COMMENTS TO THE AUTHOR:
amt-2020-517

GENERAL COMMENTS

This manuscript provides an evaluation of three noise-reduction algorithms for the "raw" equivalent black carbon (eBC) mass concentration data of the new 5-wavelength microAethelometer model MA200. This has been submitted for the special issue: New developments in atmospheric limb measurements: instruments, methods, and science applications (AMT/ACP inter-journal SI).

Below are general comments on the manuscript as a whole:

Appropriateness:

1. *To the scope of the special issue (SI):* If this is not an error, it must be aptly justified why this is submitted in this SI. The SI call is specifically for "new developments in atmospheric limb measurements" focusing on the stratosphere. This manuscript has very little relevance to the scope of the SI and is an outlier among the other preprints included in this SI. I understand that the MA200 has been and may be used in vertical profiling. However, the dataset used here were from ground-based mobile measurements in an urban area. In my opinion, this manuscript does not belong in this SI.
2. *To the scope of AMT:* On the other hand, the manuscript does fall within the scope of AMT, in general. However, it lacks the detailed discussion on the technical aspects that is common with AMT publications, particularly if the presented "decision tree" is something the authors would like others to employ. **The entirety of the comments on this manuscript are based on this manuscript fitting the scope of AMT and not of the SI to which it was submitted to.**

Scientific Significance and Quality:

The idea behind this investigation is understandably important for some users of the MA200, particularly those who use the read out directly. However, the following issues were insufficiently addressed in the manuscript:

1. The motivation for noise-reduction, in general, was not sufficient. There is a part of the community who prefer the data as it is (since the instrumental noise cancels out when averaged), and focus instead on making sure the measurements are set-up correctly to prevent artificial peaks in the data (Cai et al., 2013; Alas et al., 2019).
2. It was mentioned in the text that the MA200 has an "on-board signal-processing that reduces the noise" of the MA200 raw data. Note that upon going through the user manual and quickly searching the MA200 website, this internal post-processing of the raw data is not mentioned (please, correct me if I'm wrong). This is an extremely important point that users need to know prior to using this instrument. This would mean that the output of the instrument is not "raw" anymore and, in the context of this study, the data has been "smoothed" out twice with the treatment of the noise-reduction algorithms being evaluated.
3. The measurement strategy was not explained in detail. Particularly, what measures were done to take into consideration the sensitivities of the MA200?
4. The results were merely presented. Deeper discussion on why the algorithms performed as they did is needed.
5. The broader significance of this study and how it relates to existing literature were not discussed.

COMMENTS TO THE AUTHOR:
amt-2020-517

Presentation Quality:
There is much to be improved with the writing of this manuscript. The main points are often times hidden in a mix of redundancy, jargon, lack of proper sentence/paragraph structures, and poor grammar. This provides so much hurdles for the readers in understanding the thought process of the author(s). Important aspects such as the criteria for evaluating the noise-reduction methods were often vague and leaves so much to interpretation or misinterpretation of the reader. Only my personal experience with the microAethalometer and mobile measurements allowed me to extract the information the author wants to give, and even then, it was with such difficulty.

The figures and tables, as well as their corresponding captions were not informative enough for the reader to understand them even after reading the manuscript (much less without). The figures and tables must be intuitive. Use of informative legends would improve the figures significantly. In addition, the parameters used to evaluate these noise-reduction algorithms were not defined prior to presentation of results. For instance, "noise reduction effect" and "negative decline rate" were not defined prior to showing up in Table 2 and how they were calculated was also missing.

Some technical comments:

- Define terms before abbreviating/ using initials (BC).
- Please consider using eBC (Petzold et al., 2013) consistently. It was anyway introduced in the text.
- Use initialisms throughout the manuscript.
- Be consistent with terminologies used.
- Please use "MA200" throughout the whole manuscript instead of switching from MA200 to "sample monitor" or "sampling equipment" every now and then.
- Use complete sentences in figure captions.
- The texts (titles, labels, legends) in the figures are not all the same size.

I strongly suggest major revisions in the writing with a native English speaker contributing on and reviewing the manuscript prior to re-submission.

SPECIFIC COMMENTS	
Line #	Comments *texts in red are suggested changes in the manuscript to improve readability.
MAIN MANUSCRIPT	
ABSTRACT	
32	"Our results showed CMA to be a good prospect…"
33	This line is a little unclear. The readability may be improved. Here's a suggestion for this sentence: "Based on the interval times used here, our results showed CMA to be a suitable algorithm to reduce the noise of raw BC mass concentration data based on the decrease of negative values and the retention of details attributable to microenvironmental changes."
34	Did you mean here "highest reduction OF peak values"?
35-36	"Furthermore, after background correction, the CMA results retained more detailed microenvironmental changes in pollutants than other methods."
38-39	"These findings provide new insights on the suitable noise reduction approach for mobile monitoring data obtained from portable BC instruments."
INTRODUCTION	
General Comments	• The jump in topics from line 48 to 49 is a bit big. I suggest to introduce at first the relevance of BC particles in air quality through its health effects.

	Then you continue with line 49. This will improve the motivation of your study. • Line 52: is it really the goal to propose a monitoring method or a method to analyze data from mobile monitoring? • It would not hurt to already introduce here the ONA method by Hagler et al. and to motivate why it is necessary to explore other means of reducing noise from the eBC datasets from portable absorption photometers. • There is insufficient motivation on the purpose/advantages of noise-reduction. • Explicitly mention that the noise-reduction algorithms evaluated here are those that are readily available in the Aethlabs Dashboard. There are members of the community who log AE51 or MA200 data independently and do not use the Dashboard. In any case, it must be justified, why the LPR and CMA are options in the Dashboard as noise-reduction algorithms.
43	"Black carbon particles with size ranging from …"
67-68	"…and simply removing negative values may introduce biases in the dataset."
	In this paragraph, it would be beneficial to inform the reader that these negative values are part of the instrumental noise, before you introduce noise reduction.
70-72	"Moreover, high-time resolution measurements of air quality at roadside are susceptible to single events (e.g. occasional passing of heavy-duty diesel vehicles or cigarette smoke) that may not be representative of the street in study. This may result in over estimation of eBC levels when averaged over time/space as they introduce peaks in the dataset."
72-74	"In addition, when the sampling equipment traverses from highly-polluted area to a low-polluted one, such as a park, the instrument produces strong negative peaks that is due to the measurement principle of the instrument and the strength of the pollution gradient between microenvironments."
75	"Therefore, the noise reduction method should also be evaluated based on the retention of actual peak concentrations and number of peak samples that are related to identifiable sources of pollution."
78	Remove "including"
84	The rationale of doing background correction in relation to the performance of noise reduction methods is not clear to me. Could you further elaborate on the statement that background concentration could affect the noise of the instrument?
87-	This study evaluated three methods for reducing the noise from the raw BC dataset obtained using the MA200 in mobile measurements along a trafficked street in an urban area. The methods investigated are ONA (please cite here Hagler et al.,), LPR, and CMA.
89-92	"From these methods, the best noise reduction approach was selected by analyzing the post-processed results based on the following criteria: (1) relative number of negative values left; (2) retention of detailed information attributed to microenvironmental changes; (3) reduction of artificial peak values (is this correct?); and (4) retention of detailed information on microenvironmental changes after background correction (is this correct?)."
METHODS	
General comments:	• Some studies cited in this section used the older model AE51. Either clarify in one sentence or two that the studies you are referring here pertain to those that used a variety of portable absorption photometers for various applications, or remove citations which used the AE51 and not the MA200.

	• What methods were taken to account for the sensitivities of the instrument? The AE51 is known to be highly sensitive to vibrations, and sudden changes in the environment like RH and temperature. It is highly likely that the MA200 would have similar sensitivities, as also shown by Düsing et al. (2019) when it comes to strong RH variability. There was no detailed description on how the instrument was handled during the mobile measurements. This is vital since one of the criteria of the study is the retention of signals due to microenvironmental changes after noise reduction and identification of "peak samples". It must be noted that the instrument may produce false peaks/signals (in either direction) as a result of vibration or sudden change in RH and T. This is even more significant given that very few or even data from only a single mobile measurement round was used in the analysis. Taking these sensitivities into account in the measurement itself, would strengthen arguments made on the retention of "peak signals/samples". • You have to mention at some point that you only used data from one wavelength of the MA200. • Are the data already compensated? Was the internal DualSpot loading compensation operational during the measurements? • In which part of the analysis did you use Measurements 1-3?
100	"In mobile monitoring, the MA200 can be used to estimate personal exposure and quantify eBC mass concentrations in different microenvironments."
108-109	"In order to reduce the noise of the data obtained with high time resolution, smoothing algorithms can be used."
112	"…this study analyzed BC data collected from…"
114	• Why is it necessary to do further noise-reduction when there is already an on-board signal-processing? What is the principle behind the on-board signal processing and how does it differ from the methods investigated in this study?
115	Remove "microAeth®" Please provide some summary statistics of this comparisons.
116-117	Please clarify. Was the intercomparison between the MA200 and AE33 done during the walks (within one walk, the Ma200 stopped in the vicinity of the AE33 for short intercomparison)? How long were the intercomparisons? I understand that these results were presented in a previous publication, but summary statistics would aid readers.
119	Delete this: "To give intercomparison between the instruments…" You may start this sentence immediately at "To demonstrate the unit-to-unit comparability between the MA200 units, we performed intercomparisons at fixed monitoring stations and during collocated mobile measurements."
138	What does "To control for relative patterns in environmental exposure" mean?
139	"…the mobile measurements were carried out on the right side of the road simulating people's common habits…""
142	I suggest either removing "air" (as it is vague) or "exposure" in this sentence.
144	Remove "with" after 4 h.
149	COV and TPRS are not yet defined prior to this.
158-165	This section could improve to briefly describe HOW the ONA reduce noise in microaethalometer data. How are the three parameters used to do this noise reduction? I believe this could greatly help readers in understanding Fig. S2 and, of course, the following analyses. LPR and CMA were aptly described in the following subsections, it would be great to elaborate a bit on ONA, too.

169-173	Please briefly describe "smoothing number", as the determination of this "smoothing number" is similar for that of CMA. How did you arrive at the values 15, 7, and 3?
182	"… the number of remaining negative values was determined." Also, what "number" of remaining negative values would imply a "good" noise reduction method? Is it simply a comparison of the treated data and whichever has the least number of negative values gets the point?
182-189	This paragraph could be greatly improved. I find the structuring of the sentence hard to understand, as well as the looping the same idea. Please simplify this and improve the writing for better readability. I suggest starting again here with the "criteria" you have for selecting the best noise-reduction approach. I understand you already enumerated them in line 145, but it was within the brief list of the process of the investigation.
207	It is unclear to me how the background estimation and correction is related to the investigation of the noise-reduction approaches.
217-220	This paragraph is better suited after the description of the noise reduction approaches. As for the 3rd criteria, it would help to specify what would make a noise reduction method "good". Is it it's ability to remove or retain these peaks? The criteria in judging which method is "good" should be crystal clear.
RESULTS AND DISCUSSION	
General comments	3.1 → please improve the structure of the sentences 3.2 → please explicitly distinguish between "peak samples" and "peak values"; and then in line 190 you also have "peak-value sample". These are all quite confusing. How are the "proportions retained" calculated? For instance, in the 5-s data, 42.1% of the raw data were negative values. After post-processing, "negative values retained 33.3% for LPR and 26.1% for CMA". Are the 33.3% and 26.1% from the total amount of negative values or from the whole dataset? Please include in your methods how these numbers are calculated. 3.4 → why is background correction not applied to the Munich dataset? As I understand, one of the criteria for choosing CMA was its robustness to background correction.
230	Please elaborate on the explanation. I find it quite insufficient, particularly, in the ONA paper of Hagler et al., 2011, they published results of applying ONA on 1-s data of SootGen, stove, and mobile monitoring. Van den Bossche et al. (2015) also used ONA on 1-s data from AE51 in field measurements. Is this an instrument issue? Or an algorithm issue?
237	I do not understand the last part of this sentence. I think, I know what you are trying to say, but it's not coming across to the reader clearly.
240	Please be cautious of using the term "significant" here, particularly, that the analyses are based on comparability of statistical analyses of the raw data. I suggest the term "strong" here in place of "significant".
242	This is not a complete sentence.
243	Change "mitigating" to "decreasing".

240-245	A deeper discussion on the differences of the 3 noise-reduction approaches could greatly improve this part. In essence, this part was merely a presentation of results which are already in Table 2.
Fig. 2	The unit should be nanograms. Am I right to assume that Fig. 2 is just same as Fig. 1 but only with the 10 s time resolution? If so, I do not see any added value in having this figure separated. The point you made in lines 240-245 is already clear in Fig. 1.
255 Table 2	It is unclear for me how the "noise reduction effect" was calculated. Please include in the methods section how these numbers are calculated and defined, including the "negative decline rate".
257	In this section, is my understanding correct? You want to evaluate two things about the "peaks": 1. # of peaks left after noise-reduction 2. Magnitude of these peaks after noise-reduction Is this right?
264	How is the "reduction effect" calculated?
263-269	It was not apparent right away that these results are already in Table 2. This could be solved by adding more information in the Table caption. Again, please give more information as to how these numbers are calculated or defined. Also, include the mean values in the table and not just the range so the readers can connect the numbers in this paragraph to the table. Do these numbers mean that CMA reduces the magnitude of the peak values greater than the other two noise reduction approaches? If so, what is the main criteria here? Do you want a noise-reduction algorithm that retains the magnitude of these peaks? Do you have a threshold where you say the algorithm diminished the peaks "too much"? A bar graph comparing raw and processed data for all your parameters would help clarify these compared to Table 2 alone.
273-274	This sentence is not clear. Did you mean to say, that based only on the # of remaining "peak samples", CMA performed better than the other approaches?
288-289	I do not understand how CMA, which "greatly reduces" the peaks (magnitude and number) is helpful in identifying "hotspots", in a sense. For instance, if this peak that is related to a source happens a few moments before or after a lower (below the COV threshold) peak, and it is greatly reduced by the CMA method, wouldn't that further blur the impact of this single source? I believe, a better criterion is a noise-reduction method that **does not greatly** reduce the magnitude of these peaks, particularly for exposure studies where every real signal is important.
Fig. 3	The statement that these "spatial peaks" (Fig. 3a) are due to traffic and street canyon configuration could be better justified with a map that has spatially averaged eBC mass concentrations along the route. This also would prove the quality of the collocated measurements of the three MA200 and assure the reader that the peaks are due to local sources and not an instrumental artifact. I mean, you already have the data (running with 3 MA200 at the same time). Please provide more information in the figure caption such as the measurement number, to inform the readers that this is data from one run only.

	Please also improve Fig. 3a by adding time stamps in the map to help readers reconcile the spatial plot with the time series.
295	This sentence can be simplified for better readability.
306	What is “minus absolute value”?
Fig. 4	It is unclear if the figure 4 a and b are background concentration or background-corrected data. Please specify in the figure caption. What is “actual detection concentration”? What are the those encircled in dash black lines mean? Are they values below 1ug/m3? If so, it would help to draw a zero-line, or magnify the scale such that the data around 0 ug/m3 would be more visible. Improve figure caption.
318	Change “certify” to “verify”.
CONCLUSIONS	
General comments	• The broader significance of this study should be explicitly mentioned here.
353-355	The first sentence is misleading. As I understand, it was not the goal of this study to “assess BC pollution”, but to determine a suitable noise reduction algorithm for the new MA200.
369	“The data is available upon request by contacting the first author of the paper.”
375	“The authors declare no conflict of interests.”
SUPPORTING INFORMATION	
Table S1	Are these numbers mean or median of the 5040 data points? Either way, please indicate and provide range, either quantiles, minimum and maximum, or standard deviation. How long were the measurements?
Table S2	Another new terminology: “peak values number” ➢ Why is there no information for measurement numbers 5 and 7, 8-10?
Fig. S1	Did you use standard major axis regression here to account for the error on both axes? Technically, none of these instruments are “reference” instruments to merit the use of simple linear regression.
Fig. S2	Improve figure caption, indicate that this is for ONA.
Fig. S3	Indicate that this is from CMA treated data.
Fig. S4	So, the measurements in Munich were not simultaneous like in Augsburg? The figure labels are too big. Why is there no analysis of the “peak values” and “peak samples” for the Munich dataset? As I understand, you were testing the applicability of the CMA method to a different dataset, but fail to run the entire series of tests which “proved” CMA to be the suitable method.

References:

Alas, H.D.C., Weinhold, K., Costabile, F., Di Ianni, A., Müller, T., Pfeifer, S., Di Liberto, L., Turner, J.R. and Wiedensohler, A. (2019). Methodology for high quality mobile measurement with focus on black carbon and particle mass concentrations. *Atmospheric Measurement Techniques* 12: 15.

Cai, J., Yan, B., Kinney, P.L., Perzanowski, M.S., Jung, K.H., Li, T., Xiu, G., Zhang, D., Olivo, C., Ross, J., Miller, R.L. and Chillrud, S.N. (2013). Optimization approaches to ameliorate humidity and vibration related issues using the microaeth black carbon monitor for personal exposure measurement. *Aerosol Sci Tech* 47: 1196-1204.

Düsing, S., Wehner, B., Müller, T., Stöcker, A. and Wiedensohler, A. (2019). The effect of rapid relative humidity changes on fast filter-based aerosol-particle light-absorption measurements: Uncertainties and correction schemes. *Atmospheric Measurement Techniques* 12: 5879-5895.

Petzold, A., Ogren, J.A., Fiebig, M., Laj, P., Li, S.M., Baltensperger, U., Holzer-Popp, T., Kinne, S., Pappalardo, G., Sugimoto, N., Wehrli, C., Wiedensohler, A. and Zhang, X.Y. (2013). Recommendations for reporting "black carbon" measurements. *Atmos Chem Phys* 13: 8365-8379.

Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J. and De Baets, B. (2015). Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. *Atmos Environ* 105: 148-161.