Author’s response to RC2
Manuscript: amt-2020-325

Anonymous Referee #2

The authors present an interesting intercomparison of filter absorption photometers at a regional background site at Hyytiälä, Finland. The comparisons of this kind are important as a full characterization of the instrumental response is lacking and is compounded by the interwoven non-linearities of the measurement, which presents themselves as measurement artifacts, or as the authors’ call this, systematic errors. The comparison of the MAAP, the AE31 and the PSAP partially addresses these shortcomings and presents new viewpoints on an urgent topic.

The manuscript fits well with the scope of AMT and can be accepted for publication after addressing the following major and specific comments.

The authors correctly point to the influence of the correction algorithm and its effectiveness on the slope of the inter-instrumental regression, which is used as the multiple scattering correction factor (C_{ref}). The loading effect and the multiple scattering are artificially separated in the correction algorithms. Additionally, the particles, embedded in the filter, cause a known cross-sensitivity of the filter photometers to the scattering, which is explicitly described in the Arnott et al (2005) algorithm. Filter photometers also feature a dependence of the sensitivity on the location/depth of the particles in the filter matrix and are their sensitivity is therefore dependent on the size distribution of the sampled absorbing particles.

Weingartner et al. (2003), Park et al. (2010), Hyvärinen et al (2013), Segura et al (2014) and Drinovec et al. (2015) have discussed different approaches to showing the magnitude of these artifacts and their dependence on the loading of the sample spot. The authors should follow the same principle and plot the attenuation and absorption coefficients, and the absorption Angstrom exponent (AAE) as a function of the loading of the sample spot, for example as a function of ATN, Tr, ln(Tr)… for all filter photometers in the study. This will also serve as a strong argument for using the MAAP as the reference.
The figures of the σ_{abs} against *ATN* and *Tr* are now presented in the supplementary material (Figs. S2 and S3) and we also included a table (Table 3 in the main text) that describes the performance of the AE31 and PSAP against MAAP for different filter loading intervals. At SMEAR II the MAAP was set to advance filter spot in 24 h interval and therefore, plotting the σ_{abs} against the MAAP filter loading gives a false impression that the σ_{abs} increases notably with increasing filter loading. This behavior is caused artificially with our data because the higher filter loading values are only reached when the σ_{abs} is high (i.e., the higher the σ_{abs}, the more accumulation on filter within 24 h period). Because of the uninformativity and possible misconceptions, we did not plot the MAAP BC against the filter loading. The absorption Ångström exponent was already plotted against Tr and ATN in Figs. 7 and 8 (note new figure numbering).

We added/modified some text to discuss the new figures and table: *Surprisingly, the not corrected AE31 (Fig. 5e) did not seem to have a significant difference in correlation coefficient compared to for example to the data corrected with W2003 or CC2010 (Figs. 5a and d, respectively). However, the relation between the $\sigma_{abs,NC}$ and*

σ_{ref} depend more on the ATN than for any filter loading corrected data, which is shown by the color coding (ATN) of the data points and in Table 3, which presents the slopes of the linear fits and R^2 values for different ATN intervals. If only data from highly loaded filter (ATN > 60 at 660 nm) were taken into account, the slopes of the linear fits were 0.97, 1.06, 0.99, 0.95, and 0.93 for W2003, A2005, V2007, CC2010, and not corrected (NC), respectively. The smallest decrease in the slope with increasing ATN determined for the loaded filter was observed for data that was corrected by V2007. Interestingly, the slopes for the loaded filter actually increased for data that was corrected by A2005, meaning that the R_{ARN} had a relatively big effect with increasing ATN. The biggest decrease in the slope determined for a highly loaded filter was observed for the NC data, as expected. This observation underlines the need for filter loading correction if one studies shorter time periods. For longer time periods (e.g., trend analysis or studies of seasonal variation) the effect of the ATN smooths out, but for shorter time periods (e.g., case studies) the changing ATN can have a notable effect on the results if no filter loading correction is applied.

The dependency of the σ_{abs} on the ATN and Tr is presented in supplementary material (Figs. S2 and S3). On average, the decrease in the σ_{abs}, which is not corrected for the filter loading ($\sigma_{abs,NC}$ and $\sigma_{abs,PSAP,ATN}$), with the increasing ATN and decreasing Tr is not clear. This effect is better seen in the results presented in Table 3. However, Fig. S3 shows that especially for the PSAP, the use of correction algorithms decreases the variation, which is a strong recommendation for using the correction algorihtms. This is also seen in the AE31 data, but the effect is less notable (Fig. S2).

According to the R^2 values presented in Table 3, the precision of the AE31 decreased with increasing ATN. For example, the data corrected with the A2005 algorithm, the R^2 decreased from 0.96 for a clean filter (ATN < 20) to 0.90 for loaded filter (ATN > 60). However, the decrease in R^2 was quite minor. Miyakawa et al. (2020) observed also rather high R^2 values between Aethalometer (model AE51) and a reference instrument (single particle soot photometer and COSMOS) when the ATN was below 70, but when the ATN exceeded 70, the R^2 decreased more rapidly.

Page 2, Lines 30 – P3, L5: There is another systematic error, not considered by the authors – the measurement of flow. The first issue is the reporting conditions of the flow: have they been unified across all instruments? If yes, please state the conditions in the respective Measurements and methods sections. The authors should include a word of caution for the instrumentation and the determination of the leakage – this is a multiplicative factor affecting the slope between instruments, which is (in the experience of the reviewer) often interpreted as being intrinsically instrumental.

We had forgotten to mention this in the original manuscript even though the flow calibrations were included in the data analysis. We have now added a paragraph describing the flow calibrations in the Sect. 2.2: *In order to perform a comparison study between the different absorption photometer, the sample flows of the instruments were calibrated. The sample flow of each instrument was regularly measured with a gillian flow meter and the flow reported by the instruments were corrected to match with the gillian measurements.*

P3, L 7-15: Please add the discussion on independent check of the correction algorithms with references to Park et al. (2010), Hyvärinen et al (2013), Segura et al (2014) and Drinovec et al. (2015).

Added: *In general, after correcting the data for the multiple scattering and loading effects, the absorption instruments agree rather well with the reference measurements (Drinovec et al., 2015; Hyvärinen et al., 2013; Park et al., 2010; Segura et al., 2014).*

P4, L 16: Please add the widths of the different "wavelengths" in the filter photometers (for example from Müller et al., 2011).

We added more description: *Here, we reported the typically used AE31 and PSAP wavelengths, which are reported in the AE31 manual and by Virkkula et al. (2005), respectively. These reported wavelengths deviate slightly from the ones measured and reported by Müller et al. (2011a) (see their Table 6). For the MAAP, we decided to use the wavelength reported by Müller et al. (2011a) since it rather commonly used and it clearly deviated from the wavelength reported by the manual.*

P4, L 21: The reference to Fig. 1 is to a very nice picture of the experimental setup (which should remain in the manuscript) and not to the missing data availability plot. This missing figure could be added to the Supplement.

We moved the availability plot in the supplement, so now the schematics of the set-up is Fig. 1.

P4, L 21-28: It is RH change that perturbs the filter measurements, not RH per se. It would be interesting to take into account the RH change rate as well. For example, plot a companion to Fig. 4 with RH change rate, same for other instruments.

We did study this. However, the change in *RH* did not seem to affect the C_{ref} at all. The correlation coefficient (*R*) between the C_{NC} and *RH* change (in units of %-points h^{-1}) was -0.03 and the *R* between the C_{NC} and relative change in *RH* (units % h^{-1}) was -0.04. We added a statement about this in the manuscript: *The effect of rate of change in RH on the C_{ref} was also studied, but there rate of change in RH did not show any correlation between C_{NC}.*

P4, L 30: Reference to Fig. 2 is in fact reference to Fig. 1.

Fixed this.

P 5, L 8-10: Add the information on the filter material used.

Added this information: *AE31 operated on quartz fiber filter (Pallflex, type 55619A), PSAP on quartz fiber filter (Pall, type E70-2075W), and MAAP on glass fiber filter (Thermo Scientific, type GF 10).*

P 5, L 17-18: This is incorrect. The intensities in PSAP and AE31 are normalized to the intensity measured under the clean part of the filter – the reference sample spot. This takes into account any possible drift in the LED intensities during the measurement period.

This was modified to: $I_{t-\Delta t}$ *and* I_t *are the measured and normalized light intensities through the filter in the beginning of the measurement period* $(t-\Delta t)$ *and in the end of the measurement period (t). The intensities are normalized by comparing them to the intensity measured through a reference spot. Normalizing the intensities accounts for the possible drifts and changes in the intensities of the LEDs.*

P5, L28: The sample spot should be measured. It changes with each spot slightly, especially due to leakage, when the filter tape is not well sealed. Was the correction for the differing values of A taken into account in this work?

Like for the flow correction, we forgot to describe the correction of filter spot size in the manuscript. This was applied already before, but now we added a description in the manuscript: *For the PSAP and AE31 we used the A values of 18.1 and 54.8 mm*^{2}*, which did deviate from the default ones that were 17.8 and 50.0 mm*^{2}*, respectively. The A used by default in the MAAP matched the measured one and therefore it was not corrected.*

P 6, L 6-12: The loss of sensitivity due to non-linear effects could be presented better. Please rewrite.

Rewrote this: *Absorbing particles induce a so-called "shadowing effect", which decreases the change in the intensity* $(I_{t-\Delta t}I_t^{-1})$ *as the filter gets more loaded (Weingartner et al., 2003). This means that the instrumental response is non-linear with increasing filter loading. The increasing filter loading has an opposite effect than the scattering of the filter fibers and particles: the absorbing particles collected on the filter decrease the optical path and therefore the reported* σ_{ATN} *for a loaded filter is lower than for a pristine filter. This non-linearity is considered in the various correction algorithms presented in Sect. 2.3.1 and 2.3.2.*

P 6, L 19-20: MAAP artifacts can be checked by a BC(*ATN*) plot, please see above. This justifies the use of the MAAP as the reference (further below, next paragraph).

As already answered before, this figure was not included in the manuscript even though it would be useful and justify the use of MAAP as you said. However, due to the filter change in every 24 h, the figure shows an artificial increase in the BC concentration with *ATN* (i.e., high *ATN* values are only reached when the concentration of BC is high), which does not represent the reality.

P6, L 29: The authors talk about the precision here, not accuracy. Accuracy, however, is the parameter which is of importance. Please see above regarding the justification of the MAAP as the reference.

Please see the reply to the latest comment. We modified the sentence a bit: *According to the uncertainty and unit-to-unit variability, the MAAP is the most precise instrument for monitoring* σ_{abs} *and black carbon (BC) concentration, which is typically derived from* σ_{abs} *measurements. Also, the backscattering measurements from the filter reduce the artefacts caused by the scattering aerosol particles and the filter loading effect making it a more accurate instrument.*

P 7, L 5: The unit-to-unit variabilities of different aethalometer types is very different - please expand and reference Müller et al. (2011) and Cuesta et al. (2020).

This was mentioned in the previous paragraph: *Müller et al. (2011) reported that the unit-to-unit variability of the PSAP, AE31 and MAAP were about 8%, 20% and 3%. It must be noted that the unit-to-unit variability is a lot smaller, about 2 % for the new AE33 model (Cuesta-Mosquera et al., 2021).*

P 7 – 11, sections 2.3.1 and 2.3.2: I disagree with the Anonymous Referee #1, these sections are important for understanding and interpretation the rest of the paper and should remain in the body of the manuscript.

We decided to keep the algorithm descriptions in the main manuscript and not to move them in the supplementary material.

P 8, L9: Please define single-scattering albedo as “omega”.

Fixed this.

P 8, L 13: Why linear dependency – compare Virkkula et al. (2007 and 2015).

I’m not sure if I understand the comment. Neither of these articles discussed the WEI2003 algorithm. Also, the effect of using different type of interpolation method should only cause a minor difference in the data.

P 8, L 13: Please define Angstrom exponent as “alpha_sca”.

Fixed this.

P 9, L 3: “… were calculated from…”: Not clear if this relates to o the Hyytiälä measurements or to the Arnott et al. (2005). Please rephrase.

This was rephrased: *To calculate the* $\tau_{a,fx}(\lambda)$*, we used the same a power law function* $\tau_{a,fx}(\lambda) = \tau_{a,fx,521} \cdot (\lambda/521\ nm)^{-0.754} = 0.2338 \cdot (\lambda/521\ nm)^{-0.754}$. $\tau_{a,fx}$ *as Virkkula et al. (2011) and the resulted values are presented in Table 1.*

P 10, L 27: Which PSAP filter do these values relate to (Ogren et al., 2017)?

The Pallflex ones that are now also mentioned in the manuscript,

P 11, L 20-21: The averaging of PM1 and PM10 values is non-trivial due to possible regional contributions to BC in the larger size fractions. This does influence the nonlinearities, which in-turn cause measurement artifacts that need to be corrected. The introduction mentions no change in the size of the sampled particles. The authors mention this briefly in section 4.1. Please add this information and provide an argument and discussion how this could influence the comparison.

As we discuss about absorption measurements with filter-based methods. When we correct the data depending on the ATN decreases because of both PM1 and PM10 and it is impossible to separate the effects of these in the loading correction. We added a sentence: *Since all the instrument measured the same sample air, combining the PM1 and PM10 data caused no discrepancies between the instruments.*

We also added discussion in the results: *Because it is impossible to separate the effect of different size cuts from a loaded filter, here the PM1 and PM10 measurements were combined and averaged together. In general, PM1 accounted for about 90% of the PM10* σ_{abs}*; for the* σ_{sca} *the fraction of PM1 was about 75%* (Luoma et al., 2019). Because absorbing particles, which is considered to consist mostly of black carbon, are typically in the fine mode (diameter < 1 µm), the σabs is not expected to deviate much between the different size cuts. However, the differing size cuts, which causes more deviation in the σsca, could have affected the σabs measurements since the particulate scattering causes apparent absorption and affect the multiple scattering in the filter. For example, the coarse particles (diameter > 1 µm) do not penetrate as deep in the filter as the fine mode particles, which could possibly influence on the Cref values. In an ideal situation the PM1 and PM10 absorption would have been measured by separate instruments.

P 13, L 10-11: Why calculate C_{NC}? It is loading dependent.

Sometimes the σ_{abs} is derived from AE31 data without any loading correction by using only some estimation for the C_{ref}. For example, WMO/GAW recommends (report 227) to apply a C_{ref} value of 3.5 and not to apply any filter loading correction to the data. This value can be compared against that. This was argued for earlier in the text: *Current recommendation by the WMO and GAW is to assume the R(ATN) unity for the AE31 and to use a* C_{ref} *value of 3.5, which was determined by a comparison study of different AE31 instruments (WMO/GAW, 2016). Therefore, we also studied "not-corrected" AE31 data for which we did not apply any R(ATN) correction or particulate scattering reduction, but only the multiple scattering correction.*

P 13, L 12-13: AAE is an absorption property, attenuation features loading effects, making AAE impossible to calculate, especially measurement at the lower wavelengths are heavily loading impacted. This paragraph needs to be extended and additional explanation on the determination of AAE provided.
When *ATN* (or σ_{ATN}) is plotted against log(λ) the plot follows linear fit. Therefore, there should not be a problem in using Ångström exponent to interpolate the *ATN* to different wavelengths. In general, Ångström exponent describes the wavelengths dependency of any optical property that is linearly dependent on the log(λ).

P 13, L 20: Please extend the description of the fit – regression, it is not completely clear, cite Eq. 19...
Rephrased: C_{ref} *was determined as the slope of a linear regression for the whole data set (linear fit for a loading corrected* σ_{ATN} *vs.* σ_{ref} *-plot.*

P 13, L 21-22: The wavelength dependence of *C* is discussed in Bernardoni et al. (2020), which can be added to the discussion below (section 4.1, especially P 15, L24), provided it is calculated here.
I am not sure if I understood this comment and I do not quite understand what calculation is referred here as Bernardoni et al., (2021) did not observe statistically significant wavelength dependency for the C_{ref}. We added discussion (including the reference to Bernardoni et al., 2021) in Sect. 4.1: *There are both studies where constant* C_{ref} *has been used and studies where wavelength dependent* C_{ref} *has been used. Others observed no significant dependency for the* C_{ref} *on the wavelength (Backman et al., 2017; Bernardoni et al., 2021; Collaud Coen et al., 2010; Weingartner et al., 2003; WMO/GAW, 2016).*

P15, L 7-8: This is the place to discuss the influence of the correction algorithm performance on the *C*.
We added a section (Sect. 2.3.3) that describes the differences of the algorithms: *The W2003 algorithm only depends on the ATN, otherwise it applies constant values and it does not consider the scattering subtraction. The A2005 is not a function of ATN but it takes the filter loading into account by summing the* σ_{abs} *of the accumulated particles on the filter spot. It does not assume a constant for the scattering reduction but determines the fraction from the wavelength dependency of* σ_{sca}*. The CC2010 algorithm is similar to A2005 in a sense that it also defines its own scattering reduction factor and determines the filter loading correction by taking into account the properties of the particles accumulated in the filter. The V2007 only depends on the difference between the last and first measurements of two filter spots and it assumes no constant coefficients. The B1999 algorithm relies heavily on constants that describe the dependency on the Tr, whereas the V2010 algorithm is an iterative process that depends on the* ω*. Both B1999 and V2010 consider the scattering reduction with a coefficient.*

Then the performance of the different algorithms were discussed more in Sect. 4.2 (not in the section about C_{ref}) because we though a deeper discussion on the algorithm differences fits better in there.

P 16, L 16-19: C_{ref} is the effective slope relative to the MAAP. Please add some discussion on the artifacts of all methods and their similarities/differences. What about size distribution artifacts? See also P17, L7.

The comment about the size distribution is answered below. Modified the sentences and added some discussion: *Since the C_{WEI}, and C_{COL} had a similar seasonal variation, it is unlikely that the seasonal variation observed for C_{NC} was caused by the lack of filter loading correction. It is rather surprising, that there was a seasonal variation for the C_{COL} as well. The algorithm by CC2010 considers the wavelength dependency of scattering and the ω of the accumulated particles. It is rather surprising that taking these parameters, which have a seasonal variation at SMEAR II (Luoma et al., 2019; Virkkula et al., 2011), does not seem to reduce the seasonality of C_{ref}. For example, the seasonal variations between the C_{WEI} and C_{COL} are surprisingly similar, even though we applied constant f values in W2003*

The seasonal variations for the C_{ARN} and C_{VIR} were less obvious than for C_{WEI}, C_{COL}, and C_{NC}. The lesser seasonal variation for the C_{ARN} might be explained by the subtraction of the scattering fraction before the loading correction was applied and the C_{ARN} was determined. The fact that the C_{ARN} has less data points than the other the C_{ref} values, might also explain part of the less amplified seasonality. For C_{VIR}, the lack of seasonal variation for was probably caused by the very strong seasonal variation of the compensation parameter (k; see Fig. 10a) as will be discussed below in Sect. 4.4. The algorithm by V2007 does not assume any coefficients but depends only on the difference between the last and first measurements of the filter spots. Therefore, it seems to adjust to seasonal changes whereas the other algorithms apply coefficients. According to our results, the V2007 and A2005 accounted well the variations in the optical properties of the particles embedded in the filter and therefore the seasonal variations in the C_{VIR} and C_{ARN} were reduced.

P 16, L23-24: Or it describes the variation of the artifact better. Is this dependence on the parametrization scheme? Averaging?

This dependency is not parametrized since the data is just “justified” over the filter spot change, that creates the automatic seasonality

P 17, L 7-8: This can be quantified, there are relevant measurements at Hyytiälä. Please provide this information.

We did a quick comparison against the volume mean diameter (VMD) and geometric mean diameter (GMD), which were calculated for the PM1 and PM10 size modes by using the DMPS and APS data (in a similar manner as in Luoma et al., 2019). However, the correlation coefficients between these parameters and the C_{ref} were very low (< 0.05). We did not add any discussion about this test in the manuscript to avoid adding too much description about size distribution measurements, which were not in the focus here. The APS and DMPS measure the size distribution of the whole aerosol population. In this situation, measurements of the soot particle size would have been useful. Unfortunately, this kind of an instrument is not operating at SMEAR II.

P 17, L 14-17: This can also be described in a more quantitative manner, please see Virkkula et al. (2015) and Drinovec et al. (2017).

We do this later in the manuscript (Sect. 4.4).

P 18, L 10: The intercept of the linear fit is the scattering artifact.

We added: *As presented in Table 3, the linear fits for the AE31 and PSAP data against the reference did not have an intercept of zero. This could be caused by the scattering artifact and the fact that the correction algorithms fail to take the scattering artefact completely into account. The intercept is the smallest for the B1999-corrected PSAP data and the largest to AE31 data. A fraction of* σ_{sca} *is subtracted in the AE31 algorithms by A2005 and CC2010. However, the data corrected with these algorithms still have a higher or similar intercept as with the not-corrected data and the data corrected by the W2003 and V2007 algorithms. Considering the intercept, the V2007-corrected data performs the best in the AE31 vs. MAAP comparison, which is slightly surprising, since it does not take the scattering subtraction into account. For the V2010-corrected PSAP data, the intercept is negative suggesting that the V2010 algorithm overestimates the apparent absorption by scattering particles.*

P 18, L 21: The data featuring low *ATN* is the one which features low loading artifacts and, therefore, a *C* with less uncertainty. This can be explored and the uncertainty as a function of the loading determined quantitatively.

Actually we tested determining the C_{ref} for different *ATN* limits and the effect on the uncertainty was minor. This is also seen in Table 3, which reports the R^2 of the linear regression for different *ATN* intervals. The R^2 decreases only slightly with increasing *ATN*.

P 18, L 27-29: The "smoothing" is site dependent and the non-corrected regression slope is always lower. The r2 of the non-corrected regression nis lower as well. Please discuss.

Modified this and added some text: *For longer time periods (e.g., trend analysis or studies of seasonal variation), the effect of the ATN on the variation smooths out, but for shorter time periods (e.g., case studies) the changing ATN can have a notable effect on the results if no filter loading correction is applied. However, not correcting for the filter loading effect, the precision of the instrument and the* σ_{abs} *or BC concentration on average are reduced, which is why applying a filter loading correction on filter-based photometers is always recommended.*

P 18, L 32: This is surprising, as one would expect that at low loading, the influence would be minimal. Is this a parametrization effect. Please discuss.

There was probably a misunderstanding here (mixing of low *Tr* to low loading). The sentence/paragraph was rephrased: *Figure 6b shows that V2010 overestimated the* σ_{abs} *especially when the loading was high (Tr was low) and the linear regression was 1.24.*

P 19, L 1-2: Please elaborate, the text is unclear.

Rephrased and elaborated: *Also, the B1999 overestimated the* σ_{abs} *slightly, but in general it performed better in comparison with MAAP and the slope was 1.07 (Fig. 6a). The linear fits in Figs. 6a and b include all the data, but Table 3 presents the slopes of the linear fits for data with different Tr limits. It is actually recommended to use PSAP data with Tr > 0.7 and if only this data is taken into account, especially the data corrected with the V2010 algorithm performs much better and has a slope of 1.01, but also the slope for the data derived with the B1999 algorithm yields a smaller slope of 1.04. Unlike for the AE31, the loading on the filter does not seem to affect to the precision of the PSAP at all as the* R^2 *values do not decrease with the increasing loading (Table 3).*

P 19, L 18: This is different than explained above, Eq. 16. It is actually much more quantitative, as it allows the selectin of "good" AAE values by evaluating the fit r2, and ignoring the AAE values with low r2. This

is used in French monitoring networks as a parameter to quality control the data and source apportionment of BC. Please use the r2 AAE selection and add this information in the manuscript.

The wrong sentence was modified to: *The* α_{abs} *was determined as a linear fit over all the selected wavelengths according to Eq. (16).*

The selection of only good AAE values according to the R^2 of the fit is very interesting. However, unfortunatelly, due to time limitations, we could not conduct the evaluation of the fits according to their R^2 now. However, we did a quicker check on the R^2 values in general for majority of the 1 h averaged data (~99 %) the R^2 > -0.90. Therefore removing AAE values with low R^2 should not make a notable difference in the data analysis.

P 19, L 31 – P 20, L 2: Please see above and Bernardoni et al (2020) and add to the discussion.

Discussion was added: *This could be done with several MAAPs operating at different wavelengths, by measuring the particles suspended in the air by photoacoustic method (Kim et al., 2019), by a polar photometer (Bernardoni et al., 2021), or by a Multi-Wavelength Absorption Analyzer (MWAA; Massabò et al, 2013).*

P 20, L 30: What was the maximum AE31 ATN for advancing the spot? Please add to the instrumental section.

We added a paragraph about this in Sect. 2.3 Absorption measurements: *At SMEAR II, the MAAP advanced the filter spot automatically once per day in 24 h intervals. The AE31 also advanced the spot automatically when the ATN reached 120 at 370 nm wavelength. The PSAP filters were aimed to be changed every second day, but due to weekends and holidays, the filters were sometimes changed only after several days. On average the PSAP filters were changed once per three days.*

We also added here: *on average the filter changed when the ATN at 660 nm ≈ 90.*

P 20, L 31: Is this an observation of the data reported here (circular reference?) or an observation of Virkkula et al. (2007 and 2015) and Drinovec et al. (2017)?

Rephrased this: *The k is often larger for the shorter wavelengths, which means that the non linearity caused by the increased filter loading is relatively stronger at the shorter wavelengths (Drinovec et al., 2017; Virkkula et al., 2007; Virkkula et al., 2015), which was also observed by this study (discussed in the next chapter).*

P 20, L 32: This is not true. The correction algorithms take care of this. AAE dependence on ATN means that the loading correction is not working well. This is crucial as it shows that, except for V2007, the loading corrections do not function well! This is surprising, as this is the only correction not taking into account the cross-sensitivity to scattering. Why is "wavelength dependent k" better than other parameterizations? Same should be done for b_abs.

Rephrased this: *According to these results, the other algorithms but V2007 do not seem to account enough the wavelength dependency of the R(ATN).*

We also added figures of the effect of loading on the absorption (see my first reply to Anonymous referee #2).

P 23, L 10-12: This depends on the rate dATN/dt, or the number of spots measured. This number can be counted and can be provided here. It is a good parameter for quality control and an important finding of the manuscript.

We added a recommendation here: *The suitable period for the running average at each site depends on the rate of change in the ATN, which determines how often the filter spots are changed. According to this study and to the study by Virkkula et al. (2015) time period that includes about 6 – 9 filter spot changes on average seems to yield good results. At SMEAR II, a relatively clean site, this period was 14 days and at SOPRES, a rather polluted site, the period was 24 hours.*