Author's response to RC3
Manuscript: amt-2020-325

Anonymous Referee #3

The manuscript covers an important topic that has puzzled researchers for decades: the need to accurately measure light absorption by aerosol particles. Light absorption by aerosol particles are fundamental when assessing the direct radiative impacts of aerosols in the air but also on snow and ice. The work investigates how these measurements differ based on which post processing method is used in the quest to determine the absolute amount of light absorption by aerosol particles. The work covers three different filter based absorption photometers and how they compare against each other. The work further extends the analysis to cover how these post processing methods affects the spectral dependence of the light absorption coefficients and how this can lead to misleading conclusions when comparing one measurement to another if not considering that the post processing method is of great significance.

General comments:
The Introduction would need a section where the goals of the study are clearly stated and then these goals should be addressed one by one in the conclusion section. This would help readers to grasp the extent of the research covered by the article.
We now combined the aims in to one paragraph in the intro section: *This study has two aims, which address the variation of the C_{ref} and the differences between the different correction algorithms. The first aim is to provide a C_{ref} value suitable for a boreal forest site and to study how the C_{ref} varies between different correction algorithms. The second aim is to present how the different correction algorithms of the σ_{abs} affect the measured optical properties of the particles.*

The manuscript has dedicated a substantial proportion to the multiple scattering enhancement factor used in the Aethalometer post processing algorithms in the quest to make them perform better against the reference instrument MAAP. It is justified to scrutinize the multiple scattering enhancement factor of the Aethalometer but no attempt is made to scrutinize the multiple scattering enhancement in PSAP filters. PSAP filters are not as optically thick as the more rigid MAAP and Aethalometer filters but multiple scattering is bound to occur in those filters too which would warrant a similar kind of investigation that is now presented for the AE31.
This indeed would be an interesting study to add in the manuscript and otherwise we would follow this recommendation and determine the Cref value also for the PSAP. Unfortunately, now the main author simply does not have time to conduct this study, since it would require quite a lot more of data analysis and rewriting the current scripts. To cover this comment, we did added discussion of the topic in the results: *If all the data was included in the comparison, as in Figs. 6a and b, the overestimation of σ_{abs} would suggest to derive the C_{ref} values also for the PSAP data. Here, we did not derive the C_{ref} values for the PSAP, since they are not typically used in a similar way as for deriving the σ_{abs} from the AE31 measurements. In general, the multiple scattering does not cause such a big artefact in filter material typically used in PSAP compared to the more thicker AE31 filters. However, if we considered only the data below Tr < 0.7 the PSAP and MAAP agree well for both correction algorithms. This result then suggests that there is no need for deriving a new C_{ref} for PSAP. Svensson et al. (2019)*

studied the multiple scattering in quartz filters and they derived the equations that can be used in determining the C_{ref} value for PSAP. Differently to AE31 correction algorithms, the C_{ref} used in PSAP algorithms is included in the coefficients of Eqs. 13 – 15 and therefore determining the C_{ref} for PSAP is not as straightforward.

I wonder if the title of the manuscript couldn' be changed to something more inviting. The focus is on which effects different correction algorithms have on the post processed data which is an important topic indeed. Could the authors consider being more specific other than saying the manuscript deals with 'effects on different correction algorithms'. E.g. Effects of different correction algorithms on absorption photometers can lead to wrong interpretations if not… or something along those lines.
We now at least rearranged the title so that the main thing, which is the effects of different algorithms, is first.

The English is generally good and it is easy to understand what the authors mean. There are however grammar errors that would need to be corrected and would improve the readability of the manuscript; e.g. definite articles and prepositions can be wrong or missing. In my specific comments I have made comments on those but the list is not exhaustive.
Thank you for putting so much effort in improving the language! We combined all the comments of language suggestions at the end of this document in order to keep the answers to bigger comments more easy to read.

After addressing these comments and the specific comments below the manuscript is within the scope and of high enough scientific quality to be published in AMT. Please do also consider the specific comments below for the revision.

Specific comments:
P2L12: The sign of the radiative forcing is mentioned but could you be a bit more specific in what those signs actually mean i.e. write out cooling and warming instead of referring to the signs.
We added: *i.e., negative sign for the cooling effect and positive sign for the warming effect* in parenthesis in the text.

P2L16-19: sigma is a measure of light absorption and scattering, so it does more than "describe" it.
"Describes" was changed to "is a measure".

P2L24: I think that they are actually more unknown or not understood than actually defined.
We rephrased this to give an angle that the correction algorithms for the nephelometer are systematically used and well accepted by the users, which is different to for example AE31 algorithms. *Correction algorithms and factors that minimize the error sources and uncertainties of nephelometer measurements are systemically used (Anderson and Ogren, 1998; Müller et al., 2011b).*

P2L34-P3L5: The discussion on Cref is focused on the different types of environments but does not address the fact that those studies cited weren't conducted in the same way. Some reference instruments were different than others which is likely to be a factor when comparing Cref values between studies E.g. a study using a MAAP as a reference instrument would yield different results compared to a study using a photoacoustic instrument as a reference measuring the same aerosols.

We added a note: *Since there is no generally accepted method for deriving the C_{ref} values, the methods between different studies vary, which can also affect the results. In this study, we derived the C_{ref} by comparing the AE31 measurements against another optical filter-based instrument.*

We also paid attention in describing the method used in the different C_{ref} values that are referred in the article.

P3L8: correct "cast a so-called shadowing effect". Something casts a shadow but not a shadowing effect.
Was rephrased: *When the filter is loaded with absorbing particles, the particle loading decreases the response of the instrument.*

P3L13-14: Here you could cite Collaud-Coen &al 2010 and Backman & al 2014 as those are relevant for what is claimed in the sentence.
References were added.

P3L30: "remarkable" does not seem to be the correct word here
Changed to "significant".

P4L6-9: Why mention CAPS if it is not used?
We removed the mention of CAPS from here but left a mention later in the text just to give an explanation on why it was included in the measurement line (Fig. 2).

P4L10: Wouldn't the period be from Jun 2013 – March 2016 when all instruments are running? Why is then the period Jan 2012 – Dec 2017 chosen with the arguments of concurrent measurements?
We now did some modifications in the text (*This period was selected to have at least two absorption instruments running in parallel.*) and also in the data (affected Figs. 4 and 5). In order to prevent any differences caused by the different time periods in the instrument comparison, in Sect. 4.2, only data from Jun 2013 – Feb 2016 was used. Also, in comparing the absorption Ångström exponents, only parallel data were used.

P5L2-3: What is the Nephelometer actually measuring? The switch between PM1 and PM10 is done every 10 minutes and the flow through the comparatively large sensing chamber is 4.3 lpm. How fast is the Nephelomter flushed after a change in the inletcut-size? It is not in seconds, but rather minutes as it does not flush evenly
There was a flushing time in the nephelometer measurements, for the absorption data the three first minutes were omitted from the data analysis. This is now described in the text: *To hinder the effect of changing inlets, the first minutes of measurements after the inlet switch were omitted. For the absorption instruments the first three minutes were omitted and for the integrating nephelometer the first five minutes were omitted.*

P5L17-20 deltaT needs to be defined as the measurement interval.
Was defined.

P6L7-8 It sounds like Weingartner is the cause of the "shadowing effect" when it is the filter and the particles that are the cause. Please rephrase.

Rephrased: *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).*

P6L17-19: There isn't a correction algorithm for MAAPs but that does not mean that they don't need one. See e.g. Müller et al 2011 for e.g. the cross sensitivity to purely scattering aerosols as a function of filter loading.

Added a note: *However, even though the MAAP was used as the reference here, it must be remembered that like all the filter-based photometers, also MAAP suffers from the cross sensitivity to purely scattering aerosol and therefore it does not the best reference instrument (Müller et al., 2011a).*

P6L26: A radiative transfer scheme is no motivation for using the instrument as a reference instrument. The uncertainty and unit to unit variability (in that order) are arguments why it could be used as a "reference" although it does not provide the absolute truth either, as it is also filter based.

Changed to: *Since the uncertainty and unit-to-unit variability of the MAAP was a lot smaller than for the PSAP and Aethalometer we used the MAAP as the reference instrument for measuring σ_{abs}.*

P7L6-7 Please be more specific than 'wavelength range is not as good'

Changed to *not as wide.*

P7L8 Problem for who? It can also be an advantage since it does not leak through the side of the filter tape.

This was rephrased: *The PSAP filters have to be changed manually by the user so the instrument is not the best option to deploy at a remote site, but then again the leakage through the filter tape is lesser than for the MAAP and AE31.*

P7L19-20 R can depend on other things too, not just ATN. R can be a function of single-scattering albedo, particle size, back scatter fraction etc. etc. This is the crux of the problem. Could be worth mentioning those things too.

Added: *The R can also depend on other factors, such as the ω, and some of the algorithms take also other parameters than ATN into account.*

P7L27-28 What were the criteria which lead you to choose these algorithms and not the others listed earlier? E.g. Schmid et al or Arnott et al are listed earlier but omitted here. Maybe they perform better and therefore warrants more investigation as more promising. You might want to state that if those were your criteria.

The selection of the algorithms was rather practical, those were the ones we had already had experience on and we had the scripts ready for most of them. We aimed to use the same algorithms as in CC2010, however, the algorithm by Schmid et al. (2006) was left out by oversight. The newer algorithms were also left out because we did not know about their existence back in the days when we started doing the analysis. Now, due to time limitations and in order to get the manuscript resubmitted, we did not have time to add these missing algorithms in the study.

P9L27-28: 14 days and filter changed on average once a day gives me 14 data points, not 9. The authors might want to rephrase a bit or write out the average filter change in days with a few decimals, like on average 1.55 days.

This was fixed.

P10L26 sigma_PSAP is not defined in the text.

The equation was modified (the σ_{PSAP} is not actually mentioned at all anymore).

P10L28 Shouldn't this equation be the Ogren 2010 ajusted equation as written out by Virkkula 2010 so that it reads sigma_ATN/(1.5557*Tr+1.0227). Or which equation did you use? The old Bond 1999 or the Ogren ajusted?

This was now modified to this formula.

P11L5: Rephrase "we agreed the results"

Rephrased: *Here, the iteration was stopped one the change was less than 1%.*

P11L17: Which data did you use? PM1 or PM10? The uncertainty is greater for PM10 than for PM1 since the more signal is truncated when bigger particles are present.

Indeed, Sherman et al., (2019) defined that the fractional uncertainty for the PM10 σ_{sca} was 9.2 % and for the PM1 σ_{sca} 8 %. We modified the sentence to: *The fractional uncertainty of the integrating nephelometer for PM10 has been reported to be ± 9 % (Sherman et al., 2015).*

P12L12 It is not a model but rather an equation that is used to make the source apportionment.

Rephrased to: *The α_{abs} is typically used in a set of empirical equations, that approximate the source of black carbon (BC) (Sandradewi et al., 2008; Zotter et al., 2017).*

P12L13 Used for what? Just say that it is important measure of the aerosols ability to interact with light.

Rephrased: *The single-scattering albedo (ω), which describes how big fraction of the total light extinction (σ_{abs} + σ_{sca}) is due to scattering…*

P12L25 less sensitive? The range for b is smaller than a_sca but how would it be less sensitive? I think you mean that the range is smaller. I suggest you remove this sentence as it is not relevant for the analysis in the manuscript.

Removed.

P13L12 why focus on WEI and COL when the biggest difference was to VIR?

The paragraph was modified to make it less focused on WEI and COL: *The smallest determined C_{ref} value was C_{NC}, which was expected. Since the σ_{ATN} decreases for a loaded filter and the filter loading correction was not applied, the C_{NC} has to be smaller than for the corrected data. Since the values of the C_{WEI} and C_{COL} were almost the same, the result suggested that on average, the loading corrections R_{WEI} and R_{COL} had on average a similar effect on the data. The highest value was determined for the C_{VIR}, which suggests that on average, the value of the R_{VIR} was the lowest (i.e., the effect of filter loading correction in V2007 was stronger).*

P13L13 Similar effects? What effects I wonder? Do you mean average or mean concentrations? Being more precise would be more informative here.

Added *on average* to the sentence.

P15L1-3 Would it be possible that the different Cref values in the mentioned studies is due to the reference instrument being something else than a MAAP?

Added: *In these studies, however, the reference instruments were not filter-based photometers like in our study and that can have a remarkable effect on the results.*

P15L7 'describes' is not the correct word here

Modified to: *This is also closer to our observations, which is explained by the fact that at SMEAR II, the observed soot particles are likely aged and coated since there are no significant local emission sources.*

P15L16 Which algorithms would be good if the reader is encouraged to use different algorithms based on their performance? At least you could state that e.g. the property derived from the AE31 should not depend on ATN after post processing. E.g. Fig. 8 shows clearly that some correction algorithms perform better than others when it comes to a_abs.

We added the following recommendation in the conclusion section: *According to our study the correction algorithms by Virkkula et al. (2007) and Arnott et al., (2005) performed the best in taking the seasonal variations of the aerosol particles into account. Also, the algorithm by Virkkula et al. (2007) produced the most stable* α_{abs} *that did not depend on the ATN, which was not the case for the other algorithms.*

P15L20-31 Can the authors say something about which studies to trust and which ones not to trust?

Not really. We added some discussion in the manuscirpt: *Because the results between the different studies vary, it is difficult to conclude whether the* C_{ref} *is wavelength dependent or not. To study the wavelength dependency of* C_{ref}*, it would be ideal to use a photoacoustic (like in Kim et al., 2019) or* $\sigma_{ext} - \sigma_{sca}$ *-methods as the reference measurements, since they are independent from the filter artefacts.*

P15L33 A linear fit does not average. Please rephrase.

Rephrased: *The different* C_{ref} *values were not only determined as a linear fit that took into account the whole time series.*

P16L26 The sentence could need rephrasing. The fact that the Cref value changes is a strong indication that it is not a constant.

Rephrased: *As indicated by the seasonal variation, the* C_{ref} *is not a constant value, but it depends on the optical properties of the particles embedded in the filter.*

P17L3 'relatively more weight' could use rephrasing. How about saying that the optical size changes? This implies that smaller particles (Rayleigh reigime) aren't necessarily adding to the behaviour.

Rephrased: *For example, the size dependent b and* α_{sca} *reach their maxima in summer and minima in winter, which indicates that in summer the fraction of smaller particles increases.*

P17L19-27 I understand that it can be hard to quantify the effect of RH on b_abs in this dataset but that is a very intresting topic. Based on your findings it appears that RH is more important than the aerosols single scattering albedo. Where was the RH measured? In the nephelometer? It is now shown in the schematics figure. How do you know that the observed RH dependence isn't from RH fluctuations in the MAAP sampling line which is not actively dried which then affects the C_NC values as if that was something to do with the Aethalometer performance.

The *RH* was determined for the MAAP by using the ambient *RH* and *T* measurements. We now have also studied how the rate of change of the *RH* affected the C_{ref}, which showed no correlation whatsoever. Like the referee already mentioned, in the scope of this manuscript it is difficult to define why the variation in the *RH* would effect the derived C_{ref}.

P19L21-27 The numbers mentioned in the text does not seem to match the figure. Please check if this is true or not. E.g. the lowest median a_abs value in the figure does not seem to be 0.85 but rater close to 1 and the highest seems to be above 1.5 when in the text it is 1.48.

Fixed this.

P20L20 Figure 8 is only discussed here and is an excellent figure which I feel could be discussed a bit more. For example, the authors could for example use the figure as an illustration to state that whether it is b_abs or a_abs, the values should not depend on ATN and is an excellent test to check if the algorithm works. An important point to raise here would be that a_VIR seems to be the algorithms that performs the best.

Added a mention in the conclusions: *The correction algorithm by Virkkula et al., (2007) was the only AE31 correction algorithm, which produced a stable* α_{abs} *for the increasing filter loading*. … and … *According to our results, applying the Virkkula et al. (2007) correction algorithm could help solving if the changes in* α_{abs} *were due to real variation or due to increased filter loading.*

P21L32-33 Please, rephrase the sentence

Rephrased: *The sizes of the particles affects their scattering properties and also on their penetration depth in the filter that again could affect the k.*

P22L28 You might want to mention that there are three different makers of tape for the AE33 and all of those have different Cref values.

This paragraph was moved to results and we added: *The filter material in AE33 is Teflon-coated glass filter tape (Pallflex type T60A20), but also the "old" filter tape (Q250F) has been used with AE33 and the recommended* C_{ref} *values to use with these filters are 1.57 and 2.14, respectively (Drinovec et al., 2015).*

P23L24 Effect of increasing filter attenuation is sometimes called shadowing effect and sometimes filter loading effect. The authors should be consequent in what they call the effects.

We used now the "loading effect" term.

Table 1 Remove 'Also, ' and 'are presented in the table' To me it does not seem to be enough to have the coefficeints in the table with only two digits. Three decimals would seem appropriate if possible.

The text was modified and the third digits were added in the table for those parameters it was possible (*a* and *f*).

Figure 2 I am curious what the setup was like during the other years as this figure only illustrates the setup for 33% of the data.

The changes in the measurement line are explained in Sect. 2.2 Instrument set-up: *Also, during this period there were only few changes in the measurement line: in March 2017 the MAAP flow was decreased from 18 lpm to 9 lpm and Nafion dryers were installed in front of MAAP; and in November 2017 one of the two Nafion dryers were removed in front of the Nephelometer.*

Otherwise, the set-up changed with the instruments (the MAAP was installed and the PSAP removed).

Figure 3 The last sentence in the caption could say that the dashed line is the median for all the data as there are already other medians shown in the figure.

Fixed this.

Figure 4 Adding Root Mean Square Errors could be a more quantitative way of expressing how well the instruments agree in addition to the correlation coefficient

We now replaced the correlation coefficient with the coefficient of determination.

Figure 7 Please explain what the whiskers are in this figure as well. In the text it says that the statistics are as in Fig 5 but there is no boxplot there and in e.g. Fig 3 the mean is shown as an o whereas in this figure using an x.

Fixed this.

Figure 8 and 9 Same thing here, explain the statistics of the boxplot. The box plots could use some text about what the whiskers and boxes represent. For some figures it is a matter of a simple copy/paste. Fig. 5 A better matrix for the performance of the vaious algorithms would be to include the Root Mean Square Error of the fits which would actually yield a quantitative value of the goodness of the fit in Mm-1. R2 in all respect, but RMSE could be a good addition to the analysis.

Fixed this. And the correlation coefficient was replaced by coefficient of determination.

Language suggestions:

P1L16 resulted to –> resulted in

P1L20 filter measurements –> filter-based measurements

P2L8 climate in global –> climate on a global

P2L10: of the particles –> of aerosol particles; scatter the light –> scatter light

P2L11: "in color" is tautology so remove it

P2L12: suggest changing "light colored" to "bright"

P2L32: depends also –> also depends

P3L10: remove "and determined coefficients"

P4L17: Remove the in 'measured the b_sca…'

P4L27: above accepted –> above the accepted

P5L15 Bouguer is needlessly underscored.

P6L1 In the filter –> In a filter

P6L28 'absorption instrument' sounds rather sloppy. Please use absorption photometer or something similar: : :

P7L1 remove 'again'

P7L2 'functional and popular' says who? Why not say widely used?

P8L11-16: Correct the grammar: e.g. remove articles before f, a and omega where not needed.

P8L16 resulted –> resulting

P10L8 Arnott &al 2005, not 2003.

P11L19 averaged for –> averaged to

P11L25-26 concentration of the particles ! –>concentration of particles, amount of the –> amount of

P12L29 amount –> amounts

P13L9 corrected by –> corrected using

P13L18 within 1% limit –> within a 1% limit

P15L15 real b_abs –> true b_abs

P17L13 correct: '…the relatively more the…'

P19L32 measurements on –> measurements at

P20L23 growed –> grew or better still increased

P20L23 As a –> In

P21L24 I suggest changing correlation to behaviour

P22L7 remove 'about the'

P22L24 is the an model –> is an old model

P22L30 we observed also –> we also observed

Figure 6 Colored be –> colored by

Table 1 (which should read Table 2) These values are reported at the Aethalometer... –> These values are reported at the MAAP...

Table 4 (should be Table 3) Remove 'the' from before k (a_k)

These were all modified in the text.