Referee #1

We thank the Referee #1 for the valuable comments on the manuscript. Please find the point-to-point answers to the comments below.

- **Careful checking of the English language and the text as a whole is necessary.**

  The manuscript has been proofread in order to check for any mistakes in grammar and spelling.

- **There are also a few misstatements which I would advise correcting.**

  Removed statement that "Black Carbon is the second highest warming agent in the atmosphere" as the newest IPCC report does not support this statement.

- **The author should change the title. It does not only examine the spatial variability of ambient black carbon.**

  The title has been changed to:

  The applicability and challenges of black carbon sensors in dense monitoring networks

- **In introduction: Please correct PM2.5 to $PM_{2.5}$.**

  Corrected.

- **Page 11, Paragraph 210: A sharp ATN change |ATN| >30 was manually identified. What could be the reason for this sharp change? Please explain it.**

  The original wording was indeed confusing and referred to the sharp change of ATN related to the filter change. As an example, if a filter has an ATN of 80 and it is changed, the new filter will have a ATN of 0. Therefore, the change in ATN is larger than 30 i.e $|\Delta ATN| > 30$

  The wording has been changed to: The filter changes were manually identified, and two hours of data were removed starting from the nearest hour before the filter change.

- **Page 11, 3.1 section: The authors used different flow rates for the same type of sensors. This is sometime hard to follow in the article. Please create a table summarizing this. How much during each campaign, etc.**

  Table 3 has been added at line 210 to summarize the flow rates used at specific date ranges with the specific sensors. A reference to the table has been added to the start of paragraph 190. Section 3.1 has been edited to prefer referring to Table 3 when necessary.

- **Page 12, Paragraph 255: The authors wrote that the 1st intercomparison has on average lower concentrations compared to the 2nd intercomparison. Is it not because of the different meterological condition? Please explain it.**

  Yes, the difference is most likely because different meteorological conditions between these periods. Also, the 1st intercomparison is at the start of the summer vacation season in Finland when traffic density can be expected to be less in the urban environment. In comparison the 2nd intercomparison is more everyday life when traffic especially on the Kustaa Vaasa road can be expected to be quite heavy.

  Additions have been made to the statement as follows:
  The 1st intercomparison has on average lower concentrations compared to the 2nd intercomparison. This is due to the difference in meteorological conditions and in traffic density during these periods.

- **Page 13, Figure 2: Please check the labels (date) at the xaxis. Please correct it.**

  The figure features a split x-axis marked with the red vertical lines and the dashes on the axis itself. I.e it consists of 1st Intercomparison on the left, missing section of approximately 3.5 months in the middle and then the 2nd Intercomparison. The labels are therefore accurate in their representation. The figure caption has been updated to:

  Figure 2. *Timeseries of both intercomparison periods a) 26.5. – 6.6.2022 and b) 16.9. – 3.10.2022. In the figure there is a split x-axis, where the period in between panels (a) and (b) marked with the vertical red lines is approximately 3.5 months. This period was the deployment phase between the intercomparisons. Data points are 5-minute averages.*

- **Page22, Paragraph 400: The authors wrote the following: With the MA-series sensors (MA200, MA350) the change of the temperature and RH caused clearly erroneous data as seen in Fig. 12. However, we cannot see the results of MA200 sensor in the Figure 12. Please include its results in the figure.**

  Figure 12 has been renamed to Figure 13 and Figure 12 has been added to show the results of the MA200. A separate figure has been used to ensure the clarity of the figures.

- **The conclusion contains some statements that need to be clarified. For istance: what is DST? please explain it.**

  Added the abbreviation DST to line 136 at page 6 (Distributed Sensing Technologies). Unfortunately, Distributed Sensing Technologies ceased operations in the end of 2023. Therefore, the conclusion paragraph has been revised to remove any mentions of the company and instead refer to the Observair sensor. For the environmental compensation, reference to the original publication Caubel et al. (2018) is added and a clarifying statement as the original publication uses the name Aerosol Black Carbon Detector (ABCD) for the Observair sensor. The paragraph reads now:

  Temperature changes significantly affected the measurements and provided a challenge in the deployment of the sensor network. Development of robust enclosures or deployment in locations

that have stable or controlled temperature is needed. Alternatively, the environmental compensation used by the Observair sensors was seen to reduce the effect of temperature changes. Unfortunately, the Observair sensors are not being produced as of the end of 2023. Therefore, a suggestion is made that the environmental compensation utilized by the Observair and outlined in Caubel et al. (2018) could be applied as a measurement method to the data via post-processing or implemented to other sensors by manufacturers as a solution to the temperature artifacts. Please note that in the publication Caubel et al. (2018) the name Aerosol Black Carbon Detector (ABCD) is used, which is the academic prototype of the Observair sensor.

Referee #2

We thank the Referee #2 for the valuable comments on the manuscript. Please find the point-to-point answers to the comments below.

1. A language check is necessary for the entire manuscript. Some segments of the text do not follow common practices of reporting data or manuscript writing, such as consistently reporting units of mass concentration or time.

The manuscript has been checked in order to catch any deviancies from the guidelines.

2. Section titles require renaming (based on the updated context), particularly in Section 2 (Methods) and Section 3 (Results).

Following section titles have been updated. Section 2.2 “Dualspot correction algorithms”. 2.4 “Deployment area”. 3.1 “Intercomparison periods”. 3.1.1 “Applicability of the dualspot corrections”. 3.2.2 “Weekly features in BC concentration”

3. I highly recommend adding more text discussing the results and making section 3 from "Results" to "Results and Discussions," which is a typical nature of manuscript framing in similar types of scientific journals.

The title has been adjusted to the suggestion.

4. The title of the manuscript can be updated as the manuscript content does not presently match with the title.

The title has been changed to:

The applicability and challenges of black carbon sensors in dense monitoring networks.

5. Use $PM_{2.5}$ everywhere

Corrected.

6. Line 71: Why is MAAP chosen as a reference device to compare with Aethalometers, which work slightly differently? AE33 could be a great reference device in this work. If

the authors have AE33's data, adding a comparison might help understand the unit-specific offsets in measuring BC. If they don't, I highly recommend providing sufficient justification (here in Line 71) or in the methods section why MAAP is used or better suitable.

Unfortunately, no AE33 data was available during this campaign. MAAP was utilized as the reference instrument as it was the highest-grade instrument available during this period.

7. Line 85: Some theoretical mistakes were identified. Please check the literature and correct it. Aethelometers measure light intensity and calculate light attenuation (ATN). From ATN measurements, ATN coefficients (b_{ATN}) are derived, followed by the absorption coefficient (b_{abs}). Drinovec 2015 explained this well.

Error fixed. "The measured variable by the instrument is the attenuation coefficient $b_{atn}(\lambda)$ [m^{-1}] calculated from the measured attenuation and the operational parameters of the instrument as described in Eq. 2."

8. Line 105: There are inconsistencies around the assumptions made. If any scattering correction is made, why are they assumed to be unity? More explanations are required. Also refer to Line 152.

Scattering correction is the $s(\lambda)\sigma_{sp}(\lambda)$ part of the general correction scheme (Eq. 3). This refers to the scattering from the aerosols and is ignored in this study which was meant by saying "assumed unity". The multiple scattering correction (C_{ref}) refers to the enhanced attenuation due to the filter fibers. For this the manufacturer recommendations are used. This statement is rephrased to:

"The aerosol scattering correction requires measurement of the scattering coefficient, which in many cases is not possible due to the lack of instrumentation. Due to this the aerosol scattering correction is often voided as in this study"

Text has been adjusted that in reference to these to corrections, aerosol scattering correction is used for the $s(\lambda)\sigma_{sp}(\lambda)$ part and multiple scattering correction for the C_{ref}.

9. Line 118: Leakage factor changes over time. Please refer to Drinovec 2015. This assumption might change the corrections, and so might the final corrected BC concentration. Please reconsider adapting to such changes.

The line referring to the leakage factor has been removed from the text and equations. This was an error in the manuscript. The leakage factor was tested during data processing of the manuscript but later removed.

10. It is not clear if the MA series device's inbuilt Dual-spot corrections were used for comparison or not. Typically, MA devices have their own correction mechanism, which is not the same as Drinovec 2015. Please check and confirm with a table of corrected data or add it to Table 2.

The MA series inbuilt corrections are used. Meaning the dualspot corrected data that the instrument gives is used as is.

Confusion to this is related that the authors were not aware of the exact correction algorithm used by the MA-series instruments. It was assumed that Dualspot refers to the correction outlined in Drinovec et al. (2015). This is now considered and refered in the modified version of the manuscript.

11. Some comments on figures have been mentioned in the attached file, which mainly focuses on the visibility of the graphics and texts.

These comments have been taken into account and the plots updated.

12. Section 3.1.2: A separate segment on the sensor calibration in methods sections is recommended. Also, mention how the calibrations are assessed (metrics used, such as slope, MAE). Reporting the calibration results are helpful for future studies. While presenting the data in "Result" section, please compare it with previous literature. A table might be helpful for reporting metrics from uncalibrated data, calibrated data, and literature data - with the type of calibration procedure adopted. All these elements will expand this section, which I believe is going to improve the quality of the manuscript.

Calibration methodology is added to section 2.5. The results of the calibration are showed in section 3.1.2

13. Generally, the results and discussion section is missing references from previous literature. Please compare the reported BC levels with different devices under this work and the levels reported in previous studies in similar regions or of similar spatial characteristics.

Some references to previous studies in Helsinki have been added to section 3.2.1.

14. Section 3.2.4: Change the section title. I recommend removing this section and adding a detailed discussion explaining the variability in previous sections. If the authors want to keep this section, will this require some detailed analysis explaining how spatial variability is captured by different devices? If there are any differences in performances? Also, some discussion was required about how spatial variability can be captured by these devices. If the true spatial variability is higher/lower than the inter-device variability studied. Finally, some recommendations /comments would be helpful for the community, such as which device performed best in what context.

This section is removed and merged with section 3.2.1.

15. After restructuring the result and discussion section - please update the conclusion accordingly. Avoid including conclusions that have not been discussed well in previous sections (For example, Line 436).

The conclusions have been updated to correspond to the results and discussion provided in the manuscript.