Measurement 225 (2024) 113974

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

An uncertainty-based quality evaluation tool for nanoindentation systems

Giacomo Maculotti *, Gianfranco Genta, Maurizio Galetto

Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

ARTICLE INFO

Keywords:
Nanoindentation
Bootstrap
Performance comparison
Uncertainty
Calibration

ABSTRACT

Instrumented Indentation Test (IIT) is a nonconventional mechanical tests allowing multi-scale mechanical characterisation. It is employed for research and quality control in strategic manufacturing fields for developing edge technologies. The state-of-the-art lacks a robust methodology to assess quality of indentations and benchmark indentation devices. This is limiting the application of IIT for specifying and verifying tolerances. This work proposes an uncertainty-based quality evaluation tool for IIT. A non-parametric uncertainty evaluation of calibration contribution is proposed. The method shows the statistical significance of indentation sets modelled by the bootstrap samples. The uncertainty is then propagated according to the law of uncertainty propagation for the evaluation of mechanical characteristics. The methodology is applied to five case studies. Results show that the uncertainty evaluation model can achieve robust and sensitive quantification of the indentation results and system quality, thus providing a useful practical tool for industrial and academic practitioners within a metrological framework.

1. Introduction

Instrumented Indentation Test (IIT) is a flexible mechanical characterisation test for surfaces. It allows multi-scale, i.e. nano to macro, estimation of materials' properties such as the Young modulus, the indentation hardness H_{IT}, creep and relaxation [1,2]. Furthermore, it allows surface mapping to quantitatively distinguish different phases or constituents of composites materials [3,4]. Moreover, when dynamic indentations are applied, the technique enables the estimation of viscous material properties [5] as well as the in-depth characterisation of multi-layered coatings without destructive cross-sectioning [6,7]. Similarly, data-augmentation through electric contact resistance (ECR) allows evaluating phase change due to mechanical stress, essential in supporting the development of semiconductor manufacturing [8,9]. Therefore, IIT has a core role in edge research and development of new materials and manufacturing processes, in strategic industrial fields, e.g. automotive, aerospace, healthcare. In fact, advanced composites and coatings aim at engineering surfaces to improve their functionality while extending service life, e.g. for cutting tools [10], essential to reduce waste that is strategic in the pathway towards sustainability. Semiconductors are extremely relevant in military, aerospace and physics for they provide enhanced optical and thermal properties for a number of applications, e.g. germanium allows improved energy efficiency in solar panel coating and augmented infrared vision [11], thus optimizing related machining and understanding the induced effect on the material properties is critical [12]. Surface mapping is not only relevant for technological surfaces but also in healthcare. In fact, IIT proved effective in estimating resistance and integration of tissues on prostheses [13] as well as to predict onset of illnesses [14,15]. Furthermore, IIT can be regarded as a semi-destructive test and in general requires limited sample preparation and no specific shape of the specimen. Therefore, it is a more sustainable alternative to tensile test to estimate elastic and plastic properties of materials [16]. Moreover, it can be carried out on the final component, thus increasing the representativeness of the quality control. In particular, most recent advancements coupled IIT with machine learning to allow determining stress–strain material parameters [17,18], deconvolute properties of coating layers and substrate [19,20]. Similarly, IIT allows to estimate residual stresses representing a convenient alternative to other conventional methods, enabling less invasive than drill hole and larger range than X-ray diffraction to be investigated [21,22].

It is henceforth relevant to ensure traceability and guarantee metrological appropriateness of the methodology to assess accuracy and precision of IIT to provide end users with confidence in the obtained results and to allow trustworthy and metrological comparison of the performances of the indentation testing devices. However, current state of the art is missing an holistic metrological framework to that aim, which is ultimately hindering adoption of IIT as a large volume quality

* Corresponding author.
E-mail address: giacomo.maculotti@polito.it (G. Maculotti).

https://doi.org/10.1016/j.measurement.2023.113974
Received 24 August 2023; Received in revised form 31 October 2023; Accepted 2 December 2023
Available online 4 December 2023

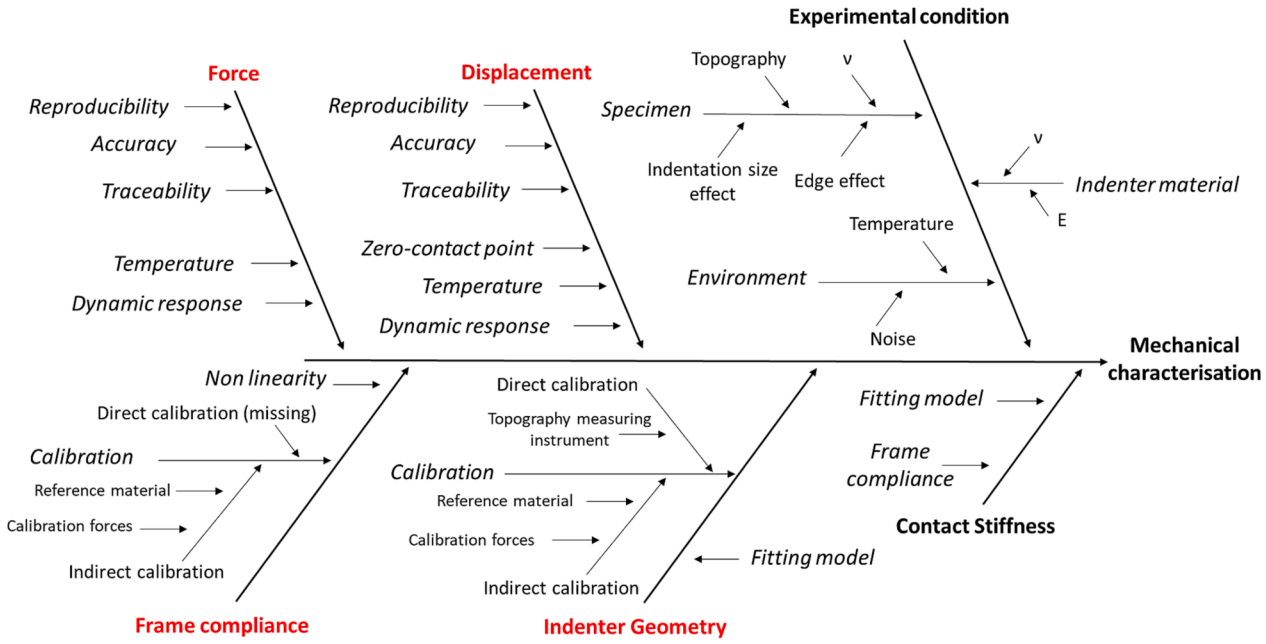

Fig. 1. Ishikawa diagram of the influencing factors to the accuracy and precision of the IIT mechanical characterisation. In read factors requiring calibration.

control tool, for tolerance and design specification [23], and has also posed some criticalities in results assessment of interlaboratory comparison [24].

This paper proposes a methodology based on non-parametric simulation to evaluate the measurement uncertainty to assess the metrological performances of IIT testing devices. The methodology will be tested and demonstrated on several practical scenarios. In the following subsections of the introduction an overview of the technique is provided along with a review of the methods to evaluate measurement uncertainty. Section 2 describes the proposed methodology and Section 3 the case studies considered for the application of the proposed method. Section 4 concludes on the findings.

1.1. Fundamentals of Instrumented indentation test

IIT consists of applying a loading-holding-unloading force cycle by an indenter of known shape to a sample that requires characterisation. During the test, the applied force F and the displacement h of the indenter in the sample due to the penetration are continuously measured. The indentation curve (IC), i.e. $F(h)$, is the most typical representation of the indentation data. From the analysis of the unloading portion of the IC, it is possible to achieve the mechanical characterisation in terms of H_{IT} and the indentation modulus E_{IT}, that estimates the Young modulus, as:

$$H_{IT} = \frac{F_{max}}{A_p\left(h_{c,max}\right)} \tag{1}$$

$$E_{IT} = \frac{1-\nu_s^2}{\frac{2\sqrt{A_p\left(h_{c,max}\right)}}{S\sqrt{\pi}} - \frac{1-\nu_i^2}{E_i}} \tag{2}$$

$$1\Big/E_r = \frac{2\sqrt{A_p\left(h_{c,max}\right)}}{S\sqrt{\pi}} \tag{3.1}$$

$$1\Big/E_r = \frac{1-\nu_s^2}{E_s} + \frac{1-\nu_i^2}{E_i} \tag{3.2}$$

where E_r is the reduced modulus, ν the Poisson ratio and the subscripts i and s indicate quantities related to the indenter and the sample, respectively. Specifically, the raw data require to be corrected, as per Eq. (4), for the elastic displacement due to the sample and the machine frame compliance, respectively $1/S$ and C_f, and for the zero error h_0, i.e. the first contact point:

$$h_c = h - h_0 - C_f F - \varepsilon\frac{F}{S} \tag{4}$$

where ε is a factor depending on the indenter geometry [1,25,26], e.g. 0.75 in the most common case of Berkovich and Vickers indenters.

The sample contact stiffness S can be evaluated from the unloading portion of the IC, as:

$$C_{tot} = C_f + \frac{1}{S} \tag{5.1}$$

$$S_m = \frac{1}{C_{tot}} = \left.\frac{\partial F}{\partial h}\right|_{h_{max}} \tag{5.2}$$

considering a spring model featuring the frame compliance in series to the sample's, where S_m is the total measured contact stiffness [27]. S_m can be evaluated by fitting according to a power law the unloading portion of the IC, according to the most commonly adopted best practice [1,28]. Conversely, the C_f requires calibration [29], and is a major contributor to the accuracy and measurement uncertainty of the resulting characterisation [26].

The essential element of IIT is the capacity of achieving the mechanical characterisation at nanoscale by overcoming limits set by conventional hardness tests that require optical measurement of the indentation. In fact, IIT relies on the area shape function $A_p = A_p(h_{c,max})$ that relates to the penetration depth the area of the contact surface between the indenter and the sample projected on the plane of the sample surface. Typically, the functional form of A_p is a polynomial, whose parameters require calibration [29,30], and is particularly impacting the metrological performances of the mechanical characterisation at the nano-scale [26,31].

1.2. Influence factors to Instrumented indentation test

Fig. 1 shows the main influence factors to the mechanical characterisation by IIT. As it can be appreciated, the framework is quite

complex, and a number of factors require calibration [1]. In particular, the force and displacement transducers can be directly calibrated, thus making negligible the related contribution to the measurement uncertainty of the characterisation. Conversely, literature has shown that the frame compliance and the area shape function are the two major contributors to the measurement uncertainty, amongst those that can be calibrated [26]. Indeed, further source of bias and high dispersion can be traced in the experimental condition, e.g. thermo-mechanical noise, poor surface finish. These tend to dominate the measurement uncertainty if not properly relieved by adequate measurement strategies and sample preparation [2,24], and should be managed based on a task-specific approach. Last, the measured contact stiffness S_m is extremely critical and liable of biasing characterisation results [32–34]. As mentioned in Section 1.1, S_m can be computed by fitting a portion of the unloading curve according to some mathematical modelling. The commonly acknowledged, and standardised, methods require either a linear or a power law fitting [1]. However, the literature has showed some limitations and proposed other approaches, e.g. based on the fitting of the numerically evaluated derivative of the unloading [32–34] which allows to constrain its mathematical properties, that is directly linked to S_m definition in Eq. (5.2). Furthermore, the literature has shown a relevant effect of the considered portion of the unloading to be considered for the fitting [32–34]. In this work, although acknowledging the several fitting methods and diverse alternative of unloading portion to be considered to evaluate S_m, the standard choice has been considered: power law fitting of the unloading from the 95 % to the 20 % of the maximum load.

Literature and the current standard ISO 14577–2 present several methods to calibrate both the frame compliance and the area shape function. As far as the area shape function is concerned, either direct methods or indirect approaches are available. The former require measuring by a surface topography measuring instrument the geometry of the indenter tip [30,35–37]. The latter rely upon indenting calibrated reference materials and by applying adequate mathematical procedures [38,39] achieve the calibration. These indirect calibration methods allows calibrating also the frame compliance, for which direct calibration method approaches are unreported [29]. Indeed, direct calibrations allow better accuracy and precision, but they are more complex and expensive, thus limiting their application to few national metrological institutes. Conversely, manufacturers and practitioners both in industry and academia tend to resort to indirect calibration methods, despite they couple the calibrated quantities and, depending on the specific algorithm, they might be more or less sensitive to experimental conditions, e.g. test forces, calibrated reference materials [39–41]. Amongst the several available alternatives, two main indirect calibration methods are worth citing. Both methods require two calibrated reference materials, typically fused silica (SiO_2) and tungsten (W) that shall be indented with replicated indentations at different maximum force F_{max}. The first alternative is included in the standard as the fourth method of the Annex D of ISO 14577–2 [29,38]. It relies on an iterative method, described in [38,40] and briefly depicted in Fig. 2, whose measurement uncertainty can be appropriately evaluated by simulative approaches [40,41]. The second approach has been proposed to provide a closed-form solution and an alternative to the iterative method by means of an Orthogonal Distance Regression (ODR) to improve robustness and mathematical modelling [39], i.e.:

$$\begin{cases} \dfrac{\pi}{4E_r^2} = \left(\dfrac{1}{S_m} - C_f\right)^2 \left\{ a_2 \left[h_{max} - h_0 - \left[C_f + \varepsilon\left(\dfrac{1}{S_m} - C_f\right)\right] F_{max} \right]^2 + a_1 \left[h_{max} - h_0 - \left[C_f + \varepsilon\left(\dfrac{1}{S_m} - C_f\right)\right] F_{max} \right] + a_0 \right\} \\ H_{IT} = \dfrac{F_{max}}{a_2 \left[h_{max} - h_0 - \left[C_f + \varepsilon\left(\dfrac{1}{S_m} - C_f\right)\right] F_{max} \right]^2 + a_1 \left[h_{max} - h_0 - \left[C_f + \varepsilon\left(\dfrac{1}{S_m} - C_f\right)\right] F_{max} \right] + a_0} \end{cases} \quad (6)$$

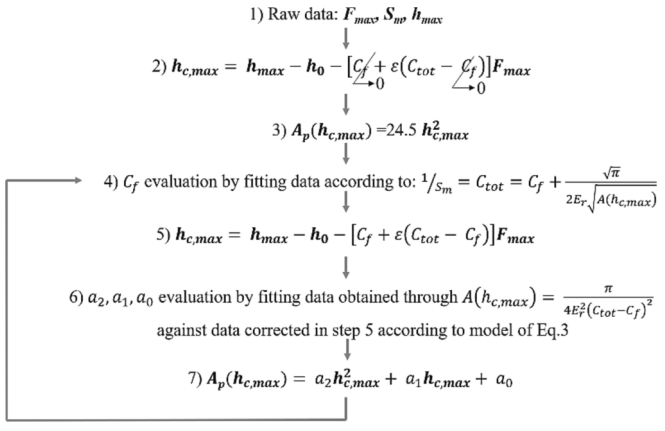

Fig. 2. Standard iterative calibration method [39].

1.3. State-of-the-art methodology for performance comparison

Current literature is scarce of benchmark methods for instrumented indentation test characterisation results. Indeed, the mechanical characterisation measurement uncertainty evaluation can be performed by propagating the uncertainty according to the Guide to the Expression of Uncertainty in Measurement (GUM) [42], as:

$$u^2(y) = \sum_{a=1}^{n} \left(\frac{\partial f}{\partial x_a}\right)^2 u^2(x_a) = \sum_{a=1}^{n} u_a^2(y) \quad (7)$$

Where y is a realization of Y, the dependent random variable, i.e., the mechanical characterisation results, dependent through a function f on the independent random variables representing the influence factors X_a (with realization x_a) and the $u^2(x_a)$ is the equivalent variance of the a-th independent random variable.

Alternatively, parametric approaches could be considered, e.g. Monte Carlo Markov Chains (MCMC) [43]. However, MCMC requires nontrivial assumption to be made on the dependency of the penetration depth and the applied force, which are correlated through the unknown sample mechanical properties. Such correlation if neglected may lead to distorted results [39], else if modelled limits the application to specific materials under prior knowledge of the mechanical system. Furthermore, other indentation quality evaluation tools proposed in the literature neglect the correction of frame compliance thus severely underestimating the measurement uncertainty, and do not propagate contributions from further influence factors' calibration [43]. Moreover, IIT testing platform performance comparison have been attempted

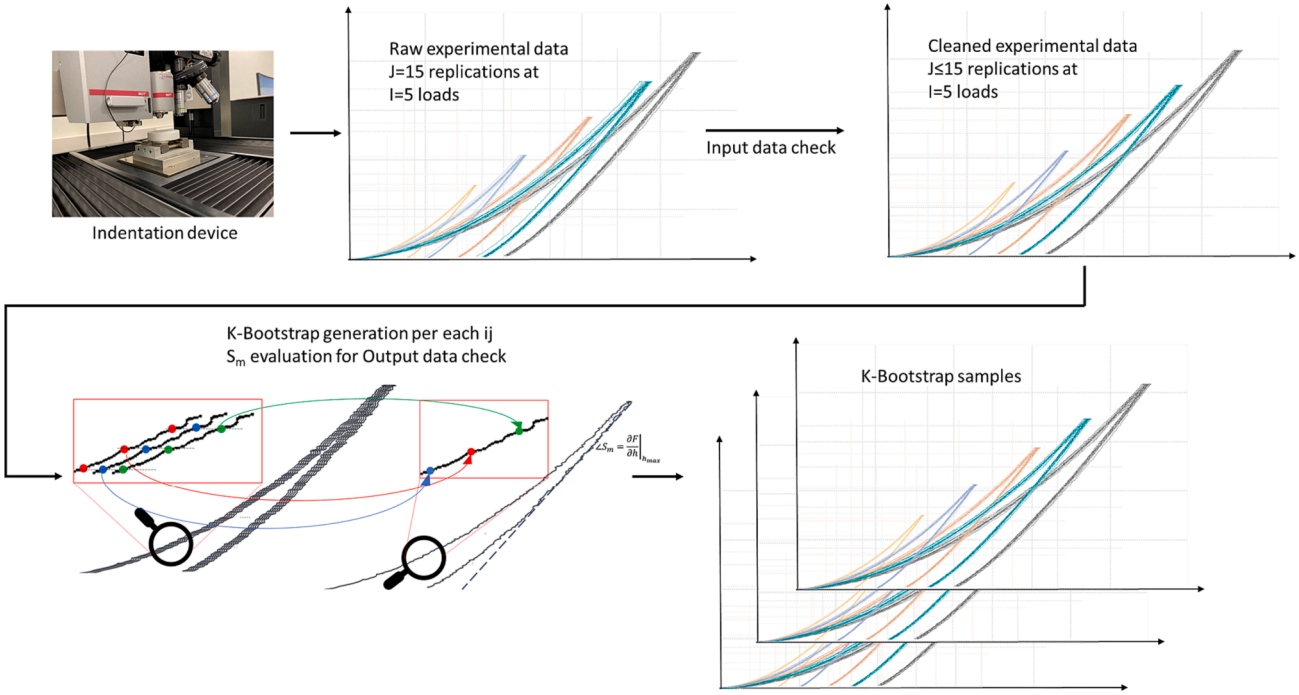

Fig. 3. Bootstrap workflow starting from experimental data collection, elimination of raw data outliers and deployment of bootstrap sample generations.

under strictly controlled conditions, thus eliminating the sample and environment effect. Although this methodology allows evaluating performances in ideal working conditions, in such a condition, the calibration of the frame compliance and the area shape function parameters becomes critical to both accuracy and precision. In fact, on one hand the accuracy depends on the calibration experimental conditions [40,41]; on the other hand, bias and improper precision estimation of the results might be introduced by inappropriate evaluation methods of the calibration measurement uncertainty, as highlighted in one of the most recent international round robin comparison [24]. Therefore, the development of statistical approaches to evaluate calibration measurement uncertainty is essential. The literature has proposed parametric simulative approaches to cope with the iterative nature of indirect calibration method [40]. However, these simulative approaches are based on MCMC simulation [39,40,43], which, as mentioned above, presents several shortcomings, and resulted in severe overestimation of the calibration measurement uncertainty for more unstable methods [39].

1.4. Scope of the work

According to the literature review previously discussed, the evaluation of measurement uncertainty of IIT results is essential to allow a practical application of nanoindentation also to establish and verify design specification and to the compare performances of nanoindentation testing devices. On the one hand performances should be compared in ideal working condition, on the other hand the procedure should be sensitive enough to be deployed in any operating condition and highlight relevant differences. In the former case, the most significant contributor to the uncertainty budget is the frame compliance and area shape function calibration whose uncertainty shall be properly evaluated. Accordingly, this work proposes a nonparametric simulative evaluation of the calibration uncertainty of the frame compliance and area shape function. The simulative approach is necessary to cope with iterative nature of some calibration methods, while allowing to test the effect of calibration dataset. This work will limit its scope to indirect calibration methods, for they are the most commonly resorted to. The evaluated measurement uncertainty will then be propagated to compare performances on mechanical characterisation results.

2. Methodology

As discussed in the Introduction, the literature proposes methods to evaluate measurement uncertainty based on MCMC [40,41]. However, MCMC, being parametric, requires performing non-trivial assumptions on the model input variables statistical distributions. Accordingly, extensive experimental plans shall be implemented to estimate these distributions' parameters. Moreover, in the case at hand, experimental input quantities, i.e., F_{max}, h_{max}, S_m, are dependent on each other, see Eq. (4) and Eq. (5). Though, to enable a simple management of the MCMC, correlations amongst these inputs were neglected. This might introduce errors in the response statistical distribution estimation and, consequently, in the evaluation of the measurement uncertainty [39].

In this work, a solution of these criticalities is addressed by exploiting a non-parametric method, i.e., Bootstrap, to estimate the measurement uncertainty. The bootstrap approach was introduced by Efron [44] and can be regarded as a particular type of MCMC. In fact, rather than performing computer experiments, i.e., simulations, on data sampled from statistical distributions a-priori determined, it resamples a pool of experimental data with replacement and re-entry. These K generated samples, i.e., the Bootstrap samples, are the input for the simulation and the estimate of a set of K model output(s). A general statement of Bootstrap method can be found in Efron [44] and relates to the problem of estimating the distribution and position and dispersion indexes of a population from which a sample $\boldsymbol{Y} = \{y_1, y_2, y_b, \cdots, y_B\}$ has been drawn. Assuming, in the most general case, a uniform sample probability distribution, i.e. $\mathbb{P}[Y = y_b] = 1/B$, sampling with replacement can be carried out to generate a new sample, i.e. a *bootstrap sample,* such that $\boldsymbol{Y}^* = \{y_1^*, y_2^*, y_b^*, \cdots, y_B^*\}$ where y_b^*, according to the sampling strategy and the uniform probability assumption, can be any observation of the original sample. The distribution of the population can be estimated relying on a MCMC by extracting K bootstrap samples and the study of the resulting position and dispersion indexes allow inference and

Table 1
ANOVA table for the estimate of the variance of regression results from a Bootstrap-based approach.

Factor	Degrees of freedom	Sum of Squares	Variance
Between	$K-1$	$s_B^2[x] \bullet (K-1)$	$\mathbb{V}ar[\hat{x}] \bullet IJ$
Within	$K \bullet (IJ-1)$	$s_W^2[x] \bullet K(IJ-1)$	$\mathbb{E}[SE^2[x]]$
Total	$KIJ-1$	SS_{TC}	$SS_{TC}/(KIJ-1)$

Table 2
Influence factors to input quantities of mechanical characterization following the calibration for measurement uncertainty of mechanical characterization.

Metrological characteristic	Input quantity			Contribution type	Distribution
	F_{max}	h_{max}	S_m		
Accuracy	±1%	±0.5 %	–	B	Uniform
Resolution	1 nN	0.04 nm	–	B	Uniform
Reproducibility	data	data	data	A	Normal

Table 3
Considered indentation devices with metrological characteristics of the force–displacement transducers. Reproducibility refers to best experimental conditions, i.e. indentations on SiO_2 at 10 mN.

Brand	Type	**Force** (resolution – noise floor – reproducibility)	**Displacement** (resolution – noise floor – reproducibility)
Hysitron	TriboScope	1 nN – 75 nN – 150 nN	0.006 nm – 0.2 nm – 1 nm
Hysitron	TI 950	1 nN – 100 nN – 100 nN	0.04 nm – 0.2 nm – 1 nm
Anton Paar	STeP6 NHT^3	20 nN – 1 μN – 3 μN	0.01 nm – 0.3 nm – 1 nm

Table 4
Calibrated reference materials set used in the experiments: average and expanded uncertainty (95% confidence level). Young modulus, reduced modulus and Poisson's ratio are calibrated with resonance frequency method, indentation hardness taken from literature and expanded uncertainty estimated as B-type contribution. Surface parameters computed on SL-surface to assess surface integrity.

Calibration set / materials		E / GPa	E_r / GPa	H_{IT} / GPa	ν / -	Sa / nm	Sq / nm
1	SiO_2	73.2 ± 0.6	70.51 ± 0.57	8.5 ± 1	0.162 ± 0.003	6.17	5.41
	W	412.9 ± 2.8	322.48 ± 1.78	8.0 ± 1	0.28 ± 0.003	7.87	10.73
2	SiO_2	73.3 ± 0.6	70.61 ± 0.54	9.5 ± 1	0.161 ± 0.003	5.48	7.24
	W	413.0 ± 2.8	322.27 ± 1.68	7.0 ± 1	0.281 ± 0.003	12.57	11.3
3	SiO_2	73.0 ± 0.5	70.38 ± 0.5	9.0 ± 1	0.163 ± 0.002	4.13	5.21
	W	414.3 ± 5.6	322.714 ± 5.6	7.0 ± 1	0.279 ± 0.005	7.7	10.44

estimation of the bootstrap distribution.

Here a procedure based on the Bootstrap method is outlined to estimate the measurement uncertainty of the calibrated parameters by indirect calibration methods. The empirical data set consists of J measured replicated indentation curves at I different maximum loads, and each curve contains B points. Therefore, the inputs result in pairs of F and h, both in $\mathbb{R}^{B,J,I}$. The Bootstrap samples will be sets of resampled $I \bullet J$ curves, i.e., pairs of F and h, both in $\mathbb{R}^{B,J,I}$. Each of them, to cater for the input correlation, i.e., $F(h)$, at the b-th, $b \in \{1, \cdots, B\}$, at the i-th load, will resample the b-th point of the IC from the sample of the J observations of this point $F(h) \in \mathbb{R}^{b,\bullet,i}$. In so doing, per each load, a maximum of J^B replicated curves may results. Considering that the bootstrap samples contain J replicated curve at each load, K is upper bounded by $J^{B\text{-}1}$. Therefore, K calibrations can be performed. Fig. 3 shows a graphical representation of the Bootstrap methodology.

Both the standard multi-step iterative approach [29,38] and the single-step method, see Eq.(6), require in input the calibrated mechanical properties of the calibration samples. These are sampled from parametric distributions and hold constant per each k-th bootstrap iteration. Conceptually, this allows including the contribution due to the traceability in the calibration method.

The K sets of calibrated parameters, let them be indicated as X, are exploited to compute their standard uncertainties as the standard deviation of a group mixture, modelled with the Analysis of Variance (ANOVA), as per Table 1. The bootstrap results, i.e. the set of K estimates of the parameters X, prior the evaluation of the average and variance, will be removed from outliers [45], identified by modified interquartile range method [46], and of non-physical values, i.e. any iteration leading to a negative estimation of the frame compliance C_f. The adoption of ANOVA modelling caters for the degrees of freedom of the estimates, as per:

$$SS_{TC} = SS_W + SS_B \tag{8}$$

where SS_{TC} is the total corrected by the mean sum of squares, SS_W the sum of squares within the groups and SS_B the sum of squares between the groups. ANOVA operates on variances of the parameter; therefore, the between iterations contribution is estimated from the variance of the mean estimate of the calibrated parameters, $\mathbb{V}ar[\hat{x}]$, i.e., variance of a sample mean, by multiplying it for the sample numerosity, IJ. Last, hypothesis test based on Fisher distribution at a conventional 95 % confidence level will be performed to test the significative contribution due to the variability between bootstrap samples on the calibrated parameters. This will highlight any possible statistical significative systematic effect due to the calibration data set.

ANOVA total variance modelling combining within and between contribution can then be employed to propagate uncertainties according to GUM, see Eq. (7), and estimate mechanical characterization results' measurement uncertainties. To this extent, the raw measured inputs, i.e., F_{max}, h_{max} are considered to be influenced by reproducibility, accuracy, and resolution of the force–displacement transducer. The respective contributions are summarised in Table 2. The standard uncertainty of S_m, which accounts for its reproducibility, is determined according to the literature [47] from measured data and associated with a normal distribution.

The propagation of uncertainty contributions for mechanical characterization parameters, e.g., E_{IT}, $H_{IT,}$ E_r, will ultimately consist of two main contributions, i.e. u_{repr} and u_{acc},

$$u_Y^2 = u_{Repr}^2 + u_{Acc}^2 \tag{9}$$

where u_{repr} is the reproducibility, u_{acc} the accuracy and Y any mechanical characterisation results. Specifically, the u_{repr} includes the influence of input parameters of Table 2 and the standard uncertainty of calibrated parameters, combined according to Eq. (7) considering the definition in Equations from Eq. (1) to Eq. (5); the u_{acc} is due to the accuracy of the measured value and is computed as the Root Mean Squared Error (RMSE) of the estimated characteristic with respect to the calibrated value and propagated with a uniform distribution, i.e. $u_{Acc}^2 = \frac{MSE}{3} J$.

Additionally, and according to the defined framework, the characterisation parameter R will be computed.

$$R = \frac{F}{S^2} = \frac{\pi}{4} \frac{H_{IT}}{E_r^2} \tag{10}$$

Such parameter is particularly convenient to evaluate the accuracy of

Table 5
Experimental conditions to gather the data set to demonstrate the methodology for estimating quality of nanoindentation devices. To avoid eliciting commercial benchmarking, nanoindentation devices ID is blinded. Dataset A comes from past interlaboratory comparison, the fifth load, i.e. 0.1 mN, was not reproducible and excluded from the analysis. Dataset E presents more loads to assess mechanical properties out of the calibrated range (1 ∼ 10) mN.

Experiment ID	Indentation device ID	Calibrated reference material set	Investigated loads / mN	Metrological room
A	a	1	(0.5, 1, 5, 10)	Yes
B	b	2	(1, 3.25, 5.5, 7.75, 10)	No
C	c	2	(1, 2.5, 5, 7.5, 10)	No
D	c	2-damaged	(1, 2.5, 5, 7.5, 10)	Yes
E	c	3	(1, 2.5, 5, 7.5, 10, 15, 20)	Yes

the frame compliance calibration method, dispensing the effect of the indenter area shape function calibration, for it is defined as $\frac{F}{S^2}$. Furthermore, because it can be estimated on the base of mechanical characterisation, it can be evaluated on calibrated samples, thus enabling a traceable assessment of the calibration performances [39,40].

3. Case study

The methodology described in Section 2 will be applied on several datasets collected on different indentation devices catering for both ideal working conditions, i.e. highly polished calibrated samples and measurements performed in metrological and environmentally controlled rooms, to evaluate best performances, and uncontrolled working conditions, i.e. rougher surfaces and non-metrological rooms. Different indentation device types and brands will be considered, which to avoid any competing interest will not be associated to the specific results.

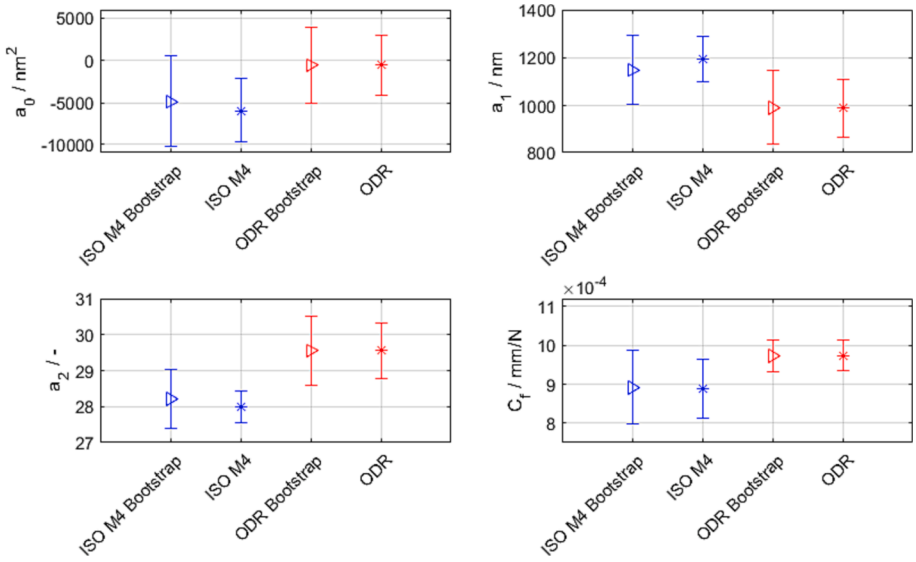

Fig. 4. Calibrated parameter results for the experimental data set A. Notice the systematic underestimation of uncertainty when Bootstrap is not applied. Notice the sometimes systematic differences between the calibration methods (ISO and ODR) as per [39].

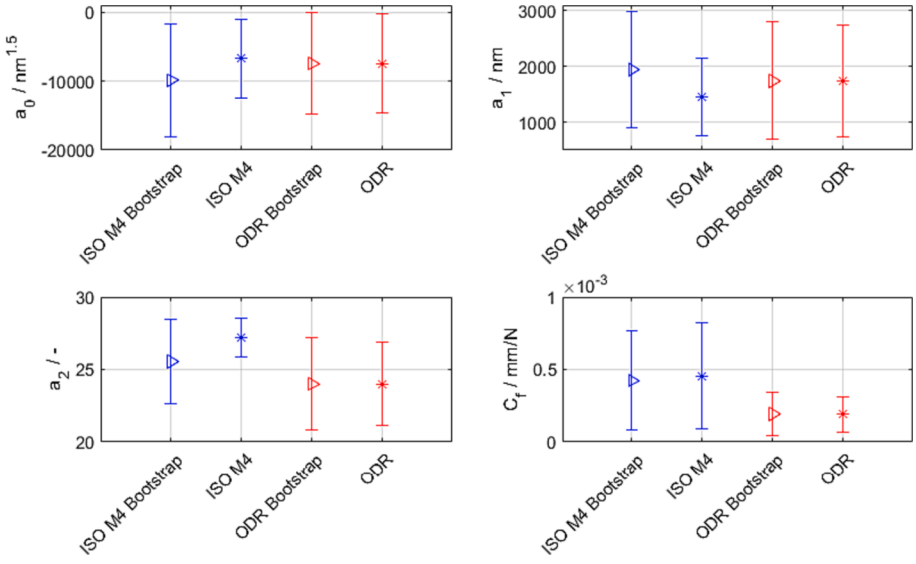

Fig. 5. Calibrated parameter results for the experimental data set E. Notice the systematic underestimation of uncertainty when Bootstrap is not applied.

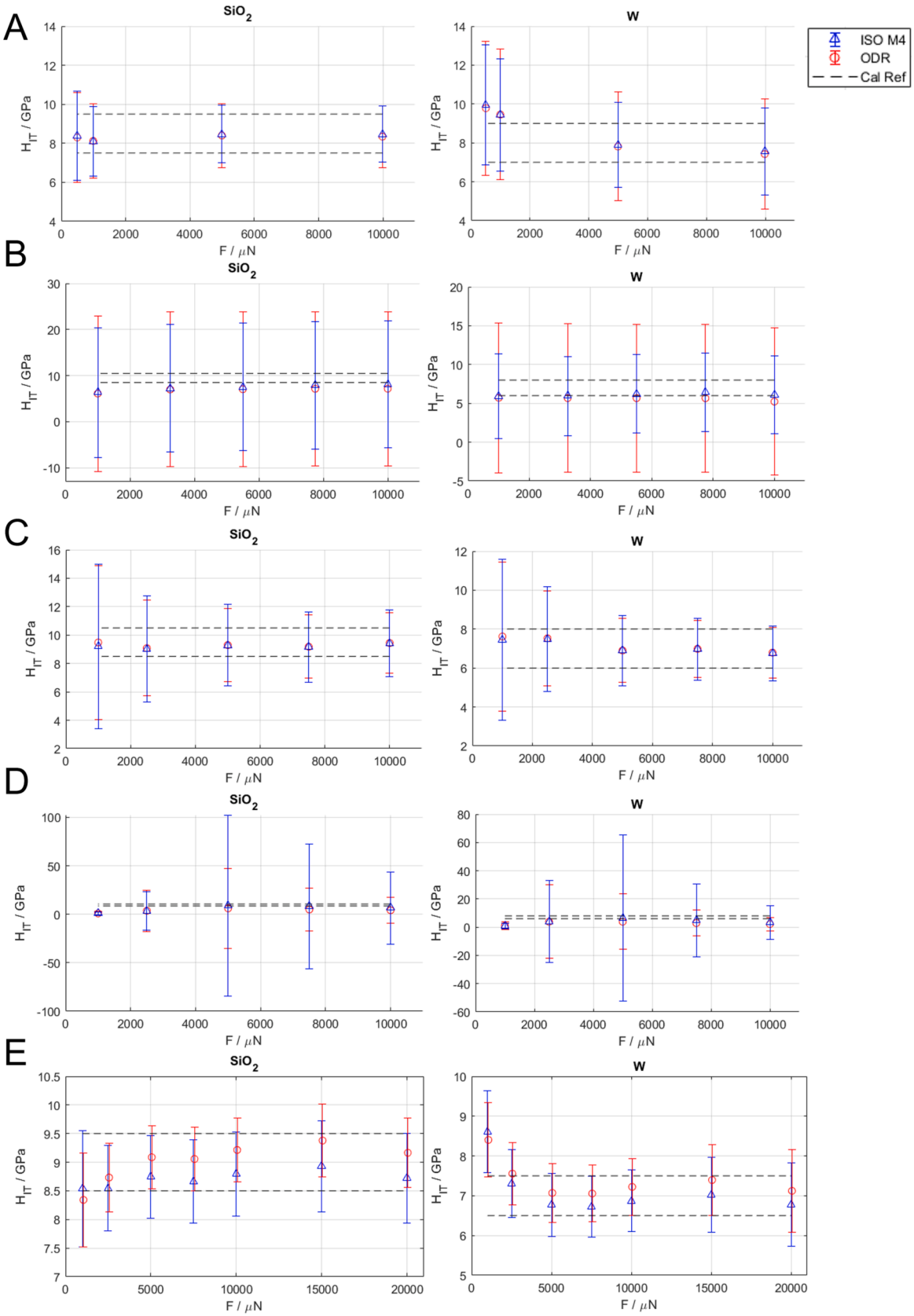

Fig. 6. Mechanical characterisation results: H_{IT}. Error bars represent expanded uncertainty (95% confidence level). Notice that only in ideal working condition, i.e. A and E, indentation size effect on W is appreciated. B: poorly maintained instrument in a non-metrological room mostly affects precision with marginal negative effects on accuracy, C adequately maintained instrument in a non-metrological room negatively affects precision, D poor surface of reference materials induces poor quality data.

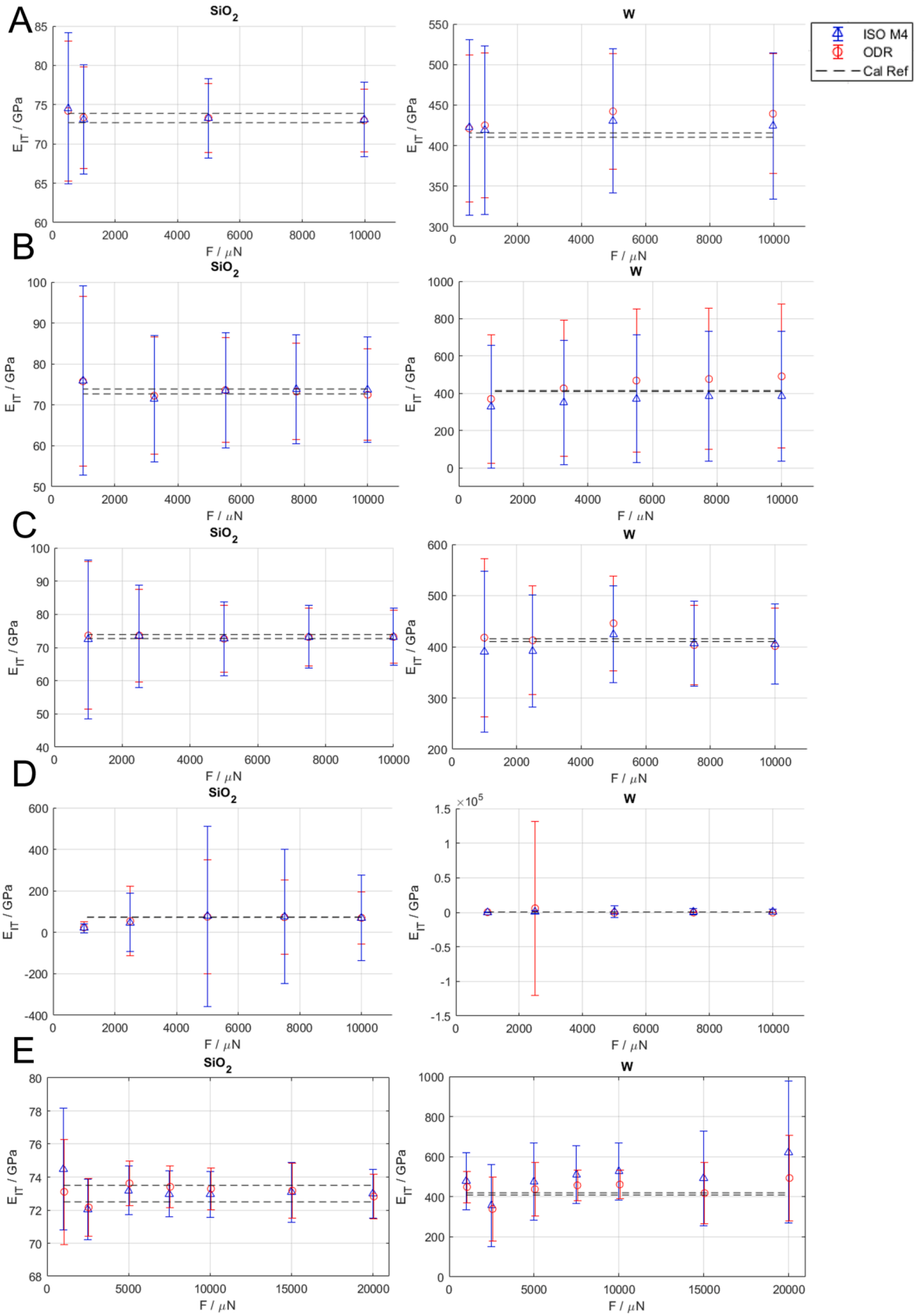

Fig. 7. Mechanical characterisation results: E_{IT}. Error bars represent expanded uncertainty (95% confidence level). In ideal working condition, i.e. A and E, ORD is more precise and accurate than ISO M4, i.e. when reproducibility does not dominate uncertainty. B: poorly maintained instrument in a non-metrological room mostly affects precision with negative effects on accuracy, C adequately maintained instrument in a non-metrological room negatively affects precision, D poor surface of reference materials induces poor quality data. E a non-significative bias can be seen in the results.

Table 6
Relative accuracy reported as percentage with respect to the calibrated reference values. Accuracy is evaluated as RMSE with respect to the calibrated reference value. Systematic errors due to indentation size effect in W have been excluded for the evaluation. Absolute accuracy is reported in the annex Table A 3.

	ISO M4			ODR		
	H_{IT} SiO_2 / W	E_r SiO_2 / W	E_{IT} SiO_2 / W	H_{IT} SiO_2 / W	E_r SiO_2 / W	E_{IT} SiO_2 / W
A	2.4 % / 4.1 %	0.8 % / 2.2 %	0.9 % / 2.9 %	2.8 % / 5.4 %	0.6 % / 1.7 %	0.7 % / 5.1 %
B	22.7 % / 10.9 %	1.9 % / 9.8 %	2 % / 13 %	27.7 % / 21.3 %	1.6 % / 9.4 %	1.7 % / 13.5 %
C	3.3 % / 2.2 %	0.6 % / 2.7 %	0.7 % / 3.8 %	2.7 % / 1.9 %	0.5 % / 2.7 %	0.5 % / 4 %
D	52.4 % / 35.8 %	35.3 % / 59.1 %	36.1 % / 104.1 %	61.6 % / 58.5 %	31.5 % / 268.5 %	32.3 % / 612.3 %
E	6.2 % / 9.1 %	0.9 % / 16.8 %	0.9 % / 26.3 %	4.6 % / 8.5 %	0.6 % / 8.6 %	0.6 % / 12.2 %

Table 7
Relative expanded uncertainty (in percentage) for the calibrated reference values. The uncertainty includes reproducibility and accuracy. Worst case, i.e. the maximum at different loads, is reported along with a more optimistic evaluation in parenthesis which either is associated with higher loads or excludes particularly noisy datasets. Expanded Uncertainty is reported in annex Table A 4. The dataset D has been obtained with a damaged surface samples set in a metrological room: surface integrity dominates the precision.

	ISO M4			ODR		
	H_{IT} SiO_2 / W	E_r SiO_2 / W	E_{IT} SiO_2 / W	H_{IT} SiO_2 / W	E_r SiO_2 / W	E_{IT} SiO_2 / W
A	27 % (17 %) / 30 %	12 % (6 %) / 17 %	13 % (6.5 %) / 24 %	27 % (19 %) / 36.5 %	11 % (5 %) / 14 %	12 % (5.5 %) / 18.5 %
B	195 % / 85 %	28.5 % (16 %) / 70 %	30 % (17 %) / 95 %	250 % / 170 %	25.6 % (14.6 %) / 60 %	27 % (15 %) / 80 %
C	63 % (25 %) / 55.5 % (21 %)	31 % (11 %) / 29.5 % (14 %)	33 % (12 %) / 40 % (19 %)	57 % (22 %) / 50 % (19 %)	28 % (10 %) / 27 % (13 %)	30 % (11 %) / 37 % (18 %)
D	500 % / 575 %	300 % / 275 %	330 % / 510 %	450 % / 450 %	230 % / 230 %	240 % / 130 %
E	12 % (8 %) / 15.5 % (11 %)	4.5 % (1.5 %) / 34 % (9 %)	5 % (2 %) / 58 % (27 %)	10 % (6 %) / 14.6 % (10 %)	4 % (1.5 %) / 33 % (6 %)	4 % (2 %) / 47 % (15 %)

3.1. Experimental setup

The considered indentation devices are reported in Table 3 along with main metrological characteristics of the force–displacement transducers. Indentations are performed on reference samples sets of silica (SiO_2) and tungsten (W) calibrated in terms of E and E_r; reference value of H_{IT} is taken from the literature associating a type B measurement uncertainty [39]. Table 4 reports calibrated values. The surface integrity of the samples is assessed in terms of Sa and Sq in Table 4. Surface parameters were measured with a Coherence Scanning Interferometer Zygo NewView9000 with a 20 × Mirau objective (numerical aperture of 0.4, pixel size of 0.7 μm) on Scale Limited-surface (S-filter nesting index 0.5 μm, least square tilt correction as F-Operator, and L-filter nesting index 0.5 mm).

Indentations have been performed with a conventional quasi-static loading-holding-unloading force-controlled cycles of (30–15-30) s with a constant strain rate of 0.017 s^{-1} [24]. Only the nano range has been investigated, performing 15 indentations at 5 different maximum forces, see Table 5, to dispense nonlinearity of the C_f [27]. The considered force range allows to characterize the indentation devices in a displacement range from 50 nm to 400 nm, respectively obtained on W at 0.5 mN and on SiO_2 at 20 mN.

Two ideal experimental conditions are considered, i.e. within a metrological environment and brand new calibrated reference materials and indenters (A, E). Furthermore, damaged surfaces reference samples (D) and measurements performed in non-metrological rooms (B, C) were considered. The damaged sample set has a Sa of 0.11 μm and Sq of 0.19 μm. It is worth remarking that due to how the different instruments were maintained and regularly calibrated, indentation device *b* is supposed to perform worse. The other two indentation devices have been regularly maintained and calibrated, and *c* has been manufactured about 10 years later than *a*. Table 5 summarizes the considered conditions. Indentation devices ID, with respect to Table 3, are blinded to avoid commercial benchmarking, which is not within the scope of this work. Indenters had a modified Berkovich geometry, i.e. a pyramid with triangular base and the nominal half-dihedral angle of 65.278°. A different indenter was used per each indentation device. Indenters' area shape functions are calibrated considering a constant term, i.e. $A_p(h_c) = a_2h_c^2 + a_1h_c + a_0$, for A dataset, and square root approximation in other cases, i.e. $A_p(h_c) = a_2h_c^2 + a_1h_c + a_0\sqrt{h_c}$. In the latter case, first equation of Eq. (6) is modified accordingly to accommodate the proportionality of a_0 to $\sqrt{h_c}$ The selection of the area shape function aims at optimizing the mechanical characterisation accuracy [31]. The different considered maximum loads are due to the fact that some data sets come from past experimental campaigns, that can be regarded as reference [24], or dismissed indentation devices.

To allow a robust numerical implementation of the bootstrap sampling methodology described in Section 2, two statistical checks on the input and output data have been added. As far as the input data are concerned, a statistical method to remove outlier indentations has been implemented, basing on the application of the interquartile range method on the maximum measured penetration depth. If any outlier indentation curve (IC) is highlighted, it is removed from the experimental dataset before it is exploited to run the bootstrap. This leaves in few situations with uneven data sets. In general, no more than a couple of indentations were removed per each load in the most critical experiments, i.e. B and D, for the presence of outlier indentations is promoted by critical experimental conditions, and smallest loads, which are less reproducible. Removing such outlier indentations is consistent with common empirical practice of disregarding indentation curves offset from replicated conditions. The verification on the output data, i.e. on the *j-i-k*-th bootstrap generated indentation curve (that is, the j-th bootstrap-generated indentation curve replication at the the i-th maximum load for the k-th bootstrap generated sample) is expedient to allow subsequent data treatment. In fact, each generated IC shall allow the evaluation of the measured contact stiffness S_m, which requires the application of a regression on the unloading portion of the IC. However, some bootstrap sampling maybe particularly critical for they might extract furthest points in the envelope of empirical IC, considering the same load. This might lead to negative Jacobian matrixes hindering regression to estimate S_m. Therefore, an output data verification has been implemented, and if such an exception is raised, the bootstrap-generated *j-i-k*-th IC is discarded and re-generated. A complete overview of the procedure is provided in the Annex, in the flowchart of Fig. A 6.

Implementation of bootstrapping and ISO calibration method is performed by an in-house developed code in MATLAB 21b, Orthogonal Distance Regression (ODR) is implemented in Python 3.7. Computations rely on parallel computing and are performed on a workstation (i7-8700, RAM 32 GB,).

3.2. Results and discussion

In the following, the results are presented and discussed. First, the relevance of the non-parametric propagation of uncertainty on the

calibrated parameters, i.e. the frame compliance C_f and the area shape function parameters a_i, is reported. Then, the mechanical characteristics of the reference materials are evaluated to assess accuracy and precision of compared methodologies. In the following, main results are reported, while complementary images and tables are presented in the Annex. Time to generate 11,000 bootstrap samples spans from 8 to 12 h, per material, depending on the quality of the input experimental data. Although the time is quite high, considering that the procedure aims at calibration uncertainty evaluation long time may be excused; in any case, bootstrap simulation is significantly cheaper and faster than empirical data collection, which per each dataset (corresponding to 1 bootstrap sample) takes about 8 h.

3.2.1. Calibration of indentation platforms parameters

Fig. 4 and Fig. 5 show the result of the calibration of the frame compliance and area shape function parameters for the data set A and E, respectively. The average is evaluated after outlier removal, and the error bars represent 95 % confidence intervals estimated from the standard deviation with a coverage factor of 2. ANOVA shows with a risk of error of 5 % that the simulations introduce a systematic difference in the estimated parameters; summary results are reported in Table A 1 and Table A 2. This sustains the methodological hypothesis presented in this paper on the necessity of simulative approaches to correctly estimate the uncertainty of calibrated parameters. Such result holds for both the considered calibration method, i.e. the ISO 14577–2 method n°4 (ISO M4) and the single-step ODR method. It is particular to notice that experimental dataset collected in critical conditions, e.g. B (poorly maintained indentation device and experiments in a non-metrological room) and D (indentation on damaged-surface reference samples), results with such a large within-the-group-variance that the bootstrap sample effect, i.e. the calibration dataset variability, becomes negligible. The greater robustness of the ODR method still allows highlighting such effect. Such ISO method instability to poor working conditions leads to estimation of the frame compliance statistically not different from 0 mm/N (see Fig. A 1, Fig. A 2 and Fig. A 3), i.e. suggesting a non-necessary correction, which might negatively impact on mechanical characterisation results. Furthermore, an indication of stability of the calibration can be provided by the analysis of the a_2 parameter, whose nominal value is 24.56 and it is correlated to the dihedral angle of the indenter. The regression does not fix its value to improve the fitting [31], thus slight, even systematic, deviation from this value might be acceptable, e.g. case A (Fig. 4), because of a compensation with other calibration parameters might take place [31,39–41]. However, the presence of large systematic deviation, i.e. case B (Fig. A 1) and D (Fig. A 3) might be good indicators of poor empirical dataset. Additionally, Fig. 4 and Fig. 5 (as well as the other figures in the Annex, i.e. Fig. A 1, Fig. A 2 and Fig. A 3) report the results of calibration when only experimental data are considered, i.e. conceptually similar to a Bootstrap simulation with only 1 bootstrap sample. In accordance with bootstrap theory, the average estimates of the parameters do not show any systematic differences with a risk of error of 5 %, and the uncertainty, in line with ANOVA results, is underestimated in this case.

3.2.2. Mechanical characterisation

Once the frame compliance and the area shape function parameters have been calibrated, mechanical characterisation can be performed according to equations Eq. (1) ~ Eq.(5) and propagating uncertainty contribution according to the law of uncertainty propagation, see Eq. (7) [42], considering the influence factors reported in Fig. 1. As far as the calibrated measurement scales, i.e. the force and the displacement transducers, are concerned, resolution, reproducibility and accuracy are propagated. The resolution is propagated as type B contribution considering the resolution the range of a uniform distribution [42]. The accuracy is propagated as type B contribution from the calibration certificate considering the standard 1 % error within the tolerance and considering as extremely conservative choice the association of such deviation to the standard deviation of a normal distribution [42].The reproducibility is propagated as a type A from the collected data, which in the worst case scenario, i.e. measurements on W at 1 mN, increases up to five times the value reported in Table 3. In so doing, the reproducibility includes the experimental condition effect, i.e. sample and environment. The frame compliance and area shape function parameters uncertainty contribution are propagated as type A from the calibration uncertainty estimated in Section 3.2.1. Last, the measured contact stiffness uncertainty is propagated as a type B contribution, assuming a gaussian distribution with standard deviation estimated as 5 % of the measured value [42]. The estimation of the contribution caters for reference literature, which showed the relevance of combining, once again with ANOVA modelling, the contribution due to the fitting of the unloading curve and the reproducibility due to different indentations [47].

In the following, Fig. 6 and Fig. 7 show the results of the evaluation of the H_{IT} and E_{IT}, error bars represent uncertainty at a 95 % confidence level (coverage factor equal to 2). Results of E_r and the area calibration-independent parameter R are reported in the Annex (Fig. A 4 and Fig. A 5), for sake of readability. Table 6 and Table 7 show the relative accuracy and relative expanded uncertainty of the mechanical characterisation, respectively.

Results show that indentation devices working in ideal condition (adequately and routinely maintained and operating in metrological controlled environments), i.e. A and E, show good accuracy, i.e. with statistically non-significant bias, and precision. In the best-case scenario, the relative accuracy on H_{IT} and E_{IT} result as low as 5 % and 0.5 % respectively, and a relative uncertainty of 5 % and 2 %. Furthermore, the method to assess indentation quality is sensitive to sample material. In fact, indentations on tungsten are less accurate and precise of one order of magnitude than those on SiO_2.

The method to assess quality of indentations is then tested on non-ideal operating conditions, i.e. a poorly maintained instrument working in non-metrological room (B), a maintained and calibrated instrument in a metrological room (C) and a maintained and calibrated instrument in a metrological room with damaged samples (D). The proposed method to estimate uncertainty and assess indentation quality allows to gather insights on the sensitivity of nanoindentation to the considered experimental conditions. Accuracy results to be mostly affected by instrument rather than working environment. In fact, a poorly maintained instrument (B) worsens of about three times the relative accuracy, whilst a regularly maintained indentation device, even in a non-metrological room (C), does not affect significantly the accuracy. As far as precision is concerned, the instrument affects the measurement uncertainty significantly more than the operating environment, which is worsened when more dispersed results are expected, i.e. in the case of the tungsten. This might pose some question on the choice of tungsten as reference material and open to other alternatives, e.g. alumina, sapphire. Furthermore, as clearly shown in Fig. 6, tungsten hardness allows appreciating the indentation size effect at very low loads. This though is only possible in ideal working conditions, i.e. A and E, whilst measurement disturbances sourced by non-ideal operating condition can hide such effect. Last, experimental condition D shows that the surface integrity ultimately dominates both accuracy and precision, i.e. the quality, of nanoindentation results by yielding the worst metrological performances.

Finally, a remark on the calibration method can be performed basing on the relative accuracy and precision. As it can be appreciated, and in accordance with reference literature [39], ODR improves the accuracy while worsening the precision, due to orthogonal distance evaluation of the RMSE, when reproducibility is poor; conversely, in ideal working condition, i.e. with good reproducibility, the fitting improvement overcomes the inherent RMSE increase due to the ODR and results in both an improvement of accuracy and a marginal decrease of the measurement uncertainty.

4. Conclusions

Instrumented indentation test is a widely resorted non-conventional hardness test for evaluating mechanical properties of surfaces. Ensuring quality of indentation is of utmost importance to allow indentation device performance comparison and to enable the prescription and verification of tolerances by nanoindentation within ISO GPS framework.

The work has proposed a method to assess the quality of nanoindentation within a metrological framework. With respect to other approaches attempted in the literature, the methodology is based on a non-parametric evaluation of measurement uncertainty of indentation device calibration. The proposed Bootstrap uncertainty evaluation method, tested by ANOVA, showed that the effect of simulation modelling different experimental dataset is critical to obtain a correct and conservative uncertainty evaluation. The method has then been successfully tested on several experimental conditions to investigate its sensitivity for indentation quality benchmarking. Results have shown that the proposed uncertainty-based quality evaluation tool can identify critical experimental conditions, thus posing as a suitable methodology to compare performances within interlaboratory comparisons. Although the conditions that were investigated are non-exhaustive, extreme laboratory conditions, i.e. ideal and worst case scenario, as well as intermediate cases were considered. Therefore, the obtained result indicates a suitable sensitivity of the methodology, which future work shall investigate further with an ad-hoc experimental design. The provided quantitative insights on the influence factors contributing metrological performances: accuracy is mostly affected by the indentation device maintenance, precision by the sample surface integrity foremostly and only at a second instance by the indentation device, hereby including the positioning in a metrological room, which is capable of halving the uncertainty for optimal sample conditions.

Future works will deploy the proposed uncertainty-based quality evaluation tool to assess performances of calibration material, investigating viable alternative to tungsten, and to compare the effect of indentation device indirect calibration methods (which are managed by the proposed tool) and other direct calibration approaches.

CRediT authorship contribution statement

Giacomo Maculotti: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Gianfranco Genta:** Methodology, Formal analysis, Validation, Writing – review & editing, Supervision. **Maurizio Galetto:** Methodology, Validation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to thank Prof. Don A Lucca from Oklahoma State University for sharing data from the CIRP intercomparison, and Proff. Daniele Ugues and Diego Manfredi and Dr. Massimo Lorusso for having allowed the access to Istituto Italiano di Tecnologia facilities. This work has been carried out within the project 22DIT01 ViDiT which has received funding from the European Partnership on Metrology, co-financed from the European Union's Horizon Europe Research and Innovation Programme and by the Participating States.

Appendix

(See Fig. A1, Fig. A2, Fig. A3, Fig. A4, Fig. A5, Fig. A6 and Table A1, Table A2, Table A3, Table A4).

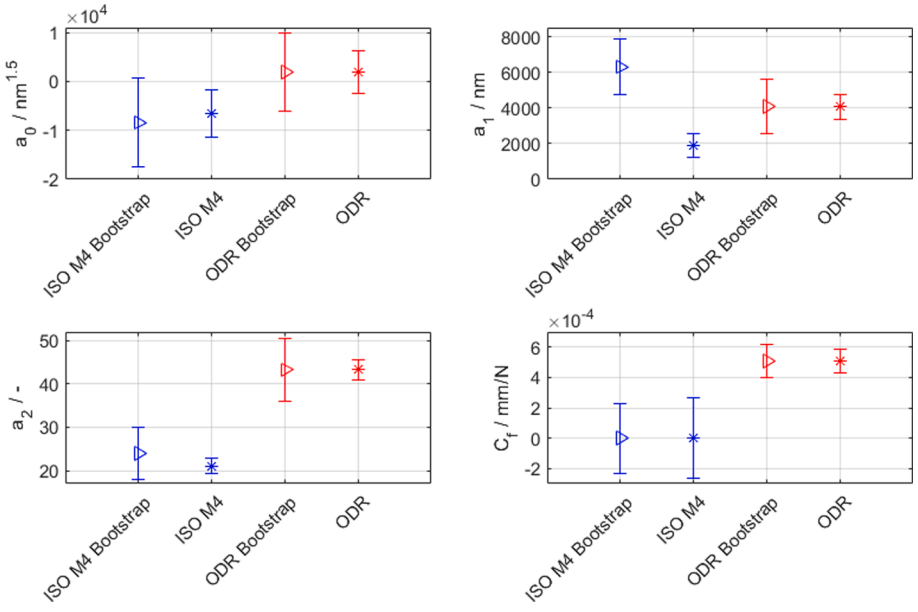

Fig. A1. Calibrated parameter results for the experimental data set B (poorly maintained indentation device and experiments in a non-metrological room). Notice the systematic underestimation of uncertainty when Bootstrap is not applied. Notice the sometimes systematic differences between the calibration methods (ISO and ODR) as per [39]. Notice ISO M4 poorer robustness resulting for C_f in worse precision [39] and C_f estimates statistically not different from 0 mm/N.

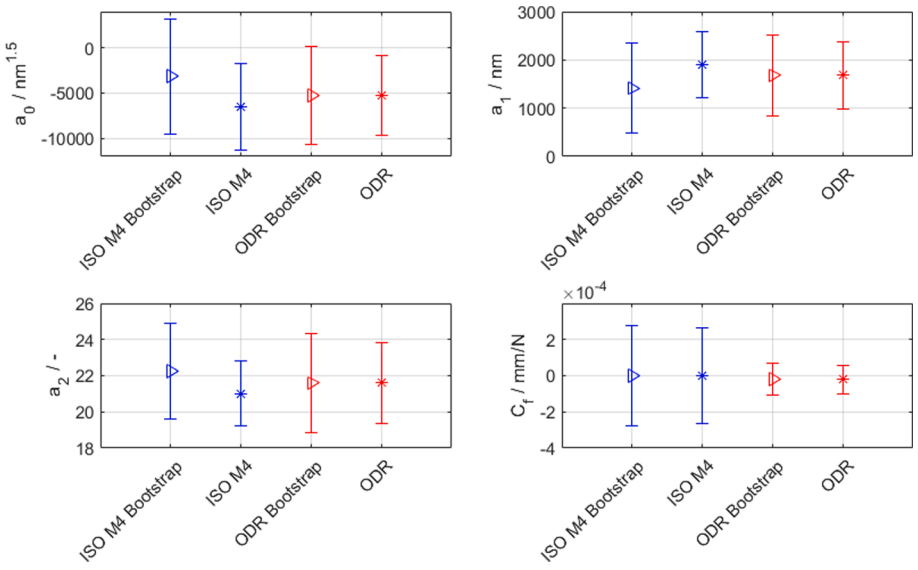

Fig. A2. Calibrated parameter results for the experimental data set B (poorly maintained indentation device and experiments in a non-metrological room). Notice the systematic underestimation of uncertainty when Bootstrap is not applied. Notice the sometimes systematic differences between the calibration methods (ISO and ODR) as per [39]. Notice ISO M4 poorer robustness resulting in worse precision [39] and estimates of C_f statistically not different from 0 mm/N.

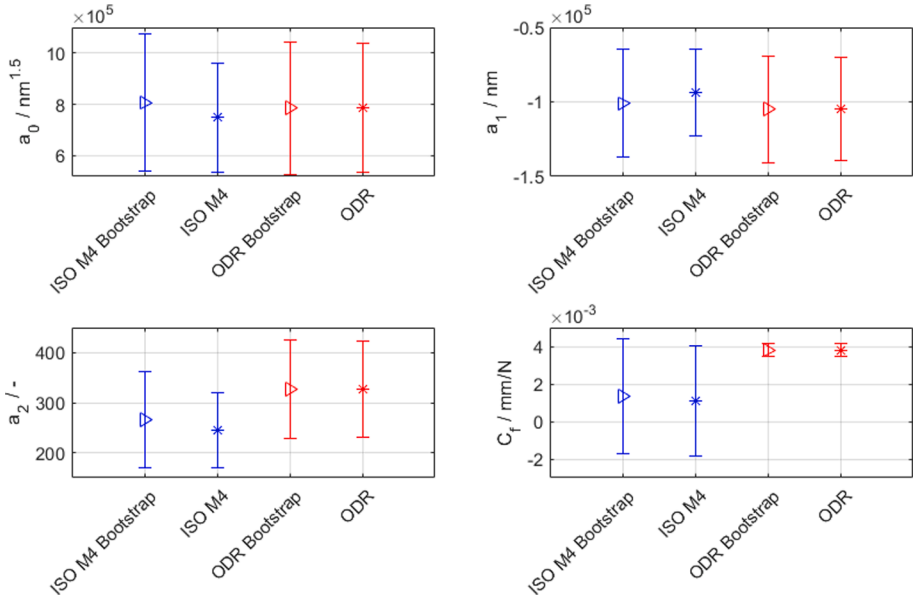

Fig. A3. Calibrated parameter results for the experimental data set D (experiments performed in a metrological room on damaged-surface calibrated reference samples). Notice the systematic underestimation of uncertainty when Bootstrap is not applied. Notice ISO M4 poorer robustness resulting in worse precision [39] and estimates of C_f statistically not different from 0 mm/N.

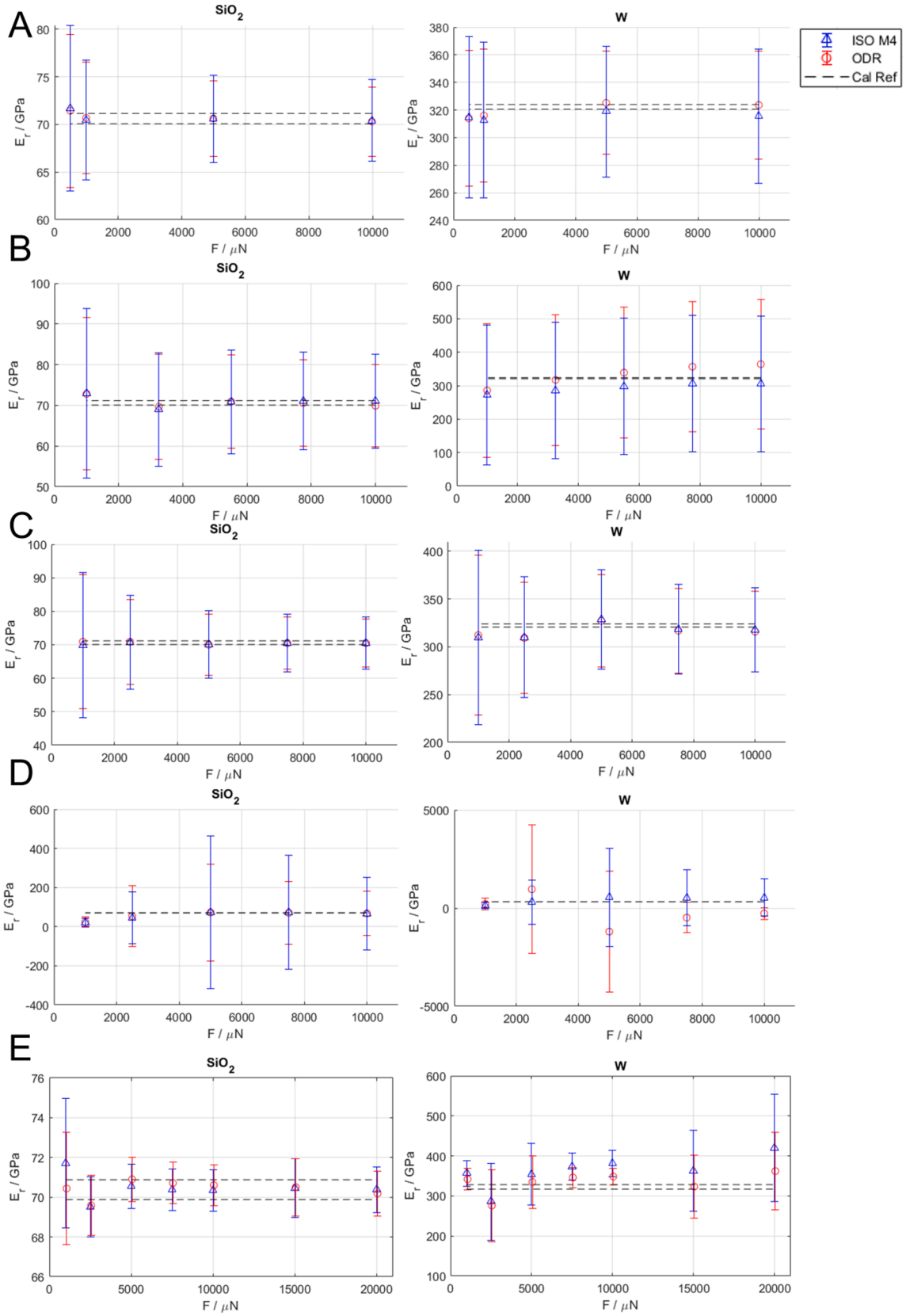

Fig. A4. Mechanical characterisation results: E_r. Error bars represent expanded uncertainty (95% confidence level). In ideal working condition, i.e. A and E, ODR is more precise and accurate than ISO M4, i.e. when reproducibility does not dominate uncertainty. B: poorly maintained instrument in a non-metrological room mostly affects precision with negative effects on accuracy, C adequately maintained instrument in a non-metrological room negatively affects precision, D poor surface of reference materials induces poor quality data. E a non-significative bias can be seen in the results. The results reflect E_{IT} (see Fig. 7).

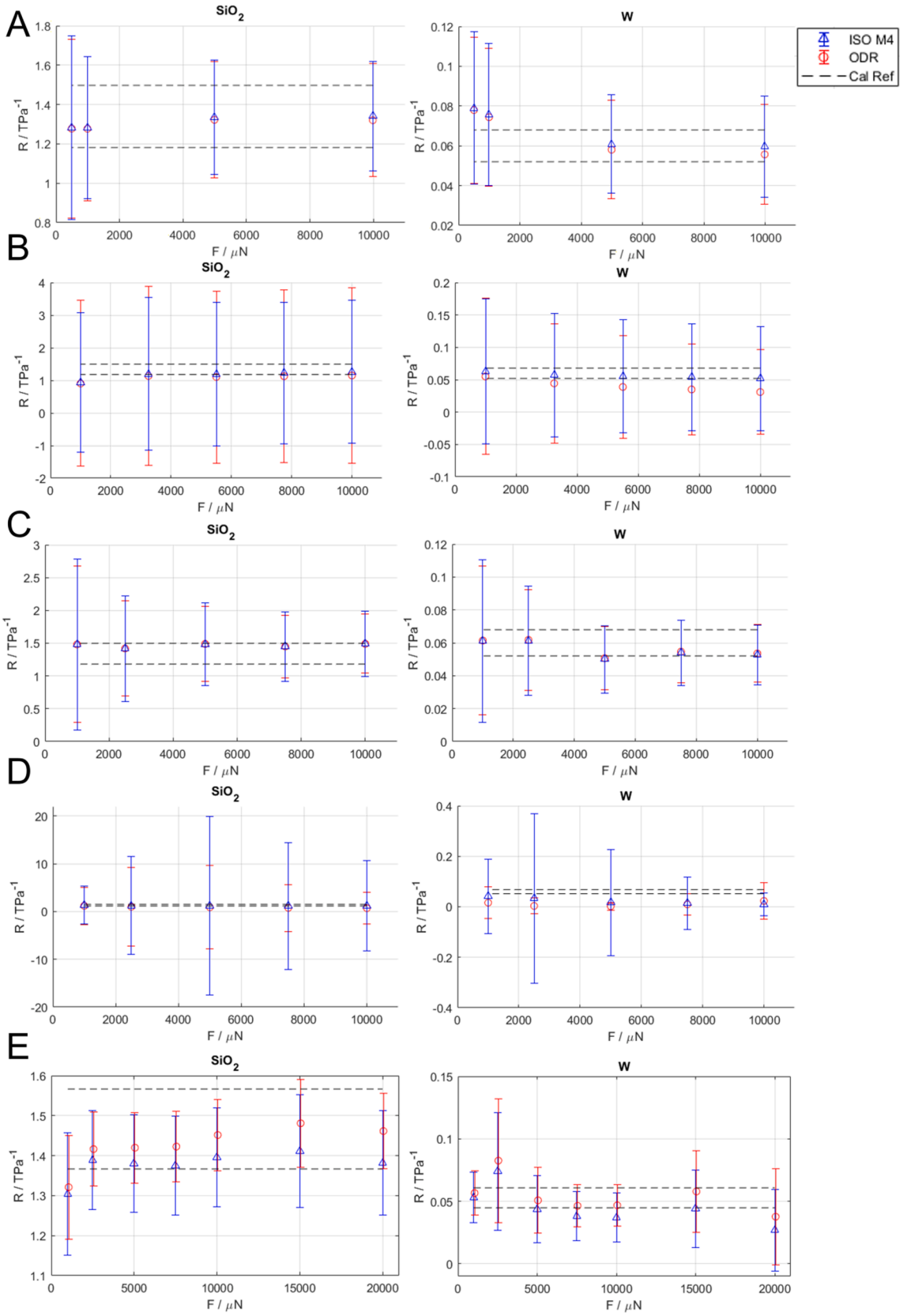

Fig. A5. Mechanical characterisation results: *R*. Error bars represent expanded uncertainty (95% confidence level). In ideal working condition, i.e. A and E, a trend is seen on W that can be traced back to indentation size effect. Otherwise and mainly, statistically non-significative trend can be seen indicating a successful calibration for the results do not depend on the area shape function. B: poorly maintained instrument in a non-metrological room negatively effects precision, condition alleviated in C (adequately maintained instrument in a non-metrological room), D poor surface of reference materials induces poor quality data showing massive dispersion and presence of some systematic differences (on W).

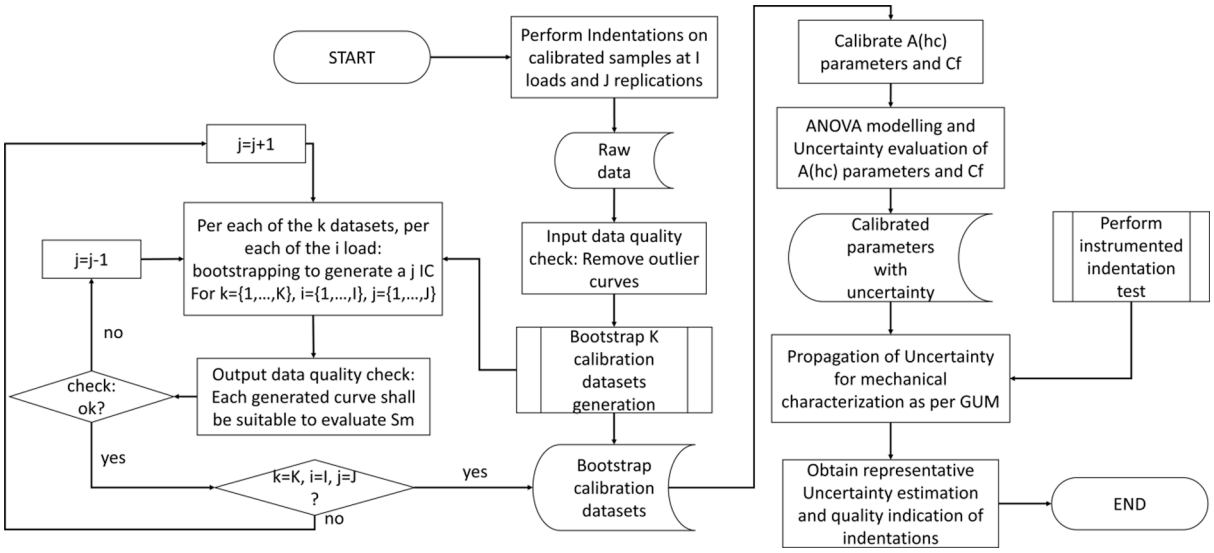

Fig. A6. Flowchart of the quality evaluation procedure based on the bootstrap non-parametric simulation. In the flowchart showing the calibration and the uncertainty propagation for any given IIT results, the subprocess based on Bootstrap simulation to robustly estimate calibration related uncertainty is shown also highlighting the input and output data quality checks.

Table A1
Summary results of the ANOVA. P-values of the Fisher hypothesis test on the significance of the between the groups variability as variance ratio is shown. ***≪0.0001 %.

Experimental Dataset	A		B		C		D		E	
Calibration algorithm	ISO	ODR	ISO	ODR	ISO	ODR	ISO	ODR	ISO	ODR
Parameter										
a_0	***	***	1	***	***	***	***	***	***	***
a_1	***	***	1	***	***	***	***	***	***	***
a_2	***	***	1	***	***	***	***	***	***	***
C_f	***	***	1	***	1	***	***	0.97	***	***

Table A2
Summary results of the ANOVA. Relative contribution of the sum of squares between the groups to the total corrected by the mean sum of squares, see Eq. (8), i.e. SS_B/SS_{TC}.

Experimental Dataset	A		B		C		D		E	
Calibration algorithm	ISO	ODR	ISO	ODR	ISO	ODR	ISO	ODR	ISO	ODR
Parameter										
a_0	25.5 %	41.4 %	0.1 %	7.1 %	49.3 %	39 %	31 %	4.2 %	31.7 %	27.8 %
a_1	36.5 %	44.6 %	0 %	8 %	54 %	39.3 %	29.5 %	3.7 %	39.9 %	31.1 %
a_2	58.7 %	42.3 %	0 %	10.7 %	59.1 %	38.6 %	31.3 %	2.8 %	70.8 %	41.1 %
C_f	47.9 %	14.3 %	0 %	22 %	0 %	21 %	5.1 %	2.4 %	23.4 %	50.5 %

Table A3
Accuracy: RMSE with respect to the calibrated reference value. Systematic errors due to indentation size effect in W have been excluded for the evaluation. Values are in GPa.

	ISO M4			ODR		
	H_{IT} SiO_2 / W	E_r SiO_2 / W	E_{IT} SiO_2 / W	H_{IT} SiO_2 / W	E_r SiO_2 / W	E_{IT} SiO_2 / W
A	0.2072 / 0.3277	0.5687 / 7.1669	0.6297 / 11.84	0.2344 / 0.4283	0.4328 / 5.4629	0.4791 / 20.9834
B	2.1586 / 0.7639	1.3283 / 31.6628	1.4710 / 53.8688	2.6304 / 1.4915	1.1308 / 30.2881	1.2528 / 55.5952
C	0.3126 / 0.1531	0.4344 / 8.8558	0.4803 / 15.6071	0.2543 / 0.1334	0.3270 / 8.5403	0.3617 / 16.3394
D	4.9755 / 2.5063	24.9104 / 190.4255	26.4738 / 429.8351	5.8552 / 4.0981	22.2453 / 865.1444	23.669 / 2529
E	0.5594 / 0.6388	0.6059 / 54.1321	0.6697 / 109	0.4109 / 0.5982	0.4002 / 27.8706	0.4489 / 50.74

Table A4
Expanded Uncertainty at 95% confidence level. Values are in GPa.

	ISO M4			ODR		
	H_{IT} SiO_2 / W	E_r SiO_2 / W	E_{IT} SiO_2 / W	H_{IT} SiO_2 / W	E_r SiO_2 / W	E_{IT} SiO_2 / W
A	2.295 (1.445) / 2.4	8.461 (4.23) / 54.82	9.516 (4.758) / 99.026	2.295 (1.615) / 2.92	7.756 (3.526) / 45.147	8.784 (4.026) / 76.387
B	18.525 / 5.95	20.124 (11.298) / 255.589	21.99 (12.461) / 392.35	23.75 / 11.9	18.076 (10.31) / 193.362	19.791 (11) / 330.4
C	5.985 (2.375) / 3.885 (1.47)	21.89 (7.767) / 95.07 (45.12)	24.189 (8.796) / 165.2 (78.47)	5.415 (2.09) / 3.5 (1.33)	19.771 (7.061) / 87.013 (41.9)	21.99 (8.063) / 152.81 (74.34)
D	47.5 / 40.25	211.83 / 886.24	241.89 / 2106	42.75 / 31.5	162.4 / 741.2	175.9 / 536.9
E	1.08 (0.72) / 1.085(0.77)	3.167 (1.056) / 109.722 (29.044)	3.65 (1.46) / 240.29 (111.86)	0.9 (0.54) / 1.022 (0.7)	2.815 (1.56) / 106.5 (19.36)	2.92 (1.46) / 194.72 (62.145)

References

[1] ISO 14577-1:2015 Metallic materials — Instrumented indentation test for hardness and material parameters - Part 1: Test method, (n.d.).
[2] D.A. Lucca, K. Herrmann, M.J. Klopfstein, Nanoindentation: Measuring methods and applications, CIRP Ann. - Manuf. Technol. 59 (2010) 803–819, https://doi.org/10.1016/j.cirp.2010.05.009.
[3] H. Engqvist, U. Wiklund, Mapping of mechanical properties of WC–Co using nanoindentation, Tribol. Lett. 8 (2000) 147–152, https://doi.org/10.1023/A:1019143419984.
[4] G. Maculotti, N. Senin, O. Oyelola, M. Galetto, A. Clare, R. Leach, Multi-sensor data fusion for the characterisation of laser cladded cermet coatings, Eur. Soc. Precis. Eng. Nanotechnology, Conf. Proc. - 19th Int. Conf. Exhib. EUSPEN 2019. (2019) 260–263.
[5] E.G. Herbert, W.C. Oliver, G.M. Pharr, Nanoindentation and the dynamic characterization of viscoelastic solids, J. Phys. d. Appl. Phys. 41 (2008) 74021, https://doi.org/10.1088/0022-3727/41/7/074021.
[6] E.S. Puchi-Cabrera, M.H. Staia, A. Iost, Modeling the composite hardness of multilayer coated systems, Thin Solid Films. 578 (2015) 53–62, https://doi.org/10.1016/j.tsf.2015.01.070.
[7] L. Settineri, R. Levi, Surface properties and performance of multilayer coated tools in turning inconel, CIRP Ann. - Manuf. Technol. 54 (2005) 515–518, https://doi.org/10.1016/S0007-8506(07)60158-9.
[8] G. Pharr, W.C. Oliver, R.F. Cook, P.D. Kirchner, M.C. Kroll, T.R. Dinger, D. R. Clarke, Electrical resistance of metallic contacts on silicon and germanium during indentation, J. Mater. Res. 7 (1992) 961–972, https://doi.org/10.1557/JMR.1992.0961.
[9] S. Ruffell, J.E. Bradby, N. Fujisawa, J.S. Williams, Identification of nanoindentation-induced phase changes in silicon by in situ electrical characterization, J. Appl. Phys. 101 (2007) 1–7, https://doi.org/10.1063/1.2724803.
[10] L. Settineri, M.G. Faga, G. Gautier, M. Perucca, Evaluation of wear resistance of AlSiTiN and AlSiCrN nanocomposite coatings for cutting tools, CIRP Ann. - Manuf. Technol. 57 (2008) 575–578, https://doi.org/10.1016/j.cirp.2008.03.103.
[11] M. Patel, A.K. Karamalidis, Germanium: A review of its US demand, uses, resources, chemistry, and separation technologies, Sep. Purif. Technol. 275 (2021), 118981, https://doi.org/10.1016/j.seppur.2021.118981.
[12] A. Zare, M. Tunesi, T.A. Harriman, J.R. Troutman, M.A. Davies, D.A. Lucca, Face Turning of Single Crystal (111)Ge: Cutting Mechanics and Surface/Subsurface Characteristics, J. Manuf. Sci. Eng. 145 (2023), https://doi.org/10.1115/1.4057054.
[13] M. Boi, G. Marchiori, M. Berni, A. Gambardella, F. Salamanna, A. Visani, M. Bianchi, M. Fini, G. Filardo, Nanoindentation: An advanced procedure to investigate osteochondral engineered tissues, J. Mech. Behav. Biomed. Mater. 96 (2019) 79–87, https://doi.org/10.1016/j.jmbbm.2019.04.042.
[14] G. Molino, A. Dalpozzi, G. Ciapetti, M. Lorusso, C. Novara, M. Cavallo, N. Baldini, F. Giorgis, S. Fiorilli, C. Vitale-Brovarone, Osteoporosis-related variations of trabecular bone properties of proximal human humeral heads at different scale lengths, J. Mech. Behav. Biomed. Mater. 100 (2019), 103373, https://doi.org/10.1016/j.jmbbm.2019.103373.
[15] T. Kwon, S. Gunasekaran, K. Eom, Atomic force microscopy-based cancer diagnosis by detecting cancer-specific biomolecules and cells, Biochim. Biophys. Acta - Rev. Cancer. 2019 (1871) 367–378, https://doi.org/10.1016/j.bbcan.2019.03.002.
[16] X. Long, R. Dong, Y. Su, C. Chang, Critical Review of Nanoindentation-Based Numerical Methods for Evaluating Elastoplastic Material Properties, Coatings. 13 (2023) 1334, https://doi.org/10.3390/coatings13081334.
[17] E.S. Puchi-Cabrera, E. Rossi, G. Sansonetti, M. Sebastiani, E. Bemporad, Machine learning aided nanoindentation: A review of the current state and future perspectives, Curr. Opin. Solid State Mater. Sci. 27 (2023), 101091, https://doi.org/10.1016/j.cossms.2023.101091.
[18] X. Long, X. Ding, J. Li, R. Dong, Y. Su, C. Chang, Indentation Reverse Algorithm of Mechanical Response for Elastoplastic Coatings Based on LSTM Deep Learning, Materials (basel). 16 (2023) 2617, https://doi.org/10.3390/ma16072617.
[19] G. Han, K.P. Marimuthu, H. Lee, Evaluation of thin film material properties using a deep nanoindentation and ANN, Mater. Des. 221 (2022), 111000, https://doi.org/10.1016/j.matdes.2022.111000.
[20] X. Long, C. Lu, Z. Shen, Y. Su, Identification of Mechanical Properties of Thin-Film Elastoplastic Materials by Machine Learning, Acta Mech. Solida Sin. 36 (2023) 13–21, https://doi.org/10.1007/s10338-022-00340-5.
[21] X. Long, Z. Shen, C. Lu, Q. Jia, C. Guan, C. Chen, H. Wang, Y. Li, Reverse Analysis of Surface Strain in Elasto-Plastic Materials by Nanoindentation, Int. J. Appl. Mech. 13 (2021), https://doi.org/10.1142/S1758825121501064.
[22] X. Long, S. Wang, Y. Feng, Y. Yao, L.M. Keer, Annealing effect on residual stress of Sn-3.0Ag-0.5Cu solder measured by nanoindentation and constitutive experiments, Mater. Sci. Eng. a. 696 (2017) 90–95, https://doi.org/10.1016/j.msea.2017.04.066.
[23] ISO 14253-1:2017 Geometrical product specifications (GPS) — Inspection by measurement of workpieces and measuring equipment, (n.d.).
[24] K. Herrmann, D.A. Lucca, M.J. Klopfstein, F. Menelao, CIRP sponsored international comparison on nanoindentation, Metrologia. 47 (2010), https://doi.org/10.1088/0026-1394/47/2/S06.
[25] W.C. Oliver, G.M. Pharr, Measurement of hardness and elastic modulus by instrumented indentation: Advances in understanding and refinements to methodology, J. Mater. Res. 19 (2004) 3–20, https://doi.org/10.1557/jmr.2004.19.1.3.
[26] G. Barbato, G. Genta, R. Cagliero, M. Galetto, M.J. Klopfstein, D.A. Lucca, R. Levi, Uncertainty evaluation of indentation modulus in the nano-range : Contact stiffness contribution, CIRP Ann. - Manuf. Technol. 66 (2017) 495–498, https://doi.org/10.1016/j.cirp.2017.04.060.
[27] J. Kholkhujaev, G. Maculotti, G. Genta, M. Galetto, Calibration of machine platform nonlinearity in Instrumented Indentation Test in the macro range, Precis. Eng. 81 (2023) 145–157, https://doi.org/10.1016/j.precisioneng.2023.02.005.
[28] W.C. Oliver, G.M. Pharr, An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments, J. Mater. Reseach. 7 (1992) 1564–1583.
[29] ISO 14577-2:2015 Metallic materials-Instrumented indentation test for hardness and materials parameters-Part 2: Verification and calibration of testing machines, (n.d.).
[30] K. Herrmann, K. Hasche, F. Pohlenz, R. Seemann, Characterisation of the geometry of indenters used for the micro- and nanoindentation method, Measurement. 29 (2001) 201–207, https://doi.org/10.1016/S0263-2241(00)00039-7.
[31] G. Maculotti, G. Genta, A. Carbonatto, M. Galetto, Uncertainty-based comparison of the effect of the area shape function on material characterisation in nanoindentation testing. Proc. 22nd Int. Conf. Exhib, EUSPEN, Genève, 2022.
[32] R. Cagliero, G. Barbato, G. Maizza, G. Genta, Measurement of elastic modulus by instrumented indentation in the macro-range: Uncertainty evaluation, Int. J. Mech. Sci. 101–102 (2015) 161–169, https://doi.org/10.1016/j.ijmecsci.2015.07.030.
[33] G. Maculotti, G. Genta, M. Lorusso, M. Pavese, D. Ugues, M. Galetto, Instrumented Indentation Test: Contact Stiffness Evaluation in the Nano-range, Nanomanufacturing Metrol. 2 (2019), https://doi.org/10.1007/s41871-018-0030-y.
[34] G. Genta, G. Maculotti, G. Barbato, R. Levi, M. Galetto, Effect of contact stiffness and machine calibration in nano-indentation testing, Procedia CIRP (2018), https://doi.org/10.1016/j.procir.2018.08.313.
[35] A. Germak, K. Herrmann, G. Dai, Z. Li, Development of calibration methods for hardness indenters, VDI Berichte. 13–26 (2006).
[36] J.L. Meneve, J.F. Smith, N.M. Jennett, S.R.J. Saunders, Surface mechanical property testing by depth sensing indentation, Appl. Surf. Sci. 100–101 (1996) 64–68, https://doi.org/10.1016/0169-4332(96)00258-9.
[37] G. Maculotti, J. Kholkhujaev, G. Genta, M. Galetto, Direct calibration of indenter tip geometry by optical surface topography measuring instruments, J. Mater. Res. (2023), https://doi.org/10.1557/s43578-023-01063-0.
[38] K. Herrmann, N.M. Jennet, W. Wegener, J. Meneve, K. Hasche, R. Seeman, Progress in determination of the area function of indenters used for nanoindentation, Thin Solid Films. 377–378 (2000) 394–400.
[39] M. Galetto, G. Genta, G. Maculotti, Single-step calibration method for nano indentation testing machines, CIRP Ann. 69 (2020) 429–432, https://doi.org/10.1016/j.cirp.2020.03.015.
[40] M. Galetto, G. Maculotti, G. Genta, G. Barbato, R. Levi, Instrumented Indentation Test in the Nano-range: Performances Comparison of Testing Machines Calibration Methods, Nanomanufacturing Metrol. (2019), https://doi.org/10.1007/s41871-019-00035-5.

[41] G. Maculotti, G. Genta, M. Galetto, Criticalities of iterative calibration procedures for indentation testing machines in the nano-range, in: Proc. 20th Int. Conf. Exhib, EUSPEN, Euspen, Genève, 2020, pp. 2–5.
[42] JCGM100:, Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM), JCGM. (2008) Sèvres, France. https://doi.org/10.1373/clinchem.2003.030528.
[43] A.C. Campbell, P. Grolich, R. Šlesinger, Niget: Nanoindentation general evaluation tool, SoftwareX. 9 (2019) 248–254, https://doi.org/10.1016/j.softx.2019.03.001.
[44] B. Efron, Bootstrap methods: another look at the jackkife, Ann. Stat. 7 (1979) 1–26.
[45] J. Muñoz-García, R. Pino-Mejías, J.M. Muñoz-Pichardo, M.D. Cubiles-De-La-Vega, Identification of outlier bootstrap samples, J. Appl. Stat. 24 (1997) 333–342, https://doi.org/10.1080/02664769723729.
[46] G. Barbato, E.M. Barini, G. Genta, R. Levi, Features and performance of some outlier detection methods, J. Appl. Stat. 38 (2011) 2133–2149, https://doi.org/10.1080/02664763.2010.545119.
[47] G. Maculotti, G. Genta, M. Lorusso, M. Pavese, D. Ugues, M. Galetto, Instrumented Indentation Test: Contact Stiffness Evaluation in the Nano-range, Nanomanufacturing Metrol. 2 (2018) 16–25, https://doi.org/10.1007/s41871-018-0030-y.