Uncertainty evaluation in atomic force microscopy measurement of nanoparticles based on statistical mixed model in a Bayesian framework

OPEN ACCESS
IOP Publishing
Measurement Science and Technology
Meas. Sci. Technol. **32** (2021) 085008 (8pp)
https://doi.org/10.1088/1361-6501/abe47f

Uncertainty evaluation in atomic force microscopy measurement of nanoparticles based on statistical mixed model in a Bayesian framework

J Pétry, B De Boeck, N Sebaïhi, M Coenegrachts, T Caebergs and M Dobre

National Standards, FPS Economy, 16 Boulevard du Roi Albert II, B-1000 Brussels, Belgium

E-mail: Thierry.Caebergs@economie.fgov.be

Received 16 December 2020
Accepted for publication 9 February 2021
Published 19 May 2021

Abstract

A major bottleneck in nanoparticle sizing is the lack of data comparability between techniques and between laboratories. However, this can be overcome by making the measurements traceable to the SI together with realistic uncertainty evaluation. In the present work, a novel approach is proposed to perform measurement uncertainty evaluation in a Bayesian framework by statistically modeling appropriately selected measurement data when no comprehensive physical model is available. The method is applied to the dimensional measurement of nanoparticles by atomic force microscopy (AFM) measurement and the calibration is performed by a multiple points calibration curve. Nevertheless, the proposed method can be applied to other microscopy techniques. The experimental data used to construct the statistical model are collected so that the influence of relevant measurement parameters can be assessed. An optimized experiment is designed under the intermediate precision conditions in order to limit the number of measurements to perform. Among the different influencing parameters, it is found that the AFM operator and image analyst do not significantly affect the measurement variability while the tip tapping force, the probe nature and the tip scan speed do. The particular case of gold nanoparticle of nominal diameter 30 nm is treated as an example of the method.

Keywords: AFM, mixed model, uncertainty calculation, Bayesian statistics, design of experiment

(Some figures may appear in colour only in the online journal)

1. Introduction

Measurements of nanoparticles can be performed with many different instruments, from microscopy to light scattering techniques, each having its own measurand. In order to compare data, the measurement should be traceable and the measurement uncertainty should be evaluated taking into account as many sources of uncertainty as possible. In the absence of a physical model of the measurement itself and the mathematical equation associated, classical approach of uncertainty calculation as described in the *Guide to the Expression of Uncertainty in Measurement* [1] might be hazardous and might lead to uncertainty underestimation by missing correlations between uncertainty sources.

At the nanometer scale, there is most of the time no physical model available for the entire measurement process. The present paper proposes a new method for the uncertainty evaluation of single particle diameter and mean particle diameter measured by atomic force microscopy (AFM) based on an extensive statistical analysis of variance. The method can be

applied to different microscopy techniques, such as transmission electron microscopy (TEM) and scanning electron microscopy (SEM).

There are currently two commonly accepted approaches to perform an uncertainty evaluation: the modeling approach and the empirical approach, which are covered by the guide of uncertainty measurement (GUM) [1] and the ISO 5725 standard [2]. In the GUM, the uncertainty evaluation is performed by propagating the various sources of uncertainty through a measurement equation. While in ISO 5725 the variability of the measurand is captured in a statistical model of random effects and a classical analysis of variance (ANOVA) is performed to evaluate the contribution of the different components. The ISO 5725 approach has been used for TEM measurements in [3] and for light scattering techniques in [4]. The two approaches are extensively compared in [5].

Concerning the particular case of uncertainty calculation in AFM, the traceability to the meter is achieved through the use of laser interferometry or through the comparison to height standards [6]. The uncertainty of the nanoparticles diameter has been evaluated so far effect-by-effect, individually, and combined using summation of uncertainties, in absence of global physical model of the measurement [7, 8]. Hybrid approach of ANOVA for repeatability and reproducibility and of classical approach for metrological AFM specific sources of uncertainties—from interferometry, typically—has been followed in [9]. The approach here followed unifies both in a common model and for a commercial AFM application.

In the following, traceability is achieved through a multiple points calibration curve and the uncertainty is calculated by an innovative method following the analysis of variance approach of ISO 5725 but extended to the use of a hierarchical mixed model to jointly consider fixed and random effects, as suggested by Deldossi and Zappa in [5]. A more general methodology of mixed models replaces the classical ANOVA and the model parameters are extracted in a Bayesian formalism by fitting to the experimental data. By using the Bayesian formalism, the size distribution of the item under measurement (nanoparticle sample or step height standard) can be estimated, instead of a single parameter estimation (average size, typically), as would be the case for the frequentist approach. The intermediate approach towards full Bayesian approach is to model the estimated distribution by a parameterized probability distribution, and apply Markov Chain Monte-Carlo (MCMC) approach for the parameter estimation.

Considering uncertainty from the calibration standards for a linear calibration curve together with detailed uncertainty estimations, effect-by-effect, a propagated uncertainty has been obtained for nanoparticle sizing by SEM [10]. Instead of using a stepwise approach to determine the influential parameters separately, an extensive experiment is designed and the measurement data are statistically analyzed to quantify the different sources of variability and compute the combined standard uncertainty, in one go. The calibration curve is established on a similar basis. The design of experiment approach here considered allows for implicitly considering correlations among the uncertainty sources, in addition to separating the contributions from the different stages of blocks for intermediate precision condition (as with ANOVA) and from fixed effects, all simultaneously in the unified framework of a hierarchical mixed model.

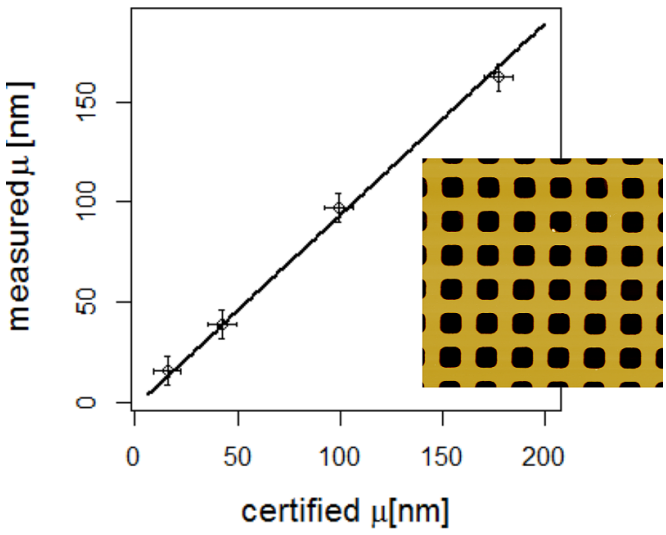

Figure 1. Example of 4-point calibration curve and a 20 µm × 20 µm topography image of a step height grid standard (z-scale of 250 nm).

2. Methods

2.1. Traceability route

Among the different AFM measurands that describe nanoparticles diameter, we use the maximum height relative to the substrate of isolated particles deposited on flat surface as defined in [11]. The measurements are performed on a commercial AFM (Oxford Instruments Asylum Research MFP-3D Infinity AFM). The metrological traceability to the SI units is obtained by comparison with step height standards in a multiple points calibration curve which is characteristics of the AFM used to perform the measurement. The calibration curve is built by comparing the certified step height values of a series of reference standards with their measured values, as illustrated in figure 1. The reference standards are certified nanogratings. In the context of this paper, the 'certified' term is to be understood as the value for which a traceability route to the SI meter is available. Depending on the step of the analysis, it can be the certificate value for a step-height grid, or the corrected value for item under measurement after the correction for instrument response by the mean of the calibration curve. In the particular case of nanoparticles measured by AFM, the measured quantity to be adjusted with the calibration curve is the height of a single particle, h, and the mean height of the nanoparticle sample, μ. The step height standards are chosen to cover the range of interest, in the present case from 10 to 100 nm. These are surface topography standards bought from VLSI Standards Inc. Their properties are listed in table 1 along with their calibrated values.

The uncertainty of the step heights and the nanoparticle size measurement are calculated in a similar way, under

Table 1. Properties of the VLSI surface topography standards used for AFM calibration.

	Product name	Nominal step height	Calibrated step height ($k = 2$)
$S_{[1]}$	STS 3 180p	18 nm	15.6 ± 1.0 nm
$S_{[2]}$	STS 2 440p	44 nm	42.3 ± 1.2 nm
$S_{[3]}$	STS 3 1000p	100 nm	99.0 ± 1.2 nm
$S_{[4]}$	STS 3 1800p	180 nm	177.4 ± 1.3 nm

intermediate precision conditions (within-lab reproducibility). It is the precision obtained within a single laboratory over a long period of time. In the present case, various operators measure on different days, at various positions on the sample, with varying AFM critical parameters: the probe, the tapping force and the scan speed. Focus is set on instrument related sources of uncertainty in the present work. Experiments have been designed to identify the significant parameters that contribute to the uncertainty and to quantify these individual contributions.

2.2. Design of experiment

The random effects under intermediate precision conditions in microscopy are typically measurement day, measurement location on the sample surface and recorded image. These measurement parameters are called factors in the statistical analysis and can take different *levels* (i.e. categories, for categorical variables). The term *level* is broadly used throughout this document with this meaning, except for a few exception in an obvious common acceptation. These levels correspond to different day, position and repeated image. As the levels of these factors are not changing independently, the design is nested as illustrated in figure 2 [12]. In the present work, the samples were measured three times consecutively, at five positions on the samples (at the center and close to the periphery) and at 5 different days.

The effect of ambient conditions is minimized by working under clean and controlled laboratory conditions: stable temperature and relative humidity ($21.8°C \pm 0.4°C$ and $43.1°C \pm 1.5°C$ RH during the measurement campaign), passive and active vibration damping and acoustic enclosure. And in absence of a drift effect, the deviations in ambient conditions merely result in normally distributed residual z-noise (approximately 35 pm, as effectively measured by the AFM).

Considering the fixed effects, the sources of fluctuations are operator-related settings (e.g. imaging force, scan range, scan speed, electronic feedback control parameters, etc) and image analysis parameters (e.g. software used, parameter choice in used algorithms). *A priori* important operator-related effects are the microscope probe (shape and material), the probe tapping force and the probe scan speed. Three different types of commercial probes classically used (Olympus AC160TS, Olympus AC240TS, PPP-NCHR) in tapping-mode AFM have been used. The tapping force has been varied in a range corresponding to soft tapping. At last, scan speed has been varied in the range of classical use. The operator effect that may be caused by the remaining subjective choices or by any physical instrument manipulations is assessed by three operators.

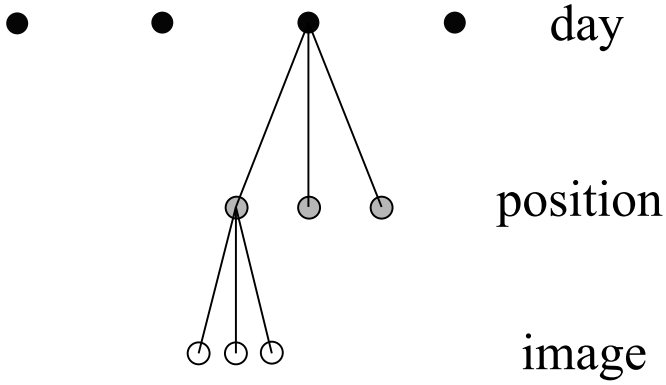

Figure 2. Nested design schema for the random variables, with its three stages of blocks.

Table 2. Fixed factors and the associated number of levels considered *a priori*.

Fixed effects	Number of levels
Probe	3
Tapping force	4
Scan speed	3
Operator	3
Image analyst	3

Regarding the image analysis, the SPIP software (Image Metrology A/S) has been selected for post-processing and analysis [13]. The first step of the image processing consists in leveling the image globally and subsequently line by line. The zero-level is fixed as the mean height of the image excluding the nanoparticle features. In a second step, in the case of nanogratings, the ISO standard 5436-1 algorithm for step heights measurement is applied [14]. In the case of nanoparticles, the maximum z-values with respect to the background for the isolated nanoparticles are reported. All parameter choices for these image post-processing and analysis steps are written down in a procedure, strongly decreasing the subjective choices to be made by the image analyst. Nevertheless, the possibly remaining image analyst effect is investigated as well for three analysts.

In summary, the fixed effects under considerations are: the probe, the tapping force, the scan speed, the operator and the image analyst, and they are treated as categorical variables of the mixed model. The categorical approach was chosen because each of the variable values is considered an appropriate condition for measurement, all on equal footing. Table 2 summarizes the different fixed effects and the number of levels for each effect considered in the hierarchical design.

In order to test all the possible combinations of these fixed effects, taking into account the repetition, the position and the day effect, 324 times 75 images would be required. Therefore, a matrix of experiment is designed in order to limit the number of images. The design is obtained with the D-optimality

criteria under the software JMP (SAS Institute) [15] with a limit of 150 images for practical reason of time limitation.

2.3. Mixed model

In absence of a measurement equation corresponding to a physical model of the measurement, the measured quantity can be modeled by an equation containing the main influencing factors. This model is chosen to be linear and contains fixed and random effects, it is a linear mixed model. Effects are considered fixed when the same value can be repeated in a subsequent experiment and random when the experimenter randomly samples the values from a population.

For the step height standards and the nanoparticles, the different factors considered to influence the measurand have been *a priori* included in the model. After running a designed experiment, the factors that influence significantly the measurand have been identified and the model has been adjusted accordingly, both for the step height standards and the nanoparticles. Using these model equations, the distribution of the measured step height and the measured nanoparticle height of a future experiment can be estimated, and by such the mean and standard uncertainty associated.

2.3.1. Definition. The variability of the measured quantity as a function of the main influencing fixed factors and the random effects is described by the mixed model equation as follows:

$$h_{ijkl} = \beta_0 + \sum_{f=1}^{m_f} \mathrm{X}_{fijk}\,\vec{\beta}_f + a_i + b_{ij} + c_{ijk} + \epsilon_{ijkl}, \qquad (1)$$

where h_{ijkl} stands for the measured height (of a nanoparticle or a step height), X_{fijk} stands for the effect-type coding matrix for effect f and for which $i = 1, \ldots, 5$ refers to the different days, $j = 1, \ldots, 5$ to the measurement positions on a given day, $k = 1, \ldots, 3$ to the repeated images for a given combination of day and position and $l = 1, \ldots, n_{ijk}$ stands for the different measurements on image k taken at position j on day i. n_{ijk} is the number of nanoparticles (or step heights) measured on the kth repetition of image taken at position j on day i. The realization h_{ijkl} of h is thus the lth observed value in the kth image taken at position j on the ith day. The total number of nanoparticles n (or step heights) measured is given by

$$n = \sum_{i,j,k} n_{ijk}.$$

The random effects $a_i \sim N(0, \sigma^2_{\mathrm{day}})$, $b_{ij} \sim N(0, \sigma^2_{\mathrm{pos}})$, $c_{ijk} \sim N(0, \sigma^2_{\mathrm{im}})$ and $\epsilon_{ijkl} \sim N(0, \sigma^2_{\mathrm{res}})$ are mutually independent for all i, j, k and l. The variance σ^2_{day} expresses day-to-day variability, σ^2_{pos} the variance between positions, σ^2_{im} the variance between repeated images and the within image variability σ^2_{res} captures the residual variance of the observed quantity within an image, representing the sample polydispersity by the squared standard deviation of its size distribution.

The fixed effects are noted $\vec{\beta}_f$, where $f = 1, \ldots, m_f$, with *a priori* $m_f = 5$ fixed effects. The intercept β_0 represents the mean response when there are no fixed effects present in the model (1). Fixed effects are categorical variables for which effect-type coding is used (X_{fijk} matrix) [12]. This coding sets the different categories into contrast while each considered on equal footing. If an effect cancels out, its resulting $\vec{\beta}_f$ is null and induces no shift between β_0 and h_{ijkl}, which is the second advantage of the effect-type coding over usual dummy coding. Every possible combination of levels of these coding variables provides an *opinion* $\mathbb{E}[h_{ijkl}|\mathrm{X}_{fijk} = \tilde{\mathrm{X}}]$ about the measured diameter (or step height). The different opinions are averaged out through a linear opinion pool [16]. The effects are analyzed and interpreted using their corresponding opinions, with effect-type coding in the present work.

Taking the linear interaction of fixed effects into account, equation (1) becomes

$$h_{ijkl} = \beta_0 + \sum_{f=1}^{m_f} \mathrm{X}_{f,ijk}\,\vec{\beta}_f + \sum_{f\neq g} \mathrm{Y}_{fgijk}\,\vec{\beta}_{fg} + a_i + b_{ij} + c_{ijk} + \epsilon_{ijkl}, \qquad (2)$$

where Y_{fgijk} is the combined coding effect matrix expressed as the tensor product of the matrices of single effect coding.

2.3.2. Significant effects. The estimation of the model is made with the restricted maximum likelihood method as it allows to estimate separately the random and the fixed effects [17, 18]. Fixed effects that are not significant at 95% level of confidence are removed from the model.

In the case of nano-gratings, although the four reference samples (listed in table 1) are processed similarly and are made of the same material (Si/SiO_2 layers coated with an uniform layer of platinum), samples $S_{[1]}$ and $S_{[3]}$ behave differently than $S_{[2]}$ and $S_{[4]}$: the probe is found to have a significant effect only for the sample $S_{[1]}$ and $S_{[3]}$. We have no clear explanation for this feature. None of the other effects are significant for all the reference samples.

In the case of nanoparticles, probe, tapping force and scan speed do have an effect while the operator and the analyst do not. The number of significant fixed effects is reduced from $m_f = 5$ to $m_f = 3$. Moreover, it is found that the interaction between the tapping force and the scan speed and the interaction between the tapping force and the probe are significant.

2.4. Bayesian approach

In the Bayesian approach, nanoparticle height or grating step height are statistically modeled either by an unknown probability density function (PDF) or by assuming an *a priori* distribution with parameters to be estimated by a Bayesian method. The latter is opted for here. On the market, samples are generally characterized by a summary value (mean, mode, standard deviation, etc) and not by its full PDF, arguing for the chosen approach. The Bayesian approach better takes into account correlations between the parameters by considering the interactions within the MCMC sampling while the frequentist one considers them globally. Despite the conceptual difference between frequentist approach and Bayesian parameter estimation, their results are expected to be similar in simple cases.

The h_{ijkl} are assumed to be normally distributed ($\sim\mathcal{N}(\mu,\sigma)$) and follow the model of equation (2). The parameters β_0, $\vec{\beta}_f$, σ^2_{day}, σ^2_{pos} and σ^2_{im} are obtained by running an Hamiltonian Monte Carlo algorithm in RStan with the n measured h_{ijkl} as input data and non-informative (flat) priors. Being the default algorithm of RStan, Hamiltonian Monte Carlo algorithm proved to be more stable, with reduced correlation in time than standard Metropolis–Hastings algorithm. Faster convergence is observed. Better exploration of the parameter space is also observed when parameters are correlated [19, 20]. The prior parameter densities are updated by the Bayesian estimation, leading to posterior densities for all the parameters in the model [17], the β's and the σ's, and the measured diameter (or step height) h_m and the mean measured diameter (or step height) μ_m. From these posterior densities, expected (measured) values ($\mathbb{E}[h_m]$ and $\mathbb{E}[\mu_m]$) and standard deviations ($SD[h_m]$ and $SD[\mu_m]$) are calculated and reported in the coming tables.

3. Application

In the following, the method described above is applied to the step height measurements in a first instance to build the multiple points calibration curve, and in a second step to the nanoparticle measurements. The nanoparticles used in this study are NIST gold particles RM8012 with nominal diameter of 30 nm [21]. Samples are prepared and deposited on modified poly-L-lysine mica substrate according to [22]. Finally, the calibration curve is applied to the distribution of the measured nanoparticle height in order to obtain the calibrated nanoparticle height.

3.1. Step height standards

3.1.1. Experiment under intermediate precision condition. Between $n = 140\,000$ and $280\,000$ step height measurements on each reference standards are performed according to ISO standard 5436-1 [14] and the Bayesian estimation with the mixed model (2) is carried on the data. The estimated variance of the respective random effects are shown in table 3 for the 4 standards. Clearly, the image repeatability does not bring any variability. For all the standards, the largest contribution comes from the within image variability but contribution from day and position cannot be discarded. Regarding the fixed effects and as already mentioned above, only the probe effect is significant and strangely enough, only for gratings $S_{[1]}$ and $S_{[3]}$, but not for gratings $S_{[2]}$ and $S_{[4]}$ (see table 4).

The expected value of the measured mean step height μ_m for the 4 standards $\mathbb{E}[\mu_m]$ and the corresponding standard uncertainty $SD[\mu_m]$ are extracted from the posterior distribution of the model (see table 5).

3.1.2. Construction of the calibration curve. The certified mean heights and standard uncertainties of the 4 reference gratings, as given in the calibration certificate, are summarized in table 1. The PDF for the certified step heights $\mu_{c[r]}$ ($r = 1, 2, 3, 4$) are assumed to follow normal distributions with mean and standard deviation set according to the certificate.

Table 3. Variance of the random effects influencing the step height measurements for each grating [nm^2].

Effect	$\sigma^2[\cdot]$ for for $S_{[1]}$	$\sigma^2[\cdot]$ for for $S_{[2]}$	$\sigma^2[\cdot]$ for for $S_{[3]}$	$\sigma^2[\cdot]$ for for $S_{[4]}$
Day	0.0075	0.0001	0.5439	1.0923
Position	0.0149	0.0016	0.1623	0.6170
Image repeat.	0.0000	0.0003	0.0002	0.0018
Within image var. (ϵ_{res})	0.0309	0.0455	0.0105	1.4797

Table 4. Standard uncertainty of the probe fixed effect for each grating.

Grating	$SD(\beta_{\text{probe}})$ (nm)
$S_{[1]}$	0.11
$S_{[2]}$	NS
$S_{[3]}$	0.25
$S_{[4]}$	NS

Table 5. Expected values and standard uncertainties for the measured mean step height μ_m for each reference standard grating.

Grating	$\mathbb{E}[\mu_m]$ (nm)	$SD[\mu_m]$ (nm)
S_1	15.91	0.13
S_2	42.15	0.01
S_3	99.06	0.54
S_4	177.04	0.70

Having the 4 calibration points $(\mathbb{E}[\mu_m], \mathbb{E}[\mu_c])_{[r]}$, the calibration curve is obtained by fitting the quadratic regression equation:

$$\mu_{c[r]} = \alpha + \gamma\,\mu_{m[r]} + \delta\,\mu^2_{m[r]} + \epsilon. \tag{3}$$

The parameter ϵ is a normal deviation from the quadratic model. In the Bayesian framework, these parameters are expressed by probabilities and the PDFs for α, γ, δ and ϵ display the variability present in the calibration points and the model uncertainty. The joint density for $(\alpha, \gamma, \delta, \epsilon)$ is approximated by sampling N times from its distribution through the following procedure:

- take N samples $(\mu_{m_j}, \mu_{c_j})_{[r]}$ (for $j = 1, \ldots, N$) from the 4 calibration points ($r = 1, 2, 3, 4$);
- calculate for $j = 1, \ldots, N$ the estimated coefficients α_j, γ_j and δ_j, and the mean squared error s_j^2 by performing N times an ordinary least squares quadratic regression with regression data $(\mu_{m_j}, \mu_{c_j})_{[r]}$ for $r = 1, 2, 3, 4$;
- sample a random value ϵ_j from $\mathcal{N}(0, s_j^2)$ (for $j = 1, \ldots, N$);
- collect the sample $(\alpha_j, \gamma_j, \delta_j, \epsilon_j)$.

The joint distribution of $(\alpha, \gamma, \delta, \epsilon)$ is approximated by merging the N samples $(\alpha_j, \gamma_j, \delta_j, \epsilon_j)$ for $j = 1, \ldots, N$. Table in figure 3 shows the results obtained for the parameters with $N = 10^6$. Given the value of δ, the calibration curve can be

parameter	$\mathbb{E}[.]$	$SD[.]$
α	$-0{,}2059\,\mathrm{nm}$	$0{,}7724\,\mathrm{nm}$
γ	$1{,}0025$	$0{,}0252$
δ	$0{,}000\,003\,\mathrm{nm}^{-1}$	$0{,}000\,135\,\mathrm{nm}^{-1}$
ϵ	$-0{,}0007\,\mathrm{nm}$	$0{,}6919\,\mathrm{nm}$

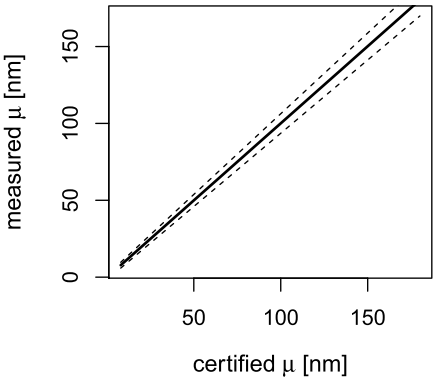

Figure 3. Expected values and standard uncertainties for the calibration curve parameters—graphical representation of the 95% confidence interval of the calibration curve.

defined as linear. The 95% confidence interval of the calibration curve and the mean calibration curve are also illustrated in figure 3.

3.2. Nanoparticles

3.2.1. Experiment under intermediate precision condition. Similarly to what has been performed for the step heights, an intermediate precision experiment is performed for the nanoparticles, following an appropriate design of experiment. To obtain a sufficient amount of particles on a single image a scan range of $3\,\mu\mathrm{m} \times 3\,\mu\mathrm{m}$ is chosen with 1024×1024 pixels, leading to pixel size of about $3\,\mathrm{nm} \times 3\,\mathrm{nm}$. Three different kinds of tapping probe are used (Olympus AC160TS, Olympus AC240TS, PPP-NCHR) with different spring constants and resonance frequencies. The effect of scan speed in the scanning direction is assessed by considering three different values (i.e. 1.8, 3.6 and $5.4\,\mu\mathrm{m\,s}^{-1}$). Also four values of probe oscillation amplitude ratio (i.e. the ratio between the tapping amplitude setpoint and the free tapping amplitude in air), related to the tip tapping force, are tested (i.e. 65%, 70%, 75% and 80%). The electronic feedback controller parameters are chosen by the operator to obtain a good tracking of the topography. These adjustments are subjective choices going into the operator effect. As previously discussed, the operator and the analyst are found to have no significant contribution to the variability of the measurements and is removed from the model.

Executing the experiments, collecting the nanoparticles heights in the different conditions and performing a Bayesian fit with model (2) as described previously, leads to the results shown in tables 6 and 7 for the random and fixed effects on the measurand variability. $n = 8259$ nanoparticle height measurements were considered for this purpose. Image repeatability is not investigated for nanoparticles because it was shown not significant for step height standards and not considering this effect reduces significantly the measurement time.

The different levels of the coding variable for the parameters probe, amplitude ratio and scan speed are expressed by posterior PDFs. The average of these level posteriors are considered to calculate the respective standard uncertainties as the standard deviation of these average PDFs, shown in table 7.

Table 6. Estimated variance of the random effects influencing the nanoparticle height measurements.

Effect	$u^2(\cdot)\,(\mathrm{nm}^2)$
Day	$\mathbb{E}[\sigma^2_{\mathrm{day}}] = 0.16$
Position	$\mathbb{E}[\sigma^2_{\mathrm{pos}}] = 0.88$
Within image variability (ϵ_{res})	$\mathbb{E}[\sigma^2_{\mathrm{res}}] = 7.61$

Table 7. Estimated standard deviation of the respective fixed effects significantly influencing the nanoparticle height measurements.

Effect	u [nm]
Probe	$SD[\beta_{\mathrm{probe}}] = 0.49$
Amplitude ratio	$SD[\beta_{\mathrm{tapping\ force}}] = 0.52$
Scan speed	$SD[\beta_{\mathrm{speed}}] = 0.39$

Table 8. Expected values and standard uncertainties for the measured particle height h_m and the mean measured height μ_m of the gold nanoparticle sample.

$\mathbb{E}[h_m]$ (nm)	$SD[h_m]$ (nm)	$\mathbb{E}[mu_m]$ (nm)	$SD[\mu_m]$ (nm)
23.39	3.18	23.40	1.19

The main effects of the fixed factors are given in this table, although interaction terms (amplitude ratio · scan speed and probe · scan speed) are also included. The combination of the posterior PDFs of all these parameters in model (2) gives rise to a PDF for the measured height h_m of a single nanoparticle and for the measured mean height μ_m. The expected value $\mathbb{E}[h_m]$ and the corresponding uncertainty $SD[h_m]$ for the measured height h_m of a single particle, and $\mathbb{E}[\mu_m]$ and $SD[\mu_m]$ for the measured mean height μ_m are given in table 8.

Results from Bayesian approach have been compared to estimation by the frequentist approach (REML). Fixed effects and random effects are compatible between the two approaches and within their uncertainties. Bayesian approach generally results in smaller uncertainties, which might be explained by correlations better accounted for.

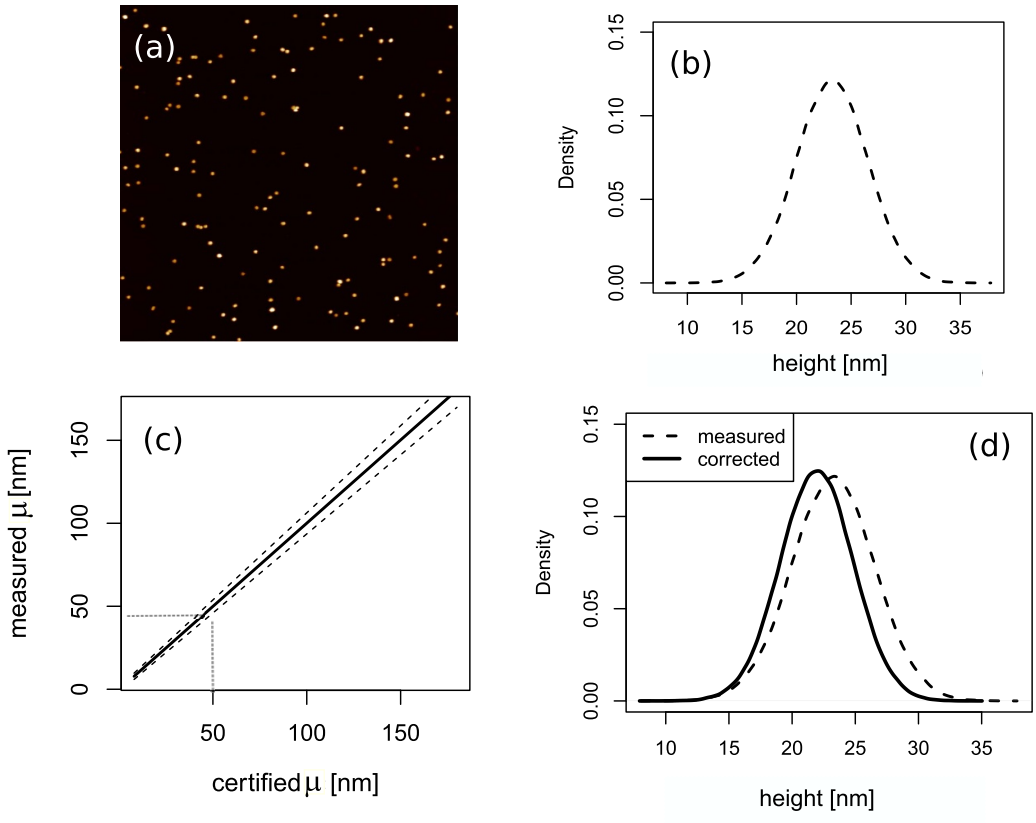

Figure 4. Example of an AFM image of nanoparticles (a), corresponding height probability density function (b) and illustration of the use of the calibration curve (c) to get the corrected and certified height probability density function (d).

Table 9. Expected values and standard uncertainties for the certified particle height h_c and the certified mean height μ_c of the gold nanoparticle sample.

$\mathbb{E}[h_c]$ (nm)	$SD[h_c]$ (nm)	$\mathbb{E}[\mu_c]$ (nm)	$SD[\mu_c]$ (nm)
23.24	3.28	23.25	1.44

3.2.2. Application of the calibration curve. The regression model (3) expresses the relation between the measured and the certified mean step height of the reference standards, so adopting this relation to obtain a calibration curve leads to

$$q_c = \alpha + \gamma q_m + \delta q_m^2 + \epsilon, \tag{4}$$

where q_c is the certified quantity of interest and q_m is the measured quantity of interest. The certified PDF of the measurand q_c is obtained by running Monte Carlo through equation (4) with the measured q_m ($q_m = h_m$ and $q_c = h_c$ for the particle height and $q_m = \mu_m$ and $q_c = \mu_c$ for the mean particle height).

The consecutive steps to perform traceable dimensional measurement of nanoparticles are illustrated in figure 4: (a) the particle heights are collected through AFM measurements under given conditions, (b) PDFs are constructed with model (2) and (c), (d) these PDFs are corrected through the application of the calibration curve. The PDFs summary parameters for the special case of the gold nanoparticles under investigation are listed in table 9. Reference value for RM8012 provided in NIST report of investigation is 24.9 nm with an expanded uncertainty of 1.1 nm [21]. The uncertainty detailed in the present work overlaps with value from NIST but is larger. The mean value is slightly smaller than NIST mean value. This might partially be explained by the deformation of the nanoparticles at the contact with substrate as described in other works [9, 23]. Taking this effect into account could be an improvement of the present work, for which the focus was set on investigating instrumental and operation effects.

4. Conclusions

In this paper, we propose an innovative statistical method for the traceable measurement of nanoparticle dimension, based on a multiple points calibration and an extended statistical analysis of design of experiment data with hierarchical mixed model in a Bayesian framework. Extensive statistical methods are particularly recommended when there is no full measurement equation available. The particular case of gold nanoparticles measured with AFM is discussed.

The traceability of the nanometer scale dimension is achieved through comparison to step height reference standards in a multiple point calibration curve. A complete measurement uncertainty evaluation of the measured dimensional quantities of the reference standards and the nanoparticles is performed. It consists in statistically modeling experimental measurement data in a mixed model containing random and fixed effects.

These data are obtained from a nested design of experiment performed under intermediate precision conditions. The random effects of the model are measurement day, position and image repeatability (for step height standards only); the fixed effects are the instrument main parameters, the operator and image analyst. For the sample of gold nanoparticle under investigation it has been demonstrated that among the fixed effects, only the scan speed, the amplitude ratio and the type of probe have a significant effect on the nanoparticle height measurements.

Bayesian inference is performed on the mixed model with non-informative prior and posterior PDFs are obtained for all the parameters of the model. From these PDFs, summary values such as expected and standard deviation can be extracted. In particular, the expected values and standard deviations of single nanoparticle height and mean height are extracted to get the certified SI traceable corresponding values.

Acknowledgments

This work was partially carried out in the context of the 17NRM05 EMUE (Examples of Measurement Uncertainty Evaluation) project which has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme (Funder DOI: 10.13039/100014132).

ORCID iDs

J Pétry https://orcid.org/0000-0003-4678-9839
T Caebergs https://orcid.org/0000-0001-6985-0909
M Dobre https://orcid.org/0000-0002-5193-0238

References

[1] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2008 *Guide to the Expression of Uncertainty in Measurement* JCGM 100:2008 (GUM 1995 with minor corrections) (BIPM)
[2] ISO 5725:2004 Accuracy (trueness and precision) of measurement methods and results
[3] De Temmerman P-J *et al* 2014 Size measurement uncertainties of near-monodisperse, near- spherical nanoparticles using transmission electron microscopy and particle-tracking analysis *J. Nanopart. Res.* **16** 2628
[4] Braun A *et al* 2012 Validation of dynamic light scattering and centrifugal liquid sedimentation methods for nanoparticle characterisation *Adv. Powder Technol.* **2** 766–70
[5] Deldossi L and Zappa D 2009 ISO 5725 and GUM: comparison and comments *Accreditation Qual. Assur.* **14** 159–66
[6] Korpelainen V and Lassila A 2007 Calibration of a commercial AFM: traceability for a coordinate system *Meas. Sci. Technol.* **18** 395–403
[7] Delvallée A *et al* 2016 Toward an uncertainty budget for measuring nanoparticles by AFM *Metrologia* **53** 41–50
[8] Meli F *et al* 2012 Traceable size determination of nanoparticles, a comparison among European metrology institutes *Meas. Sci. Technol.* **23** 125005
[9] Misumi I, Sugawara K, Takahata K, Takahashi K and Ehara K 2018 Size measurements of standard nanoparticles using metrological atomic force microscope and evaluation of their uncertainties *Precis. Eng.* **51** 691–701
[10] Crouzier L *et al* 2019 Methodology to evaluate the uncertainty associated with nanoparticle dimensional measurements by SEM *Meas. Sci. Technol.* **30** 085004
[11] Meli F 2004 Lateral and vertical diameter measurements on polymer particles with a metrology AFM *Nanoscale Calibration Standards and Methods* ed G Wilkening and L Koenders (Weinheim: Wiley-VCH) ch 27, pp 361–74
[12] Montgomery D C 2005 *Design and Analysis of Experiments* 6th edn (New York: Wiley)
[13] Image Metrology A/S 2021 *SPIP 6.7.4* (available at: www.imagemet.com)
[14] ISO 5436:2001 Geometrical Product Specifications (GPS)—Surface Texture: Profile Method; Measurement Standards—Part 2: Software Measurement Standards
[15] SAS Institute Inc. 2018 *JMP®* v.14
[16] Dietrich F and List C 2017 *Probabilistic Opinion Pooling* vol 1 (Oxford: Oxford University Press) ch 37, pp 787–814
[17] Molenberghs G and Verbeke G 2000 *Linear Mixed Models for Longitudinal Data Linear Mixed Models for Longitudinal Data* (New York: Springer)
[18] Lindstrom M J and Bates D M 1988 Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data *J. Am. Stat. Assoc.* **83** 1014–22
[19] Stan Development Team 2019 *RStan: The R interface to Stan* R Package Version 2.19.2 (available at: http://mc-stan.org/)
[20] Betancourt M 2017 A conceptual introduction to Hamiltonian Monte Carlo (arXiv:1701.02434)
[21] Small J A and Watters R L J 2017 Reference material® 8012. Gold nanoparticles, nominal 30 nm diameter *Report of Investigation* (National Institute of Standards and Technology (NIST)) p 10 (available at: www-s.nist.gov/srmors/certificates/8012.pdf) (Accessed 14 June 2017)
[22] Grobelny J *et al* 2010 Size measurement of nanoparticles using atomic force microscopy *Methods in Molecular Biology* ed N J Clifton (Totowa, NJ: Humana Press) pp 71–82
[23] Crouzier L *et al* 2019 Development of a new hybrid approach combining AFM and SEM for the nanoparticle dimensional metrology *Beilstein J. Nanotechnol.* **10** 1523–36