



Non-equivalence of different evaluation algorithms to derive mean particle size from dynamic light scattering data

Katrin Franks · Vikram Kestens · Adelina Braun · Gert Roebben · Thomas P. J. Linsinger

Received: 5 March 2019 / Accepted: 7 August 2019 / Published online: 28 August 2019
© The Author(s) 2019

Abstract Different manufacturers of dynamic light scattering (DLS) instruments implement various evaluation algorithms. Users often implicitly assume that the various algorithms yield the same mean value and deliver results of the same precision. This assumption was tested using results obtained by cumulants analysis, CONTIN and Non-Negative Least Squares fitting (NNLS) of distributions, and frequency analysis on near-spherical silica and polystyrene latex particles obtained in one laboratory and results on near-spherical silica obtained in other laboratories. Particles have been measured in the size range of 10 to 200 nm. The results show that the average particle diameters from different algorithms differ significantly and that these differences increase with an increasing polydispersity of the material. Also, precision, both within-laboratory (repeatability) as well as between-laboratory standard deviations (reproducibility), differs between the algorithms investigated, especially for the more polydisperse

silica materials. The results obtained from the cumulants method usually show the best repeatability and the lowest between-laboratory standard deviation. The results also show that the conversion of intensity-weighted results to volume-weighted results increase the between-laboratory standard deviation, confirming the theoretical expectations that conversion increases the variation significantly.

Keywords Nanoparticles · Particle diameter · Dynamic light scattering · DLS evaluation algorithms · Time domain (correlation function analysis) · Frequency domain · Reference material · Instrumentation

Introduction

Dynamic light scattering (DLS) has become one of the most widely used technique to determine the size of nanoparticles, mainly because of its rather simple and straightforward instrument operation and because it provides results within a short time. It is based on the analysis of the time-dependent fluctuations of intensity due to the interference between scattered and reference light caused by the Brownian motion of the light-scattering particles.

Many different mathematical algorithms to retrieve particle size and size distribution information from raw DLS data have been developed and proposed (Finsky et al. 1992), but only a limited number have been implemented by commercial instrument manufacturers. These algorithms can be grouped based on whether the

K. Franks · V. Kestens · A. Braun · G. Roebben · T. P. J. Linsinger
European Commission, Joint Research Centre, Geel, Belgium

K. Franks (✉)
Federal Office of Consumer Protection and Food Safety, Berlin, Germany
e-mail: katrin.franks@bvl.bund.de

A. Braun
Merck KgA, Darmstadt, Germany

G. Roebben
European Commission, Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs, Brussels, Belgium

evaluation of the measurement signal is performed in the time domain (correlation function analysis) or the frequency domain. Presently, only the algorithm outlined by Trainer and colleagues (Trainer et al. 1992), which is based on the analysis of the power spectrum (the Fourier transform of the autocorrelation function), has been implemented by commercial manufacturers for evaluation in the frequency domain. For the sake of simplicity, this algorithm will be further referred to as ‘frequency analysis’. Correlation function analysis can be split into algorithms that only give the central moment and a measure of the width of the distribution (cumulants analysis (Koppel 1972)) and algorithms that attempt to model the complete particle size distribution by first deconvoluting the autocorrelation function using inverse Laplace transformation and then obtaining the central moment and dispersity from the modelled distribution, e.g. Non-Negative Least Squares (NNLS) (Grabowski and Morrison 1983; Lawson and Hanson 1974) or CONTIN (Provencher 1982). For simplification purposes, these distribution calculation algorithms are collectively referred to as ‘inversion methods’ in this manuscript.

For perfectly spherical, monomodal and monodisperse particles, all evaluation algorithms should give equivalent particle size results because they are all based on the same signal (variation of the scattered light due to the Brownian motion of particles) and because such materials closely meet the definitions of idealized particles.

When the sample deviates from the ideal monodisperse population of spherical particles, all the particle size analysis methods including DLS deliver results, which can be significantly affected by the different physical principles of the methods and by the various assumptions made by different evaluation algorithms. Therefore, for real-world particles, results of particle size are known to be operationally or method defined, and they can only be compared effectively if the measurand, that is the quantity intended to be measured, is unambiguously defined (Kestens et al. 2016). While quite some literature exists on comparing results from DLS with results from other methods (e.g. Meli et al. 2012), surprisingly little information is available on whether the various DLS evaluation algorithms give equivalent results. Stock and Ray compared the results of different algorithms in 1985 (Stock and Ray 1985), comparing the influence of noise on the results of the methods of cumulants, CONTIN, histogram,

exponential sampling, and subdistribution. They concluded that constrained regularization was the most robust algorithm against noise in the autocorrelation function. Since then, little literature is available on a systematic comparison of different evaluation algorithms. In many publications reporting DLS results (e.g. Wang et al. 2007; Tuoriniemi et al. 2014), only a reference to DLS is made without additional information on the evaluation algorithm. Nickel and colleagues compared the results of different DLS instruments on various nanomaterial suspensions and found no significant differences between instruments (Nickel et al. 2014). However, eight participants in their study used the Malvern Zetasizer Nano ZS, and only two other instruments were used. All the participants used the cumulants method for evaluation.

A second issue is that the instrument software of most commercial DLS instruments reports particle size results as scattered light intensity-weighted hydrodynamic diameters, either directly derived from the autocorrelation functions or retrieved from the converted light intensity-based size distributions. However, many technical and regulatory applications require volume- (or mass-) or particle number-based size distributions (EC 2011). Consequently, most instrument software also allows converting the light intensity results into a volume- or mass-weighted size, or even a particle number-weighted size, applying Mie light scattering theory and using the refractive indices of the material ((e.g. 1.46 for silica and 1.59 for polystyrene (Mie 1908)). This implicitly or explicitly implies making a number of assumptions on the physical properties of the particles and on the shape of the particle size distribution. As confirmed by Babick and Ullmann, converted results always have a higher uncertainty than the original results (Babick and Ullmann 2016) and the most recent version of the ISO documentary standard on DLS therefore explicitly deprecates such conversion (ISO 22412 2017).

The first ISO documentary standard on DLS was published in 1996 (ISO 13321 1996), covering only the cumulants analysis, as the other algorithms were not deemed mature enough for standardization. Frequency analysis was added in a second standard (ISO 22412 2008), and the two standards were combined into a single (revised) ISO 22412 standard (ISO 22412 2017). Since the development of the first documentary standard in 1996, applications of nanotechnology have multiplied, creating a need for algorithms giving the

complete particle size distribution. These algorithms are today often used indiscriminately, assuming that the results of all algorithms are equivalent.

This manuscript presents DLS results from both interlaboratory comparison (ILC) studies and a single- or intra-laboratory study. The results of the ILC studies, which were part of the certification campaigns of two certified reference materials (ERM-FD100 (Braun et al. 2011)) and ERM-FD304 (Franks et al. 2012)), colloidal silica with slightly different polydispersity but different particle sizes, are mainly used to investigate the effect of particle size on the performance of different evaluation algorithms. The effect of material polydispersity was further investigated during the single-laboratory study using colloidal silica and polystyrene latex CRMs. Results computed from frequency power spectra were compared with results obtained from correlation functions. The autocorrelation functions were modelled using the cumulants generating quadratic function (further referred to as cumulants method) and two distribution calculation algorithms (CONTIN, NNLS). Furthermore, the reproducibility of the conversion from intensity-based results to volume-based results for two monomodal colloids containing near-spherical silica nanoparticles is investigated based on the results of the ILC. The investigation shall elucidate if the different algorithms give significantly different mean diameters (even for rather simple materials), whether the precision of the algorithm differs, and discussed the reliability of the conversion from intensity- to volume-based results.

Materials and methods

Test materials

Monomodal colloidal silica certified reference materials (CRMs) ERM-FD100 (Braun et al. 2011) and ERM-FD304 (Franks et al. 2012), as well as the non-certified reference material IRMM-304, were obtained from the European Commission's Joint Research Centre. The two ERM-branded materials have certified particle size values for various methods, including DLS (cumulants analysis), electron microscopy (EM) and centrifugal liquid sedimentation (CLS). IRMM-304 is the same material as ERM-FD304; it contains the same particles but has a lower particle mass fraction (2.5 g/kg). All the three materials are aqueous dispersions of silica

nanoparticles, which are stabilized by their alkaline pH. The materials are ready for use and were measured without further dilution or sonication.

Two monodisperse polystyrene latex (PSL) CRMs (3100A and 3200A) were purchased from Thermo Fischer Scientific (Fremont, US). The materials are certified for their mean diameters obtained by TEM. In addition, indicative values for the hydrodynamic diameters obtained by DLS using the cumulants method have been assigned. Test samples were prepared by diluting the as-received PSL suspensions (10 g/kg) with 10 mmol/L NaCl to target concentrations of 0.2 g/kg. Relevant properties of all materials are given in Table 1.

Data collection strategies

Data were obtained from two sources, namely from two different DLS instruments employed in one laboratory and from nine different instruments used in an interlaboratory comparison. The advantage of a study in one laboratory is that no between-laboratory variation exists that complicates the evaluation, hence allowing for a better detection of potential differences. The disadvantage is that only a limited number of instruments were available, implying that the conclusions might be valid for the tested instruments only.

On the other hand, an ILC study pools data from a larger number of instruments and the evaluation results are therefore more representative for the algorithm. As a result, the interlaboratory data can lead to conclusions that are more robust. However, given the additional between-laboratory variation component, small systematic differences may not be easily detected.

Therefore, the combination of single-laboratory and interlaboratory data allows for a more robust and sensitive assessment of potential differences.

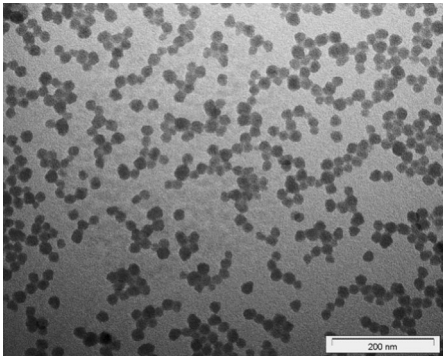
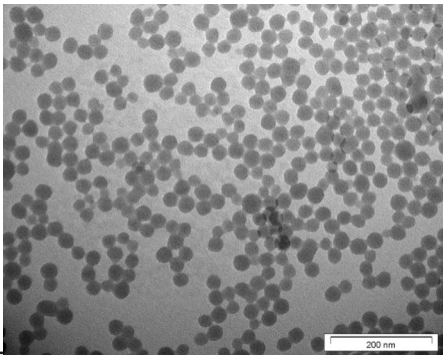
Single-laboratory study

Dynamic light scattering measurements were carried out using a Malvern Zetasizer Nano ZS particle size analyser (Malvern Instruments Ltd., Worcestershire, UK) operating in the time domain (correlation function analysis) and a Horiba LB-550 particle size analyser (Horiba Ltd., Kyoto, JP) operating in the frequency analysis domain.

Measurement results on the Malvern Zetasizer were analysed according to the cumulants method, the Malvern General Purpose NNLS (NNLS) and the CONTIN methods. The data obtained with the Horiba

Table 1 Properties of the materials used in the evaluations; polydispersity indices are averages from 18 measurement results (ERM-FD100, ERM-FD304/IRMM-304, 3100A) and 12

measurement results (3200A) obtained on a Malvern Zetasizer Nano ZS using the cumulants method

| Material code | Material | Assigned values (nm) | Polydispersity index | Representative EM micrograph |
|--------------------|------------------|-------------------------------------|----------------------|---|
| ERM-FD100 | Colloidal silica | 19.0 ± 0.6 (DLS) 19.4 ± 1.3 (EM) | 0.09 |  |
| ERM-FD304/IRMM-304 | Colloidal silica | 42.1 ± 0.6 27.8 ± 1.5 (EM) | 0.14 |  |
| 3100A | PSL | 102 ± 3 (EM) | 0.02 | |
| 3200A | PSL | 196–206 (DLS) 200 ± 6 (EM) | 0.02 | |

instrument were analysed using the in-built software for frequency analysis with 1000 data fitting iterations. Particle size distributions were expressed on both light scattering intensity and particle volume basis.

Both instruments operate in the backscattering mode. A standard quartz cuvette of 10-mm path length was used for both DLS instruments. All the measurements were performed at 25 °C ± 0.1 °C. A dispersant refractive index of 1.332 and a dynamic viscosity of 0.887 mPa s, which is typical for water, were used throughout.

Eight samples of IRMM-304 were measured on eight different days (one sample per day) using both DLS instruments. Each analysis is the average of results from three consecutive measurements.

Twelve test samples of 3200A and ten samples of 3100A were independently prepared and analysed on both DLS instruments (two test samples per day). Each result is the average of three instrument readings.

Interlaboratory comparisons

The ILC on ERM-FD100 and ERM-FD304 is described in the two certification reports (Braun et al. 2011; Franks et al. 2012). The characterization of both materials was performed in a single ILC involving the same participants and instruments with measurements at the same time. All the ILC participants received three ampoules of each material. On each of 3 days, two independent subsamples of one ampoule were analysed in triplicate, yielding six independent values per material and participant. Thirteen different instruments were used in combination with cumulants analysis (Particle Sizing Systems Nicomp DLS (1), Malvern Zetasizer Nano ZS (8), Malvern HPPS (1), Sympatec Nanophox (3)), six instruments were used with the inversion methods (Beckman Coulter Nanosizer N 4+ (1), ALV CGS-3 (1), Malvern ZS (2), Precision Detectors PDEXPERT (5), Sympatec Nanophox (6)), and four instruments were based on frequency analysis (Horiba LB-550 (3), Microtrac

Nanotracer (1)). The participants reported their results and stated the evaluation algorithms used. The number of participants for each evaluation algorithm is shown in Tables 3 and 4.

Statistical evaluation

Influence of the algorithm on the mean value

The rather wide size range covered by the materials used in the single-laboratory study (from about 20 to about 200 nm) means that variances are not homogeneous for each method. The evaluation of the single-laboratory study was therefore based on the normalized diameter, for which each result of a particular material was divided by the mean value for the material obtained by the algorithm in question. Differences in the mean diameter between groups were tested using two-way analysis of variance (ANOVA); significant differences between groups were evaluated using the Tukey honest significant difference (HSD) procedure. As the variance obtained on ERM-FD304 was higher than for the two PSL materials, the tests were performed twice, including and excluding the results on ERM-FD304, to ensure that these data do not influence the results. These tests were performed using Statistica 13 (Dell Corporation, Round Rock, USA).

Analogous to the single-laboratory study, the ILC data for each material were divided by the average of all results for this material to eliminate the influence of the particle size. The comparison of results between the different methods was based on a robust estimate of the central value, namely the median of the mean particle diameters reported, because mean diameters submitted by the laboratories did not follow normal distributions. Results of the ILCs were grouped into cumulants analysis, inversion methods (i.e. combining results from CONTIN and NNLS into one group) and frequency analysis. The median values of the three groups were compared using a Wilcoxon test on a 95% confidence level.

Influence of the algorithm on precision

For the single laboratory study, the normalized data were grouped by algorithm and material. Homogeneity of variances was tested using a Levene test (Levene 1960).

Estimations of standard deviations from ANOVA are less dependent on the assumption of normality than the comparison using the F -test. Therefore, within- (s_{within}) and between-group standard deviations (s_{between}) from the normalized data on the ILC were calculated using one-way ANOVA and were compared using Wilcoxon rank sum tests on 95% confidence levels.

Influence of conversion to volume-weighted data on precision

Differences between the variances of intensity- and volume-weighted results from the single-laboratory study were evaluated using a paired t test on the relative standard deviations of the different groups.

The precision obtained on intensity- and volume-weighted results from the data using cumulants analysis in the ILC was compared using a Wilcoxon rank-sum test on a 95% confidence level.

Results and discussion

Influence of the algorithm on the mean value

Statistical data for the single-laboratory and interlaboratory studies are shown in Tables 2, 3, and 4, and the individual data are shown in Figs. 1 and 2.

The two-way ANOVA evaluation of the single-laboratory data showed significant differences between the results obtained from the various algorithms. Evaluation of the differences using the Tukey HSD procedure showed that the results obtained by the cumulants analysis were significantly lower than those of all other methods ($p < 0.0002$). The mean value obtained using NNLS differed from the one obtained using the CONTIN algorithm on a 95% confidence level. The mean value obtained using the frequency analysis was not significantly different from the data obtained using the CONTIN algorithm but differed from both the NNLS and cumulants results. An analysis of equivalent groups puts the data from the cumulants method into a group of their own, whereas NNLS/CONTIN and CONTIN/frequency analysis were put together into two groups. The same results were obtained regardless of whether the data from ERM-FD304 were included in the analysis or not. While differences between the results from the cumulants analysis and the other algorithms

Table 2 Results obtained in the single-laboratory study. The data are based on n means of three repeated instrument readings each

| | | ERM-FD304 | | 3100 A | | 3200 A | |
|--------------------|--------------|--------------------|-----------------|--------------------|-----------------|--------------------|-----------------|
| | | Intensity weighted | Volume weighted | Intensity weighted | Volume weighted | Intensity weighted | Volume weighted |
| Cumulants | Average (nm) | 40.9 | Not applicable | 103.1 | Not applicable | 200.1 | Not applicable |
| | s (nm) | 0.3 | | 0.5 | | 1.9 | |
| | n | 8 | | 10 | | 12 | |
| NNLS | Average (nm) | 48.8 | 30.6 | 107.2 | 97.8 | 207.91 | 209.9 |
| | s (nm) | 0.2 | 0.5 | 0.6 | 0.6 | 2.2 | 2.32 |
| | n | 7 | 7 | 10 | 10 | 12 | 12 |
| Contin | Average (nm) | 48.7 | 28.2 | 106.4 | 99.7 | 204.8 | 205.9 |
| | s (nm) | 0.8 | 1.3 | 1.9 | 2.3 | 2.2 | 2.2 |
| | n | 7 | 7 | 10 | 10 | 12 | 12 |
| Frequency analysis | Average (nm) | 46.5 | 43.1 | 106.1 | 97.8 | 203.3 | 203.0 |
| | s (nm) | 1.0 | 0.6 | 2.5 | 4.3 | 3.5 | 5.1 |
| | n | 8 | 8 | 10 | 10 | 12 | 12 |

that are based on a more complete picture of the particle size distribution could be expected for the more polydisperse ERM-FD304, surprisingly, these differences are also apparent for the highly monodisperse PSL particles, even if less pronounced than for ERM-FD304. The differences (in relative terms) also seem to increase with decreasing size, being smallest for the 200-nm material (3200A), larger for the 100-nm material (3100A) and largest for the 40-nm ERM-FD304.

The relatively low number of results obtained by frequency analysis and inversion methods hampers the evaluation of differences between mean values in

the ILC. Nevertheless, Fig. 2 indicates a similar trend as observed in the single-laboratory studies, with the cumulants analysis giving the lowest results. In this case, only the differences of cumulants—inversion methods and cumulants—frequency analysis for ERM-FD304 were significant on a 95% significance level. This result was expected based on two considerations: On the one hand, the higher polydispersity of ERM-FD304 was expected to lead to larger differences between the cumulants analysis and algorithms based on the complete distribution. On the other hand, the additional between-laboratory contribution increased the overall variance, which made it more

Table 3 Results and statistical parameters from the interlaboratory study on ERM-FD100

| Algorithm | n | Median (nm) | s (nm) | s_{between} absolute (nm) | s_{between} relative (%) | s_{within} absolute (nm) | s_{within} relative (%) |
|--------------------|-----|-------------|----------|------------------------------------|-----------------------------------|-----------------------------------|----------------------------------|
| Intensity weighted | | | | | | | |
| Cumulants analysis | 13 | 18.82 | 0.24 | 0.77 | 4.09 | 0.17 | 0.89 |
| Inversion methods | 6 | 19.40 | 0.96 | 0.95 | 4.87 | 0.34 | 1.74 |
| Frequency analysis | 4 | 22.84 | 4.24 | 4.24 | 18.59 | 0.61 | 2.68 |
| Volume weighted | | | | | | | |
| Cumulants analysis | 13 | 16.78 | 1.21 | 1.21 | 7.19 | 0.23 | 1.38 |
| Inversion | 6 | 16.57 | 1.05 | 1.04 | 6.29 | 0.28 | 1.69 |
| Frequency analysis | 4 | 20.35 | 4.80 | 4.79 | 23.53 | 0.61 | 3.00 |

Table 4 Results and statistical parameters from the interlaboratory study on ERM-FD304

| Algorithm | <i>n</i> | Median (nm) | <i>s</i> (nm) | <i>s</i> _{between} absolute (nm) | <i>s</i> _{between} relative (%) | <i>s</i> _{within} absolute (nm) | <i>s</i> _{within} relative (%) |
|--------------------|----------|-------------|---------------|---|--|--|---|
| Intensity-weighted | | | | | | | |
| Cumulants analysis | 13 | 42.12 | 0.38 | 0.36 | 0.86 | 0.28 | 0.66 |
| Inversion methods | 6 | 45.44 | 2.98 | 2.96 | 6.52 | 0.77 | 1.69 |
| Frequency analysis | 4 | 46.35 | 4.27 | 4.25 | 9.16 | 1.23 | 2.64 |
| Volume-weighted | | | | | | | |
| Cumulants analysis | 13 | 34.56 | 4.87 | 4.84 | 14.01 | 1.27 | 3.67 |
| Inversion methods | 6 | 36.16 | 4.75 | 4.74 | 13.11 | 0.71 | 1.95 |
| Frequency analysis | 4 | 38.33 | 7.47 | 7.45 | 19.44 | 1.23 | 3.21 |

difficult to detect significant differences between results from different algorithms.

Nevertheless, the results of the ILC confirm the results from the single-laboratory study: Mean particle size values are algorithm-dependent, even for highly monodisperse samples. As expected, differences increase with increasing polydispersity and do not depend much on the size of the particle.

Influence of the algorithm on precision

As already indicated by visual inspection of Fig. 1, the results of the single-laboratory study did not indicate differences in the relative precision between the different materials. The Levene tests for each material grouped per evaluation algorithm did not show any significant differences between variances on a 95% confidence level. However, the variance

obtained by the cumulants analysis and NNLS was for both PSL 3100A and ERM-FD304, smaller than the ones obtained by frequency analysis and CONTIN ($p < 0.02$ for cumulants/CONTIN for ERM-FD304; $p < 0.004$ for all other comparisons). This indicates a weak tendency towards higher variances for the CONTIN and frequency analysis.

Results from the ILC obtained using the cumulants analysis had better within-laboratory repeatability (s_{within}) than those obtained using inversion methods or frequency analysis. The difference between inversion methods and frequency analysis is significant for ERM-FD100, but not for ERM-FD304.

A comparison of the between-laboratory standard deviations shown in Tables 3 and 4 reveals no statistically significant difference between the median values of the cumulants analysis and inversion methods, but statistically higher between-laboratory variation for

Fig. 1 Data obtained in the single-laboratory study. Shown are the individual results for each material and evaluation algorithm (normalized to the average of all results for that group) in boxes that encase the lowest and highest result for each group

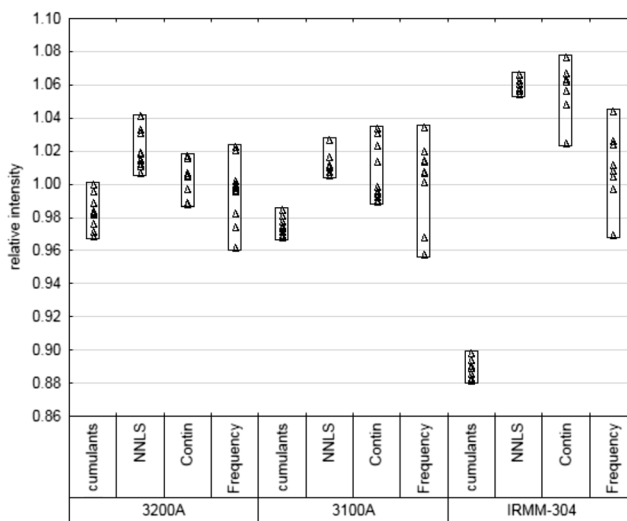
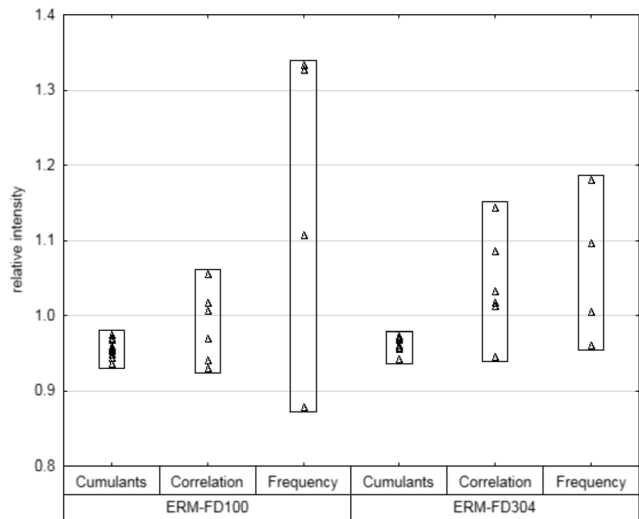


Fig. 2 Data obtained in the interlaboratory study. Shown are the mean values of each laboratory (each based on six individual results) for each material and evaluation algorithm (normalized to the average of all results for that group) in boxes that encase the lowest and highest result for each group



frequency analysis. For ERM-FD304, there is no statistically significant difference in the between-laboratory standard deviation between inversion methods and frequency analysis, but the s_{between} values of both of these methods are significantly larger than the ones obtained using cumulants analysis. The difference between the s_{between} for inversion methods between ERM-FD100 and ERM-FD304 is intriguing: This difference might be caused by the higher polydispersity of ERM-FD304, as this is the main difference (different particle size and concentration) between the two CRMs.

In both the single-laboratory study and the ILC, results obtained with the cumulants method show a better precision than the results obtained by the inversion methods or by frequency analysis. The main reason for this finding could be the robustness of the cumulants method, which

was the very reason why the technique could be standardized as early as 1996. The effect is particularly evident for less well behaved, e.g. more polydisperse samples, as the example of inversion methods shows: Reproducibility for the monodisperse ERM-FD100 is as good as cumulants analysis but is markedly worse for the slightly more polydisperse ERM-FD304.

Also here, the results from the ILC confirm the results from the single-laboratory study, namely that, as a result of its robustness against deviations from the ideal monodisperse particle size distribution, the cumulants analysis delivers more repeatable results for the main mode of a (nano-)particle size distribution than the inversion methods or the frequency analysis.

It is worth mentioning that the robustness of the results of the cumulants analysis is a result of the

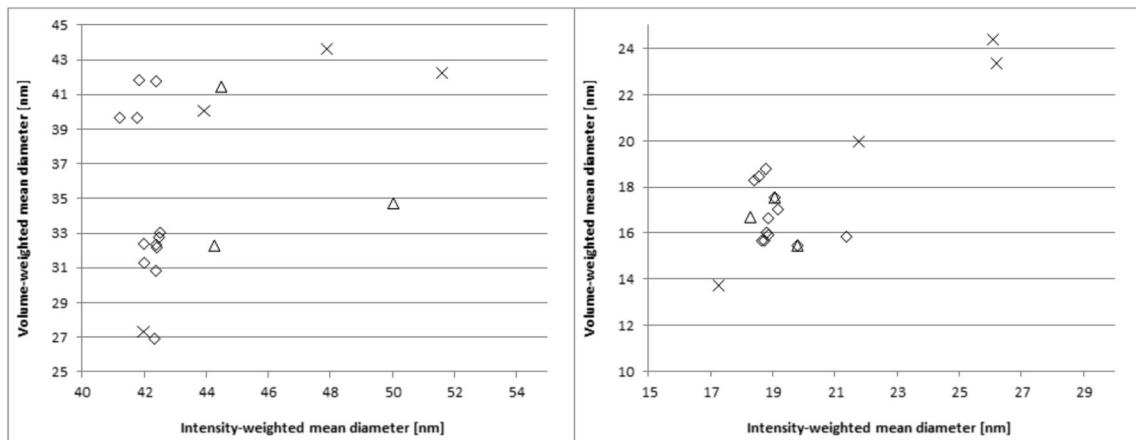


Fig. 3 Youden plots of intensity-weighted mean diameters versus volume-weighted mean diameters. Left: ERM-FD304; right: ERM-FD100. Diamonds: cumulants analysis; triangles: inversion methods; crosses: frequency analysis

assumed (Gaussian) particle size distribution of the sample. The better repeatability and reproducibility of the harmonic mean diameter, therefore, come at the cost of losing all information about the shape of the particle size distribution.

Influence of conversion to volume-weighted data on precision

In the single-laboratory study, with the exception of the frequency evaluation data for ERM-FD304, standard deviations on volume-weighted data are higher than on intensity-weighted data. The paired *t* test on the relative standard deviations shows that the difference between intensity- and volume-weighted values is significant at a 95%, but not at a 99% confidence level.

As for the data obtained in the ILC, a comparison of the s_{within} obtained using intensity-weighted and volume-weighted data showed no significant difference for inversion methods and frequency analysis, but a significantly lower s_{within} for cumulants analysis for intensity than for volume-weighted data. In examining the s_{between} , an interesting picture emerges: For the rather monodisperse ERM-FD100, only the s_{between} for cumulants analysis is significantly lower for intensity- than for volume-weighted data. However, for the more polydisperse ERM-FD304, all the methods show a significantly higher s_{between} for volume-weighted than for intensity-weighted data. It is important to point out that these data were obtained from the same measurements, i.e. raw data of the volume-weighted data are the same as the intensity-weighted data, only converted to a volume-weighted distribution. This shows that the conversion between intensity- to volume-weighted data works reasonably well for monodisperse samples, but that its reliability breaks down already at very moderate polydispersity.

These results show the considerable additional variation introduced by conversion between intensity-based results and volume-based results. This is highlighted in Fig. 3, which shows for each laboratory the result of the intensity-weighted result versus the volume-weighted one. Ideally, all these points would lie on one line, the slope of which would represent the relation between intensity- and volume-weighted results. However, data scatter significantly for results obtained for ERM-FD100 and show no correlation at all for ERM-FD304.

Conclusions

The evaluation of single-laboratory and ILC data confirms that the different evaluation algorithms of DLS are not yielding equivalent results. Precision varies between the different algorithms, with the cumulants analysis, usually showing less variation than the results obtained from the other algorithms investigated. The data also show that different mean values can be obtained, even for highly monodisperse samples. The results show that the differences may be practically negligible for certain materials, even if there is no fundamental equivalence. We, therefore, conclude that the statement of the evaluation algorithm is necessary to make DLS results comparable.

The results also confirm the theoretical expectation that the conversion of intensity-based results into volume-based results introduces a considerable additional variation of results. Consequently, special care must be taken when basing conclusions on volume-weighted results from DLS.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Babick F, Ullmann C (2016) Error propagation at the conversion of particle size distributions. *Powder Technol* 301:503–510
- Braun A, Franks K, Kestens V, Roebben G, Lamberty A, Linsinger TPJ (2011) Certification of equivalent spherical diameters of silica nanoparticles in water: certified reference material ERM®-FD100. EUR 24620, ISBN 978–92–79-18676-9. Publications Office of the European Union, Luxembourg
- EC (2011) Commission recommendation of 18 October 2011 on the definition of nanomaterial (2011/696/EU). OJ L 275:38–40
- Finsy R, De Jaeger N, Sneyers R, Geladé E (1992) Particle sizing by photon correlation spectroscopy - part III: mono and bimodal distributions and data analysis. *Part Syst Charact* 9: 125–137

- Franks K, Braun A, Charoud-Got J, Couteau O, Kestens V, Lamberty A, Linsinger TPJ, Roebben G (2012) Certification of the equivalent spherical diameters of silica nanoparticles in aqueous solution: certified reference material ERM®-FD304. EUR 25018, ISBN 978-92-79-21866-8, Publications Office of the European Union, Luxembourg
- Grabowski E, Morrison I (1983) Particle size distributions from analysis of quasi-elastic light-scattering data. In: Dahneke BE (ed) Measurements of suspended particles by quasi-elastic light scattering. Wiley, New York
- ISO 13321 (1996) Particle size analysis – photon correlation spectroscopy. International Organization for Standardization, Geneva
- ISO 22412 (2008) Particle size analysis – dynamic light scattering (DLS). International Organization for Standardization, Geneva
- ISO 22412 (2017) Particle size analysis – dynamic light scattering (DLS). International Organization for Standardization, Geneva
- Kestens V, Roebben G, Herrmann J, Jamting ÅK, Coleman VA, Minelli C, Clifford CA, De Temmerman P-J, Mast J, Junjie L, Babick F, Cölfen H, Emons H (2016) Challenges in the size analysis of a silica nanoparticle mixture as candidate certified reference material. *J Nanopart Res* 18:171
- Koppel DE (1972) Analysis of macromolecular polydispersity in intensity correlation spectroscopy: the method of cumulants. *J Chem Phys* 57:4814–4820
- Lawson LC, Hanson RJ (1974) Solving least squares problems. Prentice-Hall, Inc., Englewood Cliffs
- Levene H (1960) Robust tests for equality of variances. In: Olkin I, Ghurye SG, Hoëffding W, Madow WG, Mann HB (eds) Contributions to probability and statistics: essays in honor of Harold Hotelling. Stanford University Press, Stanford, pp 278–292
- Meli F, Klein T, Buhr E, Frase CG, Gleber G, Krumrey M, Duta A, Duta S, Korpelainen V, Bellotti R, Picotto GB, Boyd RD, Cuenat A (2012) Traceable size determination of nanoparticles, a comparison among European metrology institutes. *Meas Sci Technol* 23:125005
- Mie G (1908) Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Ann Phys* 4:377–445
- Nickel C, Angelstorf J, Bienert R, Burkart C, Gabsch S, Giebner S, Haase A, Hellack B, Hollert H, Hund-Rinke K, Jungmann D, Kaminski H, Luch A, Maes HM, Nogowski A, Oetken M, Schaeffer A, Schiwy A, Schlich K, Stintz M, von Der Kammer F, Kuhlbusch TAJ (2014) Dynamic light-scattering measurement comparability of nanomaterial suspensions. *J Nanopart Res* 16:2260
- Provencher SW (1982) CONTIN: a general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. *Comput Phys Commun* 27:229–242
- Stock RS, Ray WH (1985) Interpretation of photon correlation spectroscopy data: a comparison of analysis methods. *J Polym Sci B Polym Phys* 23:1393–1447
- Trainer MN, Wilcock WL, Ence BM (1992) Method and apparatus for measuring small particle size distributions. US Patent No 5094532
- Tuoriniemi J, Johnsson A-CJH, Holmberg JP, Gustafsson S, Gallego-Urrea JA, Olsson E, Pettersson JBS, Hassellöv M (2014) Intercomparison of the particle size distributions of colloidal silica nanoparticles. *Sci Technol Adv Mater* 15:035009
- Wang CY, Fuh WE, Lin HL, Peng GS (2007) Preliminary study on nanoparticle sizes under the APEC technology cooperative framework. *Meas Sci Technol* 18:487–495

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.