

Reviewers' comments and author's response on:

“Results and insights from the first ACTRIS intercomparison workshop on sub-10~nm aerosol sizing instrumentation”

Hartl et al. (2025)

Response deadline: 03.03.2026

Note: We thank the Reviewers' and Editors' for their effort and their comments. The numbering of the comments corresponds to the numbering as found in the received document. Unnumbered comments correspond to comments as found on the interactive review proves homepage.

1	<p><u>Anonymous Referee #1:</u></p> <p>Although the manuscript touches on various inverse methods, uncertainties from inversions are not fully discussed. For mobility spectrometers, there are significant questions regarding the charging efficiency of sub-10 nm particles and the treatment of measurement noise. It is not clear if all MPSS used the same inversion or custom inversions, which may result in additional uncertainty just from the inversion method (Wiedensohler et al., 2012). It would be good to investigate if the high noise flow of M1, M2, and perhaps M3 is due to amplification of rogue counts or due to some treatment in the inversion. A section on “inversion uncertainties” would be much appreciated.</p> <p>Hartl et al.:</p> <p>We acknowledge that inversion methods can introduce additional uncertainties, particularly for sub-10 nm particles where charging efficiency and measurement noise play a role. While investigating these effects in detail would be valuable, the primary aim of this study was to compare instrument performance, not to evaluate the inversion algorithms themselves. A full analysis would require providing identical raw signals to each inversion routine, as done in Wiedensohler et al. (2012), which is beyond the scope of this work.</p> <p>Actions taken:</p> <p>To address the referee's comment, we added a short statement in the Conclusions noting that uncertainties arising from inversion methods should be investigated in future studies.</p>
2	<p><u>Anonymous Referee #1:</u></p> <p>There is limited focus on the dependence of chemical composition. Particles other than WOx are needed to gauge the variance of performance under field</p>

conditions. Also charge is mentioned to have a potential effect on activation/particle detection (“and it is also assumed that negatively charged particles of the same composition achieve higher activation efficiencies at small sizes, regardless of the CPCs use”). A discussion section of “composition and charging uncertainties” would also be appreciated, if only to outline what types of experiments are recommended for the next workshop.

Hartl et al.:

We agree that particle chemical composition and charging state can influence activation and detection efficiencies, and that these factors are important for assessing instrument performance under realistic environmental conditions. However, Investigating the effects of different particle compositions systematically would have required additional dedicated experiments and is/was beyond the scope of the workshop Nevertheless, we have added a few sentences.

Actions taken:

We have added a short paragraph in the “Introduction” (at the end of the “Sizing Methods” section) highlighting the current knowledge gaps regarding particle composition and charging effects and the limited representation of atmospheric aerosol properties by typical laboratory calibration particles. Furthermore, in the “Conclusions,” we included a statement recommending that future intercomparison studies consider particle chemical composition and explore the use of chambers to produce calibration particles that more closely resemble atmospheric aerosols.

In “1.2 Sizing methods”

In addition to instrumental differences, several studies have highlighted how particle chemical composition and charging state can influence activation and therefore detection efficiencies, especially within the sub-10 nm size range (Kangasluoma et al., 2014, 2020). Calibration particles used in laboratory settings - such as metal oxides or salts - do not fully represent the physicochemical properties of atmospheric particles, which usually consist of a mixture of organic compounds, inorganic species (e.g. sulfates, nitrates, ammonium), black carbon or soot, sea-salt, mineral dust and other trace components (Kangasluoma et al., 2020; Kulmala et al., 2013; Canagaratna et al., 2007). These compositional and charging-related differences may therefore introduce additional uncertainties when transferring laboratory-based calibration results to ambient measurement conditions.

In “4 Conclusions”

We also advocate for the inclusion of chamber measurements as a standard component of intercomparison procedures. These offer a significantly higher degree of control and reproducibility compared to outdoor measurements and could be used to produce calibration particles that better resemble

	<p><i>atmospheric aerosols and to systematically investigate effects of particle chemical composition.</i></p>
<p>3</p>	<p>Anonymous Referee #1:</p> <p>A workshop in ice nucleation instrumentation (DeMott et al., 2018) used a blind experiment design where instrument owners were not told the composition, size, or concentration of particles sampled. Instrument operators would setup/calibrate the instrument to the best of their ability, perform the measurement, and submit their answer. Although this is sometimes challenging to do in a collaborative workshop setting, this activity tests uncertainties that might be present under field conditions. The authors might want to comment on the extent to which the experiments were blind and/or make recommendations for future workshops about the feasibility to conduct such a blind intercomparison.</p> <p>Hartl et al.:</p> <p>The intercomparison was not conducted as a blind study, as instrument operators were informed about the general experimental conditions during and prior to the workshop. We agree that fully blind intercomparison approaches, such as those described by DeMott et al. (2018), can provide valuable insights into measurement uncertainties and potential biases. However, implementing a blind protocol in a collaborative calibration setting is challenging, especially when real-time adjustments and troubleshooting of instruments are required. Despite this, blind experiments represent a useful option for future intercomparisons with a slightly adapted framework</p> <p>Actions taken:</p> <p>We have added a short note in the “Methods and Setups” section and in the “Conclusion” to clarify that the present study was not held in blind setup and to suggest that blind experimental protocols could be considered in future intercomparison workshops as a way to further assess uncertainties and reduce potential bias.</p> <p>In “2 Methods and Setups</p> <p><i>All measurements were performed in an open format, where instrument operators were aware of the general experimental conditions.</i></p>
<p>4</p>	<p>Anonymous Referee #1:</p> <p>Measurable coincidence starts to play a role at concentrations of the order of 10^3 cm^{-3} -> This depends strongly on the sample flow rate through the optical. It is my understanding that some nanoCPCs, e.g. TSI3776 and newer models reduced this flow substantially, thus pushing this value much higher. Also, later it is stated “at particle number concentrations approaching or exceeding 10^7 cm^{-3}, these systems become decreasingly precise due to limitations in single-particle counting and potential coincidence errors”, which contradicts the previous statement.</p>

In “4 Conclusions”

In addition, future experimental designs should consider particle charging state effects on instrument performance and comparability, explore blind measurement protocols as demonstrated by DeMott et al. (2018) and explicitly account for potential uncertainties arising from inversion methods.

Hartl et al.:

We agree that the onset of measurable coincidence depends strongly on the sample flow rate through the optical detection volume, as this determines the particle residence time and thus the probability of pulse overlap. Instruments with optimized flow rates (such as newer nanoCPC designs) can therefore shift the practical onset of measurable coincidence to higher particle number concentrations.

Regarding the alleged contradiction, no inconsistency was intended. The original text aimed to distinguish between

- (i) the concentration range at which coincidence effects first become measurable and
- (ii) the substantially higher concentration range at which the uncertainty of the coincidence correction begins to significantly degrade measurement precision.

While coincidence effects may theoretically become detectable at relatively low concentrations, they can typically be corrected reliably over a wide concentration range. Only when approaching the upper end of the instrument’s operating range the uncertainty of the correction increases, leading to reduced precision.

Actions taken:

We have revised the manuscript to explicitly state the dependence of coincidence effects on the sample flow rate and to acknowledge that modern designs can shift the measurable onset to higher concentrations. Furthermore, the relevant passages have been reworded to clearly distinguish between the initial appearance of measurable coincidence and the concentration regime where coincidence correction uncertainty becomes significant.

In “1.1 Condensation Particle Counter”

Coincidence effects theoretically begin as soon as more than one particle is present per detection interval. In practice, the concentration at which coincidence becomes measurable depends strongly on the sample flow rate through the optical detection volume and the associated residence time. For conventional CPC designs, first measurable effects are typically observed at concentrations on the order of 10^3 cm^{-3} , which is why CPC cutoff calibrations are usually performed at concentrations below approximately 1500 cm^{-3} . For typical CPCs, a 10% coincidence error - often taken as a practical threshold for the upper limit of reliable single-particle counting - is

	<p><i>only reached at concentrations around 104 to 105 cm⁻³ (Collins et al., 2013). Modern CPCs incorporate correction algorithms that compensate for coincidence effects over a substantial fraction of their specified concentration range, which may extend up to 106-107 cm⁻³ depending on instrument design and optical flow rate. For particle number concentrations approaching or exceeding the upper end of this range, the uncertainty of the coincidence correction increases, resulting reduced measurement precision rather than an immediate breakdown of counting capability.</i></p>
<p>5</p>	<p>Anonymous Referee #1:</p> <p>“For the calibration of the Half-Mini DMA, Tetraheptyl Ammonium Bromide (THAB) ions were generated using electrospray. In this context, term calibration refers to establishing the DMA’s voltage-to-particle diameter relationship by using well-characterized ions of known electrical mobility as references. THAB ions serve as mobility standards that allow comparison between detected ion peaks and theoretical mobility values, as described by Ude and Fernández de la Mora (2005)” -> It might be helpful to state why such a calibration is needed. The design of the traditional DMA is that of a primary calibration standard. They are not calibrated, but size is fully predicted by the sheath flow rate, DMA dimensions, and applied voltage. The sizing is typically tested and if it doesn’t agree, then the causes can be investigated. Why is this not the same for the half-mini DMA?</p> <p>Hartl et al.:</p> <p>We agree that a conventional DMA can principally , serve as a primary standard, since the electrical mobility (and thus particle size) is fully determined by the applied voltage, sheath flow rate and instrument geometry. However, this requires precise knowledge of the actual sheath flow. For compact, high-sheath-flow designs (such as the Half-Mini DMA), the sheath flow cannot always be measured directly with sufficient accuracy. Thus, small deviations in the effective flow rate may lead to systematic sizing discrepancies. To verify the effective voltage–mobility relationship under operating conditions, THAB ions were therefore used as mobility standards.</p> <p>Actions taken:</p> <p>We have revised the manuscript to clearly distinguish between the primary-standard principle of a traditional DMA and the practical limitations associated with the Half-Mini DMA design. The text now explicitly explains why experimental verification of the voltage–mobility relationship using THAB ions was necessary in this case.</p> <p>In “2.1 Experimental Setups”</p> <p><i>For the calibration of the Half-Mini DMA, Tetraheptyl Ammonium Bromide (THAB) ions were generated using electrospray. In this context, the term calibration refers to a verification and eventual adjustment of the DMA ’ s voltage-to-particle diameter relationship by using well-characterized ions of</i></p>

	<p><i>known electrical mobility as references. In principle, a traditional DMA can serve as a primary standard as a selected particle 's mobility is fully determined by the applied voltage, the sheath flow rate and the DMA 's geometry. However, this is not the case for compact, high-sheath-flow designs (such as the Half-Mini DMA), due to their actual sheath flow rates not being accessible directly with sufficient accuracy. Therefore, THAB ions serve as mobility standards that allow comparison between detected ion peaks and theoretical mobility values, as described by Ude and Fernandez de la Mora (2005). In this way, the effective voltage -mobility relationship of the instrument is experimentally confirmed.</i></p>
<p>6</p>	<p><u>Anonymous Referee #1:</u> nano meter -> nanometer</p> <p><u>Hartl et al.:</u> Agreed.</p> <p><u>Actions taken:</u> Change implemented as suggested by the Referee.</p>

<p>1</p>	<p><u>Anonymous Referee #2:</u> I miss a brief description on the status of the instruments tested, particularly age (year of manufacture) and time since last maintenance and calibration</p> <p><u>Hartl et al.:</u> A key focus of the workshop was to compare in-field used instruments, including devices whose calibration dates may have varied significantly. In addition, some operational parameters were adjusted to achieve optimal comparability of the instruments during the workshop. Therefore, the latest calibration was considered less relevant in this context, as the primary goal was to ensure the comparability of the measurement data collected during the workshop.</p> <p><u>Actions taken:</u> None</p>
<p>2</p>	<p><u>Anonymous Referee #2:</u> The first paragraph on page 3 sounds as if measurements of sub-10 nm particles would have been more or less impossible until recently. There are, however,</p>

several earlier publications on this, which should be mentioned, e.g. work by Fernández de la Mora in the late 1990s and early 2000s or the development of the nano-DMA (Chen et al., J. Aerosol Sci. 29:497-509, 1998).

Hartl et al.:

The original paragraph was intended to emphasize that measurements of sub-10 nm particles have historically been technically challenging due to limitations in detection efficiency, sizing accuracy and particle losses etc. - not that such measurements were impossible. However, we acknowledge that the wording may have unintentionally conveyed that impression.

Actions taken:

We have revised the paragraph to avoid any misleading. In addition, we now explicitly acknowledge earlier contributions (the development of the nano-DMA (Chen et al., 1998)).

In “1 Introduction”

Until recently, detecting and characterising sub-10 nm particles has been technically challenging and measurements were typically limited to laboratory-based studies and specialized instruments (Chen et al., 1998; Kangasluoma et al., 2020). Due to their small size, these particles exhibit high diffusivity and low charging probabilities, complicating reliable measurement. However, owing to innovative methods emerging over the past decade, reliable techniques for sizing aerosol particles down to 1nm have been developed.

3

Anonymous Referee #2:

Section 1.1 describes the calibration of a CPC insufficiently, as the role of the electrometer only becomes clear to someone, who already knows the procedure (and thus doesn't need the description). I suggest to add one or two sentences, describing (with proper reference, e.g. to ISO 27891) that particles are classified with a DMA and are therefore –if done right- all singly charged. Consequently, this calibration can even be a metrologically traceable.

Hartl et al.:

We agree that the role of the electrometer was not described in sufficient detail. The electrometer provides the reference concentration used to determine the CPC detection efficiency, while the DMA is used to classify particles such that they are predominantly singly charged.

Actions taken:

We have revised Section 1.1 to explicitly describe the role of the electrometer in the calibration procedure. In particular, we now clarify that particles are classified by a DMA to obtain (predominantly) singly charged particles and that the electrometer provides the reference concentration for determining the CPC detection efficiency, in accordance with ISO 27891.

In “1.1 Condensation Particle Counter”

Amongst other parameters, a CPC ’ s performance is primarily described by its detection efficiency as a function of the particle size dp . It represents the fraction of the particle number concentration per unit volume n_{tot} , $CPC(dp)$, 90 a CPC is capable of detecting in relation to the concentration as measured by a reference instrument $n_{tot,EM}(dp)$ (typically an aerosol electrometer (EM)) from the same sample. The so obtained detection efficiency $\eta(dp)$ is prone to variations over time -potentially caused e.g. by contamination of the optics, mechanical wear, improper handling or transportation - and thus requiresreevaluation during regular calibrations in an effort to ensure its validity and accuracy. The calibration follows a standardized procedure (ISO27891, 2015) in which particles of well-defined mobility diameters are selected using a DMA over a size range 95 limited, such that the fraction of multiply charged particles remains negligible (Wiedensohler, 1988). This measure poses absolutely necessary to avoid an overestimation of concentrations by the EM, as multiply charged particles would contribute proportionally more to the measured current and thus biasing its inferred particle number concentration.

4

Anonymous Referee #2:

From the description of the activation-based instruments (lines 121-125), it is not clear, why the mixed-phase droplets are not directly optically detected, but passed on to a downstream CPC. If I remember correctly, this is due to the too small size of the DEG-droplets for direct optical detection.

Hartl et al.:

The PSM serves as a pre-growth stage that activates small particles at high supersaturation, but does not grow the resulting mixed-phase droplets to sizes large enough for laser detection. Therefore, the activated particles are transferred to a downstream CPC, where further condensational growth increases the droplet size to the optically detectable range. Optical detection is then performed via scattered laser light in the CPC. This was implied in the paragraph before, but the authors agree, that it was not stated explicit enough.

Actions taken:

We have revised the paragraph to make this measurement principle more explicit. In particular, we now clearly state that the DEG droplets formed in the PSM are too small for direct optical detection and that further growth in the downstream CPC is required to enable optical detection.

In “1.1 Condensation Particle Counter”

In the saturator stage, a wicking material loaded with a volatile working fluid - typically Diethylene Glycol (DEG) - provides vapours that mix with the aerosol flow. Heating this stage increases the vapour content the aerosol's gas phase can hold, allowing equilibrium with the aerosol to be reached. As the aerosol then transitions into the condenser stage, it is rapidly cooled, causing vapour supersaturation. This process enables the aerosol particles to act as CN, causing them to grow into mixed-phase droplets. Despite the higher

supersaturation achieved in the PSM compared to a CPC, these droplets remain too small for optical detection. The aerosol is therefore passed to a downstream CPC, where the particle undergo further condensational growth and are subsequently detected and counted via scattered laser light. This serial arrangement of instruments allows for the activation of significantly smaller particles, which would otherwise remain undetected, therefore effectively lowering the cutoff size compared to a CPC alone.

5

Anonymous Referee #2:

The tests of the CPCs were conducted exclusively with negatively charged particles. The authors argue that this would result in higher activation efficiencies at small sizes but fail to discuss this implication on “real-life” measurements. When the CPC is used as a stand-alone device for the measurement of the total number concentration, the majority of the particles at the lower size end will be uncharged. Would this mean a shift of the d_{50} to larger particle sizes? Is this probably the reason, why the majority of the CPCs showed clearly smaller d_{50} values in this study compared to the nominal lower cut-sizes (figure 4)?

Authors:

We agree that the exclusive use of negatively charged particles during the workshop measurements has implications for the interpretation of the reported cutoff diameters in the context of real-world measurements. Charged particles generally exhibit higher activation efficiencies at very small sizes compared to neutral particles. This could result in a slight shift of the effective cutoff toward larger diameters under atmospheric conditions compared to laboratory conditions.

Actions taken:

We have added a clarifying paragraph in the “Results” and “Conclusions” section explicitly stating that both the workshop measurements and the manufacturer-reported cutoff diameters are determined under laboratory conditions. We also note that under environmental measurement conditions, where particles are predominantly neutral, the effective cutoff diameter may shift to slightly larger values due to lower activation efficiencies of neutral particles.

In “3.3 Detection Efficiency”

The cutoff values for all but one system, namely the TSI 1 nm SMPS (M2), were found to be 0.2 to 0.3nm below their respective nominal values, demonstrating a high level of consistency across instruments. We speculate that the origin of this systematic behavior might partly be found either in the simplistic approach used to correct for the size shift and/or in the fact that only negatively charged particles were used for the calibrations. It should be noted, that both cutoffs - the nominal value as reported by the manufacturer, as well as the value measured during the workshop - were determined under controlled laboratory

	<p>conditions employing charged particles. Under atmospheric conditions, the fraction of naturally charged sub-10 nm particles is next to negligible and the substantial majority of particles remains uncharged. Since neutral particles exhibit lower activation efficiencies compared to charged ones (see previous sections), this may effectively result in a shift of the cutoff toward slightly larger particle diameters for environmental applications. In contrast, the measured cutoff diameter for TSI 1 nm SMPS(M2) $d_{50} = (1.409 \pm 0.015) \text{ nm}$ was effectively indistinguishable from its nominal value. This difference is likely reflects the manufacturer's unique choice of using NaCl as a calibration seed. Further investigations are needed to clarify any residual differences.</p> <p>In “Conclusions”</p> <p><i>In addition, future experimental designs should consider particle charging state effects on instrument performance and comparability, explore blind measurement protocols as demonstrated by DeMott et al. (2018) and explicitly account for potential uncertainties arising from inversion methods.</i></p>
6	<p>Anonymous Referee #2:</p> <p>Figure 3: What is the difference between the red and blue “+” symbols?</p> <p>Hartl et al.:</p> <p>The red and the blue plus symbol represent two independent measurement runs performed with NiCr as particle seed and a Perez-DMA as particle discriminator. In principle, these could have been made the same colour, but colouring them differently conserves a higher level of information.</p> <p>Actions taken:</p> <p>The Figure’s caption has been adapted according to the Referee’s suggestion as to clarify the issue.</p> <p>In “3.1 Size Shift”</p> <p><i>Figure 3. Particle size shift of the 4-way-cross glowing wire generator during workshop measurements. The ordinate shows the ratio of mode diameter measured with an Airl Nais in negative ion mode to the nominal diameter (according to THAB calibration), \hat{d}_p / \hat{d}_0. The solid line is a skewed Gaussian cumulative distribution function fit (SG-CDF, Eq. 3) with parameters d_0, σ, A, α, and C. Marker types indicate measurement conditions: circles for the workshop setup; squares for NiCr particle seed; red and blue plus symbols for two independent runs performed with NiCr and a Perez-DMA; yellow cross symbols for a measurement conducted under otherwise identical conditions, but with cooling of the aerosol flow added prior to the DMA’s inlet; and grey triangles for a different source excluded from the fit. Error bars are omitted for clarity. See Sect. 3.1 for discussion.</i></p>
7	<p>Anonymous Referee #2:</p>

The last paragraph of chapter 3.3 sounds as if it was generally better, if the d_{50} of a CPC is smaller than its nominal value and that the performance of the TSI 1 nm SMPS was worst. However, the results of this device almost perfectly matched its nominal performance ($d_{50} = (1.409 \pm 0.015)$ nm compared with 1.40 nm). The authors use the term “overshooting” to describe this small difference. Considering that the other devices show smaller d_{50} than their nominal values, one could also argue that the concentrations measured with these CPCs overshoot compared to the expected concentrations (considering the nominal cut-off). I suggest describing the results in a more objective manner here and without using judgmental terms.

Hartl et al.:

We agree that the previous wording could have been interpreted as judgmental and potentially misleading regarding the relative performance of the instruments.

Actions taken:

We have revised the paragraph to present the results in a more objective manner. In particular, we now explicitly state that the TSI 1 nm SMPS’s (M2) cutoff was effectively indistinguishable from its nominal value. We further clarify that this agreement likely reflects the manufacturer’s calibration using a NaCl seed, which differs from the particles used for the other instruments in this study.

In “3.3 Detection Efficiency”

The cutoff values for all but one system, namely the TSI 1 nm SMPS (M2), were found to be 0.2 to 0.3nm below their respective nominal values, demonstrating a high level of consistency across instruments. We speculate that the origin of this systematic behavior might partly be found either in the simplistic approach used to correct for the size shift and/or in the fact that only negatively charged particles were used for the calibrations. It should be noted, that both cutoffs - the nominal value as reported by the manufacturer, as well as the value measured during the workshop - were determined under controlled laboratory conditions employing charged particles. Under atmospheric conditions, the fraction of naturally charged sub-10 nm particles is next to negligible and the substantial majority of particles remains uncharged. Since neutral particles exhibit lower activation efficiencies compared to charged ones (see previous sections), this may effectively result in a shift of the cutoff toward slightly larger particle diameters for environmental applications. In contrast, the measured cutoff diameter for TSI 1 nm SMPS(M2) $d_{50} = (1.409 \pm 0.015)$ nm was effectively indistinguishable from its nominal value. This difference is likely reflects the manufacturer ’ s unique choice of using NaCl as a calibration seed. Further investigations are needed to clarify any residual differences.

8

Anonymous Referee #2:

Figure 6: It is not clear to me, how devices M1 and M2 can show such high concentrations, when the actual concentration is zero. The authors explain that the shape of the distributions stem from internal corrections made by the two devices. However, these corrections are size-dependent factors that are multiplied with the concentrations, measured by the devices. The curves don't look as if the devices only counted individual particles, because that would have resulted in a lot less smooth curves. How can it be that the devices measured such substantial concentrations? Is this something commonly observed or a malfunction of the devices?

Hartl et al.:

The apparent particle concentrations observed for M1 and M2 during background conditions do not represent real particles in the chamber. Instead, they arise from instrument-internal correction procedures applied to very low or noise-dominated (rogue) signals (please see also Comment #1 of Referee #1). Both instruments continuously apply size-dependent loss corrections to the measured signal.

(I) In general:

When particle concentrations approach zero, statistical noise in individual size bins becomes comparable to or even larger than the signal itself. The subsequent application of multiplicative correction factors can then amplify this noise and non-zero size distributions. Correction algorithms facilitating “localization length” approaches, subsequently smooth out these curves.

(II) In specific:

- For the GRIMM PSMPS (M1), this effect is strongly influenced by fluctuations in the ion background during measurements. Since the background correction for this instrument bases on a subtraction of an averaged ion signal, elevated concentrations of ion concentrations during measurements lead to these rogue counts
- For the TSI 1 nm SMPS (M2), internal signal fluctuations caused by unwanted homogeneous nucleation of DEG remainders in the Nano Enhancer contributed to elevated baseline noise. A thorough post measurement cleansing of the instrument drastically decreased the effect and proofed this claim to be true.

Actions taken:

We have revised the manuscript to clarify this explanation and to emphasize that these signals do not represent real aerosol particles in the chamber.

In “3.5 Chamber Measurements”

Importantly, these apparent concentrations do not represent real particles in the chamber but arise from instrument-internal correction procedures applied to very low or noisy signals. These corrections, designed to compensate for diffusional losses, are continuously applied to all measured signals. When

	<p>concentrations approach zero, statistical noise in individual size bins can become comparable to or even larger than the actual signal itself. The subsequent application of size-dependent loss correction factors may then amplify this noise, leading to artificially elevated and smooth size distributions. As discussed previously for the TSI 1 nm SMPS (M2), this elevated baseline originates from internal signal fluctuations, caused by unwanted homogeneous nucleation of DEG remainders inside the Nano Enhancer, which constitutes the pre-growth stage of the system. Such nucleation events generate spurious counts that are then further modified by the instrument's internal correction algorithm. This interpretation is supported by the observation that the signal noise noticeably decreased after cleaning the Nano Enhancer of residual DEG. In contrast, the GRIMM PSMPS (M1) shows a baseline shape consistent with its ion-induced background correction method. For this instrument, the subtraction of an averaged pre-measured ion background can lead to residual artefacts, especially if ion concentrations fluctuate during the actual measurement. In combination with the applied loss corrections, this resulted in apparent particle concentrations even for the particle free chamber and is not indicative of this system malfunctioning, but reflects the limitations of background corrections to signals dominated by noise rather. The activation-based instruments recorded very low background counts, with concentrations $< 10\text{cm}^{-3}$ for A2 and A5 and $< 1\text{cm}^{-3}$ for other instruments in the sub-2 nm size range.</p>
<p>9</p>	<p>Anonymous Referee #2: Line 47: “total aerosol mass” should read “total particle mass”</p> <p>Hartl et al.: Agreed.</p> <p>Actions taken: Change implemented as suggested.</p>
<p>10</p>	<p>Anonymous Referee #2: Line 81: “...of a supersaturated working fluid” should read “...of a supersaturated vapour of the working fluid”</p> <p>Hartl et al.: Agreed.</p> <p>Actions taken: Change implemented.</p>
<p>11</p>	<p>Anonymous Referee #2:</p>

	<p>Line 194: it's --> its</p> <p>Hartl et al.: Agreed.</p> <p>Actions taken: Change implemented as suggested.</p>
12	<p>Anonymous Referee #2: Line 232: the “median of zero runs” sounds a bit awkward. I assume you mean “median concentration during runs with particle free air”.</p> <p>Hartl et al.: Agreed.</p> <p>Actions taken: Change implemented as suggested.</p>