Dear Editor,

we wish to thank the reviewers for taking their time to comment our manuscript. Several improvements have been made to it and the major concerns, which were related to the use of ancillary flow rate as well as the improper language, have been addressed. We believe that the overall quality of the manuscript is now higher, and we are looking forward to hearing your opinion. Below, you will find a compilation of our point-by-point responses (marked as **bold**) to both reviewers. The revised manuscript and supplementary material (with 'track changes' on) have also been attached to the electronic submission.

On behalf of all authors,
Joel Kuula
Atmospheric Composition Research
Finnish Meteorological Institute
joel.kuula@fmi.fi

Anonymous Referee #2

Received and published: 27 December 2019

General Comments:

This manuscript describes the development of a system to provide a flow of monodisperse particles, and the authors then use this system to evaluate several common low-cost sensors. The description and evaluation of the new system (VOAG-GP50) are clearly described. Although their conclusion that low-cost sensors mis-classify particle sizes is not new, the systematic evaluation of size selectivity in low-cost sensors is a valuable contribution to the field. However, the paper would be stronger if the authors improved the clarity of the data-processing section, address a few questions regarding sample flow rate, and polish the language.

Specific comments:

The data processing section needs some clarification. The authors discuss dividing the data into 30 size bins, but in line 146 they discuss 10 steps to produce different particle sizes. The authors need to clarify how 10 steps can yield 30 different size bins.

The 10-step program (or method) refers to the dispensing of liquids and is related to neither data processing nor the amount of different particle sizes produced. The number of used steps and the parameters assigned to them simply define the minimum and maximum particle size and the rate at which the particle size gradient evolves from the minimum size to maximum size. The word "gradient" is used to underline that a step from e.g. 2 to 3 µm does not in fact lead to a discontinuous and sudden step from one particle size to another.

The 30 bin sizes (or sections) refer to data processing in which the raw 10-second resolution data (typically ~ 300 data points altogether) was divided into 30 "sections" according to the measured CMDs. For each section, an average detection efficiency and CMD was calculated.

To make this clearer, details and a step-by-step example (with figures) of how the data was processed was added to the supplementary material.

Added to manuscript section 2.3 "Data processing": "A detailed example how the data was processed and how the valid detection ranges were calculated is shown in the supplementary material."

Added to manuscript section 2.2.2 "Sampling configuration": "It is worth underlining that the number of steps used in the GP50 dispensing program does not dictate the number of different particle sizes produced. The number of steps and the parameters assigned to them simply define the minimum (blending ratio of the first step) and maximum (blending ratio of the last step) particle size and the rate (step duration) at which the particle size gradient evolves from the minimum size to maximum size. The word "gradient" is used to note that a step from 2 to 3 µm, for instance, does not lead to a discontinuous and sudden step from one particle size to another."

In Lines 175 to 180, the authors base their discussion of valid detection ranges on a detection efficiency curve, but I could not find a discussion of how they define a detection efficiency curve. I would suggest providing an example in the supplementary material and illustrating how the upper half of this curve is defined.

The normalized detection efficiency curve is defined in Eq. 3 and the respective curves are shown in Figure 4a-f. A demonstration how the valid detection ranges were calculated has been added to the supplementary material.

I have some concerns regarding the effect of sample flowrate on the low-cost sensors. In the experimental setup, a pump draws the monodisperse particles into the sensor housing at a flow rate of 1 lpm. Figure S2 shows the sensor housing and placement of the low-cost sensors with the flow directed at the sensor inlet (mostly). The authors should consider whether this setup may be skewing their results. This is particularly important for sensors with fans that are designed to operate at a specific flowrate. It is possible that pushing a flowrate that differs from the design flow rate could alter the results. For example, the PMS sensor has a volumetric flowrate of approximately 0.1lpm (which is 10x lower than the volumetric flowrate into the sensor housing). Granted not all of the 1 lpm would flow into the sensor, but this is worth considering.

There is no clear theoretical basis as to why a different flow rate would change the way the sensor discriminates different particle sizes. Assuming the sensors function as spectrometers (which the results of this study suggest they do, at least partly), the size discrimination of such devices is predicated on the analysis of pulse height caused by the scattering light of particle (Mie-theory). Change in the flow rate would change the frequency (i.e. number concentration) and duration of pulses but not their height. Whether the measured absolute concentrations were higher than expected is trivial as the data analysis of this study was based on normalized concentrations. Although possible, the effect of particle-size dependent sampling losses was originally estimated to be negligible, and as all the tested sensors seemed to exhibit similar size discrimination characteristics as what previous studies had shown, the effect of ancillary flow rate was not addressed in any way.

However, to ensure that the ancillary flow rate did not affect the results, additional tests were conducted with flow rates of 0.5 and 2 L min^{-1}. Instead of testing all the three sensor units of the six different sensor models, only a single unit (unit #3) for each sensor model was evaluated. The results (Fig. 1, attachment) indicate that different flow rates had no meaningful effect on the responses. The SDS011 shows slightly stronger response for particles larger than ~2 – 3 µm, but this is probably resulting from operator inconsistency (or randomness) because the change is similar for both 0.5 and 2 L min^{-1} flow rates. The B5W sensor has weaker response for particle sizes larger than ~4 – 5 µm with 2 L min^{-1} flow rate which suggests that the sampling losses may have increased. However, the response is similar for 0.5 and 1 L min^{-1} flow rates (B5W was originally designed to be used with a heater resistor-induced flow which is most probably closer to 0.5 than 2 L min^{-1}). The difference in PPD42NS responses, which imply that the losses may have increased for smaller and not higher flow rates, is attributed to randomness.

While conducting the additional flow rate tests, the Sharp sensors were found to exhibit completely different characteristics to what was previously measured. The underlying reason for this is still unknown, but it is possible that a prototyping breadboard, which was used to make the required connections for the external resistor and capacitor, had loose connections which resulted in misleading bias measurements. Nevertheless, all the Sharp sensors were re-evaluated with the original 1 L min^{-1} flow rate and the unit #3 was tested additionally with the 0.5 and 2 L min^{-1} flow rates. The new valid detection range was measured to be < 0.8 µm which is no longer in an obvious conflict with the results of the previously mentioned study of Wang et al. (2015). Considering the additional tests, it appears that the different flow rates may influence the sensor response in smallest particle sizes (< 0.55 µm), but the responses with 0.5 and 1 L min^{-1} flow rates are so similar

that the stated valid detection range remains the same. Smaller flow rates are likely to better represent the original flow rate, which for the Sharp sensors, was based on plain diffusion.

Added to manuscript section 2.2.2 "Sampling configuration": "Although there is no clear theoretical basis as to why a different flow rate would affect the way the sensor discriminates different particle sizes (apart from the different particle size-specific sampling losses), additional tests were conducted with flow rates of 0.5 and 2 L min-1 to ensure that this was indeed the case (see Supplemental Figure S2)."

Section regarding the results of Sharp GP2Y1010AU0F has been revised as: "The response function of the GP2Y1010AU0F is shown in Figure 4d, and its valid detection range appears to be approximately < 0.8 µm. Like the previously discussed sensors, the GP2Y1010AU0F can be used to measure small particles (e.g., PM1) but not coarse mode particles. Several laboratory evaluations have been previously conducted for the GP2Y1010AU0F, but none of these have assessed its detection range using monodisperse test aerosols (Li and Biswas, 2017; Manikonda et al., 2016; Sousan et al., 2016). Wang et al. (2015) used atomized polystyrene latex (PSL) particles to evaluate the effect of particle size on the GP2Y1010AU0F response, but no concluding remarks can be obtained from these results. The study method utilized only three different sized PSLs; moreover, it was not designed to investigate the complete detection range of the GP2Y1010AU0F. However, according to the authors, the results implied that the sensor was more sensitive to 300 nm particles than to 600 and 900 nm particles, which is in slight disagreement with the results of this study whereby the normalized detection efficiency curve shows the highest sensitivity peak for 0.6 µm sized particles as well as a decreasing trend for particles smaller than this. There is no obvious explanation for this discrepancy, but it is worth re-emphasizing the differences in the used evaluation approaches."

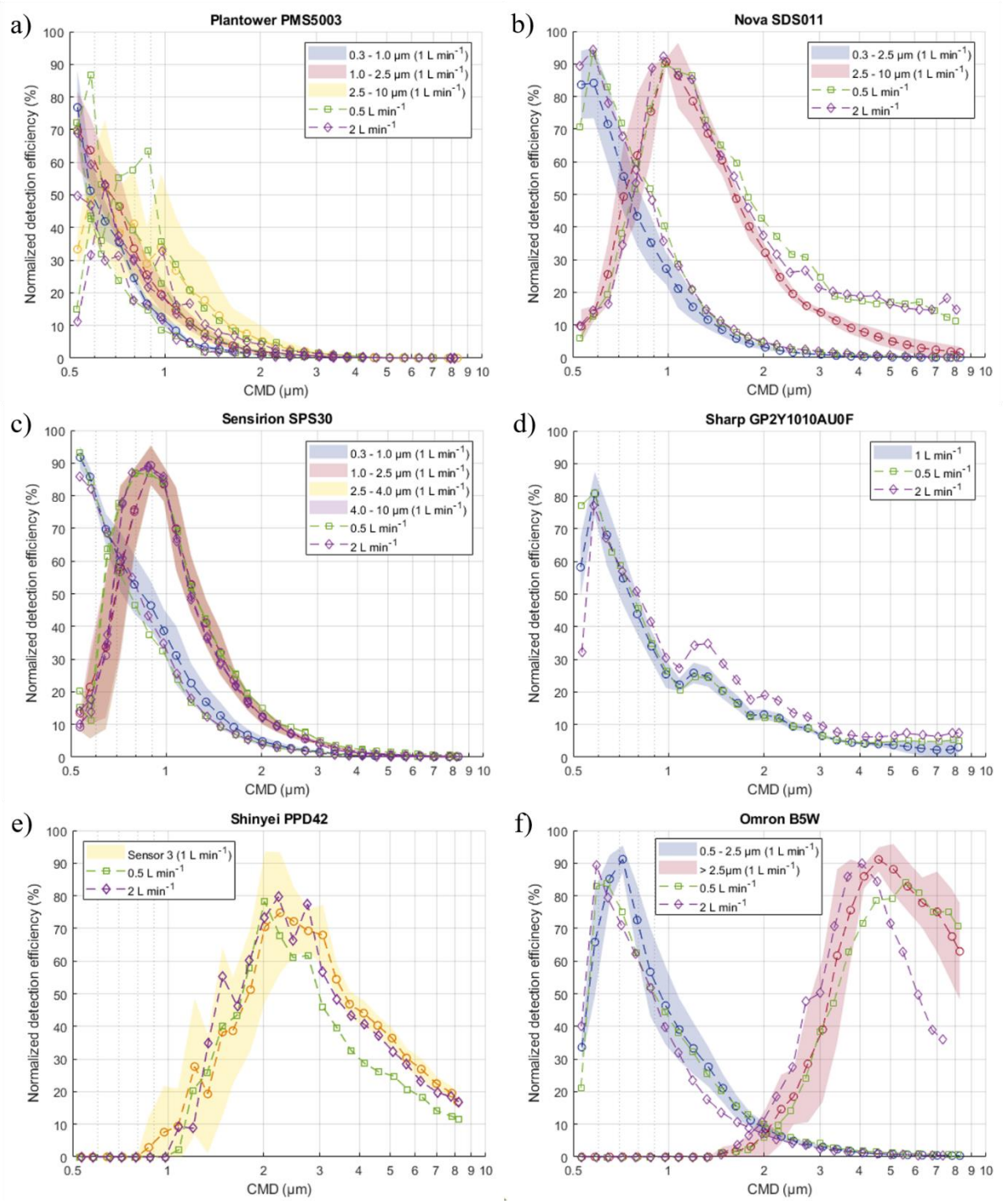


Fig 1. Results of the additional tests.

The manuscript needs a thorough review and edit by a native English speaker. The language is awkward and sometimes confusing. I am including a few examples from the abstract, but this list is not comprehensive: - “due to their prospective nature regarding spatial extension of measurement coverage”. Vague and awkward wording. - “sensors can be useful in achieving this goal”. No goal is mentioned previously. - “it is often reminded that the risk of sensor misuse”. Improper usage.

Commercial editing services were used to check and correct the language.

Technical corrections:

Line 121. Do the authors mean stable particle size distribution rather than particle size gradient? If they mean particle size gradient, this needs to be explained.

This has been rephrased as: "The novelty of the aerosol generation method used in this research is based on the observation that the particle size of the monodisperse and constant number concentration reference aerosol can be controlled by feeding solutions with different non-volatile concentrations to the VOAG, one after each other." (as suggested by Referee #1).

Line 185. The authors should clarify what the response curves are for. The 10-step(30-bin) generation of monodisperse particle sizes?

The GRIMM 1.108 was tested the same way as was all the low-cost sensors. The goal was to show how the particle size discrimination characteristics of the mid-cost, 15 bin spectrometer type instrument differed from the ones of the low-cost sensors.

Following changes were made to the manuscript: "The response curves" replaced with "The normalized detection efficiencies of the 15 bin GRIMM 1.108…"

Line 203. The authors should provide the standard deviations of the CMDs. In Figure2, the text says, "The GSD of the size distribution remains below 1.2, but line 113 says that the VOAG has relative standard deviation of less than 3 %."

The reason not to provide standard deviations of CMDs is discussed in the added supplementary material. In short, they were insignificant.

The statement "… standard deviation of less than 3 %" refers to the standard deviation of the *number concentration* of the VOAG, which was stated in the original paper of Berglund and Liu (1973), and it is in no way related to *size distributions*.

Line 113 rephrased as: "According to Berglund and Liu (1973), the output aerosol number concentration of the VOAG has a relative standard deviation of less than 3 %, and the formed particle size distribution is monodisperse having a geometric standard deviation (GSD) less than 1.014".

The authors mention the size limitation of the APS as being a limiting factor in the analysis (Figure 2), but in line 119 they mention that the VOAG cannot reliably generate particles smaller than 0.55 um. This limitation should be mentioned in Figure 2 in addition to the APS limitation. It would also be worth mentioning this important limitation in the abstract.

Added to line 19 (Abstract): "(from ~ 0.55 to 8.4 µm)"

Added to Figure 2 caption: "Along with the lower detection limit of the APS, another limiting factor of the study was the smallest producible particle size, which was approximately 0.55 µm."

Figure 1 – The GRIMM is not shown. Where does the GRIMM draw its sample? How are the flows distributed symmetrically between the APS and GRIMM since the GRIMM's flow is 1.2 lpm whereas the GRIMM is 1 lpm?

The GRIMM drew its sample from where the sensor enclosure is now shown. The isokinetic flow splitter was designed for equal (1 L min^{-1}) flow rates, but this was not considered problematic as there was never intention to evaluate the GRIMM other than cursorily. Furthermore, the Figure 3 shows that the unequal flow distribution probably had little to no effect on the response.

Figure 1 caption rephrased as: “Figure 1: Schematic of the sensor evaluation setup. The GRIMM 1.108 drew its sample from where the sensor enclosure is now shown.”

Anonymous Referee #1

Received and published: 3 February 2020

In this work, Kuula et al. present an investigation of counting efficiency as a function of size for a series of low-cost light scattering devices. The research question that the authors seek to answer (whether low-cost sensors can correctly assign particles to the appropriate PM mass fractions) is an important one. The work appears well thought out and carefully conducted, although the manuscript is below average from an English language perspective. There are dozens of grammatical and stylistic mistakes that will need attention prior to publication. However, the authors present a novel experimental procedure and some interesting results, so I am supportive of publication if this and other concerns are addressed in a revised manuscript. Comments on how the presentation of the methods and the interpretation of the results could be improved are below.

Major Comments:

One major drawback of this work is that the authors appear to have used a "forced flow" of 1 L/min through each sensor, which may not be in accordance with the manufacturer's intended use. The use of a non-standard flow rate (i.e., one that is different from the manufacturer's recommendation) could compromise a sensor in a number of ways. For example, it could lead to differential inertial losses of particles as they transmit through the sensor. Or, it could lead to variable pulse widths of particles during detection that deviate from the expected pulse widths during calibration. The authors need to verify that this method of testing under a forced flowrate did not affect sensor response. One way would be to repeat these tests at flow rates of 0.5 and 2 Lpm to see if the response factors change. If they do change, then the results presented here might lead to misinterpretation...I do not support publication of this manuscript until this issue is investigated further and/or resolved.

Below is the copy of our response to Referee #2 (concerning the same question):

There is no clear theoretical basis as to why a different flow rate would change the way the sensor discriminates different particle sizes. Assuming the sensors function as spectrometers (which the results of this study suggest they do, at least partly), the size discrimination of such devices is predicated on the analysis of pulse height caused by the scattering light of particle (Mie-theory). Change in the flow rate would change the frequency (i.e. number concentration) and duration of pulses but not their height. Whether the measured absolute concentrations were higher than expected is trivial as the data analysis of this study was based on normalized concentrations. Although possible, the effect of particle-size dependent sampling losses was originally estimated to be negligible, and as all the tested sensors seemed to exhibit similar size discrimination characteristics as what previous studies had shown, the effect of ancillary flow rate was not addressed in any way.

However, to ensure that the ancillary flow rate did not affect the results, additional tests were conducted with flow rates of 0.5 and 2 L min^{-1}. Instead of testing all the three sensor units of the six different sensor models, only a single unit (unit #3) for each sensor model was evaluated. The results (Fig. 1, attachment) indicate that different flow rates had no meaningful effect on the responses. The SDS011 shows slightly stronger response for particles larger than ~2 – 3 µm, but this is probably resulting from operator inconsistency (or randomness) because the change is similar for both 0.5 and 2 L min^{-1} flow rates. The B5W sensor has weaker response for particle sizes larger than ~4 – 5 µm with 2 L min^{-1} flow rate which suggests that the sampling losses may have increased. However, the response is similar for 0.5 and 1 L min^{-1} flow rates (B5W was originally designed to be used with a heater resistor-induced flow which is most probably closer to 0.5 than 2 L min^{-1}). The difference in

PPD42NS responses, which imply that the losses may have increased for smaller and not higher flow rates, is attributed to randomness.

While conducting the additional flow rate tests, the Sharp sensors were found to exhibit completely different characteristics to what was previously measured. The underlying reason for this is still unknown, but it is possible that a prototyping breadboard, which was used to make the required connections for the external resistor and capacitor, had loose connections which resulted in misleading bias measurements. Nevertheless, all the Sharp sensors were re-evaluated with the original 1 L min^{-1} flow rate and the unit #3 was tested additionally with the 0.5 and 2 L min^{-1} flow rates. The new valid detection range was measured to be < 0.8 µm which is no longer in an obvious conflict with the results of the previously mentioned study of Wang et al. (2015). Considering the additional tests, it appears that the different flow rates may influence the sensor response in smallest particle sizes (< 0.55 µm), but the responses with 0.5 and 1 L min^{-1} flow rates are so similar that the stated valid detection range remains the same. Smaller flow rates are likely to better represent the original flow rate, which for the Sharp sensors, was based on plain diffusion.

Added to manuscript section 2.2.2 "Sampling configuration": "Although there is no clear theoretical basis as to why a different flow rate would affect the way the sensor discriminates different particle sizes (apart from the different particle size-specific sampling losses), additional tests were conducted with flow rates of 0.5 and 2 L min-1 to ensure that this was indeed the case (see Supplemental Figure S2)."

Section regarding the results of Sharp GP2Y1010AU0F has been revised as: "The response function of the GP2Y1010AU0F is shown in Figure 4d, and its valid detection range appears to be approximately < 0.8 µm. Like the previously discussed sensors, the GP2Y1010AU0F can be used to measure small particles (e.g., PM1) but not coarse mode particles. Several laboratory evaluations have been previously conducted for the GP2Y1010AU0F, but none of these have assessed its detection range using monodisperse test aerosols (Li and Biswas, 2017; Manikonda et al., 2016; Sousan et al., 2016). Wang et al. (2015) used atomized polystyrene latex (PSL) particles to evaluate the effect of particle size on the GP2Y1010AU0F response, but no concluding remarks can be obtained from these results. The study method utilized only three different sized PSLs; moreover, it was not designed to investigate the complete detection range of the GP2Y1010AU0F. However, according to the authors, the results implied that the sensor was more sensitive to 300 nm particles than to 600 and 900 nm particles, which is in slight disagreement with the results of this study whereby the normalized detection efficiency curve shows the highest sensitivity peak for 0.6 µm sized particles as well as a decreasing trend for particles smaller than this. There is no obvious explanation for this discrepancy, but it is worth re-emphasizing the differences in the used evaluation approaches."

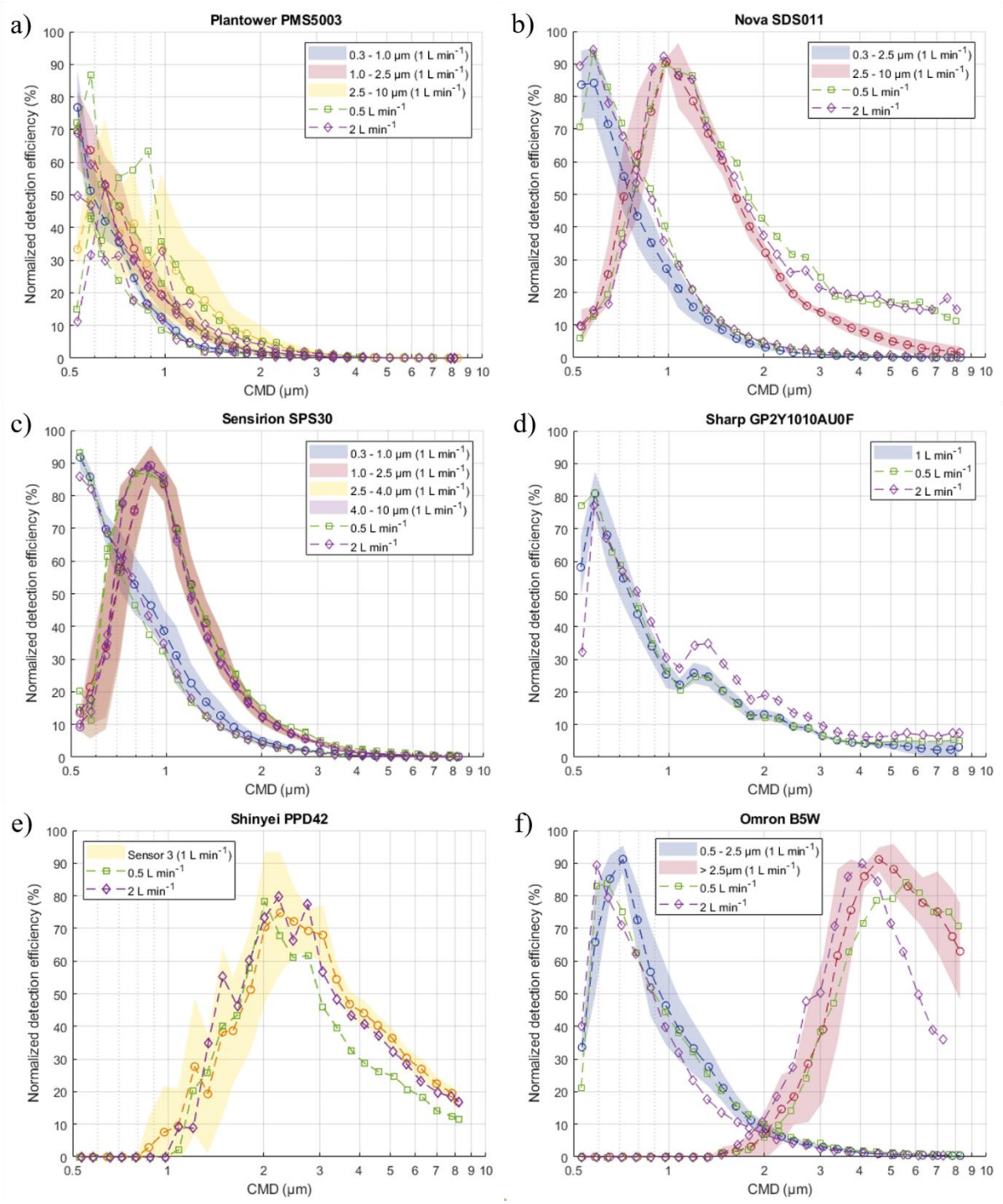


Fig 1. Results of the additional tests.

I think the VOAG method for generating a sequence of monodisperse particles of varying size is both novel and useful to the community. However, I would appreciate more detail on how the method is implemented so that others can properly reproduce this method. Such details could be enumerated as supplemental material.

All necessary running parameters of the VOAG and GP50 have been presented in the supplementary material Tables S1-2.

Comments on the Methods section:

The Methods section could be organized better and was missing key information needed to reproduce the work.

See our previous comment. Improvements to the Methods section have been made according to the referee comments listed in the next sections.

Comments on technical issues:

1. Table 1: The "number of size bins" column is problematic. It's not appropriate to compare the number of PM mass fractions reported by the low-cost sensors to the number of size bins that the GRIMM 1.108 classifies particle into. Perhaps the authors could remedy this problem by creating two columns: one labeled "number of mass fractions reported" and "number of particle size bins". For example, the SPS30 reports data in 4 mass fractions and 5 particle size bins.

Due to space limitations, the column "Number of size bins" was rephrased as "Number of mass fractions reported" but no additional column was added. Column "Sensor output" was modified for the GRIMM as "All fractions individually".

2. Line 76: Why did the authors remove the mechanism that drives sample flow through some of the sensors but not others? Specifically, why did they remove the air heating resistors from the PPD42 and B5W but not remove the fans from the PMS5003, SDS011, and SPS30? This discrepancy seems like it could lead to an unfair comparison between different sensors. A sensor's output is partially determined by how quickly and efficiently particles are drawn into the sensing zone. If the flow rate through the sensor is modified, the output of the modified sensor might differ from the output of the "as-purchased" sensor. For example, results presented by Tryner et al. (doi:10.1039/C9EM00234K) indicate that varying the flow rate through a modified PMS5003sensor changes the output. Since the authors modified some of the sensors used in this study, the results presented here might not be directly transferrable to unmodified sensors being used in the field.

All decisions regarding the used sampling setups were based on two factors. Most importantly, the configuration had to be such that the reference aerosol remained representative and valid throughout the test run. A secondary factor was that the configuration had to resemble the original configuration as closely as possible.

When relied on convection, the sensor must be orientated in an upright position, and the sample needs to be fed to the sensor from the bottom. This poses practical issues which are hard to overcome. For example, it is unclear how a fresh and representative sample could be provided to the proximity of the resistor without disturbing the temperature gradient and/or the flow rate of the convection itself. The feeding flow rate would have to be very small in order to avoid any disturbance, but this would simultaneously compromise the explicitness of the reference aerosol. The use of primary standard reference particles was one of the most important elements of this study. Regarding the PPD42NS specifically, if the resistor-demanded orientation is used, the lenses of this sensor are facing towards the incoming flow (forward angle scattering) which, consequently, exposes them to fouling (drift in sensitivity). This is a major problem, which we encountered in our previous paper (10.3390/s17122915), and there is no clear solution for this apart from reversing the flow direction. On a general note, the convection-based flow is poorly suited for any kind of sampling due to its susceptibility to external disturbances, such as the dynamic effects of wind and changes in ambient temperature.

Removing the fans from the PMS5003 and SPS30 would most probably lead to significant sampling losses due to their 90-degree elbows (and no mechanical flow through the sensor), and the obvious way to account for this would be to relocate the intake (drill an alternative hole to the sensor body) so that the sample is introduced directly to the detection zone. However, this is most certainly not the way the sensors are typically used. The SDS011 would probably function acceptably even without its fan, but it was considered better to retain the original configuration.

We agree that the testing setups are not perfect and that currently, some of the sensors are probably used in a manner that they were not designed for. However, creating a uniform experiment configuration for sensors which all entail different type (and poorly designed) layouts and flow rates will always be a compromise of some sort. To the extent that the custom sampling setups may have affected the outcome, we want to underline the main finding of this study is that the sensors do not measure particle sizes which their technical specifications imply, and that the ancillary flow rates, which were discussed earlier, did not have a significant impact on the results.

Study of Tryner et al. is not helpful as the particle size discrimination of PMS5003 was not investigated; it likely that the concentrations increased across all outputs respectively.

3. Lines 85-89: The PMS5003 reports each mass fraction two ways: "CF=1, standard particle" and "under atmospheric environment". Which values were used in the data analyses presented here?

Added to Table 1 notes: "Standard particle (CF=1) output was used."

4. Lines 164-165: Please state the fraction of raw measurement points that were disregarded.

Added: "(~ 2.1 % of the data)"

5. Line 170: Equation 3 is not clear. What is the maximum Sensor/APS ratio? Was this the maximum ratio measured during the entire 60-minute long test run?

Maximum Sensor/APS ratio refers to the maximum Sensor/APS ratio measured during a single test run.

Added: "Maximum Sensor/APS ratio refers to the maximum ratio measured during a single test run."

6. Line 175: One issue is that the APS is known to have poor counting efficiency for liquid droplets above about 5 microns in diameter. Did the authors account for this artifact? See https://doi.org/10.1016/j.jaerosci.2005.03.009

This artifact was not accounted.

Added to section 2.2.2 "Sampling configuration": "Although the reference instrument APS is known for having decreased counting efficiency for liquid droplets over ~5 µm in size (Volckens and Peters, 2005), no additional corrections were used."

7. Line 179-180: This definition of "valid detection ranges" is not clear. Do the authors mean that the valid detection range was the range of CMDs for which the normalized detection efficiency was greater than 50%? If so, please consider also rephrasing the text on lines 209-210 as something like "(the valid detection range was defined as the range of particle sizes for which normalized detection efficiency was > 50%)".

Definition of > 50 % was not used because the obtained detection efficiency ranges were different for different sensor models (e.g. not from 0 to 100 %) and the normalized detection efficiency does

not describe the proportion of measured concentrations. A detailed example how the upper half of the normalized detection efficiency curve was calculated has been added to the supplementary material.

Added to section 2.3 "Data processing": "A detailed example how the data was processed and how the valid detection ranges were calculated is shown in the Supplementary Material".

Comments of organization:

1. Lines 75-84: This information should be moved to the "Sampling configuration" section. These methods are difficult to understand or justify without additional knowledge of the system used to pass monodisperse aerosol through the sensor.

This paragraph was moved to section 2.2.2 "Sampling configuration".

2. Lines 85-82: This information should be moved to the "Data processing" section.

It is assumed that the referee means lines 85-89 (otherwise it contradicts the comment above). This information was moved to section 2.3 "Data processing".

3. Lines 89-90: Please move the sentence that starts "Three units of each sensor..."to the end of the first paragraph in Section 2.1.

Corrected.

4. Lines 90-93: Please start a new paragraph here that contains the information on the GRIMM 1.108.

Corrected.

5. Line 95: Section 2.2.1 should start with a sentence that says, "The aerosol sampled by the low-cost sensors was generated using a Vibrating Orifice Aerosol Generator3450 (VOAG, TSI Inc., USA)." That way readers will know why the information contained in this section is relevant.

Corrected.

6. Lines 101-102, "This aerosol generation method...": This sentence is neither relevant nor helpful. Please delete it.

Deleted.

7. Lines 121-125: The second sentence in this paragraph is more helpful than the first. I suggest rephrasing as "The novelty of the aerosol generation method used in this research is based on the observation that the particle size of the monodisperse and constant number concentration reference aerosol can be controlled by feeding solutions with different non-volatile concentrations to the VOAG, one after each other."

Rephrased as suggested.

8. Lines 161-162: Please show at least one example panel from Figure S2 in the main text.

PMS5003 panel added to manuscript: "An illustration of the PMS5003 sampling arrangement is shown in Figure 3."

9. Lines 176-177, "The decision to divide the data into 30 bins...": This sentence is not necessary. Please consider deleting it.

Deleted.

10. Lines 186-187: The sentence that begins “The previously-described data processing technique...” should be moved to the “Data processing” section.

Moved.

Comments of the Results section

Comments on technical issues:

1. Overall, the authors interpretation of their experimental results seems to be based on the tenuous assumption that the low-cost particulate matter sensors all function as aerosol spectrometers rather than as nephelometers. In the GRIMM 1.108 aerosol spectrometer, sample air is aerodynamically focused and passes through a narrowly focused laser beam so that the detector sees the light scattered by just one particle at a time. Thus, the GRIM1.108 can count and size individual particles. The low-cost sensors evaluated in this study do not aerodynamically focus the sample air and, as a result, it is unlikely that the detectors in these sensors see light scattered by individual particles. Instead, the mass concentrations reported by the low-cost sensors are most likely determined by the intensity of light scattered by a group of particles, and not from a measured particle count and size distribution.

Whether the sensors function as spectrometers, nephelometers, or as a combination of these two is unknown, and a thorough reverse engineering of the electronic design (and software) would be required in order to reach ultimate conclusion. According to the results of this study, the sensors exhibit some kind of size discrimination capability and this observation in an agreement with the broader scientific community which uses sensors to measure size-specific mass fractions, such as the PM2.5. Whether the sensors function solely as spectrometers, or whether they are switching between single particle detection and total scattered light intensity measurement, remains unknown. It is worth noting that nephelometers, by design, do not have a mechanism (apart from size-selective inlets) which would allow them to size discriminate individual particles, and thus, the fact that many of the sensors evaluated in this study showed little to no response in larger particle sizes would be unexplainable. According to Mie-theory, the total intensity of the scattered light is proportional to the sixth power of particle diameter.

To avoid particle coincidence, the optical detection layout does not need to be complicated. For example, a detection volume of 1 cm^3 (e.g. 10 x 10 x 10 mm cube) would imply that, statistically, the concentration limit for particle coincidence is 1 cm^{-3}. With detection volume of 0.125 cm^3 (e.g. 5 x 5 x 5 mm cube) the respective concentration limit is 8 cm^{-3}. Typical number concentrations for coarse mode (1–10 μm) particles are relatively low, for instance, 7 ± 9 cm^{-3} in urban Beijing (Wu et al. 2008 10.1016/j.atmosenv.2008.06.022) and 2 ± 3 cm^{-3} in regional China (Shen et al. 2011, 10.5194/acp-11-1565-2011), and therefore, it is reasonable to assume that a fairly simple layout capable of measuring individual particles (at least occasionally) without additional optical or aerodynamic lenses is possible to engineer and design.

Results presented by Kelly et al. in 2017 (doi: 10.1016/j.envpol.2016.12.039) suggested that, for Plantower sensors, “the allocation of light scattering to PM1, PM2.5, and PM10 is based on a theoretical model rather than a measurement” (pp. 495-496) and “the size distribution provided by the PMS is based on a theoretical model rather than a measurement” (pp. 497). Claims in the literature that sensors such as the PMS5003 do function as optical particle counters are dubious.

The conclusion by Kelly et al. is most probably right and it is in an agreement with the results of this study. But it does not provide evidence that the PMS works solely as a nephelometer.

2. Lines 207: I would not say that "the particle sizing of the different bins is not working properly." See the comment above for justification. I think it would be appropriate to say, "it is apparent that the PMS5003 does not accurately distinguish between the PM1, PM2.5, and PM10 size fractions." Do the ratios of the masses in the PM1, PM2.5, and PM10 change as the sensor are exposed to different particle sizes?

Line 207 corrected as suggested. Whether the ratio of mass changes or stays constant does not determine the difference or similarity in size discrimination capability. For example, two outputs with identical size discrimination could have different mass ratios if their relation is exponential (e.g. PM1= PM2.5^2). For the PMS, the ratios were mostly the same; PM1/PM2.5 (3.7±0.4), PM2.5/PM10 (1.6±0.4), PM1/PM10 (5.8±0.9).

3. Line 223: Levy Zamora et al. evaluated the PMSA003 sensor, so results from that study might not be directly comparable to results from this study.

This reference was removed.

4. Lines 231-235: The authors state that they expect the SDS011 to measure PM2.5 more accurately than the PMS5003 based on differences in 2.5 – 10 um mass fractions reported by the two sensors. I don't see how differences in the 2.5 – 10 um mass fractions would affect the reported 0.3 – 2.5 um mass fractions. The two sensors have similar detection efficiencies for the 0.3 – 2.5 um mass fraction. How does the accuracy of PM2.5 measurements reported by the SDS011 (as reported by Badura et al. and Liu et al.) compare to the accuracy of PM2.5 measurements reported by the PMS5003(see, for example, Malings et al., doi: 10.1080/02786826.2019.1623863)?

The statement is based on the notion that in practice, the SDS011 has two different outputs for ~ 0.3 – 2.5 µm range (approximately < 0.8 and 0.7 – 1.7 µm) whereas the PMS has only one (all three outputs seem to measure similar size fractions; < 0.7, < 0.8, and < 1 µm). By utilizing both outputs of the SDS011 (for example, by calculating the ratio of the bins), a more accurate approximation for the mass distribution in 0.3 – 2.5 µm range, and thus higher accuracy, is possible to obtain. Of course, the user of the sensor must be aware of the particular size discrimination characteristics of the SDS011 (e.g. PM10 does not measure PM10).

To make this clearer, the statement has been rephrased as: "For example, by calculating the ratio of bins 1 and 2, it is possible to approximate the distribution of mass in the 0.3–2.5 µm size range, thus using an additional correction factor to obtain more accurate results."

5. Lines 240-241: The SPS30 product datasheet states that the SPS30 is calibrated using "a defined potassium chloride particle distribution".

(www.sensirion.com/fileadmin/user_upload/customers/sensirion/Dokumente/0_Datasheets/Particulate_Matter/Sensirion_PM_Sensors_SPS30_Datasheet.pdf).

To be precise, the documentation states: "PM2.5 accuracy is verified for every sensor after calibration using a defined potassium chloride particle distribution". It does not disclose how PM1, PM4 or PM10 have been calibrated. We asked Sensirion about this and the unofficial answer was that PM1 is calibrated with KCl and PM2.5, PM4 and PM10 with Arizona dust. PM4 and PM10 are just extrapolations of the PM2.5 response (which is evident according to the results of this work).

6. Lines 246-247: The R^2 values reported here provide no indication of sensor accuracy. Accuracy is related to the absolute difference between the concentration reported a low-cost sensor and a reference monitor. It's possible for the mass concentrations reported by a low-cost sensor to be very

different in magnitude from the concentrations reported by an FEM monitor but for the concentrations reported by the two to still be strongly correlated. Saying that the fraction of variance explained by a linear model is high doesn't imply that the model has an intercept of zero and a slope of one.

Word "accuracies" replaced with "coefficient of determinations".

7. Lines 281-288: Again, the Rˆ2 values reported here provide zero information on sensor accuracy. These values only imply whether a linear model explains the relationship between concentrations reported by a reference monitor and concentrations reported by the sensor. Measures of accuracy include mean error, mean absolute error, and mean percentage error.

Word "accuracy" replaced with "performance".

8. Lines 309-317: See comment #1 on the results.

See the respective response.

Comments on organization:

1. Figure 3: This figure could be improved to help readers interpret it. A title above the legend that says "Particle size bins" would be nice. I also suggest revising the first sentence in the caption to something like "Normalized detection efficiency of the 15particle size bins as a function of the count median diameter of the reference aerosol." Finally, please also move the sentences from lines 185-186 ("For the sake of clarity...") and line 187-188 ("Bins 14 and 15...") to the figure caption.

Legend title added. Caption rephrased as suggested: "Figure 4. Normalized detection efficiency of the 15 particle size bins as a function of the count median diameter of the reference aerosol. Consecutively increasing and decreasing response curves indicate that the particle sizing of the instrument is functioning correctly. For the sake of clarity, degrees of measurement variation have been excluded from the figure. Bins 14 and 15, which correspond to 10–15 and 15–20 µm, respectively, are not shown as they did not produce any response (as expected)."

2. Figure 4: The caption for this figure should be more informative. I suggest "Normalized detection efficiency of discretized PM mass fractions reported by the low-cost sensors as a function of the count median diameter of the reference aerosol." I also suggest moving the text on lines 201-204 ("The colored circles represent.... ranges stated by the corresponding manufacturer.") to the caption.

Caption rephrased as suggested: "Figure 5. Normalized detection efficiency of discretized PM mass fractions reported by the low-cost sensors as a function of the count median diameter of the reference aerosol. The colored circles represent the calculated average responses of the three sensor units, and the shaded background areas represent the respective standard deviations. Standard deviations of the average CMDs were negligible due to the reliable and reproducible test method. Figure legends correspond to the bin size ranges stated by the corresponding manufacturer."

Editorial comments:

1. Line 13: The PM acronym should come before the word 'sensors'.

Corrected.

2. Lines 17-18: I disagree with the statement "This implies that there are underlying reasons yet to be characterized which are causing inaccuracies in sensor measurements." The underlying reasons for

why light-scattering sensors are inaccurate (regardless of cost) has been known among the aerosol research community for decades; there is a vast literature on the limitations and biases associated with this measurement method.

This statement refers to sensor type measurements specifically, and the results of this study imply that multiple studies have been conducted without understanding what the sensors are truly measuring. We do not consider it appropriate to take a hostile stance in reasoning our study objectives and background.

3. Line 29: “enabling” should be “enable” and “higher resolution” should be “higher-resolution”.

Corrected.

4. Lines 30-31: It might be helpful to refer readers to some more recent studies demonstrating applications of sensor networks. See, for example, Feinberg et al., 2019 (doi: 10.1016/j.atmosenv.2019.06.026) and Rickenbacker et al., 2019 (doi:10.1016/j.scs.2019.101473).

Study of Feinberg et al. added. Study of Rickenberg et al. was not accessible with our institute credentials and was excluded.

5. Line 54: I would use a phrase like "size discrimination" instead of "size selectivity" in this context. The term “size selectivity” is most often used in the aerosol literature to discuss inlet aspiration/transmission performance, such as that of a PM10 or PM2.5inertial separator.

Corrected.

6. Line 59: “cursory” should be “cursorily” or “concurrently”, depending on what the authors mean by this sentence.

Corrected as “cursorily”.

7. Line 82: Rephrase as “However, the more stable sample flow system (i.e., fan instead of convection) might help compensate for the sub-optimal layouts of these sensors.”

Corrected.

8. Line 85: “analogue” should be “analog”.

Corrected.

9. Line 86: For clarity, please change “sensor outputs shown” to “PM mass fractions listed”.

Corrected.

10. Line 92: Change “being” to “to be” and delete the word “other”. The GRIMM1.108 does not measure PM mass directly. Also, please specify which metric reported by the GRIMM 1.108 (or derived from values reported by the GRIMM 1.108) has a comparable accuracy to the filter weighing method.

Corrected and added: “(mass of C-factor adjusted total suspended particles)”

11. Line 126: Delete the word “produced”.

Corrected.

12. Lines 227-228: Consider replacing “two more clear different” with “two clearly different”.

Corrected.

13. Line 242: Revise as “and neither Web of Science nor Scopus literature research showed...”.

Corrected.

14. Line 249: Replace “align” with “alignment”.

Corrected.

15. Line 293: Replace “prominent” with “promising”.

Corrected.

16. Line 296: Add the word “Neither” before “Web of Science”.

Corrected.

17. Line 307: Replace “pertain” with “retain”.

Corrected.

18. Figure 2: The title for this graph is “Sample aerosol”, but the aerosol sampled by the ow-cost sensors is referred to as the “reference aerosol” throughout the manuscript. Please use consistent terminology.

Corrected.

References

GRIMM 1.108 manual: https://wmo-gaw-wcc-aerosol-physics.org/files/opc-grimm-model–1.108-and-1.109.pdf

Added.

Laboratory evaluation of particle--size -selectivity of optical low-cost particulate matter sensors

Joel Kuula[1], Timo Mäkelä[2], Minna Aurela[1], Kimmo Teinilä[1], Samu Varjonen[3], Óscar González[4], and Hilkka Timonen[1]

[1]Atmospheric Composition Research, Finnish Meteorological Institute, Helsinki, FI-00560, Finland
[2]Climate System Research, Finnish Meteorological Institute, Helsinki, FI-00560, Finland
[3]Department of Computer Science, University of Helsinki, Helsinki, FI-00560, Finland
[4]Fab Lab Barcelona, Institute of Advanced Architecture of Catalonia, Barcelona, 08005, Spain

Correspondence to: Joel Kuula (joel.kuula@fmi.fi)

Abstract. Low-cost particulate matter (PM) sensors ~~(PM)~~ have been under investigation as it has been hypothesized that the use of low-cost and easy-to-use sensors could allow cost-efficient extension of the currently sparse measurement coverage~~due to their prospective nature regarding spatial extension of measurement coverage~~. While majority of the existing literature highlights that low-cost sensors can indeed be ~~useful in achieving this goal~~a valuable addition to the list of commonly used measurement tools, it is often reminded that the risk of sensor misuse is still high~~,~~ and that the data obtained from the sensors is only representative of the specific site and its ambient conditions. This implies that there are underlying reasons yet to be characterized ~~which are causing~~behind inaccuracies in sensor measurements. The objective of this study ~~was~~ is to investigate the particle--size- selectivity of low-cost sensors. Evaluated sensors were Plantower PMS5003, Nova SDS011, Sensirion SPS30, Sharp GP2Y1010AU0F, Shinyei PPD42NS, and Omron B5W-ld0101. The investigation of size- selectivity was carried out in the laboratory using a novel reference aerosol generation system capable of steadily producing monodisperse particles of different sizes (from ~ 0.55 to 8.4 µm) on-line. The results of the study show~~ed~~ that none of the low-cost sensors adhered ~~exactly~~ to the detection ranges declared by the manufacturers~~, ; and~~ moreover, cursory comparison to a mid-cost aerosol spectrometer (GRIMM 1.108) indicate~~d~~s that the sensors ~~could~~ can only achieve independent responses for ~~1-2~~one or two size bins, whereas the spectrometer ~~could~~ can sufficiently characterize particles with 15 different size bins. These observations provide insight and evidence to the notion that particle--size- selectivity ~~may have~~has an essential role in the analysis of the sources of errors ~~source analysis of~~in sensors.

1 Introduction

The ~~Recent~~ recent emergence of low-cost sensors has enabled new possibilities in traditional air quality monitoring (Kumar et al., 2015; Morawska et al., 2018; Snyder et al., 2013). As a result of low unit costs and compact size, sensors can be deployed to the field in much higher quantities than before ~~and~~ thus enabling higher--resolution spatiotemporal data. Few

studies have demonstrated applications of sensors networks (Caubel et al., 2019; Feinberg et al., 2019; Gao et al., 2015; Jiao et al., 2016; Popoola et al., 2018; Yuval et al., 2019). Distributed sensing of air quality can be seen as an important progression towards a more comprehensive understanding of city-scale air quality dynamics as air pollution, and ~~especially~~ particulate matter (PM), in particular, may have highly localized concentration "hotspots" in urban areas. Practical limitations, such as expensiveness and bulkiness, constrains the use of conventional instrumentation in monitoring networks; ~~and~~ therefore, low-cost sensors could have an essential role in the spatial extension of measurement coverage.

Numerous field studies have been conducted previously~~,~~; ~~and~~ the majority ~~of these~~ have underlined the potential usefulness of optical particulate matter sensors (Karagulian et al., 2019; Rai et al., 2017). ~~In spite of this~~However, the literature~~it~~ has also emphasized that the risk of sensor misuse is still high~~,~~ and that some external factors, such as relative humidity, may produce significant measurement ~~artefacts~~ artifacts to the data (Jayaratne et al., 2018; Kuula et al., 2018; Liu et al., 2019). In comparison to gas sensing, PM measurements are notably more challenging ~~considering that~~when ambient particle sizes~~,~~ and their respective ~~size~~ distributions~~, may~~ vary significantly from source to source and from location to location~~,~~. ~~and a~~Along with size, particle physical properties such as shape and refractive index also ~~have an effect on~~affect the sensor output. Several studies have pointed out that along with dynamic adjustment for meteorological parameters, on-site calibrations are required in order to achieve higher levels of accuracy and precision (Zheng et al., 2018). However, when considering advanced calibration techniques, Schneider et al. (2019) have~~s~~ raised a valid point noting that it may be unclear whether the sensor data resulting from complex correction and conversion processes (e.g. machine learning) is still a legitimate and independent product of the sensor measurement and not a combination of secondary data and statistical model prediction. This is an important remark when evaluating the usability of sensors as it~~, and moreover,~~ highlights the need to identify the ~~underlying~~ reasons ~~causing~~ behind inaccuracies in low-cost sensor measurements.

While field evaluations are a natural step towards understanding and developing sensors, they provide limited information about the detailed sensor response characteristics. In particular, less attention has been paid to the investigation of particle size discrimination ~~selectivity~~ of sensors. Although a few studies have noted that the detectable particle ~~–~~size ranges of sensors may be significantly different from the ones declared in their respective technical specification sheets (Budde et al., 2018; Levy Zamora et al., 2019), this factor is~~it is~~ not a commonly considered ~~factor~~ when assessing sensor accuracy. Thus, more research is needed. The objective of this study was to investigate and characterize the size ~~-~~selectiveness of some of the optical low-cost sensors commonly appearing in the literature. The evaluated sensors were Plantower PMS5003, Nova SDS011, Sensirion SPS30, Sharp GP2Y1010AU0F, Shinyei PPD42NS, and Omron B5W-ld0101. Along with these low-cost sensors, a mid-cost optical aerosol spectrometer (GRIMM model 1.108, GRIMM Aerosol Technik GmbH., Germany) was evaluated cursorily to highlight the differences between the responses of ~~the~~ low-cost and mid-cost devices. The investigation of size ~~-~~selectivity was carried out in the laboratory using a novel reference aerosol generation system capable

of steadily producing monodisperse particles of different sizes. Sensor responses were compared to a reference instrument (APS, Aerodynamic Particle Sizer 3321, TSI Inc., USA), and detectable particle size ranges of the sensors were obtained.

2 Methods

2.1 Evaluated sensors

The sensors evaluated in this study, and their main detection properties, are listed in Table 1. The optical detection configurations of these sensors ~~are~~ were arranged in either a 90- or 120--degree scattering angle, and either a red laser or an infrared (IR) light--emitting diode (LED) ~~is~~ was used as a light source. Sensors utilizing an LED ~~are~~ were equipped with additional light focusing lenses. The optical chamber itself ~~is~~ was composed~~constructed~~ of an ~~injection~~ injection-molded plastic body which ~~is~~ was placed onto an electronic circuit board. The PMS5003, SDS011, and SPS30 use fans to generate sample flow, whereas the PPD42 and B5W utilized natural convection resulting from a heating resistor. The ~~Sampling~~ sampling of the GP2Y1010AU0F ~~is~~ was based on diffusion. The optical configurations and plastic body layouts are shown in Supplemental Figure S1. Three units for each sensor model were evaluated in order to assess their inter-unit variation.

~~All sensor units were in original condition except the PPD42 and B5W sensors which had their air heating resistors removed. The evaluation platform used in this study did not necessarily require independent means of sample flow. Furthermore, holes were drilled to the plastic body of the PPD42 in order to ensure that the sample aerosol could reach to the optical detection volume. The inlet of the PPD42 was originally designed to be on top of the plastic body (facing towards the electronic circuit board), and therefore, when the electronic circuit board of the sensor was orientated in parallel with the sample stream, majority of the particles would have bypassed the sensor. In general, along with the PPD42, the plastic body layouts of the PMS5003 and SPS30 are susceptible to inertial deposition losses due to their 90 degree elbows in particle stream pathways. However, the sub-optimal layouts of these sensors are better compromised by the more stable sample flow system (i.e. fan instead of convection).~~

~~The PMS5003, SDS011, and SPS30 sensors have digital outputs whereas the others are analogue based. Along with the sensor outputs shown in Table 1, the PMS5003 and SPS30 sensors also output particle number concentrations, but these signals were not used as the response comparison to the reference instrument was carried out using only mass concertation values. This decision was based on the observation that low-cost sensors have been predominantly used to measure mass concentration and not number concentration. Three units for each sensor model were evaluated in order to assess their inter-unit variation.~~

The mid-cost GRIMM 1.108, which was used here for demonstration purposes, is an optical aerosol spectrometer with 15 size bins (from 0.23 to 20 µm). Previous evaluations of the GRIMM 1.108 has shown its response to be similar to the APS

(Peters et al., 2006~~), ~~); ~~and~~ furthermore, its accuracy (mass of C-factor adjusted total suspended particles) ~~being~~ is comparable to ~~other~~ mass measurement methods~~,~~ such as the filter weighing method (Burkart et al., 2010).

2.2 Reference aerosol

2.2.1 Vibrating orifice aerosol generator and gradient elution pump

The aerosol sampled by the low-cost sensors was generated using ~~The operating principle of~~ a Vibrating Orifice Aerosol Generator 3450 (VOAG, TSI Inc., USA). The operating principle of the VOAG is based on the instability and break-up of a cylindrical liquid jet. Mechanical disturbances of a resonance frequency vibration disintegrate~~s~~ the cylindrical jet into uniform droplets, which are dispersed into an aerosol flow system with appropriate dilution air. Dispersed droplets evaporate before significant coagulation occurs, and form particles from the non-volatile solute dissolved in the volatile liquid. If the droplet liquid is non-volatile, the particle diameter and droplet diameter are equal. Otherwise, the produced particle size is calculable from the volumetric fraction of the non-volatile solute, as shown in Eqs. 1-2~~.~~: ~~This aerosol generation method and an apparatus (predecessor of the TSI's VOAG) was first introduced by Berglund and Liu in 1973.~~

$$D_d = \left(\frac{6Q}{\pi f}\right)^{1/3} \quad (1)$$

~~Where~~ where D_d is the generated droplet diameter, Q is the solution feed rate, and f is the disturbance frequency.

$$D_p = (C + I)^{1/3} * D_d \quad (2)$$

~~Where~~ where D_p is the diameter of the formed particle, C is the volumetric concentration of the non-volatile solute in the volatile liquid (typically 2-propanol or purified water), and I is the volumetric fraction of impurity in the volatile liquid.

According to Berglund and Liu (1973), ~~The~~ the output aerosol number concentration of the VOAG has a relative standard deviation of less than 3 %, and the formed particle size distribution is monodisperse having a geometric standard deviation (GSD) less than 1.014 ~~(Berglund and Liu, 1973)~~. These, and particularly the capability to produce highly monodisperse size distribution, are important features regarding sensor size-selectivity evaluation; while polydisperse aerosol can be used, for instance, to estimate ~~e.g.~~ response stability and linearity to varying concentration levels (Hapidin et al., 2019; Papapostolou et al., 2017; Sayahi et al., 2019a), the presence of multiple different sized particles prevents the distinction between sensor response and specific particle size. The ~~greatest~~ most significant deficiency of the VOAG (and the main limitation of this study) is that its smallest producible particle size is in practice limited by the impurity within the carrier liquid~~, and is in practice limited~~ to approximately 0.55 µm.

The novelty of the aerosol generation method used in this research is based on the observation that the ~~blending of two liquid solutions with different non-volatile concentrations produces a stable particle size gradient, respective of the concentrations of the blending solutions. In other words, the produced~~ particle size of the monodisperse and constant number concentration reference aerosol can be controlled by feeding solutions ~~of~~ with different non-volatile concentrations to the VOAG, one after ~~each~~ the other. Such aerosol generation technique was first utilized by Kuula et al. (2017), who accomplished the solution blending with a supplementary syringe pump and a manually operated ~~3~~three-way valve. In this study, however, the solution feeding was done with a gradient elution pump typically used ~~typically~~ in ion-chromatography (GP50, Dionex Inc., USA). The GP50 gradient pump has four different eluent channels~~,~~ and is capable of dispensing liquids with high pressure (max. 5000 psi) and accurate volume flow rate (0.04 – 10.0 mL min-1 in increments of 0.01 mL min-1). The four eluent channels can be mixed with a resolution of 0.1 % (combined output of the four channels always 100 ~~%), %~~); ~~and, moreover~~furthermore, the GP50 has a user-interface ~~which~~ that enables the operator to generate parameterized eluent dispensing programs. In essence, the utilization of the GP50 allows the user to freely choose and produce monodisperse aerosols of desired particle sizes without the tuning of VOAG running parameters or manual alternation of the liquid concentrations. ~~Furthermore~~Additionally, the preconfigured dispensing programs are fully automated making the comparison of consecutive test runs more reliable.

2.2.2 Sampling configuration

A schematic figure of the used test setup is shown in Figure 1. Reference aerosol was generated using the VOAG-GP50-system as described in the previous section. Dioctyl sebacate (DOS, density of 0.914 g cm-3) was used as a non-volatile solute in a 2-propanol solvent (> 99.999 %, Sigma-Aldrich), and the formed particles were ~~thereby~~ transparent oil droplets. Although the reference instrument APS is known for having decreased counting efficiency for liquid droplets over ~5 µm in size (Volckens and Peters, 2005), no additional corrections were used. Running parameters of the VOAG and GP50 are shown in Supplemental Table S1. The three different DOS concentrations (A–C) refer~~s~~ to the four different eluent channels of the GP50 (the use of three channels was sufficient for this study).

The GP50 was operated in a ~~Method~~method ~~-~~mode, meaning that an automated program was used to dispense the liquids. A program ~~(i.e. method)~~ ~~constitutes~~ involves a number of consecutive time steps in which the blending ratios of eluent channels, step durations, and volumetric flow rate of the liquid can be defined separately ~~for each time step~~. Executing the program means that the GP50 dispenses the liquids according to the settings determined in each step. The program used in this evaluation consisted of 10 steps in which the produced particle sizes were logarithmically distributed from 0.45 to 9.78 µm. The calculated blending ratios and the respective particle sizes are shown in Supplemental Table S2. Step duration of 5 minutes was used~~;~~ ~~and thus~~ a single test run thus lasted approximately 60 minutes. Dead volumes in the GP50 and VOAG slightly extend the theoretical run time duration. A complete test run can be performed as quickly as in 15 minutes, ~~but this~~which though results in fewer measurement points and weaker statistical power. An example of the produced reference

aerosol number size distribution measured with the APS is shown in Figure 2. It is worth underlining that the number of steps used in the GP50 dispensing program does not dictate the number of different particle sizes produced. The number of steps and the parameters assigned to them simply define the minimum (blending ratio of the first step) and maximum (blending ratio of the last step) particle size and the rate (step duration) at which the particle size gradient evolves from the minimum size to maximum size. The word "gradient" is used to note that a step from 2 to 3 µm, for instance, does not lead to a discontinuous and sudden step from one particle size to another.

Formed particles were neutralized in the dispersion outlet of the VOAG, and further fed into a flow splitting section where the reference aerosol was symmetrically directed ~~symmetrically~~ to both the reference instrument (Aerodynamic Particle Sizer 3321, TSI Inc., USA) and sensor. The sensors were encapsulated in ~~3d~~ 3D-printed air-tight enclosures with an external pump connected to it in order to ensure appropriate sample flow through the sensor. The ~~symmetrical~~ sample flow rate was set to be 1 L min-1 — ~~as this was~~ the aerosol flow rate of the APS (sheath flow of the APS taken from the laboratory air). Although there is no clear theoretical basis as to why a different flow rate would affect the way the sensor discriminates different particle sizes (apart from different particle size-specific sampling losses), additional tests were conducted with flow rates of 0.5 and 2 L min-1 to ensure that this was indeed the case (see Supplemental Figure S2). For the PMS5003 and SPS30 sensors, an exhaust deflector was used to prevent unwanted sample mixing resulting from the fan outlet, which for these sensors, was situated right next to the sensor inlet. An illustration of the PMS5003 sampling arrangement is shown in Figure 3. A schematic figure of all the sampling arrangements is shown in Supplemental Figure ~~S2~~S3.

All sensor units were in the original condition except for the PPD42 and B5W sensors which had their air heating resistors removed. The evaluation platform used in this study did not require independent means of sample flow. Furthermore, holes were drilled to the plastic body of the PPD42 to ensure that the sample aerosol could reach to the optical detection volume. The inlet of the PPD42 was originally designed to be on top of the plastic body (facing towards the electronic circuit board); therefore, when the electronic circuit board of the sensor was oriented in parallel with the sample stream, the majority of the particles would have bypassed the sensor. In general, along with the PPD42, the plastic body layouts of the PMS5003 and SPS30 are susceptible to inertial deposition losses due to their 90-degree elbows in particle stream pathways. However, the more stable sample flow system (i.e., fan instead of convection) might help compensate for the sub-optimal layouts of these sensors.

2.3 Data processing

The output signal of the evaluated sensor and APS was measured synchronously using a 10-~~-~~second time resolution and moving average. Any raw measurement point which had GSD (calculated from the APS data) exceeding 1.2 was disregarded (~ 2.1 % of the data), but typically the GSD values ranged between ~~were within~~ 1.04 —and 1.08 ~~range~~. The sensor bias was set to zero by sampling clean air for 10 minutes (60 data points) and then subtracting the clean air response from the test

aerosol response. The bias correction was only relevant for the GP2Y1010AU0F~~f~~ and B5W sensors. In order to prevent arbitrary unit comparisons, the sensor response was normalized using Eq. 3:

$$Normalized\ detection\ efficiency = \frac{\frac{Sensor_i}{APS_i}}{\max\left(\frac{Sensor}{APS}\right)} \quad (3)$$

~~Where~~ where i is the ith measurement point, $Sensor$ is the sensor signal, and APS is the APS total mass concentration. The maximum Sensor/APS ratio refers to the maximum ratio measured during a single test run.

The normalized 10-~~-~~second resolution data was divided into 30 logarithmically distributed size bins (from 0.45 to 9.73 µm) according to the count median diameters (CMD, aerodynamic) measured by the APS. An average sensor response as a function of average CMD was then calculated for each size bin. ~~The decision to divide the data into 30 bins was based on the clarity of the produced figures and statistically sufficient number of measurement points belonging to each bin.~~ This process was completed for three different sensor units, and a combined (average and standard deviation) sensor response was calculated. Valid detection ranges, which were defined as the upper half of the detection efficiency curve, of the sensors were linearly interpolated from the average response functions. A detailed example of how the data was processed and how the valid detection ranges were calculated is shown in the supplementary material. The cursory evaluation of the GRIMM was conducted using the same data processing method. The size bins of PMS5003, SPS30, SDS011, and B5W were discretized so that no overlapping signals were obtained. For example, the outputs of the SDS011 were used as PM2.5 and PM10-2.5 (PM10-2.5 calculated as PM10 - PM2.5) instead of PM2.5 and PM10.

The PMS5003, SDS011, and SPS30 sensors have digital outputs whereas the others are analog-based. Along with the PM mass fractions listed in Table 1, the PMS5003 and SPS30 sensors also output particle number concentrations, but these signals were not used because the response comparison to the reference instrument was carried out using only mass concertation values. This decision was based on the observation that low-cost sensors have been predominantly used to measure mass concentration and not number concentration.

3 Results and discussion

3.1 GRIMM model 1.108

The ~~response curves~~normalized detection efficiencies of the 15 bin GRIMM 1.108 are shown in Figure 3. ~~For the sake of clarity, the degrees of measurement variation have been excluded from the figure. The previously described data processing technique was used, and the comparison to APS was conducted with mass concentration. Bins 14 and 15, which correspond to 10 – 15 and 15 – 20 µm, respectively, are not shown here as they did not produce any signal (as expected).~~ The

normalized detection efficiency of 70 – 90 % results from the average efficiency from multiple data points and, in this case, does not imply that the GRIMM would systematically underestimate particle mass concentrations. The same applies ~~for~~ to the respective sensor response figures (next section).

The response characteristics of the GRIMM are in line with its technical specifications showing that each size bin only corresponds to its specific detection range. A flat response curve would indicate that the strength of the output signal remains unchanged regardless of the particle size, ~~and thus~~which would show that the size bin is unable to make a distinction between different particle sizes. Some mismatch between the particle sizing of the APS and the GRIMM can be observed as a result of different particle sizing techniques (time of flight and optical), but this is trivial, considering the objective of this study. The purpose of this figure is to highlight how an aerosol measurement device with several particle sizing bins should respond to the evaluation method used in this study.

3.2 Low-cost sensors

Response functions of the evaluated sensors are shown in Figures 4a – f. ~~The coloured circles represent the calculated average responses of the three sensor units and the shaded background areas respective standard deviations. Standard deviations of the average CMDs were negligible due to the reliable and reproducible test method. Figure legends correspond to the bin size ranges stated by the corresponding manufacturer.~~

Plantower PMS5003

According to Figure 4a, it is apparent that the PMS5003 does not accurately distinguish between PM1, PM2.5, and PM10 size fractions ~~particle sizing of the different bins is not working properly~~. The first and the second bin (supposedly corresponding to 0.3 – 1.0 and 1.0 – 2.5 µm) are similar, ~~having~~ with valid detection ranges of approximately < 0.7 µm and < 0.8 µm, respectively (valid detection ranges were defined as the upper half of the detection range~~,~~; see the section "Data processing"). ~~It is possible that t~~The lower cut-points of these bins may reach~~es~~ close to 0.3 µm, as stated by the manufacturer~~,~~; however, ~~but~~ this could not be confirmed using the VOAG-GP50 system. As the larger standard deviations indicate, ~~The~~ the third bin is noisier ~~as indicated by the larger standard deviations,~~ and ~~is~~ significantly off of its stated detection range (2.5 – 10 µm).

Based on the test, the PMS5003 cannot be used to measure coarse mode particles (2.5 – 10 µm~~),~~); ~~and~~ furthermore, its ability to measure PM2.5 ~~is dependent~~depends on the stability of the ambient air size distribution~~;~~: for example, if the proportions of mass in ~~e.g.~~ < 0.8 and > 0.8 µm fractions change~~s~~ significantly, the PMS5003 is susceptible to inaccuracies~~.~~ ~~This is~~ because its valid detection range cannot account for changes occurring in parts of the size distribution that~~which~~ it ~~essentially cannot~~can hardly observe. However, if the ambient size distribution is stable, the PMS5003 can be adjusted to measure PM2.5 with reasonable accuracy (Bulot et al., 2019; Feenstra et al., 2019; Magi et al., 2019; Malings et al., 2019).

Similarly, the validity of PM10 measurements can only be ensured ~~only~~ when the proportion of mass in > 0.7 or > 0.8 µm size fraction is either constant or negligible with respect to the total PM10 mass. In reality, this is rarely the case, which poses ~~and therefore~~ a high risk of sensor misuse ~~is posed~~. This observation is in line with the findings ~~of~~ from previous studies (Laquai, 2017b~~; Levy Zamora et al., 2019~~; Li et al., 2019; Sayahi et al., 2019b) which show~~ed~~, for ~~example~~instance, that the PMS5003 could not detect a substantial dust storm episode while deployed in the field. The most accurate and reliable results are most likely achieved for the PM1 size fraction by using either bin 1 or bin 2 signals.

Nova SDS011

The ~~Response~~ response function of the SDS011 is shown in Figure 4b. Contrary to the PMS5003, the SDS011 exhibits two ~~more clear~~clearly different detection ranges~~;~~: the first bin (0.3–2.5 µm) corresponds approximately to < 0.8 µm, and the second bin (2.5–10 µm) corresponds approximately to 0.7–1.7 µm. Similarly to the PMS5003, the SDS011 is not suitable for the measurement of coarse mode particles, and the measurements of PM10 can be grossly inaccurate. ~~Previous studies have~~ as also noted ~~this~~ by ~~(~~Budde et al.~~,~~ (2018) and~~;~~ Laquai~~,~~ (2017a). However, due to the clearer difference between ~~the two~~ bin 1 and bin 2 detection ranges, the SDS011 has the potential to measure PM2.5 more accurately than the PMS5003. For example, by calculating the ratio of bins 1 and 2, it is possible to approximate the distribution of mass in the 0.3–2.5 µm size range, thus using an additional correction factor to obtain more accurate results.~~Moreover, an artificial correction factor approximating changes in the measured size distribution may be possible to calculate from the ratios of bin 1 and bin 2 signal strengths.~~ Previous studies have shown that the SDS011 can be reasonably accurate in the measurements of PM2.5 (Badura et al., 2018; Liu et al., 2019).

Sensirion SPS30

The ~~Response~~ response function of the SPS30 is shown in Figure 4c. The valid detection range of the first bin (0.3–1.0 µm) is approximately < 0.9 µm. The second, third, and fourth bin (supposedly corresponding to 1.0–2.5, 2.5–4.0, and 4.0–10 µm) are nearly identical, with ~~having~~ valid detection ranges of approximately 0.7–1.3 µm. The identical detection ranges indicate~~s~~ that these bins may have been factory calibrated using the same test aerosol. The SPS30 is a relatively new sensor (introduced to the markets in late 2018), and neither Web of Science nor Scopus ~~literature research~~ showed any existing studies as of September 2019. However, the South Coast Air Quality Management District (SCAQMD) has conducted a preliminary field test where three SPS30 units were compared to three different federal equivalent method (FEM) monitors (SCAQMD, 2019). The results of this test showed that the SPS30 sensors had very low cross-unit variability (~1, 1.3, and 2.4 % for PM1, PM2.5, and PM10, respectively) and, more importantly, the ~~accuracies~~ coefficient of determinations for the measurement of PM1, PM2.5 and PM10 decreased from R2 ~ 0.91 to 0.83 and further down to 0.12, respectively. These observations ~~are in high agreement~~strongly align with the results of this study~~,~~; ~~and~~ furthermore, they illustrate how a sensor with limited operational range may exhibit a near regulatory grade ~~accuracy~~ performance if the measured size fraction is in alignment with the valid detection range of the sensor (< 0.9 µm and PM1). On the other hand,

the severity of data misinterpretation is apparent when the sensor measurement is extended to cover particle sizes ~~which~~ that it cannot observe.

Sharp GP2Y1010AU0F

The ~~Response~~ response function of the GP2Y1010AU0F is shown in Figure 4d, and its valid detection range appears to be approximately ~~1.8 – 6.5~~< 0.8 µm. Like the previously discussed sensors, the GP2Y1010AU0F can be used to measure small particles (e.g., PM1) but not coarse mode particles. ~~This is problematic considering the typically monitored PM2.5 and PM10 parameters as the sensor reacts to particulate mass belonging to both < 2.5 and > 2.5 µm size fractions. Consequently, the measurement output can be particularly difficult to interpret if this detail is not known. Nevertheless, the GP2Y1010AU0F has been used in variety of different applications (Alvarado et al., 2015; Zuidema et al., 2019).~~

Several laboratory evaluations have been previously conducted ~~previously~~ for the GP2Y1010AU0F, but none of these have assessed its detection range using monodisperse test aerosols (Li and Biswas, 2017; Manikonda et al., 2016; Sousan et al., 2016). ~~A study of~~ Wang et al. (2015) used atomized polystyrene latex (PSL) particles to evaluate the effect of particle size ~~to~~ on the GP2Y1010AU0F response, but no concluding remarks can be obtained from these results ~~(Wang et al., 2015)~~. The study method utilized only three different sized PSLs~~, ; and~~ moreover, it was not designed to investigate the complete detection range of the GP2Y1010AU0F ~~to begin with~~. However, ~~some of~~ according to the authors, the results implied that the sensor was more sensitive to 300 nm particles than to 600 and 900 nm particles, which is in slight disagreement with the results of this study whereby the normalized detection efficiency curve shows the highest sensitivity peak for 0.6 µm sized particles as well as a decreasing trend for particles smaller than this. There is no obvious explanation for this discrepancy, but it is worth re-emphasizing the differences in the used evaluation approaches.~~could detect particles as small as 0.3 µm, which is in significant conflict with the results of this study. There is no obvious explanation for this.~~

Shinyei PPD42

Response functions of the three PPD42 sensor units are shown in Figure 4e. Contrary to the other sensors, a combined response function was not calculated as the three units exhibited significantly different response characteristics. The circles and shaded background areas ~~in this case~~ represent average responses and respective standard deviations of the individual sensor units (calculated from the ~ 300 raw data points). The valid detection range of the first unit is 1.0 – 2.1 µm, and it is likely to be best suited for PM2.5 measurements. However, the low detection efficiency of < 1.0 µm sized particles may considerably hinder its accuracy ~~considerably~~. Valid detection ranges of the second and third units are > 5.9 and 1.5 – 4.9 µm indicating preferable applicability to coarse mode particle measurements. Previous laboratory evaluations have noted that the PPD42 output is a function of particle size but could not provide a more detailed analysis of the complete detection range (Austin et al., 2015; Wang et al., 2015). A study of Kuula et al. (2017) reported a valid detection range of approximately 2.5 – 4.0 µm, which is in the same range as the third unit of this study ~~(Kuula et al., 2017)~~.

Due to the apparent inter-unit inconsistency in valid detection ranges, it is evident that the response characteristics of the PPD42 have to be quantified case-by-case before reliable measurements can be achieved. Accordingly, the inconsistent response characteristics may also contribute to the ~~notion~~ fact that previous field evaluation studies have achieved varying results regarding the ~~accuracy~~ performance of PPD42; ~~studies of~~ Bai et al. (2019) and Holstius et al. (2014) reported R2 values of 0.75 and 0.55 – 0.60, respectively, for the measurement of PM2.5, whereas ~~studies of~~ N. E. Johnson et al. (2018) and K. K. Johnson et al. (2018) reported more modest values of 0.36 – 0.51 and 0 – 0.28, respectively (Bai et al., 2019; Holstius et al., 2014; Johnson et al., 2018a, 2018b). On the other hand, ~~studies of~~ Kuula et al. (2017, 2018) showed that higher levels of accuracy can be achieved if the measured size fraction is targeted to correspond to the characteristic response function of the PPD42 (R2 = 0.96 and R2 = 0.87, respectively) ~~(Kuula et al., 2017, 2018)~~.

Omron B5W

The ~~Response~~ response function of the B5W is shown in Figure 4f. The two size bins exhibit two ~~clearly~~ different detection ranges (0.6 – 1.0 and > 3.2 µm, respectively) ~~which~~ that are reasonably close to the ones declared by the manufacturer (0.5 – 2.5 and > 2.5 µm, respectively). In fact, out of all sensors, the B5W appears to be the most promising~~nent~~ sensor for the ambient monitoring of PM2.5 and PM10-2.5 size fractions. In comparison to ~~e.g.~~ SDS011 and SPS30, for instance, the usability of the B5W may be hindered by its temperature gradient~~-~~based sampling method, which ~~as it~~ is not as reliable as the respective fan~~-~~based method. Nonetheless, it is the only sensor capable of measuring both fine and coarse fraction particles. Neither Web of Science nor Scopus ~~literature review~~ showed existing studies for the Omron B5W.

Conclusions

According to the results obtained in this study, ~~optical~~ low-cost optical sensors exhibit widely varying response characteristics regarding their size-selectivity (from < 0.7 to > 5.9 µm, Table 2). However, none of the sensors ~~had~~ have ~~exactly~~ precisely the same response characteristics ~~as~~ stated by their manufacturers~~.~~, which ~~This~~ provides ~~some insight and~~ evidence to the fact~~notion~~ that particle~~-~~size ~~-~~selectivity may ~~have~~ play an essential role in the analysis of the sources of errors in sensors~~error source analysis of sensors,~~ and ~~furthermore,~~ underlines that scientists, as well as manufacturers ~~for that matter~~, need to acknowledge the limitations related to this~~.~~: ~~Respectively, it is worth noting that~~ attempts to artificially extend the operational range of sensors beyond their practical capabilities using complex statistical models ~~can~~ may be unreasonable and ~~may~~ lead to misleading conclusions. Empirical corrections for known ~~artefacts~~artifacts, such as ~~the~~ humidity, can be justifiable~~,~~; however, ~~in general~~ sensor data and advanced modelling techniques should be merged cautiously in order to retain ~~pertain~~ both the validity and representativeness of the data.

A ~~Cursory~~ cursory comparison to a mid-cost aerosol spectrometer (GRIMM 1.108) shows that low-cost sensor development is still considerably behind its more expensive alternative~~;~~: while the GRIMM 1.108 could sufficiently characterize particle sizes with up to 15 different size bins, the low-cost sensors could only achieve independent responses for ~~1-2~~one or two ~~to~~ bins~~.~~, ~~This~~ which is a ~~major~~ significant weakness, considering that the ability to ~~correctly~~ measure particle size correctly is at the foundation of accurate mass measurement (mass α dp3). The ~~Development~~ development of low-cost sensors should focus on increasing the number of size bins, and more importantly, making sure that each size bin is calibrated correctly. Improperly configured bin sizing poses a significant risk of data misinterpretation~~,~~ and will inevitably lead to inaccurate measurements. A ~~Low~~ low number of size bins limits the valid operational range of sensors~~,~~; however, it is unclear how the ~~amount~~ number of advanced measurement features and low unit cost should be reconciled.

The VOAG-GP50 aerosol generation system described in this study~~,~~ introduced a novel approach to the quick and efficient evaluation ~~in how~~of aerosol measurement devices ~~can be evaluated quickly and efficiently~~. The use of a GP50 gradient pump eliminates much of the manual labo~~u~~r that ~~which was~~ previously was inseparable part of the VOAG operation, ~~and~~ thus ~~makes~~ making the generation of reference aerosols more consistent and reliable. Its automated dispensing programs allow~~s~~ for highly repeatable testing~~,~~; ~~and~~ furthermore, the four different eluent channels enable~~s~~ the operator to pick and choose the desired particle size to be produced freely ~~pick and choose desired particle size to be produced~~. Along with saving manual labo~~u~~r and time, this is also a cost-saving feature as traditionally used polystyrene latex (PSL) particle~~s~~ are not needed. Considering these matters, the VOAG-GP50 system can potentially be scaled to an industrial level operation, which is an intriguing feature when considering the mass deployment of sensors and their respective quality assurance and control.

Author contribution

JK and TM designed the experimental setup, and JK carried out the tests. KT had an important role in refurbishing the gradient elution pump. SM and OG provided some of the sensors. JK was responsible for the data analysis, although all co-authors provided valuable feedback, particularly TM. JK wrote the manuscript with the help of all co-authors.

Conflict of interest

The authors declare no competing financial interest.

Acknowledgements

This study was funded by the Urban innovative actions initiative of the European Regional Development Fund (project HOPE; Healthy Outdoor Premises for Everyone, project no: UIA03-240), and by the European Union's Horizon 2020 research and innovation program under grant agreement No 689954 (project iSCAPE; Improving the Smart Control of Air Pollution in Europe).

References

Alvarado, M., Gonzalez, F., Fletcher, A. and Doshi, A.: Towards the development of a low cost airborne sensing system to monitor dust particles after blasting at open pit mine sites, Sensors, 15(8), 19703–19723, doi:10.3390/s150819667, 2015.

Austin, E., Novosselov, I., Seto, E. and Yost, M. G.: Laboratory evaluation of the Shinyei PPD42NS low-cost particulate matter sensor, PLoS One, 10(9), 1–17, doi:10.1371/journal.pone.0137789, 2015.

Badura, M., Batog, P., Drzeniecka-Osiadacz, A. and Modzel, P.: Evaluation of low-cost sensors for ambient PM2.5 monitoring, J. Sensors, 2018, doi:10.1155/2018/5096540, 2018.

Bai, L., Huang, L., Wang, Z., Ying, Q., Zheng, J., Shi, X. and Hu, J.: Long-term field Evaluation of Low-cost Particulate Matter Sensors in Nanjing, Aerosol Air Qual. Res., 1–12, doi:10.4209/aaqr.2018.11.0424, 2019.

Berglund, R. N. and Liu, B. Y. H.: Generation of Monodisperse Aerosol Standards, Environ. Sci. Technol., 7(2), 147–153, doi:10.1021/es60074a001, 1973.

Budde, M., D. Schwarz, A., Müller, T., Laquai, B., Streibl, N., Schindler, G., Köpke, M., Riedel, T., Dittler, A. and Beigl, M.: Potential and Limitations of the Low-Cost SDS011 Particle Sensor for Monitoring Urban Air Quality, ProScience, 5(3rd International Conference on Atmospheric Dust (DUST2018)), 6–12, doi:10.14644/dust.2018.002, 2018.

Bulot, F. M. J. J., Johnston, S. J., Basford, P. J., Easton, N. H. C. C., Apetroaie-Cristea, M., Foster, G. L., Morris, A. K. R. R., Cox, S. J. and Loxham, M.: Long-term field comparison of multiple low-cost particulate matter sensors in an outdoor urban environment, Sci. Rep., 9(1), 7497, doi:10.1038/s41598-019-43716-3, 2019.

Burkart, J., Steiner, G., Reischl, G., Moshammer, H., Neuberger, M. and Hitzenberger, R.: Characterizing the performance of two optical particle counters (Grimm OPC1.108 and OPC1.109) under urban aerosol conditions, J. Aerosol Sci., 41(10), 953–962, doi:10.1016/j.jaerosci.2010.07.007, 2010.

Caubel, J. J., Cados, T. E., Preble, C. V. and Kirchstetter, T. W.: A Distributed Network of 100 Black Carbon Sensors for 100 Days of Air Quality Monitoring in West Oakland, California, Environ. Sci. Technol., 53, 7564–7573, doi:10.1021/acs.est.9b00282, 2019.

Feenstra, B., Papapostolou, V., Hasheminassab, S., Zhang, H., Boghossian, B. Der, Cocker, D. and Polidori, A.: Performance evaluation of twelve low-cost PM2.5 sensors at an ambient air monitoring site, Atmos. Environ., 216(February), 116946, doi:10.1016/j.atmosenv.2019.116946, 2019.

Feinberg, S. N., Williams, R., Hagler, G., Low, J., Smith, L., Brown, R., Garver, D., Davis, M., Morton, M., Schaefer, J., Campbell, J.: Examining spatiotemporal variability of urban particulate matter and application of high-time resolution data from a network of low-cost air pollution sensors, Atmos. Environ., 213(December), 579-584, doi: 10.1016/j.atmosenv.2019.06.026, 2019.

Gao, M., Cao, J. and Seto, E.: A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China, Environ. Pollut., 199, 56–65, doi:10.1016/j.envpol.2015.01.013, 2015.

GRIMM model 1.108 aerosol spectrometer user manual, [online]. Available from: https://wmo-gaw-wcc-aerosol-physics.org/files/opc-grimm-model–1.108-and-1.109.pdf. (Accessed February 11, 2020). GRIMM Aerosol Technik GmbH., Germany.

Hapidin, D. A., Saputra, C., Maulana, D. S., Munir, M. M. and Khairurrijal, K.: Aerosol chamber characterization for commercial particulate matter (PM) sensor evaluation, Aerosol Air Qual. Res., 19(1), 181–194, doi:10.4209/aaqr.2017.12.0611, 2019.

Holstius, D. M., Pillarisetti, A., Smith, K. R. and Seto, E.: Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California, Atmos. Meas. Tech., 7(4), 1121–1131, doi:10.5194/amt-7-1121-2014, 2014.

Jayaratne, R., Liu, X., Thai, P., Dunbabin, M. and Morawska, L.: The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog, Atmos. Meas. Tech., 11(8), 4883–4890, doi:10.5194/amt-11-4883-2018, 2018.

Jiao, W., Hagler, G., Williams, R., Sharpe, R., Brown, R., Garver, D., Judge, R., Caudill, M., Rickard, J., Davis, M., Weinstock, L., Zimmer-Dauphinee, S. and Buckley, K.: Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States, Atmos. Meas. Tech., 9(11), 5281–5292, doi:10.5194/amt-9-5281-2016, 2016.

Johnson, K. K., Bergin, M. H., Russell, A. G. and Hagler, G. S. W.: Field test of several low-cost particulate matter sensors in high and low concentration urban environments, Aerosol Air Qual. Res., 18(3), 565–578, doi:10.4209/aaqr.2017.10.0418, 2018a.

Johnson, N. E., Bonczak, B. and Kontokosta, C. E.: Using a gradient boosting model to improve the performance of low-cost aerosol monitors in a dense, heterogeneous urban environment, Atmos. Environ., 184(August 2017), 9–16, doi:10.1016/j.atmosenv.2018.04.019, 2018b.

Karagulian, F., Barbiere, M., Kotsev, A., Spinelle, L., Gerboles, M., Lagler, F., Redon, N., Crunaire, S. and Borowiak, A.: Review of the Performance of Low-Cost Sensors for Air Quality Monitoring, Atmosphere (Basel)., 10(9), 506, doi:10.3390/atmos10090506, 2019.

Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L. and Britter, R.: The rise of low-cost sensing for managing air pollution in cities, Environ. Int., 75, 199–205, doi:10.1016/j.envint.2014.11.019, 2015.

Kuula, J., Mäkelä, T., Hillamo, R. and Timonen, H.: Response characterization of an inexpensive aerosol sensor, Sensors, 17(12), doi:10.3390/s17122915, 2017.

Kuula, J., Kuuluvainen, H., Rönkkö, T., Niemi, J. V., Saukko, E., Portin, H., Aurela, M., Saarikoski, S., Rostedt, A., Hillamo, R. and Timonen, H.: Applicability of Optical and Diffusion Charging-Based Particulate Matter Sensors to Urban Air Quality Measurements, Aerosol Air Qual. Res., 122, 1–16, doi:10.4209/aaqr.2018.04.0143, 2018.

Laquai, B.: Impact of Particle Mass Distribution on the Measurement Accuracy of Low-Cost PM-Sensors, , (October), doi:10.1093/brain/123.6.1161, 2017a.

Laquai, B.: Particle Distribution Dependent Inaccuracy of the Plantower PMS5003 low- cost PM-sensor Particle Distribution Dependent Inaccuracy of the Plantower PMS5003 low- cost PM-sensor, , (October) [online] Available from: https://www.researchgate.net/publication/320555036_Particle_Distribution_Dependent_Inaccuracy_of_the_Plantower_PMS5003_low-cost_PM-sensor (Accessed 10 September 2019b), 2017.

Levy Zamora, M., Xiong, F., Gentner, D., Kerkez, B., Kohrman-Glaser, J. and Koehler, K.: Field and Laboratory Evaluations of the Low-Cost Plantower Particulate Matter Sensor, Environ. Sci. Technol., 53(2), 838–849, doi:10.1021/acs.est.8b05174, 2019.

Li, J. and Biswas, P.: Optical characterization studies of a low-cost particle sensor, Aerosol Air Qual. Res., 17(7), 1691–1704, doi:10.4209/aaqr.2017.02.0085, 2017.

Li, J., Mattewal, S. K., Patel, S. and Biswas, P.: Evaluation of Nine Low-cost-sensor-based Particulate Matter Monitors, Aerosol Air Qual. Res., 1–17, doi:10.4209/aaqr.2018.12.0485, 2019.

Liu, H. Y., Schneider, P., Haugen, R. and Vogt, M.: Performance assessment of a low-cost PM 2.5 sensor for a near four-month period in Oslo, Norway, Atmosphere (Basel)., 10(2), 41, doi:10.3390/atmos10020041, 2019.

Magi, B. I., Cupini, C., Francis, J., Green, M. and Hauser, C.: Evaluation of PM2.5 measured in an urban setting using a low-cost optical particle counter and a Federal Equivalent Method Beta Attenuation Monitor, Aerosol Sci. Technol., 0(0), 1–13, doi:10.1080/02786826.2019.1619915, 2019.

Malings, C., Tanzer, R., Hauryliuk, A., Saha, P. K., Robinson, A. L., Presto, A. A. and Subramanian, R.: Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation, Aerosol Sci. Technol., 0(0), 1–15, doi:10.1080/02786826.2019.1623863, 2019.

Manikonda, A., Zikova, N., Hopke, P. K. and Ferro, A. R.: Laboratory assessment of low-cost PM monitors, J. Aerosol Sci., 102, 29–40, 2016.

Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F., Christensen, B., Dunbabin, M., Gao, J., Hagler, G., Jayaratne, R., Kumar, P., Lau, A. K. H., Louie, P. K. K., Mazaheri, M., Ning, Z., Motta, N., Mullins, B., Rahman, M. M., Ristovski, Z., Shafiei, M., Tjondronegoro, D., Westerdahl, D. and Williams, R.: Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?, Environ. Int., 116, 286–299, doi:10.1016/j.envint.2018.04.018, 2018.

Papapostolou, V., Zhang, H., Feenstra, B. J. and Polidori, A.: Development of an environmental chamber for evaluating the performance of low-cost air quality sensors under controlled conditions, Atmos. Environ., 171, 82–90, doi:10.1016/j.atmosenv.2017.10.003, 2017.

Peters, T. M., Ott, D. and O'Shaughnessy, P. T.: Comparison of the Grimm 1.108 and 1.109 portable aerosol spectrometer to the TSI 3321 aerodynamic particle sizer for dry particles, Ann. Occup. Hyg., 50(8), 843–850, doi:10.1093/annhyg/mel067, 2006.

Popoola, O. A. M., Carruthers, D., Lad, C., Bright, V. B., Mead, M. I., Stettler, M. E. J., Saffell, J. R. and Jones, R. L.: Use of networks of low cost air quality sensors to quantify air quality in urban settings, Atmos. Environ., 194, 58–70, doi:10.1016/j.atmosenv.2018.09.030, 2018.

Rai, A. C., Kumar, P., Pilla, F., Skouloudis, A. N., Di Sabatino, S., Ratti, C., Yasar, A. and Rickerby, D.: End-user perspective of low-cost sensors for outdoor air pollution monitoring, Sci. Total Environ., 607–608(Supplement C), 691–705, doi:10.1016/j.scitotenv.2017.06.266, 2017.

Sayahi, ., Kaufman, D., Becnel, T., Kaur, K., Butterfield, A., Collingwoodd, S., Zhang, Y., Gaillardon, P.-E. and Kelly, K. E.: Development of a calibration chamber to evaluate the performance of low-cost particulate matter sensors, Environ. Pollut., 255, doi:10.1016/j.envpol.2019.113131, 2019a.

Sayahi, T., Butterfield, A. and Kelly, K. E.: Long-term field evaluation of the Plantower PMS low-cost particulate matter sensors, Environ. Pollut., 245, 932–940, doi:10.1016/j.envpol.2018.11.065, 2019b.

Schneider, P., Bartonova, A., Castell, N., Dauge, F. R., Gerboles, M., Hagler, G. S. W., Hüglin, C., Jones, R. L., Khan, S., Lewis, A. C., Mijling, B., Müller, M., Penza, M., Spinelle, L., Stacey, B., Vogt, M., Wesseling, J. and Williams, R. W.: Toward a Unified Terminology of Processing Levels for Low-Cost Air-Quality Sensors, Environ. Sci. Technol., acs.est.9b03950, doi:10.1021/acs.est.9b03950, 2019.

Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G., Shelow, D., Hindin, D. A., Kilaru, V. J. and Preuss, P. W.: The changing paradigm of air pollution monitoring, Environ. Sci. Technol., 47(20), 11369–11377, doi:10.1021/es4022602, 2013.

Sousan, S., Koehler, K., Thomas, G., Park, J. H., Hillman, M., Halterman, A. and Peters, T. M.: Inter-comparison of low-cost sensors for measuring the mass concentration of occupational aerosols, Aerosol Sci. Technol., 50(5), 462–473, doi:10.1080/02786826.2016.1162901, 2016.

South Coast Air Quality Management District (SCAQMD): Field Evaluation Sensirion SPS30 Evaluation Kit Background, [online] Available from: http://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/sensirion-sps30-evaluation-kit---field-evaluation.pdf (Accessed 5 November 2019), 2019.

Volckens J. and Peters T. M.: Counting and particle transmission efficiency of the aerodynamic particle sizer, J. Aerosol Sci., 36, 1400-1408, doi: 10.1016/j.jaerosci.2005.03.009, 2005.

Wang, Y., Li, J., Jing, H., Zhang, Q., Jiang, J. and Biswas, P.: Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement, Aerosol Sci. Technol., 49(11), 1063–1077, doi:10.1080/02786826.2015.1100710, 2015.

Yuval, Magen Molho, H., Zivan, O., Broday, D. M. and Raz, R.: Application of a sensor network of low cost optical particle counters for assessing the impact of quarry emissions on its vicinity, Atmos. Environ., 211, 29–37, doi:10.1016/j.atmosenv.2019.04.054, 2019.

Zheng, T., Bergin, M. H., Johnson, K. K., Tripathi, S. N., Shirodkar, S., Landis, M. S., Sutaria, R. and Carlson, D. E.: Field evaluation of low-cost particulate matter sensors in high-and low-concentration environments, Atmos. Meas. Tech., 11(8), 4823–4846, doi:10.5194/amt-11-4823-2018, 2018.

~~Zuidema, C., Sousan, S., Stebounova, L. V, Gray, A., Liu, X., Tatum, M., Stroh, O., Thomas, G., Peters, T. and Koehler, K.: Mapping Occupational Hazards with a Multi-sensor Network in a Heavy-Vehicle Manufacturing Facility, Ann. Work Expo. Heal., 63(3), 280–293, doi:10.1093/annweh/wxy111, 2019.~~

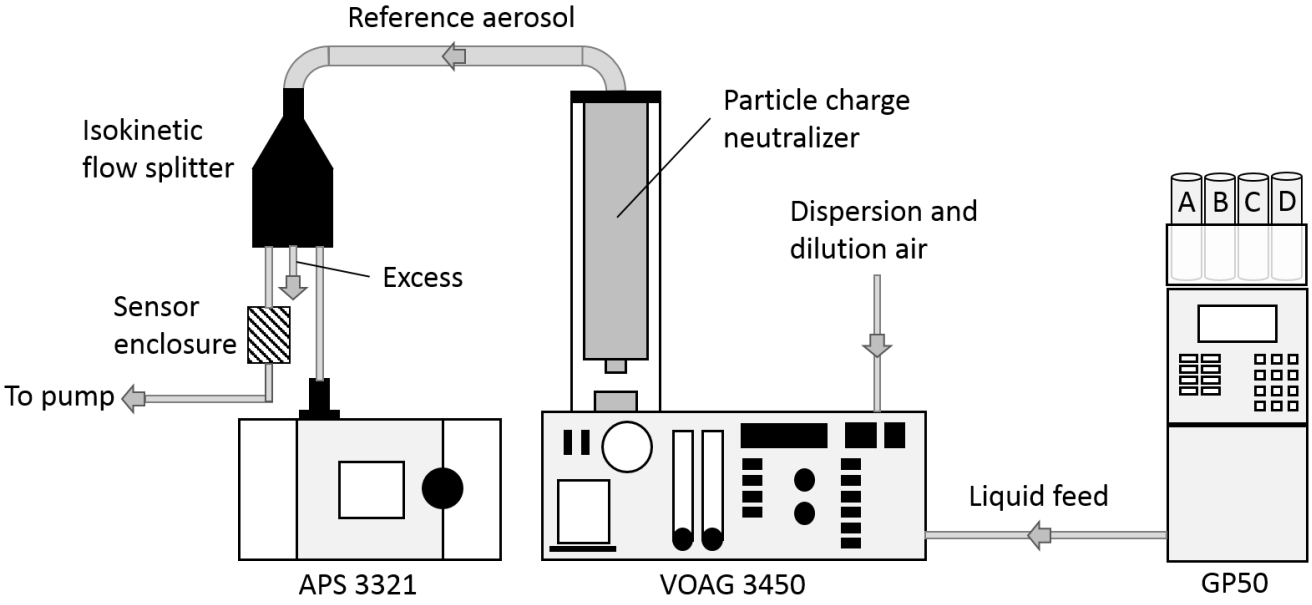
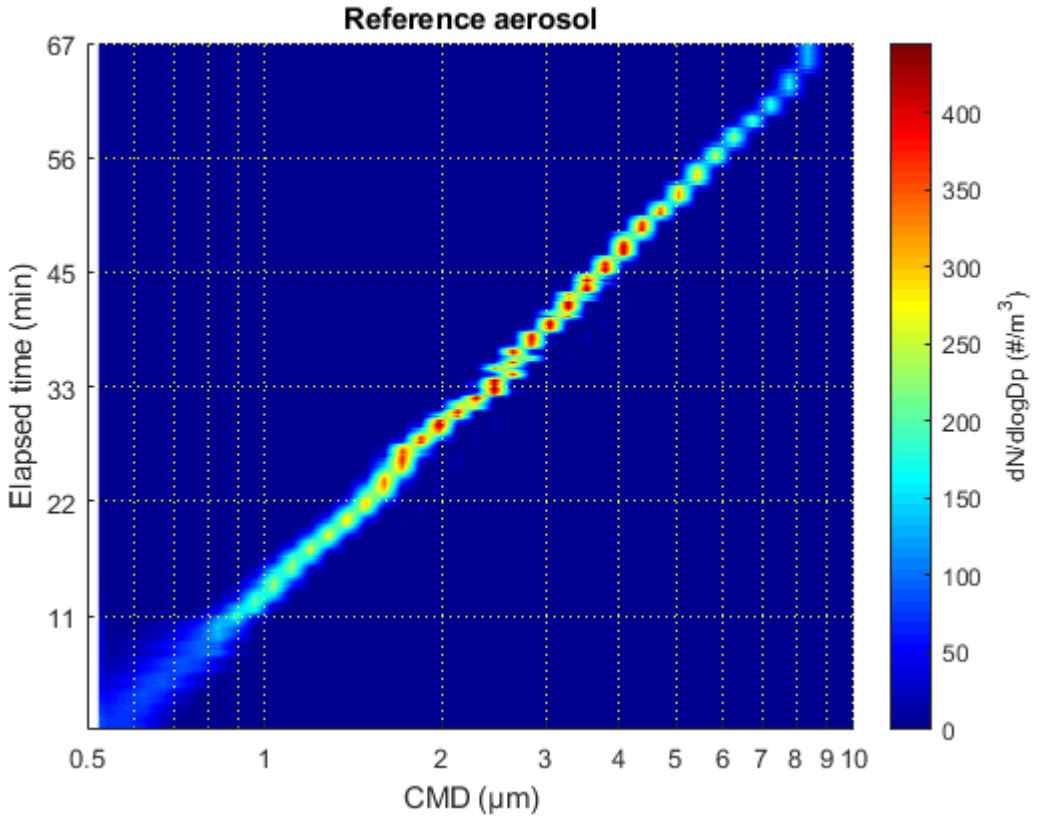


Figure 1: Schematic of the sensor evaluation setup. The GRIMM 1.108 drew its sample from where the sensor enclosure is now shown.


Reference aerosol
Elapsed time (min)
67
56
45
33
22
11
0.5
1
2
3
4
5
6
7
8
9
10
CMD (μm)
400
350
300
250
200
150
100
50
0
dN/dlogDp (#/m3)

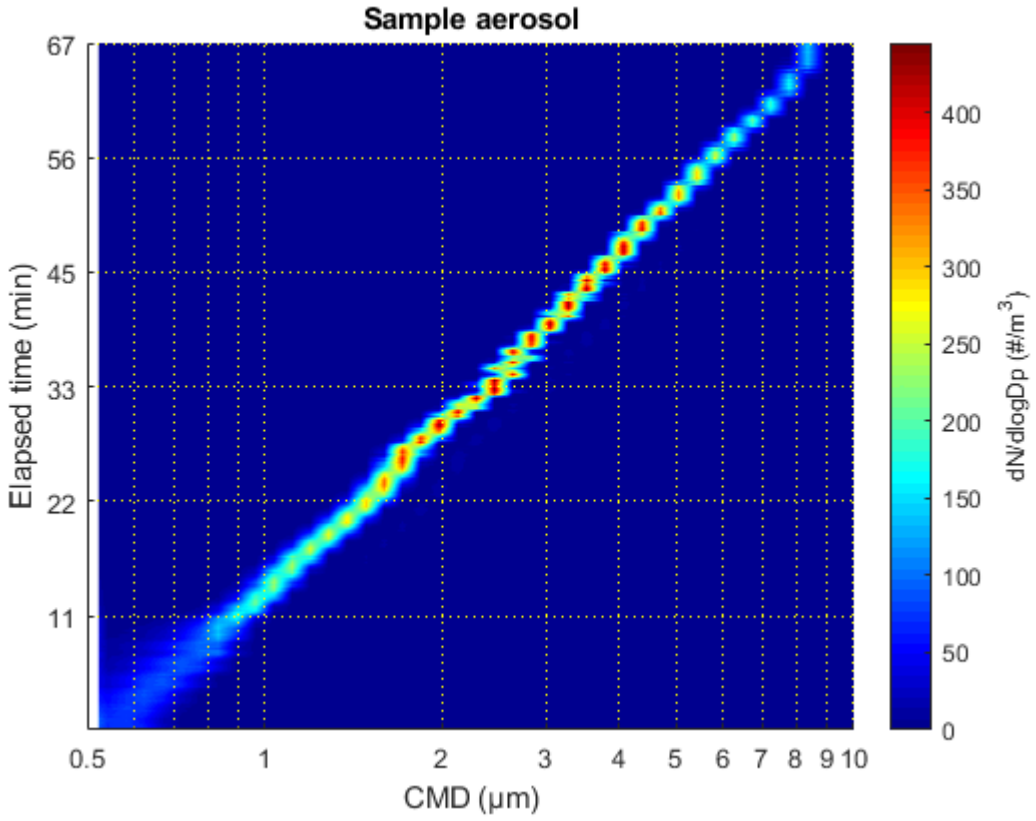


Figure 2. An example of ~~a~~ the produced reference aerosol. Decreasing number concentrations below 1 µm and above 5 µm result from ~~the~~ approaching the lower detection limit (0.5 µm) of the APS and ~~from the~~ increasing inertial deposition losses in the sampling lines, respectively. This had, however, no effect on the evaluation results as the sensor response was normalized against the concentration measured by the APS. Along with the lower detection limit of the APS, another limiting factor of the study was the smallest producible particle size, which was approximately 0.55 µm. The GSD of the size distribution ~~remains~~ remained below 1.2.

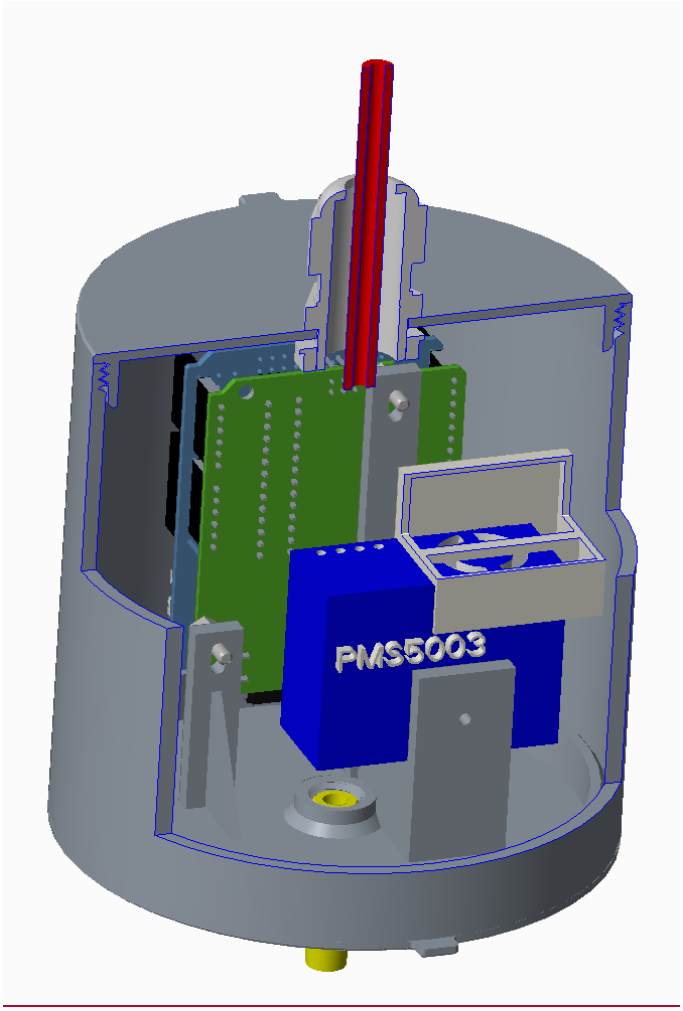


Figure 3. A cross-section view of the sampling arrangement of the PMS5003 sensor. The flow deflector is shown in white, the inlet pipe in red, and the exhaust connector of the housing in yellow.

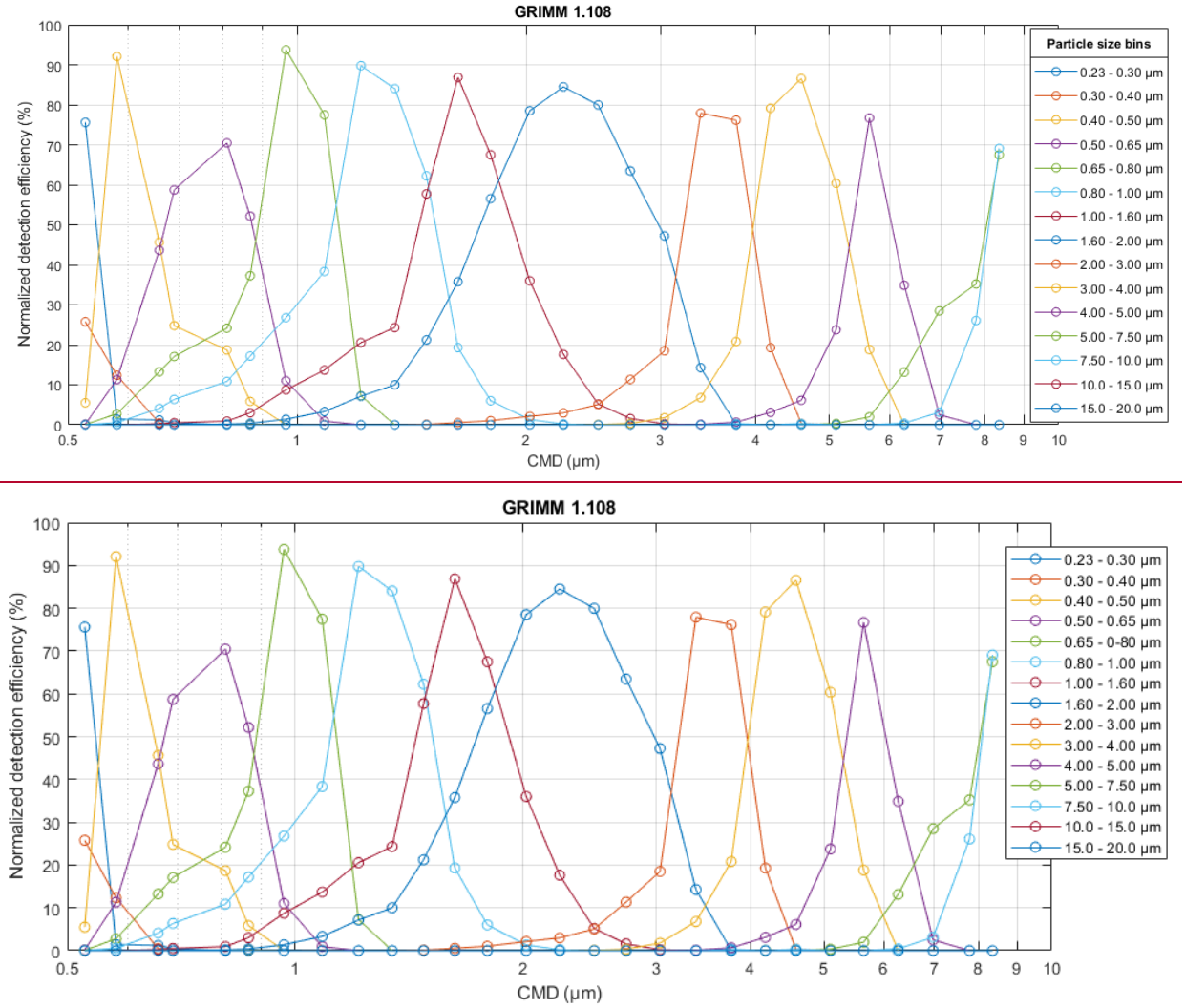
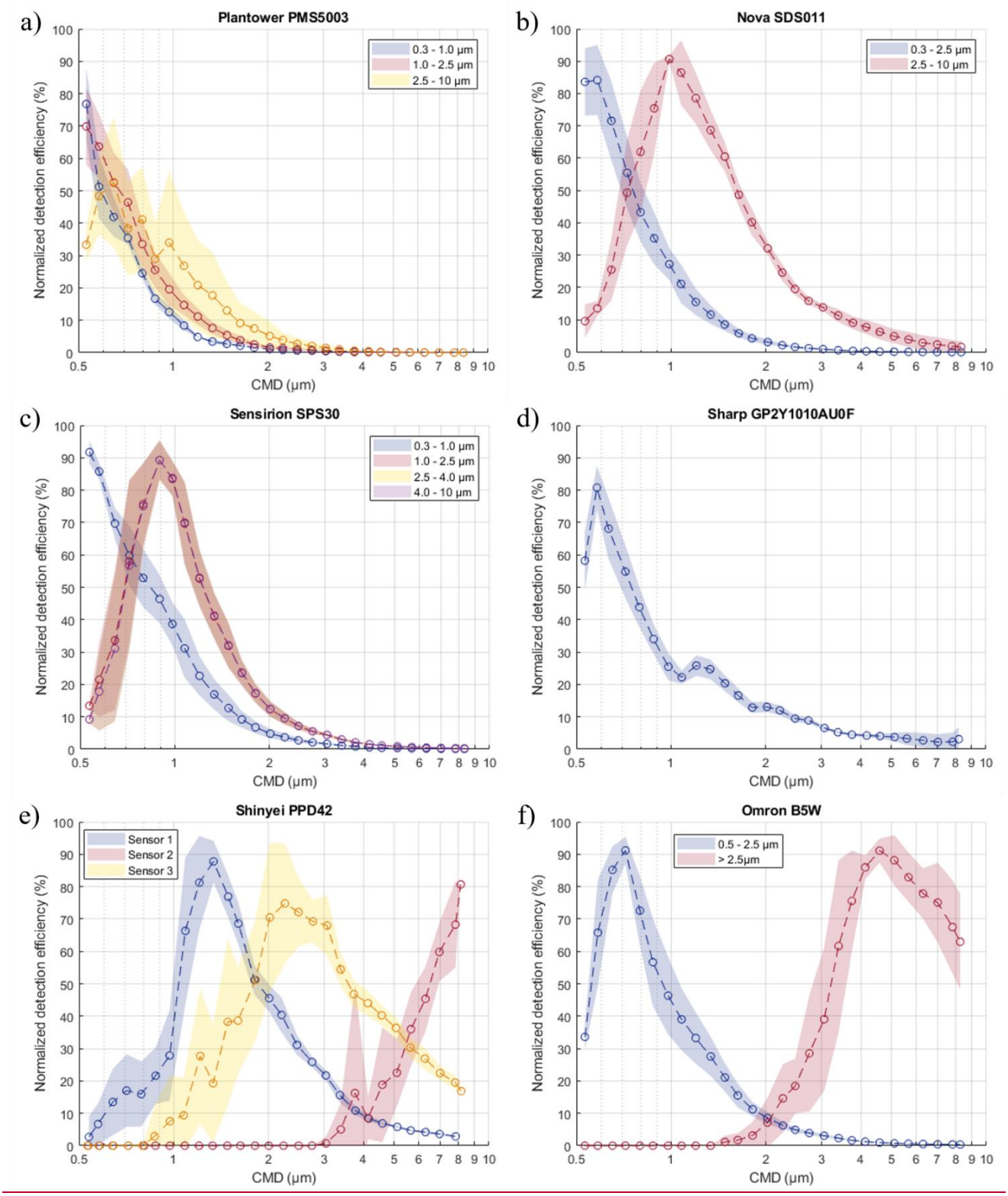


Figure ~~34~~. Normalized detection efficiency of the 15 particle size bins as a function of the count median diameter of the reference aerosol.~~Illustration of the response characteristics of the GRIMM 1.108 aerosol spectrometer.~~ Consecutively increasing and decreasing response curves indicate~~s~~ that the particle sizing of the instrument is functioning ~~properly~~correctly. For the sake of clarity, degrees of measurement variation have been excluded from the figure. Bins 14 and 15, which correspond to 10–15 and 15–20 µm, respectively, are not shown as they did not produce any response (as expected).


a)
Plantower PMS5003
0.3 - 1.0 µm
1.0 - 2.5 µm
2.5 - 10 µm
Normalized detection efficiency (%)
CMD (µm)
b)
Nova SDS011
0.3 - 2.5 µm
2.5 - 10 µm
Normalized detection efficiency (%)
CMD (µm)
c)
Sensirion SPS30
0.3 - 1.0 µm
1.0 - 2.5 µm
2.5 - 4.0 µm
4.0 - 10 µm
Normalized detection efficiency (%)
CMD (µm)
d)
Sharp GP2Y1010AU0F
Normalized detection efficiency (%)
CMD (µm)
e)
Shinyei PPD42
Sensor 1
Sensor 2
Sensor 3
Normalized detection efficiency (%)
CMD (µm)
f)
Omron B5W
0.5 - 2.5 µm
> 2.5µm
Normalized detection efficiency (%)
CMD (µm)

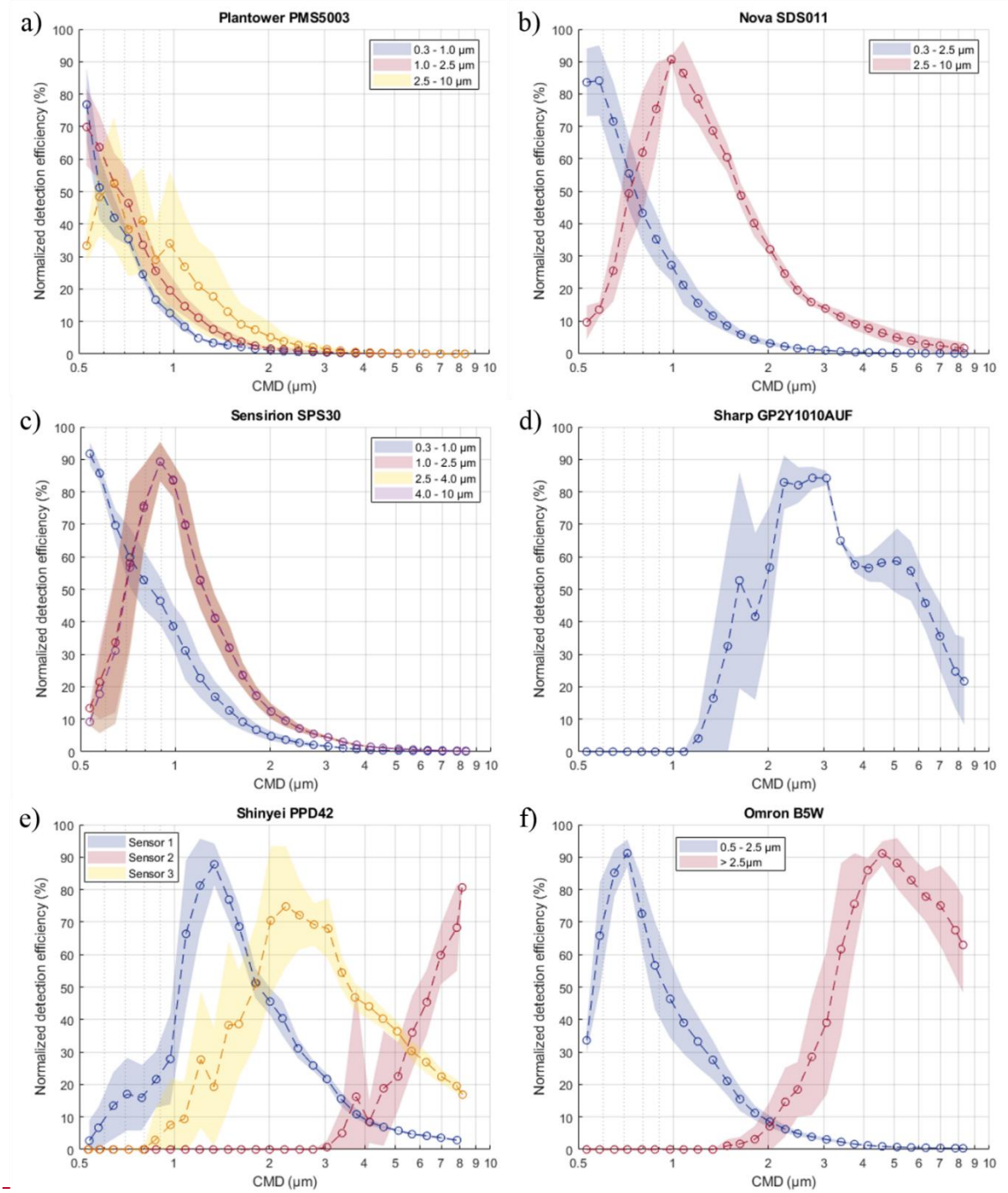


Figure 4~~5~~. ~~Measured particle size response functions of the low-cost sensors.~~Normalized detection efficiency of discretized PM mass fractions reported by the low-cost sensors as a function of the count median diameter of the reference aerosol. The colored circles represent the calculated average responses of the three sensor units, and the shaded background areas represent the

respective standard deviations. Standard deviations of the average CMDs were negligible due to the reliable and reproducible test method. Figure legends correspond to the bin size ranges stated by the corresponding manufacturer.

560 **Table 1. Basic features of the evaluated sensors declared by the manufacturers.**

Low-cost sensor	Detectable size range (µm)	Number of mass fractions reported~~size bins~~	Scattering angle	Wavelength	Sensor output
Plantower PMS5003*	0.3–10	3	90°	Red (laser)	PM_1, $PM_{2.5}$, PM_{10}
Nova SDS011	0.3–10	2	90°	Red (laser)	$PM_{2.5}$, PM_{10}
Sensirion SPS30	0.3–10	4	90°	Red (laser)	PM_1, $PM_{2.5}$, PM_4, PM_{10}
Sharp GP2Y1010AU0F	n/a	1	120°	IR (LED)	Voltage level
Shinyei PPD42	> 1	1	120°	IR (LED)	PWM-signal
Omron B5W-ld0101**	> 0.5	2	120°	IR (LED)	Pulse count (> 0.5, > 2.5 µm)
Mid-cost monitor: GRIMM 1.108	0.23–20	15	90°	780 nm (laser)	~~Number, surface, and mass conc.~~All fractions individually

* Standard particle (CF=1) output was used.

** Manually adjusted threshold voltage was set to 0.5V as recommended by the manufacturer.

Table 2. Valid detection ranges of the evaluated sensors. Symbols of "greater than" or "smaller than" refers to cases where the other end of the size cut-point was outside of the particle size range producible by the VOAG-GP50 system (0.45—9.73 µm). Units are in µm.

Sensor	Bin 1	Bin 2	Bin 3	Bin 4
Plantower PMS5003	< 0.7	< 0.8	< 1.0 (noisy)	-
Nova SDS011	< 0.8	0.7—1.7	-	-
Sensirion SPS30	< 0.9	0.7—1.3	0.7—1.3	0.7—1.3
Sharp GP2Y1010AU0F	~~1.8—6.5~~< 0.8	-	-	-
Shinyei PPD42*	1.0—2.1	> 5.9	1.5—4.9	-
Omron B5W	0.6—1.0	> 3.2	-	-

* Valid detection ranges of the individual sensors, not bins.