (MC): We thank the reviewer for pointing this out. We have added the following lines in the main text to explain the importance of coating on BC particles. (**L62-L67**)

In addition, the mixing state of aerosol can influence the light absorption measurements of bulk aerosol (Bond and Bergstrom, 2006; Saleh et al., 2015; Healy et al., 2015). At lower wavelengths (near-UV), light absorption measurements were found to be enhanced (by a factor >2) due to the presence of brown carbon (BrC) during wildfire-affected aerosols in Canada (Healy et al., 2015). The same study identified minimal light absorption enhancements due to the lensing effect at higher wavelengths (near-IR) regions. As such, light apportionment-based BC measurements and their source apportionment can be further complicated by the bulk aerosol's source and mixing state.

Line 56: did you mean to write soot?

Author's Response (MC): We thank the reviewer for catching this error. Yes, we meant soot and not aerosol. We have made this change in the manuscript. (**L59**).

Fossil-fuel-based aerosol sources (e.g., diesel vehicles) generate soot, which tends to absorb light uniformly across the spectrum (Bond and Bergstrom, 2006)

Line 145: no mention of lensing effects? scattering of *light by* loaded aerosols

Author's Response (MC): We thank the reviewer for pointing this out. We have updated **Section 2.3** (**L156-L163**) where we include discussion on lensing effects.

Filter-based light absorption techniques are subject to measurement artifacts due to scattering on the filter, scattering of light aerosols loaded on the filter surface or due to some particles being shadowed by others (Weingartner et al., 2003). In addition to these measurement artifacts, current aethalometer real-time correction algorithms do not consider light absorption enhancement occurring from the lensing effect, particularly at lower wavelengths due to light-absorbing organic components. Present designs of a stand-alone aethalometer are not equipped to estimate light absorption enhancement in real-time, as they can not distinguish the aerosol mixing state and focus on bulk aerosol properties. Therefore, aethalometers require proper estimation of loading compensation factors and multiple scattering factors for accurate measurement (Virkkula et al., 2007; Weingartner et al., 2003; Virkkula et al., 2015).

Line 169: It is important to note that MA300 corrections do not include any filter leakage parameters and use higher MAC values as compared to AE33. – why?

Author's Response (MC): We are using the manufacturer provided MAC values for both MA300 and AE33. Wavelength-specific MAC values are usually provided by the manufacturer based on laboratory tests carried out while developing the instrument. It is worth noting that the MA300 devices used in this study operated on firmware v1.10, which did not include leakage parameters for BC estimation. We anticipate that this will be addressed in future firmware upgrades. In the updated manuscript, we remove this line and modified the text as we found this line redundant. (**L187-L189**)

Even though the MA300 includes flow measurements in the raw data files, it does not consider any lateral filter leakage parameters and flow values in their correction algorithm. Wavelength-dependent MAC values for both MA300 and AE33 models (Table S1) were taken from the user manuals provided by their respective manufacturers.

Line 234: This is a strange metric and not used anyway, so delete.

Author's Response (MC): We have considered this suggestion and removed bR2 estimates from the manuscript and SI.

L266: "calculated", not "estimated" (which means an approximation)

Author's Response (MC): We thank the reviewer for pointing this out. We have updated in the manuscript accordingly.

L275: doesn't this simply mean that the MA300 performs better at higher concentrations (aka higher filter loadings)? You could calculate the filter loading (microg/cm2) at the different flow rates (5 lpm for AE33 vs 0.15 lpm for MA300).

Author's Response (MC): This is a great point and we further investigated our work to reframe the error matrices used in this work. In our previous calculation, we calculated normalized error as, $NRMSE = \frac{RMSE}{mean(AE33)}$ for individual periods. However, we recognize that choosing mean value for individual periods does not take into account the variability in the concentration. Upon further investigation, we found that the range of hourly eBC concentration during the Wildfire period is much wider than that of Regular period. In such cases, mean based NRMSE calculation may not be sufficient. Hence, we adopt a range based NRMSE estimation, which allows for a more meaningful comparison of MA300's performance in both Regular and Wildfire periods. The normalization process is based on the range measured by each individual device and has been explained in detail in the updated manuscript (**supplementary Section E**), and previously mentioned in the response to major comment #4. We calculated average NRMSE of MA300 units as 8.5% during the regular period and 12.5% during the wildfire period. We hypothesized that the increasing relative error during the wildfire period could be related to the filter loading effect. As suggested by the reviewer we calculated the filter loadings (in μg cm^{-2}) for individual instruments and compared the results with the errors calculated (**Table S6**). Since AE33 and MA300 samples at different flowrate, we normalized the filter loading estimates by their set flow. We verified that larger errors in Wildfire periods are related to higher filter loading and low face velocity and updated the text accordingly (**L316-L329**).

Increased absolute error during the Wildfire period can be attributed to the higher observed absolute concentration of eBC which results in more potential for large absolute differences with respect to the reference measurement. Average normalized errors (NRMSE) for the three MA300s were calculated as 8.5% and 12.5% during the Regular and Wildfire period, respectively. Higher MAE and NRMSE during the wildfire period indicate that the MA300's errors have increases in both absolute and relative terms. This means, MA300's relative accuracy can deteriorate in a highly polluted environment. Since, the aerosol sampling process between the AE33 and MA300 can differ (filter mechanism, flowrate), it is possible that, MA300's measurement errors are associated with filter loading interactions. Differences in sampling flowrates (for MA300 150 mL min^{-1} and for AE33 5000 mL min^{-1}) can change differences in face velocity, and hence change the particle penetration depth into the filter (Moteki et al., 2010). Further, from the real-time estimates of eBC concentration, we calculated the device-specific filter loadings (in μg cm^{-2}) for the corresponding flow rate of measurement (see Table S6) and found that Wildfire periods were subjected to higher loadings (almost 3.5 times than Regular period). When filter loadings of MA300s and AE33 were normalized with their set flow rate, we found that MA300s were experiencing significantly higher filter loading (0.054 μg cm^{-2} mL^{-1}) than AE33 (0.006 μg cm^{-2} mL^{-1}) during the wildfire period. Higher filter loading and lower face velocity can lead to large measurement errors in MA300's results, particularly during wildfire period.

L280 not absolute, but variability in

Author's Response (MC): Mentioned variability in AE33 measured eBC concentration in the updated manuscript. (**L331**)

The linear fit results in Figure Figure 1(a) indicate that the variability in AE33 reported eBC concentrations were well captured by the MA300 units during the whole campaign.

L289: was the shown plot 1b for SD of average MA300 in each AE33 bin and S8 plot is for unaveraged SD for each bin? The latter seems more representative - though perhaps you should (a) normalize response of each MA300 (as that is a known bias, not noise) and then calculate SD for normalized response.

Author's Response (MC): This comment is linked to Major Comment #4 and has been addressed.

L309-310: "The absorption enhancement in the UV channel reflects the elevated contribution of organic compounds originating from wildfire smoke" -- only if it is normalized. Absolute increases in UV absorption can simply reflect absolute increases in BC.

Author's Response (MC): We thank the reviewer for pointing this out. We computed normalized b_{abs} values from the AE33 for all channels and for both the periods and showed the enhanced absorption during the Wildfire period in UV channel in Figure S8. In all the wavelengths, we performed Welch's t-test and found the enhancements during Wildfire period were statistically significant ($p < 0.05$). The updated discussion can be found in **L383-L387**.

To further check the differences in relative spectral absorption, for the Wildfire and Regular period, we compared the normalized b_{abs} from the AE33 (Figure S8). We used Welch's t-test on the regular and wildfire data separately for all the channels and found the observed spectral light absorption enhancement to be statistically significant ($p < 0.05$). The UV had the largest absorption enhancement during the wildfire period, reflecting the elevated contribution of organic compounds originating from wildfire smoke (Healy et al., 2019; Laing et al., 2020).

Figure S8 from Supplementary Information

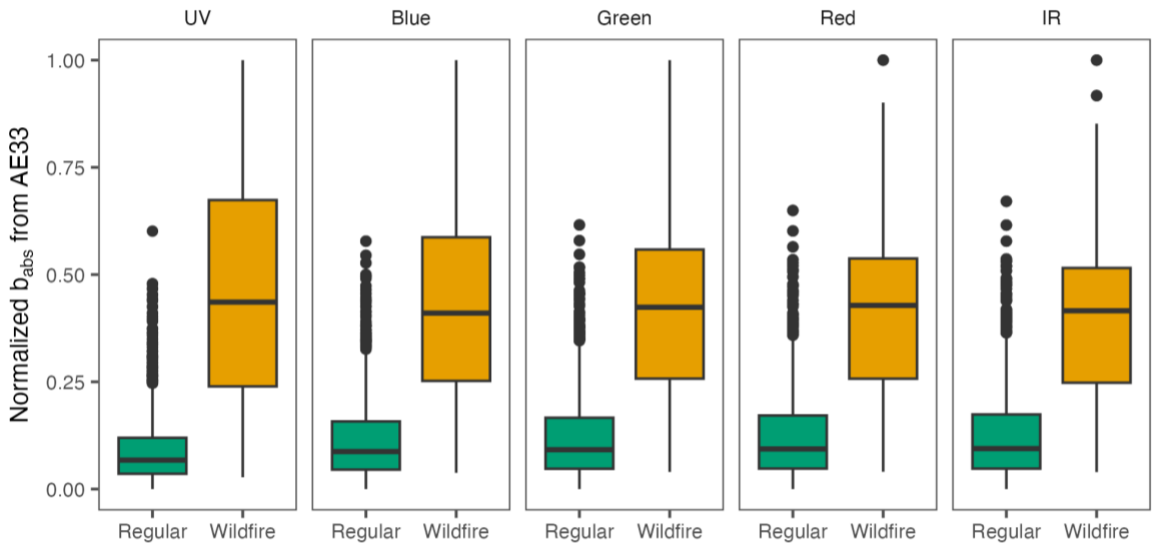

Figure S 2: Box plot of normalized b_{abs} from AE33 over the five channels categorized by the measurement period. Increased median values during the Wildfire period can be identified from the horizontal lines of each boxplot.

L319-320: However, the overestimation percentage decreases to 10% for lower wavelength light absorption (babs,UV) during the Reg period and interestingly changes to underestimation during the WF period by 9% -- has this been reported before? seems like a significant finding - if it is statistically significant.

Author's Response (MC): We thank the reviewer for highlighting the importance of our findings. We could find enough literature on this particular device. However, Li et al., (2021) showed that, for less absorbing BrC, MA200 (sister model of MA300) overestimates light absorption measurement. In the updated manuscript (**Section 3.2.2 L398-L404)** we have included results of statistical significance test (unpaired t-test) with a significance level of 0.05. t-test results revealed most of the findings are statistically significant ($p < 0.05$), except for $b_{abs,UV}$ during the Wildfire period ($p = 0.075$). However, we mention is as a borderline significance – i.e. underestimation is real but with slightly lower probability.

We find that the MA300-reported absolute measurements overestimated $b_{abs,IR}$ by 20% and 18% during the Regular and Wildfire periods, respectively, when compared against AE33. However, the overestimation percentage decreased to 10% for the UV channel ($b_{abs,UV}$) during the Regular period and interestingly switched to a 9% underestimation during the Wildfire period. Unpaired t-test results revealed most of these findings to be statistically significant ($p < 0.05$) except for $b_{abs,UV}$ during the Wildfire period ($p = 0.075$). The underestimation of $b_{abs,UV}$ during the Wildfire period appears to be borderline statistically significant with a p-value of 0.075, indicating that there is some evidence to suggest that the underestimation is real; however, the result does not reach the usual threshold for statistical significance.

Figure 3: binning these data like Fig 1b might be more robust? And maybe use different scales for different wavelengths, as the magnitudes vary.

Author's Response (MC): We thank the reviewer for this suggestion. However, in the updated manuscript, this figure has been removed. The unit-to-unit variability has been explained with the help of Figure 2,

instead, which shows the slope of linear fit of normalized b_{abs} values from MA300 units and from AE33. We determine the wavelength specific unit-to-unit variability comparing the coefficient of variation (CV) of slope values determined from Figure 2.

L331: just typing out "regular" and "wildfire" might improve readability.

Author's Response (MC): We thank the reviewer for pointing this out. We have updated from "Reg" to "Regular" and "WF" to "Wildfire" in the manuscript.

L336: Lower variability in the high pollutant period can further be explained by the errors estimated in the linear fits of MA300 vs AE33. -- this sentence is confusing.

Author's Response (MC): We thank the reviewer for mentioning this. This line was in reference to old Figure 3, which has now been removed and our analysis with normalized b_{abs} values have been presented.

L337: but the AE33 - descendant of AE31 - is your reference method! So this is not a good comparison. What was the reference instrument used by Muller et al.?

Author's Response (MC): In the updated manuscript (**L418-L422**), we removed the comparison with Muller et al, and included a new comparison study done on multiple units of AE33 by (Cuesta-Mosquera et al., 2021).

Previously, Cuesta-Mosquera et al. (2021) tested 23 units of AE33 in both laboratory and ambient settings, assessing the instrument's performance before and after maintenance. They found that, after maintenance, AE33 tends to slightly underestimate (slopes slightly reduced from 1) for ambient aerosol measurements at wavelengths 590, 660, and 880 nm, but any wavelength dependency of the unit-to-unit variability of AE33 was not reflected.

L354: How did you verify the flow rate? You earlier mentioned the Drinovec et al. algorithm includes the effect of flow rate, so should not these fluctuations be accounted for by that method?
Finally, the flow rate fluctuations should impact all wavelengths equally, and yet the two methods are about equally better than the raw data for the (noisier!) UV. I am not sure this interpretation is well supported by these results, though it is fair to say that empirically, the onboard correction is a better technique.

Author's Response (MC): This comment was captured in Major Comment #3 and was addressed above.

L360: "noiser"?

Author's Response (MC): This line has been removed in the updated manuscript.

L364: Stronger than in the AE33?? Where was this shown?

Author's Response (MC): We have performed the loading effect calculations for individual devices and included the results in supplementary information (Table S6). Our results show that MA300s were experiencing significantly higher filter loading (0.054 μg cm^{-2} mL^{-1}) than AE33 (0.006 μg cm^{-2} mL^{-1}) during the wildfire period. To improve clarity, we have cross-referenced Table S6 where we discuss this point. (**L506**)

In our study, all three MA300s were subjected to strong loading effects (Table S6) in addition to RH changes (45% to 95%) during the sampling periods.

L365: "wildfire smoke affected aerosols can become hygroscopic with aging as compared to freshly emitted soot particles from diesel vehicle emission. Hence, the combined effect of imposed noise due to filter loading correction, and highly loaded hygroscopic aerosol" -- very speculative

Author's Response (MC): In the updated manuscript, in Section 3.3 (**L442-L447**) we include discussion the presence of aged fossil fuel derived particles with the variations of AAE values. And then compare the understanding of aerosol hygroscopicity with respect to wildfire emissions in the discussion of Source apportionment results in Section 3.5 (**L5487-492**)

From Section 3.3

The peak of AE33's α distribution was found to be 1.69, similar to Zotter et al. (2017) for SA calculations. In Figure 3, we present each device's hourly mapped α values calculated by UV-IR and Blue-IR pairs. Day-time α values for both wavelength pairs during the Regular period were found to be lowest and closer to unity, representing aerosol sources from traffic sources Healy et al. (2019); Bernardoni et al. (2021). In contrast, nighttime α values were found to be highest during the regular period, which could be attributed to the local wood-burning sources (Healy et al., 2019).

From Section 3.5

Being a near-road emission measurement site, our measurements captured complex aerosol mixtures of various mixing states. During the regular period, local traffic during the daytime contributes to fresh BC-enriched aerosols, which can be hydrophobic in nature (Sarangi et al., 2019; Wang et al., 2020) and by night-time, these fresh BC-enriched aerosols can evolve by ageing and change their morphological and optical properties. In contrast, during the wildfire smoke-affected days, the measurement site experienced enhanced quantities of aged aerosols through long-range transport from the Pacific North-West. These claims align with our calculated α values, as shown in Figure 3.

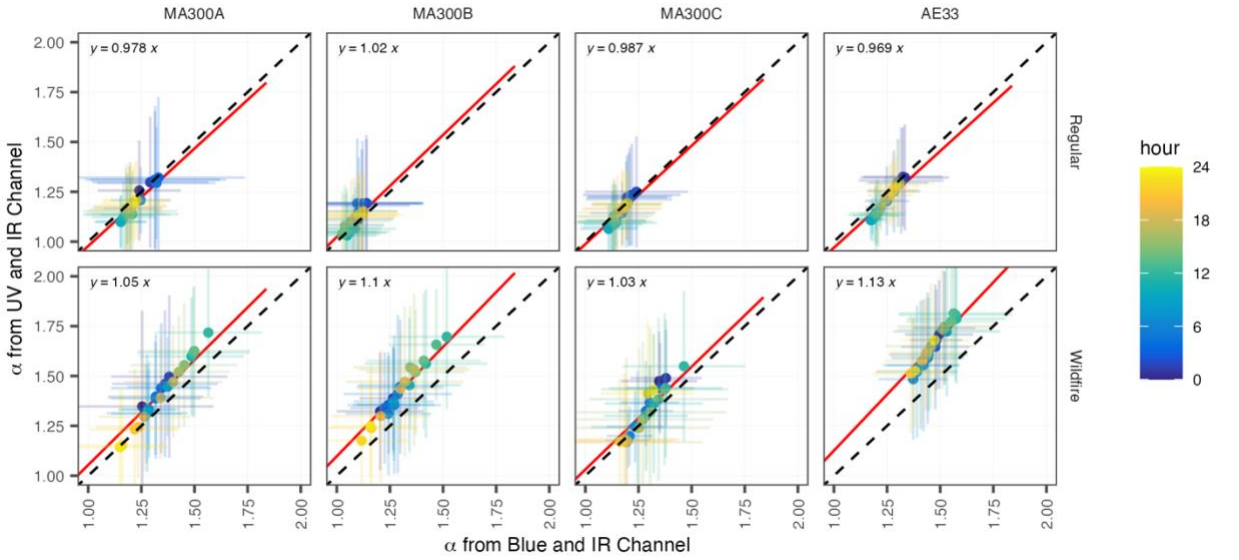

Figure 3. Ångström Exponent (α) (by the hour of the day) measured by different aethalometers during the Regular and Wildfire period. Average hourly AAE values derived from the Blue-IR wavelength pair (on x-axis) and the UV-IR wavelength pair (on y-axis) with error bars representing respective standard deviation. Red lines represent the linear relationship (forced through the origin), and the dashed line is 1:1.

L380: daily eBCf f concentrations were in the range of 0.6 – 1.9 µg m−3 during Reg period and in 1.0 – 3.8 µgm−3 during WF period. – lensing

Author's Response (MC): In the updated manuscript, we highlight the effect of lensing elsewhere (**L520-L526**).

During the Wildfire period, the enhanced eBC concentration was heavily influenced by biomass burning components, eBCbb; however, we hypothesize that the enhancement of eBCff may be due to two major factors. Firstly, there may have been a real increase in the number of heavy-duty vehicles during the early morning and evening hours that coincided with the Wildfire period, increasing the eBC emission from fossil fuel-based sources. Secondly, during wildfire smoke-affected days, aerosols can be a complex mixture of fresh and highly aged components, and the presence of BrC (Wang et al., 2019) and the lensing effect (Healy et al., 2015) have been found to impact the bulk aerosol light absorption measurements, particularly in the lower wavelengths.

L391: The *estimated" contribution. Why not run the source apportionment for five days on either side of the wildfire period?

Author's Response (MC): We thank the reviewer for this suggestion. We ran a source apportionment analysis as mentioned previously in the response to Major Comment #7.

L393: "be lower"

Author's Response (MC): We thank the reviewer for this correction. We have incorporated the change.

L394: In the case of biomass burning-based source contribution, eBCbb, 22% underestimation was identified during the WF period using Blue-IR based SA instead of UV-IR. Utilizing Blue- IR method based source apportionment resulted, 41% overestimation of eBCbb component -- these sentences are contradictory. I am having a hard time following this argument.

Author's Response (MC): We thank the reviewer for highlighting this. Our results identify that overestimation of eBC_{bb} during the Regular period using Blue-IR pair can be flawed (Figure S5 and Figure S11 (Panel A)). However, we choose to report our observations and discuss the failure of Source Apportionment results with low b_{abs} signals with the help of sensitivity analysis performed. These were discussed in **Section 3.5.1, L540-L548.**

It is important to note that derived absolute eBC_{ff} and eBC_{bb} components are dependent on the absolute b_{abs} inputs in the Aethalometer model and inherently, $b_{abs,UV}$ measurements are higher than $b_{abs,Blue}$. The effect of different input levels of b_{abs} and α pair on the Aethalometer model was explored through a sensitivity analysis and has been presented in Figures S10 and S11. For test purposes, we used AE33's mean b_{abs} concentrations for UV, Blue and IR channels from the Regular and Wildfire period with α_{bb} range 1.6 -- 3.0 and α_{ff} range 0.8 -- 1.5. Sensitivity analysis results show that apportioned $b_{abs,bb}$ and $b_{abs,ff}$ can often get negative values or even higher than the input b_{abs} values, which is an established flaw of Aethalometer model Grange et al. (2020). For a clean environment, lower b_{abs} input can cause large errors in the estimates of $b_{abs,bb}$ and $b_{abs,ff}$. Hence, we claim that source apportionment should not be conducted below the MDL of the black carbon concentration, which has been found as 0.21 μg m^{-3} for MA300.

Figure 6: if the Drinovec correction is not appropriate, why show it here?

Author's Response (MC): We thank the reviewer for this suggestion. In the revised manuscript only MA300's instrument data were included (Figure 6). Results from modified Drinovec correction have been moved to supplementary information (Figure S13).

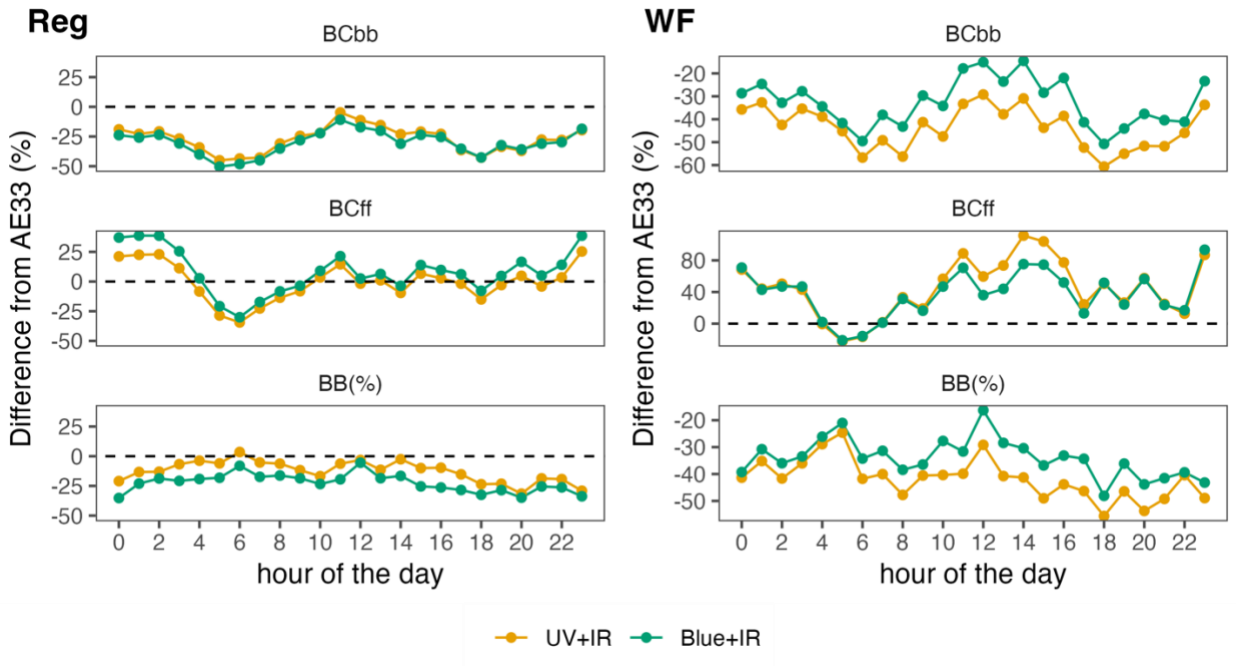

Figure S13 Percentage difference of hourly SA results in between MA300 and AE33. For MA300s, the average of onboard corrected data for the two periods (Regular and Wildfire). SA results from UV-IR and Blue-IR wavelength pair have been evaluated separately

L441: "lower"

Author's Response (MC): Updated in the revised manuscript.

L449: However, we find that MAE of eBC measurement from MA300 (compared to the AE33) can be in the range of 0.44 – 0.98 µgm−3 depending upon the measurement period. We observe larger MAE during high pollution conditions (e.g., WF period in this study). Based on these findings, caution may be required when MA300s are used to capture spatial or temporal differences in eBC below the 0.98 µg m−3 threshold. -- no, that's not the MDL.

Author's Response (MC): We agree with the reviewer that this is not the MDL. Here, we were emphasizing on capturing the spatial or temporal differences in absolute eBC measurements using MA300s. However, this cannot be treated as a limit of detection for MA300, but a concentration level to identify meaningful differences in measurement. We have now added a sentence that clarifies that the MDL should also be taken into account in experimental design. As discussed in Major Comment #8, we have now calculated the MDL following (Hyslop et al., 2022).

L454: I thought it was 1-minute data.

Author's Response (MC): We thank the reviewer for pointing this mistake out. Yes, it is 1 min data. Updated to 1 min in the revised manuscript.

L457: We were able to eliminate the drifted signals through post-processing the raw data using statistical outlier detection method. – this seems odd.

Author's Response (MC): Yes. This was highlighted in Section 2.4 (**L198-L200**) with the aim to raise awareness about the LED performance in MA300 units.

We observed that the MA300 sensor data was affected by both drift and post-filter- change transient effects. The drift in the photometer data was removed by calculating statistical outliers before calculating ATN and has been explained in detail in the Supporting Information (Section C).

L468: we hypothesize that an inconsistency is flow fluctuation in MA300, which is a key variable in deriving the real-time loading correction parameter. – a key roadblock.

Author's Response (MC): We thank the reviewer for this suggestion. We have updated the line (**L621**) accordingly.

We hypothesize that inconsistency in flow fluctuation in MA300 is a key roadblock in deriving the real-time loading correction parameter.

L470-476:
More like 21%,
these are reference methods for you, come on,
sounds like an LED problem, not a filter loading problem.

Author's Response (MC): We incorporated the suggestions made by the reviewer and updated the text (L630-L636)

Characterizing unit-to-unit variability can speak to instrument precision and may be particularly important for use cases where multiple MA300s are simultaneously deployed to measure a pollution event. We reported the precision of MA300 eBC in terms of unit-to-unit variability (based on normalized responses) as 5%. This value is slightly higher than that reported for other aethalometer models: 4.3% for AE31 (Müller et al., 2011), 0.5% for AE33 (Cuesta-Mosquera et al., 2021). This variability can increase to 21 % when absolute measurements are considered, reflecting individual instruments' sensitivity and noise characteristics affecting precision. For the multi-wavelength babs, the highest unit-to-unit variability was found in the UV channel (8%) with large instrumental noise, which is consistent with previous studies on multi-wavelength aethalometers (Cuesta-Mosquera et al., 2021; Müller et al., 2011). The unit-to-unit variability in the UV channel was not identified to be significantly varying with filter loading impacts and hence could be occurring due to problems associated with LED light sources or detectors.

L479: That should be semi-volatile organic compounds (because they condense on the filter material otherwise), but I think it is not a good hypothesis - it is possible there were denuder losses. Not going through Vecchi to verify, sorry.

Author's Response (MC): We have considered the presence of semi-volatile organics in the updated manuscript. (**L650**).

Light absorption measurements in the UV channel can also be sensitive to interference from the volatile to semi-volatile organic compounds on the filter tape Vecchi et al. (2014) or from the other light-absorbing non-BC combustion particles, which affect lower wavelengths more than higher ones.