

Final reply to all reviewer comments

Dear Jose, in the following, I will reply to all reviewer comments. In case I changed the text, I will copy the revised paragraph into this document and will mark all changes in the article with a **yellow background**. I will mark all my comments and revised text here in *italic* and

Then I would like to thank all reviewers for the very constructive comments to improve this article. I especially thank Mark Stolzenburg for his very detailed review and the fruitful discussions afterwards.

In the beginning, I would like to state that we first moved the section 4.4. It is now section 3.3, and we included also an estimation of enhanced diffusional losses in 90° bends.

Second, we added another section to chapter 4 related to the results of the ACTRIS intercomparison workshop held in September/October 2011 in Leipzig. There, we describe the link of the WCCAP reference SMPS to the total number concentration and the unit-to-unit variability between the three spectrometers. We replaced the existing Figure 9.

Reviewer: Mark Stolzenburg

General Comments

This is a very well-written and reasonably comprehensive paper on a topic of growing significance. In the past and still to a significant extent today, many researchers have concentrated on aspects of the number size distribution beyond the simple dry distribution. In particular, they have looked at the hygroscopicity and volatility as a function of size. This has been of great value for isolated studies that do not involve a whole network of stations. However, as particle mobility size spectrometers become more reliable, user-friendly and potentially somewhat less expensive, they will more often be deployed at multiple sites in long-term monitoring programs and their intercomparability will become much more important. Modelers are likely to make liberal use of this data and it is important that the quality and comparability of it is assured.

The authors here have laid a fairly solid groundwork as a first step in achieving this goal. I applaud the no doubt heroic efforts they made to get all these disparate researchers and their instruments together in one place, not once, but three times. For those organizing all of this (Wiedensohler) it must have taken an incredible amount of time and dedication to schedule and run the workshops as well as collating all the necessary data from the different groups to write this paper. For this incredible effort I am quite grateful. It has made this paper the landmark work that it is.

Thank you very much Mark

That said, I would like to see the extra effort put in to make this as good as it can be. Because of this, I have been very rigorous in my evaluation below. Do not be put off by the length of my comments. There is very little that will require significantly more work.

We did our best to improve the article; however, we cannot repeat the workshops. I will explain later in detail.

However, I am disappointed/concerned about the lack of agreement above 200 nm and below 20 nm. Even the degree of agreement within in this range is overstated (see below). My personal experience is that it is significantly more difficult than one might think to obtain +/-1% uncertainty in flows even under controlled laboratory conditions, much less in the field.

The lack of agreement in the inversion routines (Fig. 2) above 200 nm is particularly disturbing. Unlike what the authors suggest, I feel that this must be due to differences in how the multiple charge correction is applied in the various routines. (A different possibility that just occurred to me is that the disagreement at the upper end may have something to do with use of splines that carry over information from the lower size bins to the upper size bins. Different size bins and splines could cause variations in the carry over effect.) Whatever the problem is, it should be a fairly straightforward problem of resolving and correcting it (if it is not a spline problem). I think the authors could probably decide on a best practice in applying the correction or at least guidelines as to what method to use in which situation. As this is all computer work, this should be resolvable without the need of gathering the instruments and the investigators all together in one place again. Furthermore, if the Fig. 5 has been stored according to the recommendations in this paper, then it should be possible to rerun the processing of the data for Fig. 5 using the corrected routines, though it looks doubtful that it would make much visible difference given the way it is plotted here (i.e. linear vertical scale). Given all this, I think this problem of disagreement in the inversion routines above 200 nm should be resolved and incorporated into this paper before publication.

The lack of agreement in the measured distributions below 20 nm is a very different matter and will probably be much more difficult to resolve. The possible oversimplification of modeling losses (especially in the inlet and outlet of the DMA) as those in a tube of equivalent length may have something to do with the disagreements in this lower range as the measurement of losses and subsequent determination of the equivalent length may have been done using significantly larger particle sizes. Note that the losses in sequential tubes of the same length with mixing (i.e. re-establishment of a radially uniform aerosol concentration profile) between the tubes is not the same as a single tube of any equivalent length. Thus, this method of approximating losses by an equivalent length is even more crude and less accurate for the more complicated plumbing of the DMA aerosol inlet and outlet regions. This and other considerations make it too difficult to resolve this problem within the scope of this paper. However, I believe the topic of this disagreement deserves a somewhat more thorough discussion about the possible causes in the text of this paper.

Personally, I was also disappointed about the results for particles smaller than 20 nm. It seems that there are undetermined losses, which we cannot explain yet. This is a task in the new infrastructure project ACTRIS. In terms of the increased uncertainties of particles larger than 200nm, we cannot really explain this. We tried to calculate the run#4 results with the IFT inversion routine, which did not give better results. The second possibility is that in inversion routines might not be correct for multiple charge correction. This is not directly visibly from the inversion routine exercise. There the differences were smaller than 5%. The last option would be that some bipolar chargers were not able to provide the charge equilibrium,

especially for multiple charged particles. Here, small deviations can lead to the observed +/- 30% uncertainty.

Specific Comments

page 5525, line 12: Define "long-term".

Changed to:

The first atmospheric long-term (period of several years) measurements

p. 5530, ln. 6-11: In this paragraph, the authors cite a few references for the various elements involved in calculating the electrical mobility distribution from SMPS (DMPS) data. Key original theoretical references have been omitted in favor of more recent references emphasizing practical application for the bipolar charge distribution and the DMA transfer function. However, the theory is the basis of these practical applications and those appropriate references need to be cited as well. At minimum this would include Fuchs for bipolar charging theory and Knutson & Whitby and Stolzenburg for DMA transfer function theory.

Paragraph changed to:

(Fuchs, 1963; Wiedensohler, 1988; Baron and Willeke, 2005; see also Appendix A.5) is used. A DMA-transfer function (Birmili et al., 1997; Knutson and Whitby, 1975, Stolzenburg, 1988) might be also employed to include losses in the DMA directly in the data inversion. Birmili's method is based on a triangular ideal transfer function, which may lead to small uncertainties. Later, the number size distribution can be corrected by the CPC counting efficiency function (e.g. Wiedensohler et al., 1997; Hermann et al., 2007) estimations of internal losses by diffusion.

Also, the Wiedensohler fit for charge distribution is only good for a specific set of ion parameters. These may change with humidity and/or the level of pollutant gases in the sample flow.

There are basically no other bipolar charge distributions used for atmospheric measurements. It does not make sense to apply different bipolar charge distributions, since the concentration of trace gases are unknown. The relative humidity should be kept low anyway.

The Birmili et al. method of determining the transfer function relies on two DMAs having identical transfer functions. This can only be assured when the transfer functions are ideal, producing the minimum spread in the TDMA response function. There is no assurance that two nominally identical non-ideal DMAs will underperform to the same degree (e.g. same mixing characteristics). The suggested test for this in Birmili et al. is not very sensitive so the equal assignment of the non-ideality to both DMAs may not be correct. Thus, though the methods/formulae of the references cited here may be the most practical to use they are not necessarily accurate in all cases.

Changed to:

... (Fuchs, 1963; Wiedensohler, 1988; Baron and Willeke, 2005; see also Appendix A.5) is used. A DMA-transfer function (Birmili et al., 1997; Knutson and Whitby, 1975, Stolzenburg, 1988) might be also employed to include losses in the DMA directly in the data inversion. Birmili's method is based on a triangular ideal transfer function, which may lead to small uncertainties. Later, the number size distribution can be corrected by the CPC counting efficiency function (e.g. Wiedensohler et al., 1997; Hermann et al., 2007) estimations of internal losses by diffusion.

p. 5533, ln. 12: "These (temperature and absolute pressure) are preferentially monitored at the aerosol inlet of the DMA." Why there? The aerosol or sheath air out would actually be more representative of the temperature in the DMA as the temperature of the sample in flow must equilibrate with the station temperature. Depending on the length of plumbing between the station wall (or ceiling) and the DMA aerosol inlet, the temperature may not fully equilibrate until it meets the DMA sheath flow. Of course, this is also true of the relative humidity of the sample flow. The measured absolute pressure may be the closest to that of the internal DMA pressure in any of the four flow lines entering or exiting the DMA, depending on which has the least pressure difference between the sensor and the classification section of the DMA. The aerosol outlet line may very well have the smallest pressure drop.

The temperature and humidity are additionally measured at the excess air (see figure1). The absolute pressure we suggested to monitor directly at the aerosol inlet of the DMA, which should not give a significant difference in pressure compared to inside of the DMA. It might be 1 hPa.

p. 5533, ln. 14ff: Why store particle number concentrations when the actual particle count per channel is needed to calculate uncertainties of measured concentrations? "Raw data" should be raw, number concentration is not raw. Store counts (# of particles) and delta time per channel as well as the CPC sample flow. In some SMPS systems, the CPC inlet flow rate is not the same as the DMA aerosol out flow rate. It may also be different from the actual CPC internal sample flow rate such as in ultrafine sheathed CPCs. Thus, four flow rates should be recorded: DMA sheath and aerosol flows and CPC inlet and sample flows. Some CPCs report these latter two flows (e.g. TSI 3025) such that they could be actively monitored during sampling. Thus, these flows should be recorded for every distribution.

To store the raw counts is indeed the real raw data, but we agreed to use the concentration due to following reasons. 1) Everybody can use this data to recalculate the number size distribution without knowing anything about the specific aerosol flows and counting times etc. 2) There are DMPS systems that have variable counting times as function of the concentration, and this cannot anymore be considered in the metadata. Therefore, we would like to keep this format.

p. 5533, ln. 20: "The DMA dimension (...) sheath and aerosol and flow rates ... should be supplied in the meta-data." The sheath and aerosol flow rates are being monitored and recorded with each distribution; there is no need for them in the meta-data. In the underlined

portion of this sentence, perhaps the second “and” is a typo or perhaps you meant to say “... and CPC flow rates ...”. Please clarify and note that the CPC flow rates should be monitored and recorded with the distribution data, not in the meta-data. Some SMPS systems collect both upscan and downscan data. This parameter should also be recorded with the distribution data. The DMA to CPC detector delay time should also be recorded in the meta-data.

The think that the nominal flow rates should be given in the meta-data and the actual ones in the raw data level 0.

Corrected and move to the end of chapter 5:

The DMA dimensions (rod diameter and cylinder inner diameter and length), the nominal sheath and aerosol (and CPC) flow rates, and possibly the serial numbers of the DMA and CPC should be supplied in the meta-data.

p. 5535, ln. 3: Upon close inspection of the lower plot of Fig. 2b and noting the 5% deviation of the “GRIMM ISO 15900” and “TSI ISO 15900” curves, I believe a more accurate description would be: “Apart from “Old Grimm”, all inversion routines agreed within 5% over the size range 5 – 300 nm. Only toward the upper end of the size range (300 - 500 nm), some discrepancies occur probably due to uncertainties by of transferring ...”. Your following suggestion as to the cause of these discrepancies does not make sense. Why would there be any uncertainties in transferring the data to different formats? It ought to be a straightforward, clearly-defined process. If there were some sort of problem, why would it only affect the upper end of the size distribution? It would seem more likely that this may be due to differences in implementation of the multiple charge correction. If this is the case, this problem needs to be addressed.

Corrected to:

To see small difference as function of particle size, we plotted the ratio of the results of any individual inversion routine against the IFT one (lower plots in Fig. 2a and b). Apart from “Old Grimm”, all inversion routines agreed within 5% over the size range 5-300 nm. Only towards the end of the size range, some discrepancies occur probably due to splining the EUSAAR data to obtain individual data formats for the individual inversion routines.

p. 5538, Sec. 4.3.1: The DMA is a primary standard. That is, we calculate the mobility size from the basic parameters of dimensions, flows and voltage. When we measure PSL of a certified size with the DMA, we are checking that the DMA is performing properly and that the calculated size is the same as the PSL size. We are NOT calibrating the DMA size. We either accept the theoretically calculated size because it agrees closely with the PSL size, or, when it doesn't, we find the source of the problem and fix it. We do NOT create a “measured” size to voltage relationship based on the PSL measurements. That would be a “calibration”. We simply conduct a performance check of the DMA sizing capability. It either works or it doesn't. (This is not strictly correct for the ultra-high sheath flow DMAs of de la Mora and co-workers. There the sheath flow is initially unknown and it is determined by fitting the calculated size to a known molecular size. Still, this is a calibration of the sheath flow, not the particle size.)

Corrected to

For a DMA size check, monodisperse PSL spheres of one or more particle sizes should be used..

p. 5538, ln. 14: "... +/-1% ...". See note for Fig. 4 below.

Corrected to:

*Taking into account an additional minimum uncertainty of the sheath air flow rate of +/-1% for the individual spectrometers, and a maximum uncertainty of +/-2.5% of the actual PSL size, we defined that a **maximum** deviation of +/-3.5% from the nominal diameter of the PSL particles is tolerable for an instrument to pass the sizing quality test.*

p. 5540, ln. 7: Clearly, some of the curves in Fig. 5 deviate outside the shaded +/-10% range between 20 and 200 nm. In particular, as the size decreases, the IFT TDMPs and UHEL curves deviate more than 10% starting at about 60 nm and the PSI curve at about 40 nm. Over much of the range from 20 to 60 nm there are deviations by 20% or more. Thus, you are considerably overstating the agreement between the distributions in this range. You must find a more accurate way to describe the agreement in the curves here and in the abstract.

Corrected to:

The shaded area in Fig. 5 represents a +/-10% range around the number size distribution measured by the EUSAAR reference instrument. Visually, most of number size distributions agree from 20 to 200 nm within the shaded +/-10% range of the reference instrument. The two DMPS (IFT and UHEL) are slightly outside of this range. Also the PSI-SMPS is too high in the range below 40 nm. Outside the 20-200 nm diameter range, the discrepancies increase. Two spectrometers were specifically designed to measure the number size distribution of nucleation mode particles (IFT-TDMP and FMI). They deviate significantly below 10 nm. The reason for these differences could not be clarified during this work. It will be investigated in the infrastructure project ACTRIS.

p. 5540, ln. 13: The sharp decline of the NILU and TNO size distributions below about 15 nm is clearly not normal. Both systems use CPCs that detect 80% of the particles at 10 nm so this is not a CPC counting efficiency issue assuming they continued to perform the same at the third workshop compared to the measured efficiencies at the second workshop. Any other losses would not drop so sharply. A possible explanation is that the DMA voltages were not properly maintained at this very low voltage. However, TSI personnel say that this is definitely not normal behavior for the standard TSI 3034 and that the DMA voltage should be steady and accurate in this range. Was the IFT scanning software used by TNO with this instrument also controlling the voltage? If so, then perhaps there was some problem with that. For both the NILU and TNO instruments this is a glaring problem in their performance in the lower size range and most likely indicates some sort of problem that can be fixed if the time were taken to diagnose the cause. As the TNO instrument is a commercial product, it is important that you clearly state in the text that this is not normal behavior for the original instrument as shipped by TSI. Similarly, I assume the NILU instrument has been used in a

number of studies and the performance you have measured in this workshop is not necessarily representative of its performance in other studies. Of course, the best thing would be to determine the cause(s) of these problems so that it can be demonstrated that this is not normal behavior. But I assume it is too late for that. In any case, you must address in some manner in the text this abnormal behavior. It should not be simply lumped in with the characterization of the other sub 20 nm deviations. In addition to losses and uncertainties in DMA voltage, low counting statistics (due to the very low charge fractions) below 20 nm may play a role in the variability of the size distributions from the other SMPS systems.

Corrected to:

In general, the deviations below 20 nm mobility size spectrometers are probably due to unconsidered additional losses and increased uncertainties. As found out later, The TNO and NILU spectrometers had problems with the high voltage power supply with an offset, we did not realize. This led to zero or too low voltages for particles below 20 nm during this workshop.

p. 5541, ln. 10-12: Measuring at the peak of the distributions in Fig. 8, the shaded area represents +/-20% (not 25%) range around the reference spectrometer. Replace "25" with "20" twice in this range of lines. Also, the sentence beginning on line 10 should begin as "This time the shaded area ..." since this is the first time the 20% bounds are used.

Corrected to

Again, the shaded area represents a $\pm 20\%$ range around the particle volume size distribution of the reference system. All spectrometers are within this 20% range at the volume peak (250 nm) beside the TNO instrument ($\sim 40\%$). Towards the end of the size range, the discrepancy of two mobility spectrometers increases up to 70% compared to the IFT reference 15 system. Since there is no absolute standard of aerosol volume distribution, we cannot really judge the correct value. We re-run the data with the IFT-inversion routine, but this did not significantly improve the values. It does not seem to be the multiple charge correction of the inversion programs used. For the TNO system, the IFT-inversion was used also before. One reason can be that the bipolar charge equilibrium was not completely reached. However, we are not able to firmly conclude what the underlying reasons for the observed divergences are. On the other hand, mobility size spectrometers are not designed to determine the volume distribution.

p. 5541, ln. 24: From Fig. 9 at 40 nm I calculate that the mean of the shaded area is at about 4750 and the curve for the LB system is at about 5000, which is about 5% above the mean, not 3% as written on this line in the text.

Corrected to:

The maximum deviation of all different mobility size spectrometers was found to be +/-5% from the average at the number peak (40 nm) of the size distribution. The light grey shaded area represents the +/-10% range around the average of the five mobility size spectrometers. Additionally, we plotted the calculated volume size distributions of all mobility size spectrometers. The dark grey shaded area is again the +/-10% range around the average of all systems. All spectrometers are within this range. This unit-to-unit variability of identical

mobility size spectrometers represent the accuracy that can be achieved under controlled laboratory conditions at the current the state of knowledge and technology.

p. 5549, ln. 21: "... (+/-1%) ...". See note for Fig. 4 below.

Corrected to

The measured mean particle size should be within 3.5% of the nominal PSL size including uncertainties of the PSL spheres (+/-2.5%) and the sheath air flow (+/-1%).

Figure 4: Latex sphere (200 nm) calibration of the modified mobility size spectrometer to determine the sizing accuracy under laboratory conditions and supervised by experienced users. The size distributions are normalized to the peak concentration of the reference mobility size spectrometer. The 3.5% uncertainty range includes the uncertainties of the latex particles from the nominal size (2.5%) and the sheath air flow rate (+/-1%). The ratios given in the legend are aerosol to sheath air flow rate.

Table 2: Can you provide any references for the inversion routines which currently do not have any?

There are no other references

Table 2: The way in which the calculations are presented makes it difficult to compare and contrast the different inversion methods. Part of the problem seems to be that consistent language is not used to describe the same thing in different inversion routines. For instance, some routines say $dN/d\log Z_p$ using real/ideal area of the DMA transfer function where others say something about linear inversion. Essentially, it seems that most of the routines use the usual monodisperse approximation in evaluating the integrated response function used to calculate the CPC concentration. Thus, $dN/d\log Z_p$ only depends upon the area of the transfer function. Others (NILU, UHEL/FMI, PKU?) seem to imply that they may do something else. But this is not at all clear from the descriptions as they are now and it should be. In the case of the PSI routine, the first steps listed are the obvious ones that none of the other routines even bother to mention. Then when it comes to the important part of the PSI inversion routine, the description is too brief. Since the TNO system (as noted in Table 3) uses the IFT scanning software the same as the IFT-SMPS except for a switch in charge polarity, it should be included in the description of the IFT routine with an indication of the polarity difference. This makes it obvious that they are the same routine such that the TNO routine is not shown separately in Fig. 2. All in all, I think Table 2 could benefit from a total rework, possibly including a change of format. Much of the material in this table could be readily displayed in a table with columns corresponding to particular inversion routines and rows corresponding to particular features of the Input or Calculation. The entries in the table would then be simply which choice of each feature was used. Special cases that don't fit this pattern could be explained in notes. Thus, the options for the scan data input would be (number concentration or CPC count rate or CPC counts & count time) vs. (mobility diameter or DMA voltage). Features for the calculation could be

- whether or not the standard calculation of $dN/d\log Z_p$ is used with notes for those that do not use it

- what transfer function or ideal vs. real area is used
 - charge correction option
 - use, or not, of interpolation/spline and rebinning in log mobility space
- I think the conversion from Z_p to D_p is the same for all routines and need not be mentioned.
- use, or not, of interpolation/spline and rebinning in log mobility diameter space
 - which corrections are applied: CPC efficiency, diffusion losses

We simplified table 2 for the transfer functions to minimize the information to the transfer function, bipolar charge distribution, and

	Transfer function	Bipolar charge distribution	Interpolation (new sampling points/bins)
IFT	Real transfer probability Birmili et al. (1997)	Wiedensohler (1988)	Linear (no new discretizing)
NILU (Fiebig et al., 2005)	Stolzenburg (1988)	Wiedensohler (1988)	Yes (finer discretizing)
UHEL/FMI	Stolzenburg (1988)	Wiedensohler (1988)	NA
PSI	Real transfer probability	Wiedensohler (1988)	Yes (discrete logarithmic equidistant bins, 2^n per decade)
ULUND (Zhou, 2001)	Real transfer probability	Wiedensohler (1988)	Spline (discrete logarithmic equidistant bins)
ISAC	Real transfer probability	Wiedensohler (1988)	Cubic or linear spline (discrete logarithmic equidistant bins)
LAMP	Ideal transfer function (triangle)	Wiedensohler (1988)	NA
JRC	Ideal transfer function (triangle)	Wiedensohler (1988)	Linear spline (discrete logarithmic equidistant bins)
UMN	Stolzenburg and McMurry (2008)	Wiedensohler (1988) & Fuchs charging theory (Fuchs 1963, Hoppel	NA

		and Frick 1986)	
PKU	Ideal transfer function	Wiedensohler (1988)	Linear spline in scale of logDp
TNO (IFT inversion)	Real transfer probability	Wiedensohler (1988)	(discrete logarithmic equidistant bins)
UBIR (TSI inversion routine)	Ideal transfer function	Wiedensohler (1988)	NA

Table 2: Why did some routines correct for CPC counting efficiency while others did not, even though the routine had the option to do so? Particularly, why is the IFT routine correcting for neither CPC efficiency nor diffusion losses when the IFT-TDMPS system, which uses this routine, supposedly measures down to 3 nm where these corrections must be quite significant?

In some routines, counting efficiency curve can be included. For the IFT systems, corrections for diffusional losses and CPC counting efficiency are always done by post-processing. We will harmonize the procedure so that all correction should be in a post-processing stage. Anyway, as shown in the new table, we do not mention the CPC correction anymore.

Table 5a,b: As noted above, CPC inlet and sample flow rates need to be recorded (mandatory). Also as noted above, raw CPC counts should be recorded. If this is not done, then Standard Deviation – raw number concentration (in cm⁻³) should be recorded. Due to low count rates, this is often of far greater importance than Standard Deviation – median particle mobility diameter (in nm).

At this point, only the sample flow rate is recorded. We assume that the sample flow rate will significantly change, if the CPC flow rate changes in a closed-loop system (common now). Furthermore, we wanted to avoid another capillary between the DMA and CPC to avoid additional losses.

It is correct that the variability of the particle concentration across the set point, e.g. recorded at 10 Hz resolution, will contribute a much larger uncertainty to the result than the uncertainty of the median particle diameter selected at the set point. However, it is a rather philosophical question what the fundamental time resolution of a DMPS or SMPS is and how many counts should be collected for stating a significant sample. The raw 10 Hz particle concentration will always be dominated by Poisson statistics independently of the actual, atmospheric variability. The Poisson uncertainty can always be re-inferred from the reported concentration and the sample flow rate. The recommendation focuses on the atmospheric variability, which can be deduced from the particle size distributions collected during an hour.

Table 5a,b: For the last optional system parameter, Sheath air status, the recirculating flow is often driven by a blower as in the TSI SMPS and the error condition indicates the flow is not stable. I suppose “not critical” could be applicable to a sheath flow driven by a pump preceded by a critical orifice but I don’t see how it applies when using a blower.

Good point, we should change the wording to “not critical / unstable” and “critical / stable”

Changed in table 5a,b to:

Sheath air flow status, 0 – ok, 1 – flow not critical or unstable/low/high

Table 5b: Since there are essentially two independent SMPS's (except for a common inlet and synchronized scan periods) in a dual-DMA spectrometer, the Level-0 raw data should consist of essentially two independent sets of data. The start and end times are in common but all other parameters and measurements should be independently recorded for each SMPS system. In particular, there are independent temperatures and pressure for DMA1 and DMA2 as well as independent saturator and condenser temperatures for CPC1 and CPC2. These parameters have not been properly allocated in Table 5b. The only possible exception might be the DMA absolute pressure if it is measured at the aerosol inlet to the DMA. Then, one might expect there to be negligible difference between DMA1 and DMA2 such that only one sensor may be needed. Also, absolute pressure sensors are expensive. More important even than these system parameters, there should be no elimination of size bins in the overlap region of the two SMPS's at this level and the number of size bins in each SMPS data set should be recorded separately. It would not be uncommon to decide after the fact that the data from one SMPS is compromised while data from the other can be trusted. Also, the agreement, or lack thereof, of the two SMPS's in the overlap regime is an important indicator of the quality of the data. All in all, I think the data format would be more user-friendly at this level if the data from the two SMPS's were recorded totally separately as two distinct sets of data of the form given in Table 5a.

*Also here, the question whether to consider a twin-DMPS or twin-SMPS to separate collocated instruments or one instrument seems to be rather philosophical. Both options are possible, and valid arguments can be found for both of them. This article favors to see a twin-system as **one instrument** since the data are processed in the inversion routine together, to avoid the two datasets become separated, and to avoid issues with lack of synchronization.*

Table 6: Following the lead from Level-0 in the case of a dual SMPS system, I think the Level-1 data should consist of two distinct size distributions with associated system parameters plus a third data set in which the two size distributions have been combined with some sort of smoothing in the overlap region. Again, this would allow investigators to observe the degree of agreement in the overlap region and make their own decision about how much of the data to trust. This same procedure would then be followed if SMPS size distribution data were combined with size distribution data from another type of instrument such as an optical particle counter. Also, uncertainty information about the size distribution sizes and concentrations should at least be optional at this level as noted in the text at the bottom of page 5545. Are you thinking that whatever re-binning of the size distribution data there is to be done is complete at this level? This should be made clear.

We do not agree with this. 1) The two data sets are inverted together (this is different to the TSI approach, which leads to a wrong multiple charge correction of the nano-SMPS (DMPS)). 2) There is only one size distribution, and this one is then given in level-1. This is not really comparable to combining size distributions of DMPS/SMPS with OPC or APS measurements where the fundamental measurement principle differs and introduces

much larger systematic uncertainties than in a twin-DMPS/SMPS that used the same measurement principle throughout. Also, the data center doesn't perform any re-binning, but uses the same size bins reported in the level 0 file (for NRT data processing) or as reported by the data provider.

Table 7: As in Table 6, the uncertainties in size and concentration of each size bin should be listed as optional here. Note that these uncertainties are different from what is indicated by the percentiles already listed.

As before, a philosophical question whether you consider each 10 Hz CPC value at a given set-point one sample, or whether you consider one scan as one sample. For the purposes of a ground-based monitoring network monitoring aerosol properties for climate purposes, the variability within one scan is essentially irrelevant, but may be reproduced from the level 0 data if need arises. The specifications therefore include only the percentiles over one hour averaging period, which quantify atmospheric variability.

Table 8: The appropriate references should be included in the caption here. I hope you have thoroughly checked the values in this table this time.

The references are included in caption of Table 8:

Approximation coefficients for the formula describing the bipolar charge distribution for the charging states n (-2, -1, 0, +1, +2) (Wiedensohler, 1988 and Baron and Willeke, 2005)

Fig. 1: In many situations there will be a significant temperature difference between outside and inside the field station. Care must be taken to assure the sample aerosol flow has acclimated to the inside temperature prior to drying and measuring its flow and temperature. Nothing has been said about this here or in the text. The RH/T sensor in the sample aerosol flow will help determine that this acclimation has occurred; however, perhaps it should be upstream of the Nafion dryer or, at least, the flow measurement which is temperature dependent. As noted earlier, the RH/T sensor in a DMA outlet flow (sheath out here) should be used to represent conditions inside the DMA. Thus, the sheath out sensor should not be moved to the sheath in line as the aerosol in and sheath in flows may not have the same RH and T such that the RH and T inside the DMA is something in between these two.

We agree, the sensor should be in the excess air. We changed Figure 1.

Fig. 1: For a recirculating sheath flow, the dryer in that loop is not really needed since all the sheath air ultimately comes from the already-dried aerosol sample flow. By putting a dryer in this loop there is the potential of the entering sheath flow being drier than the entering aerosol flow. Thus, a particle may shrink slightly while traversing the DMA, resulting in some uncertainty in its size. Of course, if the system is started up with moist air somewhere in it, it will take longer to get all the flows dry if the sheath air dryer is eliminated. Drying of the HEPA filters, which are often hygroscopic, also needs to be considered here. These things should be addressed to some extent in the text and reasons provided for preferring one dryer configuration over the other.

Text was changed to:

The recommended set-up includes dryers to reduce relative humidity in the aerosol sample and sheath air flows. The dryer in the sheath flow helps to avoid to startup measurements with moist air somewhere in the DMA and the relative humidity become more stable. Furthermore, it would take longer to get all the flows and HEPA filters dry if the sheath air dryer is eliminated.

Fig. 3: These efficiency curves do not show a lower tail that is typical when using DMA aerosols for the calibration as they are not really monodisperse. Were these curves somehow corrected for the finite dispersion of the DMA challenge aerosols?

No, they were not correct for the DMA dispersion. We added following sentence. We did no correction for particle dispersion in the DMA.

Fig. 3: The TNO SMPS (TSI 3034) uses an internal CPC that is almost identical to a TSI 3010, as indicated. However, according to the manufacturer, the dT for this CPC is 25, not 17. This is clearly in agreement with the grouping of the TNO CPC performance curve with those of the other CPCs with $dT=25$, rather than with the curves of the CPCs with $dT=17$.

The temperature difference is corrected in the Figure. We modified the text to: Four CPCs were set to a temperature difference of 25°C between saturator and condenser, while the other CPCs were operated at the nominal, factory-set temperature difference of 17°C or 25°C for the TNO-CPC.

Fig. 4: Your reasoning about the acceptable spread in the peak locations in this graph is incorrect. First of all, curves (possibly quadratic) should be fitted to the regions around the peaks of the data curves to get a more accurate estimate of the true peaks of the responses. Then, note that all spectrometers see the same PSL and therefore their peaks should register within $\pm 1\%$ of their mutual mean according to your estimate of instrument uncertainty. That mean, in turn, should lie within 2.5% of the nominal PSL size according to the manufacturer of the PSL. Leaving aside the TNO instrument, the mean of the other curves appears to meet the latter condition. However, it appears that the spread in the peaks, even after determining the peaks more accurately, will significantly exceed the estimated $\pm 1\%$ instrument uncertainty. Thus, you need to reevaluate the level of instrument uncertainties, particularly the sheath flow calibrations.

We could not reap the flow calibration and don't know why they were of by max 3%, since they were done by the individual users. We corrected to test to following:

Inverted number size distributions of 200 nm PSL spheres are plotted in Figure 4. As indicated above, the width of the measured latex particle distribution depends mainly on the ratio of the aerosol to sheath air flow rate (see legend of Figure 4) and on the standard deviation of the size of the PSL particles. The centroid diameters of all mobility size spectrometers were within the uncertainty range of 3.5%. However, if the sheath air flows of all instruments were within $\pm 1\%$, all peak diameters should be also within this range. However, it seems to be that there were uncertainties up to $\pm 3\%$ of the sheath flows. The TNO instrument clearly underestimated the 200 nm PSL size due to an erroneous sheath air flow rate, which we recognized and adjusted after the experiments..

Fig. 4: The widths of the peaks in this graph appear to be scattered somewhat randomly according to theory. For ideal DMAs (i.e. no mixing) the width of each peak should be somewhat greater than the width of the corresponding transfer function of the DMA that produced the peak. The full relative width at half maximum of the transfer function in mobility space is just the ratio of the aerosol flow to the sheath flow. For most of the DMAs here that ratio is either 0.1 or 0.2 while that for the TNO instrument is 0.25. Given that 200 nm is in the transition regime, the corresponding relative widths of the transfer functions in diameter space should be a little more than half of that in mobility space. The relative widths of the peaks in this graph should then be a bit more than that according to the width of the PSL distribution. In particular, the UHEL instrument has a flow ratio of 0.2 so its peak in this graph should have a relative width of a little more than 0.1. For the unsmoothed version of the curve here, it looks like the width may be fairly close to that predicted. The UHEL curve has the smallest width of all the measured peaks here. However, theoretically the FMI, UBIR and PSI curves should be even narrower because their flow ratios are half that of the UHEL instrument. Yet, they are significantly wider. Other instruments with flow ratios of 0.2 also have response curves that are significantly wider than that of the UHEL instrument. This graph, then, is clearly giving us a measure of the non-ideality (i.e. mixing) of these DMAs similar to what a TDMA experiment would do. This should be discussed to some degree in the text as well as its implications for measurements of size distributions.

I cannot explain this. So I don't discuss this in the article.

Fig. 4: As noted earlier, this measurement of certified PSL by these DMAs is not a calibration of size for the DMAs and should not be referred to as such in the caption of this figure. Also, you refer to "modified ... spectrometers". What modifications are you referring to?

We corrected the to ro following:

A mandatory part of a performance check of a mobility spectrometer concerns the sizing accuracy. This task can be accomplished using monodisperse PSL (polystyrene latex) spheres, whose diameters are certified by the manufacturer to be within +/-2.5% of the nominal diameter. For a DMA size check, monodisperse PSL spheres of one or more particle sizes should be used

Fig. 8: As noted earlier, "The shaded area marks the +/-20% range ..."

corrected

Technical Corrections

page 5522, note 8: Table 9 lists this organization as "University of Lund"; make them the same.

corrected

p. 5526, line 10: "...TSI Inc. (Shoreview, MN, USA) and or Grimm GmbH ..."

corrected

p. 5526, In. 17: Define acronym here. "... World Meteorological Organization (WMO) issued ..."

corrected to:

(WMO-GAW Report 153, 2003; World Meteorological Organization-Global Atmosphere Watch)

p. 5526, In. 19: Add date to citation. "... (WMO-GAW Report 153, 2003), but ..."

see above

p. 5526, In. 23: This reference is missing. "... (Dahmann et al., 2001)."

Added:

Dahmann, D., G. Riediger, J. Schlatter, A. Wiedensohler, S. Carli, A. Graff, M. Grosser, M. Hojgr, H.-G. Horn, L. Jing, U. Matter, C. Monz, T. Mosimann, H. Stein, B. Wehner and U. Wieser (2001). Intercomparison of mobility particle sizers (MPS). Gefahrstoffe-Reinhaltung der Luft 61(10), 423-428.

p. 5527, In. 5: If possible, please give an English user-friendly reference for this. "... a custom-made TDMPS (Twin-DMPS) ..."

corrected to:

TDMPS (Twin-DMPS; based on Birmili et al, 1999)

p. 5529, In. 2: This needs a reference. "... the equilibrium ion pair concentration, which is believed to be approximately 10^7cm^{-3} ."

I don't have a real reference to it.

p. 5529, In. 8: "... annular slit closed to the outer electrode ..."

??

p. 5529, In. 18: "... to an annular slit ..."

corrected

p. 5530, In. 5: The year for the Schladitz et al. reference conflicts with that given in References.

Corrected to:

(Birmili et al., 2008; Schladitz et al., 2009).

p. 5530, In. 7: Correct the spelling of "Knutson".

Corrected

p. 5530, In. 20: Some SMPS inversion routines also calculate particle diameter from DMA voltage. Reword this sentence to accommodate calculation in either direction.

Corrected to:

It is also essential to calculate the DMA voltages for the actual temperature and pressure conditions encountered at the field station (or particle mobility from a given voltage).

p. 5531, In. 1: "Ambient air samples can increase considerably their RH ..."

p. 5531, In. 7: "...minimizes diameter changes due to hygroscopic growth to typically ..."

corrected to:

which minimizes diameter changes due to hygroscopic growth to typically less than 5%

p. 5532, In. 8: I have difficulty properly interpreting this sentence. "... the aerosol flow should not deviate systematically more than 5% from the set-point on daily average." Perhaps you could reword it.

Corrected to:

As a guideline for quality control, the continuously recorded aerosol flow should not deviate more than 5% from the set-point

p. 5534, In. 3: The correct year of the Wiedensohler reference is (1988).

Corrected to:

This bipolar charge distribution is based on the approximation formulas and the Gunn equation (1956) in Wiedensohler (1988) with corrections to two approximation coefficients published in Baron and Willeke (2005) and also listed in Appendix A5.

p. 5534, In. 26: "... around the results of the IFT inversion routine."

Corrected

p. 5535, In. 1: "To see small differences as a function of particle size, ..."

Correted

p. 5535, In. 6: "Note: the "Old Grimm" inversion routine is not any used anymore ..."

corrected to:

Apart from "Old Grimm", which is not used anymore, all inversion routines agreed within 5% over the size range 5-300 nm.

p. 5536, In. 4: "... in the frame of the third EUSAAR DMPS/SMPS intercomparison workshop ..."

corrected

p. 5536, ln. 5: "... are shown here as an example."

corrected

p. 5536, ln. 8: "... we achieve after the DMA ..."

corrected

p. 5536, ln. 9: This sentence would be clearer if written as "Due to the relatively small degree of polydispersity of the aerosol from the furnace ..."

Changed to:

Due to the relatively small degree of polydispersity of the aerosol from the furnace generator

p. 5537, ln. 3: "... may also differ by a few nanometers ..."

corrected

p. 5537, ln. 15: The acronym ACCENT has been previously defined and adding "programmes" at the end makes it read better. "... in the frame of the EUSAAR, ACCENT, and WMO-GAW programmes." You should really decide if you are going to use American or British English spellings of words. e.g. program/programme appears both ways.

I harmonized the text to American English and added programs

p. 5538, ln. 2: "... mobility size spectrometers ..."

corrected

p. 5539, ln. 1: "... and the TNO instrument works well again."

Reformulated to:

The TNO instrument clearly underestimated the 200 nm PSL size due to an erroneous sheath air flow rate, which we recognized and adjusted after the experiments.

p. 5542, ln. 1: "... represents the accuracy ..."

corrected

p. 5542, ln. 2: "... the current the state of knowledge ..."

changed to:

This unit-to-unit variability of technical identical mobility size spectrometers represents the accuracy that we are able to achieve the current state of knowledge and technology under controlled laboratory conditions

p. 5543, In. 17: Give the model number(s) of the ^{85}Kr bipolar chargers. I assume they are from TSI. They have a variety of configurations for their ^{85}Kr chargers having very different losses I would imagine.

Changed to

particle losses for sub-10 nm particles across ^{85}Kr bipolar chargers (IFT custom-made)

p. 5549, In. 15: What is the “absolute detection efficiency” of a CPC?

Changed to:

the plateau detection

p. 5549, In. 16: “... checked on a monthly basis.”

corrected

p. 5550, In. 2: “... is not available, also the total number ...”

changed to:

If a reference spectrometer is not available, also the total number concentration measured by a CPC can be compared to the number integral of the size distribution.

p. 5550, In. 10: “... for several size distributions.”

corrected

p. 5552, In. 2,3: “... maximum variation around setpoint: +/-_% ...”

corrected

p. 5553, In. 14: “... median values of the averaged the particle number size distributions ...”

corrected to:

median values of the averaged particle number size distributions

p. 5555, In. 19: “... counts can either been accumulated ...”

corrected to:

can either be accumulated

p. 5556, In. 1: “... determined by the geometric mean ...”

corrected

p. 5556, In. 6: “... data format indicated as particle number ...”

corrected

p. 5556, In. 9: "... for the institutes or and universities ..."

corrected

p. 5562, Table 1, third entry line: The reference cited here "(Jiang et al., 2011)" has a different year than in the References.

Its 2011

p. 5562, Table 1, fourth and fifth entry lines: Given that the nominal lengths of these straight Permapure Nafion dryers are 24" and 12", respectively, I am surprised that their equivalent lengths are more than four times longer. Is this correct? In that the equivalent lengths scale directly with the actual length, this would indicate that the losses occur within the tubes rather than at the entrances and exits. I would not think the slight curvature of the Nafion tube within its sheath would be enough to enhance the losses that much. Were these measurements possibly done with charged particles because there are probably some significant static electric fields associated with the Nafion?

Its 1.25 and 2.5 m

p. 5562, Table 1, last entry line: Give the type (e.g. 85Kr) and model number of this charger.

Added:

IFT custom-made

p. 5563, Table 2, IFT Input: According to Table 3, the IFT-SMPS (scanning) and the IFT-TDMPS (stepping) both use the IFT software. Therefore, shouldn't there be an IFT input corresponding to "- Selection: Scanning or Stepping"?

No, we have the same "raw data" format and can just use the same inversion software

p. 5563, Table 2, IFT Input: "- Consider of simple dynamic shape factor ..."

Table was simplified

p. 5563, Table 2, IFT Calculation: The order of adjectives makes a difference here. "Equidistant" operates on the "logarithmic" mobility scale so it should come first. Also, doesn't the interpolation come first followed by splitting into bins? Thus, "- Interpolate and split into discrete equidistant logarithmic mobility ... bins". This also applies to similar entries in the ULUND, ISAC, JRC and TNO sections of Table 2.

Table was simplified

p. 5563, Table 2, NILU Calculation: "- Total DMA DMPS? counting efficiency ..." DMAs don't count.

Table was simplified

p. 5563, Table 2, NILU Calculation: “- Charge correction (Wiedensohler, 1988)”

Table was simplified

p. 5563, Table 2, NILU Calculation: Write as “- Problem formulated ... by discretizing (finer) in particle size independently of number of measured size bins” or put a closing parenthesis at the end of this entry.

Table was simplified

p. 5564, Table 2, UHEL/FMI Input: It is customary to write (dependent variable) vs. (independent variable). Thus, “- DMA Number concentration versus DMA voltage”

Table was simplified

p. 5564, Table 2, UHEL/FMI Calculation: The inversion method is a choice; it belongs in the Input section. A note explaining a bit about how these different inversion methods work or differ would be helpful.

Table was simplified

p. 5564, Table 2, PSI: In the Input section it is indicated that the CPC efficiency correction was not applied for this study. However, the Calculation section indicates that the raw measurements were corrected for the CPC detection efficiency. Which is correct?

Table was simplified; It was not applied in this study

p. 5564, Table 2, **Note: The note is not an input; it should not be preceded by a dash “-”.

Table was simplified

p. 5564, Table 2, *Note: A note corresponding to “*” needs to be included. Apparently, it marks the selection used in this paper.

Table was simplified

p. 5565, Table 2, ULUND Input: Are the charge and spline type actually choices (i.e. negative or positive, and cubic or linear, respectively), or are they fixed at the indicated values and therefore not really inputs? The same question applies to the last four entries of the JRC Inputs on the next page.

Table was simplified

p. 5565, Table 2, ULUND Calculation: It is indicated that $dN/d\log Z_p$ is calculated using the ideal width of the DMA transfer function. However, the usual method of calculating $dN/d\log Z_p$

assumes that the width is negligible compared to the variation of the other factors within the integrated response function for the CPC concentration and therefore it does not enter into the final expression for $dN/d\log Z_p$. Thus, use of the ideal width should be irrelevant to the calculations unless they are not using the standard equation for $dN/d\log Z_p$. If the latter is the case, then more information is needed here.

Table was simplified

p. 5565, Table 2, ISAC Input: Insert comma between “count number” and “counting time” for the first entry.

Table was simplified

p. 5565, Table 2, ISAC Input: The last entry in this section is actually a note that applies to the entire column. It should be placed at the bottom of the column as “* Selected for ...” with no dash “-” in front of it.

Table was simplified

p. 5565, Table 2, LAMP Calculation: “Integrate between set mobility’s transfer functions”. This does not make any sense. Instead of “Integrate”, do you mean “Interpolate”? What would be the purpose of integrating between transfer functions? Nor does it make sense to integrate across an ideal transfer function as that is already a known quantity.

Table was simplified

p. 5566, Table 2, UMN Calculation: Was the transfer function actually “experimentally determined” or was it actually just determined to be within experimental error of the ideal width for larger particles? Again, only the area or penetration is used in the standard linear inversion method; other details of the shape of the transfer function (e.g. width) are not used.
p. 5566, Table 2, PKU Calculation: “Iterative” is misspelled.

Table was simplified

Table 3: Table 3 also has a problem with consistency of notation. The three critical dimensions – R1, R2, L – should be given for all DMAs, including the commercial ones. There are many “Hauke-type” DMAs listed here with identical, or nearly so, dimensions. Are all these actually manufactured by Hauke or some other single entity? If so, is there a model number associated with these that could be referenced for all these different systems that seem to be using the same DMA? If so, then perhaps the dimensions need only be given once.

The Hauke-DMA used in the European network are always custom-made copies manufactured by different commercial or institute machine shops. I guess the dimensions are the same. DMA dimension are added

Table 3: Though most are listed in Fig. 3, the dTs for all CPCs should be listed in Table 3 since many use a non-standard dT.

added

Table 3; If the 85Kr bipolar chargers are from TSI, give the model numbers. If any of the other chargers are commercially manufactured, give those model numbers as well.

changed

p. 5568, Table 3, IFT-SMPS: "... heat exchanger for closed loops, ..." (only one loop)

corrected

p. 5568, Table 3, NILU-DMPS: Is this DMA really 285 mm long? Just 5 mm different from that of the IFT, UHEL, FMI, JRC and ISAC setups? Otherwise, this would seem to be the same DMA.

corrected

p. 5569, Table 3, UHEL-DMPS: There is no entry for Software. Table 2 indicates that it is the same as the FMI-TDMPS system. This should be listed here.

added

p. 5569, Table 3, PSI-SMPS: This DMA appears to a TSI 3081 with a slightly shorter length. The radial dimensions given are actually the inner and outer radii, not diameters as stated. To be consistent with the other entries of DMA dimensions the true diameters (x2) should be given. Also, "... heat exchanger for closed loops, ..." (only one loop)

corrected

p. 5569, Table 3, FMI-TDMPS: Judging by the length, it appears that DMA1 may actually be a Hauke 3/150. If that is the case, please give that model number. Also, " FUG positive power supply (x2) (up to ...)"

all corrected, however no model number

p. 5570, Table 3, JRC-DMPS: "... heat exchanger for closed loops, ..." (only one loop)

corrected

p. 5570, Table 3, LAMP-DMPS: This DMA appears to be a TSI 3081 as the dimensions are exactly the same except for the precision given. If that is the case, state this model number. Also, use the same precision as the manufacturer in stating the radii and give them as diameters to be consistent with the other entries in the table.

corrected

p. 5570, Table 3, NUIG-DMPS: I believe the manufacturer of the high voltage power supplies is actually “Bertan” as I could find no listing of the given company name on-line.

cancelled

p. 5571, Table 3, ISAC-DMPS: The radial dimensions of the DMA should be given as diameters for consistency. Also, “Other hardware: aerosol flows: Nafion dryer; ...” (Presumably only the sample aerosol flow (singular) has a Nafion dryer.)

corrected

p. 5571, Table 3, TNO-SMPS and UBIR-SMPS: Though these are commercially-built systems the information given about them should be as complete as possible, including the measured size range, the DMA dimensions, the fact that the CPC in the 3034 is essentially a TSI 3010 as it is listed in Fig. 3 and details of Other Hardware that is part of the base TSI system.

corrected

p. 5576, Table 7: The column for the 15.87 percentile should read “Particle size Bin 01 dN/dlogDp 15.87 percentile (cm-3)”

corrected

p. 5578, Table 9: This table should also provide information about the location of each of these entities. Most of them, though not all, have corresponding entries in the footnotes providing affiliations of the authors on pages 5521-5522. Perhaps this could be used somehow to reduce the duplication of information. The footnote number could perhaps be included in this table with additional footnotes on this page for those entities not associated with any author. As mentioned earlier, the University of Lund or Lund University should be listed the same way in both locations. Also, the acronym TNO should be defined.

All corrected

p. 5579, Fig. 1: The symbols used for the critical orifice and pump are not obvious in terms of which is which. Therefore, the labeling of these should clearly indicate that the orifice comes before the pump. Also, the absolute pressure measurement between these two as noted in the text should be shown in this figure.

improved

p. 5580, Fig. 2: The acronym “BOL” in the key of Fig. 2a needs to be defined. The symbols used on some of the non-IFT curves in the upper graphs should also be used on the corresponding curves in the lower graphs, making it easier to distinguish one curve from another. The last sentence of the caption should read “The lower graphs show the ratio of the ...”. These graphs are rather small and some of the grid lines are fairly faint such that I had to use a magnifying glass to see what I wanted to see.

Corrected and improved

p. 5582, Fig. 4: More use of symbols on the curves would make it easier to distinguish between those that have very similar colors, such as IFT REF 1 and LAMP. Also, in Fig. 3 it is IFT SMPS whereas here it IFT REF 1. Are these actually different instruments? If not, use the same notation.

corrected

p. 5583, Fig. 5: “Intercomparison of the measured particle number size distributions ...”

corrected

Reviewer Fred Brechtel

Overall this is a very good paper describing the key issues involved with performing high quality electrical mobility size distribution measurements. Some very practical solutions are offered and the suggestions will likely find broad use in the community of users of this technology.

Thanks

On page 5536 related to the discussion of the CPC size dependent counting efficiency method: when the location of the electrometer on the manifold is switched with one of the 10 CPCs, is the 'correction' required to the electrometer data the same to match the defined 100% CPC efficiency at 40 nm?

We checked if the position of different CPC or the electrometer would change something. It did not.

The magnitude of the correction that needed to be applied to the electrometer data for the CPC efficiency studies should be stated.

Added to this sentence:

Herewith, we determined a gain error of the electrometer of 7%, which was used to determine the CPC counting efficiencies for the size range smaller than 40 nm.

Discussion of maintenance, especially cleaning of the DMA column and flow measurement components?

Correted to:

The DMA and the laminar flow element to determine the aerosol flow rate have to be cleaned once per year.

The results reported on page 5540 of the ambient measurement intercomparison are disappointing. After all the effort expended to upgrade the various instruments according to the proposed standards, the poor intercomparison for sub-20 nm particles and particles larger than 200 nm suggests there is something we still do not understand about operating mobility spectrometers for long-term monitoring applications. Is it possible that the different aerosol neutralizers produced different charge size distributions? This may not have been evident in the PSL studies since the PSL are 'monodisperse' and presumably a multiple charge correction is not applied.

Fred, you are completely right, we have been also disappointed. However, we are also happy that we have been successful for the size range 20-200 nm. In the next infrastructure project ACTRIS, we will concentrate on the nucleation mode. It is likely that the method for the correction of diffusional losses is not sufficient enough. We added a correction for enhanced losses in bends. For the accumulation mode, we checked if the individual inversion routines may do slightly different multiple charge corrections. This is not the case,

might be that different neutralizers may do not produce the same charge distribution for larger particles.

In the processing of our mobility spectrometer data, we propagate measured uncertainties in air flow rates, voltage, temperature, and pressure through the equations defining the size selection process in the DMA to try and establish the 'actual' measurement uncertainty in particle size. Scanning mobility spectrometers are unique in that detected counts by the CPC must be related to a correct size by understanding the plumbing time between the DMA and CPC and CPC smearing time constant. These parameters were not discussed - especially the CPC smearing time constant/delay time should be mentioned. How was this measured?

Since we recommend performing 200nm PSL scans to obtain the size accuracy, one can see uncertainties either in the sheath air flow rate or plumbing time. Also the scan time should be discussed.

Wrote following now:

Appendix:

In case the 200 nm PSL sizing check results in a deviation larger the 3.5%, either the sheath air flow or the plumbing time (in case of a scanning mobility particle size spectrometer) is wrong and has to be carefully checked again.

Technical description:

For scanning mobility particle size spectrometers, we recommend to have the scanning time (up or down scan) minimum 2 min to avoid smearing effects in the particles counters with a relatively slow responds time. These smearing effects can cause e.g. significant false measurements of the slope towards larger particle in the accumulation mode range.

On page 5541, five systems were intercompared and much better agreement found. Was the only difference between this study and the previous one the more careful air flow rate calibrations? Was the same charge neutralizer design with similar age sources used on all 5 systems during the Aug 2010 study? Were different neutralizers of different source ages used on the instruments in the previous studies where the intercomparisons were not as good?

We changed this chapter completely now. We introduced the reference mobility size spectrometers and their accuray

Alternatively, even though the different inversions were tested on the same input distribution, it may be that the way in which the multiple charge correction is implemented in the different systems could cause some of the poor measurement intercomparisons. The magnitude of the error resulting from an incorrect charge correction algorithm will likely depend on the shape of the size distribution. For example, if the particle counts in the measurement intercomparison had relatively more large particles compared to the distribution used during the inversion test exercise, then perhaps a problem with how multiple charging is dealt with in the processing could have a relatively larger influence on the measurement intercomparison.

Using the same inversion routine did not provide better results

On page 5546 line 25 I would recommend: 'Most important are regular system maintenance and checks and calibration of ...'

Corrected

On page 5547 line 20 I would recommend: '..we can reach uncertainties around 10% with...' (the word 'that' should at least be changed to 'than')

corrected

On page 5548 line 22 it is stated that mass flow meters on the aerosol flow should be calibrated at least twice per year but previously it is stated that mass flow meters should not be used on the aerosol flow - please be consistent.

Corrected to:

Pressure transducers to measure the aerosol flow rate or mass flow meters used to determine the sheath air flow rate have to be calibrated at least twice a year.

Grammatical corrections:

page 5535 line 6 should read "Note: the 'Old Grimm' inversion routine is not used in newer software revisions."

Now:

Apart from "Old Grimm", which is not used anymore in newer software routines, all inversion routines agreed within 5% over the size range 5-300 nm.

page 5536 line 5 should read '...are shown here as an example...'

corrected

page 5537 line 26 spell out 'incl.'

done

page 5539 line 24 should read '...following section are corrected for ...'

now

To take into account particle losses by diffusion in the mobility size spectrometers, we agreed on the standardized method of the "equivalent length" as described above. All size distributions shown in the following section are corrected for particle losses following this method.

page 5546 line 21 should read '...occasionally higher than...'

corrected

page 5548 line 15 should read '...adopted by manufacturers and ...'

done

page 5549 line 8 remove the word 'Please'

done

page 5566 table 2 under 'PKU' listing, the work 'Iterative' is misspelled

Table was completely changed

Reviewer 1:

This paper reports on laudable efforts to harmonize DMA measurements of particle size distributions in long-term, ground-based, stationary measurements. It represents the results of European groups to reduce ambiguity in aerosol measurements, and may be considered as a reasonable first start at developing a truly international approach to DMA measurement standardization. That said, the harmonization standards reported here are incomplete, and will ultimately work against efforts to standardize or harmonize DMA measurements in general. The approach is highly prescriptive, which may be appropriate when all measurements are being performed for similar reasons. However, DMAs are used far more broadly than in long-term stationary measurements. Some of the prescriptive elements, such as measuring the size distribution at low relative humidity or time averaging data to report 1 hour average size distribution measurements will, if applied too broadly, confound efforts to understand important aspects of the atmospheric aerosol. Others are convenient approximations that may be incorrect in some important applications of DMAs, e.g., in source characterization studies, or in characterizing aerosols in a number of technological and fundamental studies where the gas is not air. I support the effort that this manuscript represents, but hope that it is a predecessor to a more open consideration of the full range of applications of mobility measurements. In the discussion below, I raise a number of points in the order that they appear in the manuscript. I finish with a number of points that are omitted entirely from the paper.

Thank you very much for your critical comments. We are completely aware of many other applications of mobility size spectrometers. However, this paper is not meant to give an overall recommendation for all kind of applications. As the title already emphasizes; the recommendations are only valid for long-term atmospheric observations e.g. at GAW- or European super-sites to determine the number size distributions in a comparable manner.

One critical issue is that, in addition to the references that may be of immediate use according to the proposed prescription, primary references should be cited. For example, on p. 5530, l. 8, Wiedensohler's fit to the Fuchs/ Hoppel & Frick model is cited, but not the primary references. The Wiedensohler fit is a very powerful tool which I also use, but it only applies to dry air under relatively mild conditions. Variations in relative humidity, temperature, pressure, or gas composition may alter the distribution which, throughout the paper, is incorrectly labeled an equilibrium charge distribution.

The philosophy to determine comparable number size distribution is to measure them at dry conditions and at room temperature as described in the recommendations. The only question left is if another pressure would change the bipolar charge distribution. There is no paper in the literature, which describes such empirical study. However, the mobility of the ion is a function of $1/p$. The "Gunn bipolar charge distribution" uses e.g. the ratio mobility of positive to negative ions to calculate the fraction of the different charging states; and this would not lead to another bipolar charge distribution since the pressure dependency is cancelled out. Presently, it is thus common to use the known bipolar charge distribution for different atmospheric pressures. Recently, we employed a SMPS at the new GAW site at 5200 using the standard bipolar charge distribution, and the closure between total number CPC and

SMPS were still within 10%. This indicates that the bipolar charge distribution can be used also for lower pressures

It is actually a representation of a steady-state charge distribution which differs dramatically from that which one would predict at thermodynamic equilibrium, especially in the low nanometer size range. The charge distribution is described as an equilibrium distribution at several points in the text, especially on p. 5528. This will suggest to some readers that the Boltzmann equilibrium distribution should be applied, exacerbating a long standing problem in the interpretation of aerosol mobility measurements. The authors should err on the side of caution in their choice of terminology, especially in a paper that is attempting to improve the quality of data and data reporting.

Thanks for this comment. Although the Boltzmann bipolar charge distribution is not used in mobility size spectrometers since 20 years anymore, we will mention in the text that the Fuchs and Gunn distribution is meant.

Added:

This bipolar charge equilibrium can be described by the work of Fuchs (1963) and Gunn (1956).

For CPC counting efficiency curves, only Birmili and Hermann are cited; why not Agarwal and Sem, Stolzenburg and McMurry, and the many others who have provided data, and several of which are cited elsewhere in the paper.

Added:

For the DMA transfer function, only Birmili is cited. Birmili provides a mathematically convenient form of the transfer function that is not based upon the physics of the instrument. Others have provided important physically based models that can actually be used in a predictive sense, especially Mark Stolzenburg's thesis, whose model has been placed in the journal literature by Hagwood, and by Zhang and Flagan.

Added in the text and references

Another noteworthy physically-based DMA transfer function model is that of de la Mora whose analysis has motivated much of the work on optimizing DMAs for measurements in the low nanometer size regime.

You are correct, but this kind of DMA is not build for atmospheric long-term operation, and therefore it is not considered here.

The focus of this manuscript appears to be on accumulation mode aerosols, but much future work will address the nucleation mode, so the paper should provide guidance for researchers in that field.

The focus here was indeed on the Aitken and accumulation mode range from 10 to 800nm. Presently, there are only very few instruments in permanent use at observatories, which

measure also the nucleation mode range. This measurement requires an additional nano-DMA.

The authors contend on p. 5527 that no comprehensive intercomparison of custom-built mobility spectrometers has been published to date. There have, in fact, been many intercomparisons – some led by authors to this paper. The statement may be correct if one takes an exceptionally narrow definition for the word comprehensive, limiting it to studies that include all of the instruments used in some narrowly defined range of studies. This study is, in fact, not comprehensive in that a number of DMAs that are used by groups around the world are not included. There is no need for such claims to be first, or to cover new ground. This study reports a rigorous intercomparison of a wide range of instruments. That would be a sufficient claim for this rapidly changing field.

I think the reviewer is not correct in this case. There was never such an effort of a network bringing mobility size spectrometers three-times together. Standardizing them and finding a way to have them comparable in the size range larger than 20 nm. We cancelled the word comprehensive to avoid a misunderstanding.

The paper also claims that there have been no comparison of inversion routines. There have also been comparisons of data inversion algorithms, although they may not have been focused entirely on mobility analyzers. Many papers describing new algorithms include comparisons with the prior algorithms. I find little value in the intercomparison of inversion algorithms presented in this paper since no useful information is provided about the algorithms. They are merely black boxes supplied by different manufacturers or used by different groups. The nature of the algorithms is not discussed at all. This is an important issue.

The reason to do this intercomparison was not to highlight this inversion intercomparison. The reason was to find out if differences between different mobility size spectrometers with different inversion routines produce the same size distribution by providing all of them with the same measured mobility distribution. There may have been intercomparisons before, but never in this extent. We will rephrased the sentence saying that there was “no such comparison study” before.

In other fields where particle measurements are key, many manufacturers produce instruments that do the inversion internally and only report the inferred size distribution. The reported size distributions are often suspiciously free of apparent noise and remarkably consistent in the shapes of size distributions. Users cannot gain access to the raw data, nor can they examine the details of the inversion algorithms. Blind reliance on the manufacturers to produce quality instruments and data analysis has hampered the advance of the science in these fields. That same marketing philosophy may well enter the aerosol field if care is not taken to ensure that all users of aerosol instruments have full access to the raw data and data analysis methods. The recommendation that the raw data be reported is a step in the direction of ensuring such access, but more is needed. I will return to this point later.

A key conclusion of this paper is that DMA measurements should be performed only at relative humidities below about 40%. For a particular study of the type that motivated this

paper, this may make sense, but DMA measurements are made for many reasons. Measurements made at such low humidity may not, for example, be directly comparable to impactor measurements. Nor will they directly relate to optical measurements – at least, not without making assumptions about the hygroscopic behavior of the particles being measured. That said, the authors are certainly correct in stipulating that the relative humidity should be measured and reported, as should all other operating parameters.

Again, this paper focuses only on comparable long-term observations. This recommendation is not meant for intercomparison using an impactor with another humidity in an intensive field campaign. If somebody wants to do this additionally, it should be not a problem.

Figure 1 suggests an approach for generating the dry aerosol, and for dealing with a number of other operating details that the harmonization standard recommends, e.g., using a Nafion™ drier or a silica based diffusion drier, placement of RH and temperature sensors, flow meters, etc. One point in that figure that I find confusing is the inclusion of a mass flow meter in the recirculating sheath loop when they correctly note in the text that a volumetric flow measurement is preferable since it is volumetric flow rate, not mass flow rate that determines the DMA sizing characteristics.

It is mentioned in the paper that the flow meter has to be calibrated. In the case of a mass flow meter, the volumetric flow has to be calibrated as function of the mass flow at the specific site.

Recirculating sheath flow is, in theory, a good idea, but there are problems that are only obliquely addressed in the manuscript. Notably, many of the readily available pumps, and most of the blowers have a potential for leaks which can, as noted in the text, lead to biases in the measurements. The manuscript should note that, if a recirculating flow is to be used, the pump/blower must be sealed and leak testing should be part of the regular maintenance schedule for the instrument.

Added:

In case of a closed-loop instrument, the pump/blower must be sealed and leak testing should be part of the regular maintenance schedule for the instrument.

The authors do note that a critical orifice can be used to control the exhaust and sheath flows, but the figure suggests that it be used in a recycle loop. Most use of critical orifices for flow control is in an open loop operating mode which can be just as precise as the recirculating loop. When used in a recirculation loop, the temperature and absolute pressure at the entrance to the critical orifice must be known since the mass flow rate through the orifice is proportional to the gas density of the entering gas. If I interpret Fig. 1 correctly (the location of the orifice is ambiguous), this would require the addition of temperature and absolute pressure transducers between the pump and the orifice. Perhaps some discussion of that mode of operation should be included, along with recommendations about ensuring that the operation meets the desired standards.

The temperature and pressure is known from the DMA pressure and the T-measurement in the excess flow. However, since there is always a calibrated mass flow meter or a capillary, the flow is known.

Perhaps this is not a problem in Europe, but it is a problem in the USA: reference to ISO standards is of limited use to the research community since these are not accessible without paying an unreasonably large fee. ISO standards documents are not generally available in university libraries, so researchers may not have access to them. Any aspect of the cited ISO standards that is required as part of this harmonization standard should be reproduced in the paper or cited as an open literature reference.

The formulas and constants taken from the ISO 15900 are given in the appendix.

Section 4.1: Inversion routines. This section addresses a critical, and often overlooked part of DMA measurements, but provides remarkably little detail, and that is provided only for the noncommercial versions. The lack of transparent data analysis methods leaves many questions about the measurements. The discussion in this section provides a broad-brush treatment, but will be of little use to a new researcher trying to learn how to analyze their data properly. The comparison provided looks promising, but was it the result of a blind comparison, or were the inversions tuned to reproduce the desired results? I would argue that, like the details of the measurement, the details of the inversion are critical to the evaluation of any data and should be open.

There was no tuning of the data at all. All owners of the comparison got the same mobility distribution incl. TSI and Grimm, and they provided the resulting size distribution.

To this end, I would suggest that part of the harmonization should be a requirement that all data inversion algorithms be open source, readily accessible to all users and readers, and fully documented and validated. Furthermore, software version and all data analysis parameters need to be reported. Lacking this, the only way that data can be compared is to reanalyze the raw data, even if the data analysis routines are comparable under some circumstances. This will hold commercial instrument makers and users of many custom-built instruments to a higher standard than that to which they have been accustomed, but many artifacts have been generated with poorly documented and, often, inaccessible, data analysis software. Too often, even the group that produced a data set cannot reproduce the data analysis some time later. The CPC counting efficiency curves and charging probabilities are critical components of the measurements. This is only part of the CPC characterization that should be reported. The counting efficiency curve is an adequate characterization of CPC performance only for DMA operation in the slow stepping mode. Proper interpretation of measurements from scanning mode operation (SMPS) or fast stepping mode operation (the so-called sequential method of GRIMM) require knowledge of the time response of the detector since particle counts that nominally are attributed to one particle mobility will, in actuality, result from that channel and several previous ones.

I fully understand the concerns of the reviewer, but I am afraid this requirement is beyond what we can provide so far. We are at the beginning of a process which also includes sustainable capacity building. This means also to keep the expertise at each data provider.

Our philosophy is also to compare instruments in intercomparison or on-site workshops as part of the quality assurance.

Sizing performance can be validated using authentic size standards such as PSL. The approach proposed makes sense, including the emphasis on relatively large particles. However, some instruments may be optimized for size ranges that make the specification of an arbitrary size standard such as 200 nm PSL inapplicable to some instruments. Rather than a prescriptive statement that imposes a particular size standard, the paper should be descriptive in the approach. It may remain prescriptive in terms of documentation and the reporting of all relevant parameters, operating conditions, and data.

Again, we focused on mobility size spectrometers for the size range 10-800 nm. The decision to use 200 nm is practical. 200 nm latex particles are produced by a nebulizer. So, if we don't want to have an interference with the salt peak below 100 nm, the particles have to be larger, but still to be small enough to have a sufficient number concentration.

Number concentration standards remain the weak link in DMA measurements so the discussion of the particle number concentration is welcome. The need for a reference instrument is critical. The approach outlined for harmonizing these models will not ensure that the data are correct, but will ensure that the biases and errors are consistent. This does not solve the problem, it only postpones addressing it.

Exactly

The discussion of particle losses is reasonable for the most part. Losses in the DMA misses a critical point: at the transition from high voltage to grounded inlet or outlet, losses due to electrophoretic migration augment those due to diffusion (KOUSAKA et al. J Chem. Eng. Japan (1986) 19: 401-407; Zhang and Flagan. J Aerosol Sci. (1996) 27: 1179-1200). This is neglected in the discussion of losses in the DMA. Some, but not all, instrument designs take these additional losses into account.

The reviewer is not correct at this point. All empirical loss functions of the DMAs suggested here include losses due to the transition from high voltage to grounded inlet or outlet automatically.

To the list of data that should be included in the archival data records, I suggest some additions. First, the raw data that are recorded should be truly raw data, not data that has been processed through any kind of inversion software. Ideally, the absolute counts recorded in each channel, the flow rate through the detection region of the CPC, and the counting time interval would be recorded. This enables one to assess the role of counting statistics in the reported noise.

I understand this point, but this we agreed on to avoid this. Especially, since some DMPS programs use a variable counting time for different size depending on the concentration. Therefore, we agreed on the number concentration in each bin. This is a decision we would not like to change. The advantage of this harmonized "raw data" is that everybody can

recalculate the number size distribution without knowing any settings of the individual system.

Reported CPC concentrations are less useful since they mask key parameters needed for proper data inversion. Basically, the recorded raw data should be sufficient to allow reanalysis of the data from scratch, thereby enabling validation of whatever data analysis/inversion routines are employed. In stepping mode operation, the delay time between voltage steps and the start of counting should be recorded since finite detector response time may lead to data analysis biases that can be corrected by data deconvolution. The response time of the plumbing (and electronics in the case of electrometer detection) between the DMA outlet and the detection point within the CPC or other detector may also lead to biases, so these data are also needed for proper data analysis. CPC dead time corrections and methods, and firmware and software versions for all instruments and data acquisition and instrument components should be included in the metadata.

Everything what is useful should be in the metadata, I agree. Since we recommend frequent latex particle scans, the plumbing time should not be a problem. If there is an error, the flow has to be checked. This is mentioned in the recommendation.

References to archival literature that provide such data may be appropriate.

Where multiple instruments are used in parallel, redundant channel data should not be removed from the raw data. I further question the recommendation that redundant channel data be removed from Level-1 EBAS data. This may mask problems with the data.

Since level-1 are final and combined size distribution (inverted at once), there should not be redundant channels.

Specific points: 1. Check all references and reference citations. At least one reference (Dahmann et al., 2001) is missing from the bibliography.

added

2.p. 5536, l. 9: "Due to the relative small polydisperse aerosol ..." is unclear. What is small?

Changed to

Due to the relatively small degree of polydispersity of the aerosol from the furnace generator (geometric standard deviation of 1.3-1.4),

3.p. 5538, l. 28: suggest changing "200 nm PSL size, which was" to "200 nm PSL size; this was"

new sentence:

The TNO instrument clearly underestimated the 200 nm PSL size due to an erroneous sheath air flow rate, which we recognized and adjusted after the experiments.

4.p. 5539, l. 6: spectrometers (plural)

corrected

5.p. 5541, l. 10: delete “again.” This is a different range than the previous shaded band.

corrected

6.p. 5541, l. 20: references missing

New paragraph

7.p. 5543, l. 10: Delete section title 4.4.1, and combine next sentence with the previous paragraph. This is a continuation of the previous train of thought. Put the paragraph break before “A more widely ...”

new paragraph

8.p. 5544, l. 6: I suggest indicating that a manufacturer’s calibration curve “can be applied with caution.”

Changed to:

If experimental data on the counting efficiency of a particular CPC are not available, the manufacturer’s calibration curve can be applied with caution.

9.p. 5545, l. 1: I suggest adding after parameters: “not process through any for of data analysis or inversion software.” This would ensure that raw data would be available for reanalysis. Partially inverted data may not be recoverable.

added

10.p. 5549, l. 23: The suggestion that “Generally, the delay time is correct of up-and down-scans show the same result” can lead to misinterpretation of data if fast scanning or stepping is performed with a slow response detector. This comparison could be correct if performed on data that has been inverted using the full knowledge of the transient response of all aspects of the measurement. See Collins et al. (Aerosol Sci. Tech. 36: 1-9, 2002).

Changed to:

Generally, the pluming time is correct if up- and down-scans show the same result, if the scan time is long enough to avoid smearing in the CPC due to its slow responds.

Reviewer #4

This study presents intercomparison of a number of SMPS carried out in instrument workshops. The authors developed technical standards for design and operation of SMPS, and data archive. These technical standards, when applied in future long-term surface-based measurements, will help improve the reliability and comparability of the SMPS dataset, which are important for improved understanding of the impact of aerosols on climate and air quality. The paper is quite comprehensive, and the topic is well suited for the journal of Atmospheric Measurement Techniques.

Thank you

General comments:

One important consideration when inverting SMPS data is the time required for particles to travel from the exit of DMA to the CPC detector. For laminar flow, this travel time is non-uniform and its distribution is often described using two time constants: a fixed delay time and a mixing time (Collins et al., 2002 and Russell et al., 1995). If the distribution of the travel time is not taken into consideration, inverted size distributions may be skewed, especially for faster scans (1-2 minutes). The fixed delay time and the mixing time can be derived by comparing up scan and down scans. (If both up scans and down scans are used in SMPS, one additional data quality check is to compare up and down scans; they should be identical when appropriate fixed delay time and mixing time are employed in data inversion). For some SMPS, such as the commercial SMPS from TSI, the mixing time is not taken into consideration. Therefore, it is important that these SMPS are operated at very slow scanning speed (e.g. > 5min), which should be sufficient for surface-based measurements. I would suggest the authors to include in the manuscript the measurement speeds of SMPS during the study, if and how the correction of travel time distribution is implemented in the data inversions, and whether both up and down scans are used in the SMPS measurements.

Added to the recommendation

For scanning mobility particle size spectrometers, we recommend to have the scanning time (up or down scan) minimum 2 min to avoid smearing effects in the particles counters with a relatively slow responds time.

The intercomparisons show substantial differences among SMPS measurements when particles are smaller than 20 nm or larger than 200 nm. Ideally these differences should be fully resolved with additional experiments. I think it would also strength the paper just including some analyses and discussions to narrow down the potential causes for these discrepancies. If the distribution of the particle travel time between DMA and CPC detector (as described above) is not fully considered in some of the inversion programs, the skew size distribution may contribute to the some of the discrepancy at $D_p > 200$ nm, especially for fast scans and when only up scans are used.

We will try to narrow down the reason for the increased uncertainty above 200 nm. First, we tried to invert all measurements with the same routine to find out if the multiple charge corrections are correct, although they were visible in the comparison of the inversion

routines. Again, the delay time is not the problem since each up- or down scan was larger than 2 min.

Figure 8 shows large difference in size distributions measured NILU and IFT REF1 at large particle sizes. Both NILU and IFT REF1 use Hauke-type DMA and TSI, 3010 CNC. The DMAs were also operated at the same flows (1:5). The IFT inversion routine uses “real area of DMA transfer function” whereas NILU uses Stolzenburg (1988) DMA transfer function. Could the differences be reduced if both routines use the same DMA transfer function? This could be easily verified as it does not require additional experiments.

No, it could not be reduced

The difference could also be due to different neutralizers used in the two systems (TSI Kr85 vs. Ni63). We occasionally observed differences in measured size distributions when different neutralizers were used in SMPS. This could be easily tested by using identical neutralizers in both systems (but unfortunately would require new experiments).

Yes, this is true. Unfortunately, we cannot repeat this experiment easily.

Regarding the discrepancies for particles smaller than 20 nm, one potential cause is the treatment of particle diffusion losses. It is not clear from the manuscript whether the size dependent particle loss in each system was characterized experimentally or was estimated based on tubing length and flow rate. The other potential reason for the differences may be related to the counting efficiency of CPCs.

All diffusion los corrected were determined by the method of the equivalent length, and the actual CPC efficiencies were applied, too.

In the new infrastructure project ACTRIS, we will try to understand better what the reason of the larger uncertainties are.

It is very surprising to see that the NILU and TNO showed no sensitivity at all for particles less than 15 nm, while Figure 3 shows both NILU and TNO CPCs have detection efficiencies greater than 80% at 10 nm! It is also not clear from Table 2 if CPC counting efficiency was corrected in inversion routine for TNO. It is worth examining the raw data to find out if any particles less than 15 nm were detected at all for the NILU and TNO SMPS. If not, this might suggest major problems/malfunctions of the systems.

Afterwards, we found out that the DMA voltage for these systems were wrong.

Specific comments:

Page 5524, Line 3, Please change “a wide application” to “a wide range of applications”.

Changed to:

Particle mobility size spectrometers often referred to as DMPS (Differential Mobility Particle Sizers) or SMPS (Scanning Mobility Particle Sizers) have found a wide range of applications in atmospheric aerosol research.

Page 5524, Line 10, Please add “s” after “distribution”.

corrected

Page 5524, Line 18, Please change “analyser” to “analyzer”.

done

Page 5528, Line 7, “This work was : : : Evaluation Program)”. It may be more appropriate to move this sentence to Acknowledgements.

I left the sentence there to introduce the acronyms

Page 5531, Line 21 -23, To ensure aerosols size distribution are under “dry conditions”, the RH needs to be even lower than 40%. Some studies suggest the efflorescence RH of ammonium sulfate is $\leq 37\%$ or lower.

This is true, but I would rather like to keep this 40% RH. Pure ammonium sulfate is rather seldom in the atmosphere.

Added:

However, one has to be careful, if the aerosol consists mainly of hygroscopic material.

Page 5531, Line 25-27, “Generally, a dry : : aerosol particles”. Wet aerosol can also reach steady state charging distribution in a bipolar charger. The key is to ensure no changes in particle size after steady state charge distribution is reached, such that it can be conveniently accounted for in inversion routine. Please rephrase this sentence.

Changed to:

A dry aerosol sample is needed to ensure the correct bipolar charge equilibrium and thus sizing downstream of the aerosol neutralizer in the DMA.

Page 5535; Line 5-6, What are the uncertainties? Not enough significant digits in the data format? Please clarify.

Changed to:

Only towards the end of the size range, some discrepancies occur probably due to splining the EUSAAR data to obtain individual data formats for the individual inversion routines.

Page 5540, Line 13-14, Were particle diffusion losses characterized experimentally, or estimated using flow rate and tubing length? What are the “unconsidered additional losses”?

Changed to

The deviations below 20 nm mobility size spectrometers might be due to additional losses, which are not considered by the method of the equivalent length.

Page 5540, Line 14, What are DMA classifying voltages for 20 nm particles?

cancelled

Page 5540, Line 15-16, How long were the size distributions (shown in Figure 5 and 8) averaged for? The (averaged?) size distributions appear to be quite smooth, suggesting counting statistics was sufficient. Please provide the uncertainties for each size bin (due to counting statistics) using raw counts detected. This will help narrow down the potential causes of the discrepancies.

Yes, the averaging time was long enough.

Page 5541, Line 13, please change “beside” to “except”.

done

Page 5543, section 4.4.2, for neutralizers other than TSI Kr 85, how were the particle diffusion losses calculated?

We assumed no big difference.

Page 5545; Line 16-17, what are the automated data quality checks? Are they checks of instrument status parameters?

I cancelled this sentence

Page 5545, Line 25-26, how are the data uncertainties propagated? How are the uncertainties in particle size and concentration derived? I think these are the important features of the new data format, and should be detailed.

I cancelled this item

Page 5563, Table 2, row “IFT”, column “calculation”, what does “using real area of the DMA transfer function” mean? Does this mean using experimentally determined transfer function?

We use a simpler table now

Page 5565, Table 2, row “ULUND”, column “calculation”, what is the ideal width? Width of non-diffusing DMA transfer function?

Simpler table now

Page 5566, Table 2, For PKU and TNO, is CPC counting efficiency corrected in inversion routines?

Simpler table now

Page 5568, Table 3, Only three systems are described, how about other systems used in the 3rd inter-comparison workshop?

All systems are described

Page 5581, Figure 3, TNO CPC 3010, operated at DT=17 degree, showed similar high counting efficiency as other CPCs operated at DT=25 degree. Was there any special modification of this CPC?

It's an in-built CPC. It might be different to a TSI 3010. Its operated with 25k temperature difference. Mentioned in the figure now.

Reference:

Russell, L. M., R. C. Flagan, and J. H. Seinfeld (1995), Asymmetric Instrument Response Resulting from Mixing Effects in Accelerated DMA-CPC Measurements, *Aerosol Sci. Technol.*, 23, 491-509.

Collins, D. R., R. C. Flagan, and J. H. Seinfeld (2002), Improved inversion of scanning DMA data, *Aerosol Sci. Technol.*, 36, 1-9.

added