GENERAL COMMENTS

The study by I. Ylivinkka et al. analyzes three long-term (2005–2020) ground-based particulate matter (PM) datasets collected at a rural boreal forest site, with the aim of comparing the respective measurement techniques and evaluating temporal trends. This work builds upon the earlier preprint by Kersiken et al. (2020), which was submitted to Atmospheric Measurement Techniques Discussions. Compared to that manuscript, the current one attempts to improve the analysis of seasonal dependencies, to extend the dataset, and to strengthen the overall conclusions. The manuscript is generally clear and well written, and the research objectives are clearly stated. However, certain aspects of the study could be improved, further clarified, or made more robust before the manuscript is suitable for publication.

We thank the referee for taking the time to evaluate our manuscript and pointing out improvements to make the manuscript suitable for publication. Below the referee comments are in black, and our responses shown in red color.

SPECIFIC COMMENTS

- Statistical significance vs. measurement uncertainty: While the identified trends may be statistically significant based on the calculations, it remains unclear whether the instrumental accuracy is sufficient to attribute these trends confidently to environmental changes rather than to potential instrumental drift. Hence, the measurement uncertainties (e.g., line 468) should be explicitly discussed to support the conclusion that the observed reduction in PM is attributable to a decline in emissions, rather than to instrumental artifacts.
  In this manuscript we compare three completely independent measurement techniques. Although some discrepancies were identified in the data, the very good overall agreement between the measurements implies that the observed trends are not caused by instrumental drift. To further understand where the discrepancies stem from, we revised the correlation plots (Fig. 3 and S1a) by coloring the markers with PM1 to PM10 ratio, which can indicate changes in the correlation due to change in the ratio. Mostly the scattered data points are independent of the ratio, except the PM10 impactor data, where the scattered data points have lower PM1 to PM10 ratio, which suggests that the PM10 concentrations were overestimated. Additionally, we made a new plot (Fig. S1b) of monthly PM2.5 to PM1 and PM10 to PM2.5 correlation to see if there is clear shift in the PM fractions. Mostly the measurements are aligned with the 1:1 line, except a few impactor data points, further implying of the overestimated PM10 values of the impactor. The corresponding analysis has been added in section 3.1. Moreover, the instruments are calibrated regularly to ensure stability of the measurements. This information has been added to section 2.

- Introduction and motivation: The Introduction emphasizes the health effects of particulate matter and references the Ambient Air Quality Directive. However, the measurements are conducted at a remote background site located more than 50 km from the nearest urban areas where the majority of the population resides. While this context may be clear to the authors, the connection between public health concerns and observations at such a remote site requires further clarification. In this context, the comparison of PM10 concentrations at the SMEAR II site with those in a highly urbanized and polluted environment such as Beijing (lines 378–386) appears inappropriate and potentially misleading. I recommend removing this paragraph.
  We thank the referee for the comment. We removed the comparison to polluted environment and added comparison to European background sites (l. 428-432).

- Humidity control and sampling methodology: A more detailed discussion is needed regarding the methods used to control or reduce humidity in the sampled air across the different measurement techniques. For instance, in Section 2.2, are the filters used for gravimetric analysis conditioned

after sampling? If so, under what temperature and humidity conditions? How might this conditioning influence the comparison with other techniques? Similarly, what humidity control mechanisms are implemented in the DMPS+APS system? Additionally, can the height of the inlet affect the comparability or representativeness of the measurements, especially in case of vertical gradients in aerosol properties?

After sampling, the impactor filters are brought to a clean laboratory room (regular room temperature and humidity, but they are not controlled) inside the sealed impactor. The filters are then dried in a laminar flow hood for at least two hours before weighing. APS inlet line is heated to 35 C, similarly as the SHARP inlet line. The sheath flow of the DMPS system is dried with a silica diffusion dryer. The RH of the sheath flow is kept under 40 %. Hence, in all cases the sample is dried, which reduces the possible bias related to the humidity. However, without specific test sampling, it is not possible to quantify the impact of humidity on our measurements or the correlation between different methods. We acknowledge that the information of heating / drying was missing for other measurements except SHARP, and we have now added the information in section 2.

The inlets are situated next to each other at the roof of the measurement cottage (Fig. 1b). All inlets are within the forest canopy and inside the boundary layer. Hence, there is not any additional sources or sinks of particles, which would create a vertical profile of particles affecting our measurement.

- Density assumptions (lines 196 and 276): The assumption of constant particle density may not be strictly necessary, as particle density can vary with size and season (and composition, if analyzed). A discussion of how incorporating size-dependent or seasonally varying density values might affect the results would be valuable.

  Referee is correct, the density of the particles likely varies seasonally, since also the composition of the particles varies seasonally (e.g. Heikkinen et al., 2020), and size-dependently (e.g. Kannosto et al., 2008). Estimating the authentic seasonal and size-dependent density variation at SMEAR II is, however, out of the scope of this manuscript since we do not have them properly determined previously, and hence the estimation would completely rely on assumptions.

  However, to estimate the impact of the constant density assumption, we calculated the PM masses using lower and upper limit (1.1 and 2.0 g cm-3; Kannosto et al., 2008). These results are presented in Table S1. The average difference ranges from 6 to 29 % compared to the original, and it is larger for smaller particles. The analysis is added in section 3.1 (l. 307-314) and described in methods section 2.4 (l.214-217).

  We further revised the equations 1-3, since we realized that the equations were written in somewhat misleading manner and were in conflict with the text. The APS diameters were first converted to mobility diameters (Eq. 1) and then, using Eq. 2, we integrated the data to get mass concentrations, as indicated by Eq. 3. The equations previously suggested that we would have calculated the mass for DMPS and APS data separately, which was not the case.

- Choice of diameter conversion approach (line 228 and Eq. 4): The decision to convert aerodynamic diameter to electrical mobility equivalent diameter (as done in Eq. 4) needs further justification. Why was this direction of conversion chosen, rather than the reverse? Given that the first two techniques report mass as a function of aerodynamic diameter, converting to mobility diameter may reduce the direct comparability between methods.

  As the referee points out, the other methods indeed measure aerodynamic diameter and hence converting DMPS mobility diameter to aerodynamic diameter would be a logical choice. However, converting the diameter this way would require accounting also slip correction factor (Seinfeld & Pandis, 2006). Slip correction factor is a function of particle size and therefore it does not have analytical solution. Moreover, trying to estimate slip correction factor introduces additional

variables and more assumptions. Hence, as the slip correction factor approaches 1 for larger particles but is large for small particles, it is more convenient and commonly used direction to convert aerodynamic diameter of APS to mobility diameter and not vice versa.

- Outlier filtering (lines 249–251): The use of the 6-MAD criterion to filter out data points should be further explained. How was this threshold determined? Are the excluded values considered to be erroneous measurements, or might they reflect real but localized events? Does excluding these points significantly improve the agreement between instruments?
  The outlier filtering targets to remove the clear outliers from the data, which are erroneous measurements. In some cases, also local contamination peaks may have been removed, but in general local contamination events are rare at SMEAR II. Due to the aim of removing only the outliers, the MAD limit was set to be high. The value of the threshold was determined with visual inspection. This is now implemented in the text in section 2.6 (l. 275-276).

- Role of size fraction dominance (line 294): The analysis could be further enhanced by exploring how the relative contribution of fine and coarse particles within PM10 influences the comparisons among instruments. Have the authors investigated whether the agreement between techniques depends on the dominance of one size fraction over the other?
  The revised Fig. 3 and S1 show PM1 to PM10 ratio as a color. Mostly the correlation is independent of the size fraction, but the scattered impactor PM10 data shows lower PM1 to PM10 ratio.

- Discrepancies in 2011–2015 (lines 331–333): The observed divergence between DMPS+APS and impactor-based trends during 2011–2015 deserves a more detailed explanation. Can the authors elaborate on potential causes, such as instrumental drift, calibration issues, or changes in aerosol density or composition, that may have contributed to these differences?
  Although the average concentrations for impactor data in 2011-2015 are somewhat higher than the values given by the DMPS+APS method, the yearly and seasonal variation of the methods exhibit similar features. This information has now been elaborated in the text. The instruments are calibrated regularly to avoid any drift in the data and with the outlier filtering we aimed to remove erroneous measurements. Detailed analysis of the possible changes in the density or composition of the particles would require a large amount of new analysis and is therefore unfortunately out of scope of this manuscript.

- Literature review: Some parts of the manuscript, particularly lines 349–362, read more like a review of previous findings rather than a direct contribution to the current analysis. I suggest either removing or substantially condensing these sections. Alternatively, if context is needed, such content could be relocated to the Introduction or incorporated into a newly structured Discussion section, where previous studies could be more directly integrated into the interpretation of the present findings.
  We modified the text to make the previous results to be more supportive for the current analysis.

TECHNICAL REMARKS

- The title could be more specific. The measurements were conducted in a rural boreal forest environment, which may not be representative of all of "southern Finland".
  We thank referee for this comment. The title is now changed to “Long-term PM trends at boreal forest site in southern Finland from three different measurement techniques”

- Abstract: The time period covered (2005–2020) should be explicitly stated at the beginning of the abstract (line 19) rather than at the end (line 28).
  The information of the time period is now moved to the first sentence of the abstract: “Three

independent particulate matter (PM) mass concentration measurements and their long-term (2005–2020) trends…".

- Lines 24–25: Pearson's correlation coefficient alone does not adequately characterize the quality of a comparison. Two datasets can be highly correlated and yet exhibit considerable differences in slope or systematic offsets. Therefore, additional metrics, such as the slope and intercept of the regression line, root mean square error (RMSE), or bias, should be included already in the abstract to provide a more comprehensive evaluation of the agreement between datasets.
  We agree with the referee that high correlation coefficient does not necessary mean the variables are well aligned on 1:1 line. However, to avoid long abstract and a list of numbers, we prefer to leave the abstract as is since the variables do not deviate largely from 1:1 line and thus do not give a misleading impression for a reader. The slopes and intercepts of these five regression lines are presented in corresponding figures (Fig. 3 and S1a).

- Lines 27 and 31–32: Since the manuscript includes a statistical analysis of trends, the abstract should mention their statistical significance.
  We added the further information and statistical significance of the trends in the abstract: "Statistically significant (Mann–Kendall test) declining annual trends were observed in DMPS+APS and impactor data in all size classes, ranging from -0.021 to -0.036 µg m^{-3} y^{-1}. While DMPS+APS method indicated statistically significant decline also in all seasons, the decline in impactor data was statistically significant only in spring and winter. SHARP data could not be used for trend estimation due to the change in inlet heating temperature, affecting the measured PM_{10} concentrations." (l. 28-34).

- Line 27: Alongside absolute variations, percentage (relative) variations should also be reported. This is especially important because mass concentration levels, and therefore their trends, can differ substantially across particle size classes. Expressing variations in both absolute and relative terms would enhance the interpretability of the findings.
  Annual trends are now presented also in percentages in Table 3.

- Line 43: The phrase "large uncertainty is related to aerosol particles" is unclear. Consider rephrasing.
  We changed the sentence to: "radiative forcing due to aerosol particles and especially due to aerosol-cloud-radiation interactions is uncertain".

- Line 63: The term "variant" does not sound appropriate in this context. Also, "long-range transported emissions" is an imprecise expression. Finally, aerosol lifetime in the atmosphere ("one week"?) is variable and depends on particle size, composition, and source region. Consider rephrasing the sentence entirely for clarity.
  We revised the sentence: "Aerosol particles have both natural and anthropogenic sources. Additionally, particles can be transported over hundreds or thousands of kilometers, since the lifetime of PM in the atmosphere is days to weeks, depending on the size, composition, and source region of the particles:"

- Lines 96–108: This paragraph contains general background information that may be redundant or too detailed for the Introduction. Consider omitting or significantly condensing it.
  We revised the paragraph and condensed it to focus on PM.

- Lines 110–118: There is some content overlap with lines 53–57. To avoid redundancy and maintain a concise Introduction, consider removing the repeated information.
  We revised the paragraphs and removed the overlapping information.

- Section 2.5: Are black carbon concentrations monitored at the site? If so, it would be valuable to explain why BC data were not included in the current analysis, especially since BC could help distinguish between different aerosol types and sources.
  We added figures of monoterpene and BC concentration to Fig. S4 in the supplementary information of the manuscript. The methods are described now in section 2.5 (l. 249-258).

- Lines 243–245: The description of the backtrajectory analysis could be improved. Please provide details such as the duration of the trajectories and the method used to define the three source sectors. This would allow readers to better understand the methodology without referring back to Räty et al. (2023).
  We thank the referee for the comment. We now added more information of the trajectories. (l. 261-267)

- Lines 270–271: The statement might suggest that comparison is sufficient for validation. Comparison does not, by itself, constitute validation. Consider rephrasing.
  We agree with the referee and removed the part referring to validation:
  "Here, we present the comparison between the different aerosol mass measurement techniques at SMEAR II." (l. 293-294)

- Figure 4 (caption): Please clarify the axis labeling by specifying that the tick marks refer to the x-axis.
  We revised the plots to make them easier to read and updated the figure captions.

- Lines 329–338: This paragraph discusses temporal trends and would be more appropriately placed in Section 3.3. Consider starting Section 3.2 at line 340, since the seasonal cycle is a more dominant feature.
  We moved the paragraph to section 3.3. Additionally, we changed the subtitle to "PM concentrations, seasonal variation, and emission events" to be more precise with the content of the section.

- Line 345–346: Are pollen or other biological particles expected to peak during summer or in other seasons (e.g. spring or autumn)?
  Pollen peaks in late spring – early summer and other biological particles peak from late spring to early autumn at SMEAR II (Manninen et al., 2014). We added this information to the sentence: "Furthermore, pollen and other biological particles contribute especially to coarse mode particle mass at SMEAR II from late spring to early autumn (Manninen et al., 2014)." (l. 394-396)

- Lines 371–373: The claim appears to be based on general patterns, but is it also supported by analyses of specific episodes or case studies?
  We revised the sentence:
  "Air mass source area analysis shows that winters with higher fraction of easterly air masses (Fig. S5) were colder and had also higher PM levels, although we acknowledge that this analysis does not reveal the actual source of the measured PM." (l. 420-422)

- I recommend adding a dedicated Discussion section to the manuscript. This could include interpretive content currently located between lines 412 and 455.
  Interpretation of the results is a large part of this manuscript. Hence, we prefer to keep the interpretation integrated with the results instead of having a separate discussion section. We have, however, revised the text to make the discussion of previous finding more supportive for the new findings.

- Line 433: Consider specifying "anthropogenic precursors" instead of just "precursors" for clarity.
  Added as suggested.
- Figure S1: It would be helpful to include Pearson's correlation coefficient in Figure S1.
  We added correlation coefficient in the figure caption.

References

Heikkinen, L., Äijälä, M., Riva, M., Luoma, K., Dällenbach, K., Aalto, J., Aalto, P., Aliaga, D., Aurela, M., Keskinen, H., Makkonen, U., Rantala, P., Kulmala, M., Petäjä, T., Worsnop, D., and Ehn, M.: Long-term sub-micrometer aerosol chemical composition in the boreal forest: inter- and intra-annual variability, Atmos. Chem. Phys., 20, 3151-3180, DOI:10.5194/acp-20-3151-2020, 2020.

Kannosto, J., Virtanen, A., Lemmetty, M., Mäkelä, J. M., Keskinen, J., Junninen, H., Hussein, T., Aalto, P., and Kulmala, M.: Mode resolved density of atmospheric aerosol particles, Atmos. Chem. Phys., 8, 5327-5337, DOI:10.5194/acp-8-5327-2008, 2008.

Manninen, H., Bäck, J., Sihto-Nissilä, S.-L., Huffman, J., Pessi, A.-M., Hiltunen, V., Aalto, P. P., Hidalgo, P., Hari, P., Saarto, A., Kulmala, M., and Petäjä, T.: Patterns in airborne pollen and other primary biological aerosol particles (PBAP), and their contribution to aerosol mass and number in a boreal forest, Boreal Env. Res., 19 (suppl. B), 383-405, 2014.

Seinfeld, J. and Pandis, S.: Atmospheric chemistry and physics: from air pollution to climate change. Second edition. John Wiley & Sons, New Jersey, USA, 2006.