Indian Journal of Pure & Applied Physics
Vol. 64, March 2026, pp. 260-269
DOI: 10.56042/ijpap.v64i3.24618

Proposals for Ensuring the Validity of Force Measurement Measures

Seif M Osman* & K M Khaled

Force and Material Metrology laboratory, National Institute of Standards (NIS), Giza 122 11, Egypt

Received: 8th October 2025; accepted: 12th January 2026

Results validity is a cornerstone in structuring ISO/IEC 17025:2017. Results validity is the backbone for measurement quality. Validity of results may be ensured through different techniques, this article is directed to interested parties in force measurements to present the importance of ensuring the quality of results obtained from force proving instruments, force testing machines, and force standard machines to assure quality of life. Different proposals for ensuring the validity of results are given in detail. Each proposal is feathered by a measurement procedure, data analysis techniques, evaluation process, and acceptance criteria considering the associated risks. The proposed procedures are supportive methods in the force measurement field. They may be considered as an initiation to set a general guide in ensuring the quality and validity of force measurement results.

Keywords: Force results, Results validity, intermediate check, Intra-laboratory comparison

1 Introduction

Ensuring the validity of results is one of the general requirements determined as a measure of the competence of laboratories according to ISO/IEC 17025:2017; results can be validated externally and /or internally[1]. Laboratories shall plan and apply a procedure to monitor and record the validity of results. The resulting data shall be archived and analysed using statistical techniques (if applicable) to detect trends to avoid risks. International organisations and associations issue standards, guides and recommendations for calibration and measurement procedures in force application fields[2-7]. Calibration procedures define how calibration intervals may be defined but it didn't specify how to perform validation of the results over the calibration period or its frequency. Several approaches and methods may be used to ensure the validity of results; laboratories may combine one or more or even simplify one of the well-known calibration or measurement methods to suit the intended application.

Results of force measurements can be validated externally and/or internally. Force laboratories can ensure the validity of their results externally by comparing their results with other laboratories. External validation techniques may be either through laboratory comparisons or proficiency testing. National measurement laboratories and designated bodies have to participate in inter-laboratory comparisons organized under the BIPM umbrella to recognise their calibration measurement capabilities[8-10], other laboratories have to participate in proficiency testing schemes organized by a competent proficiency testing provider as per ISO/IEC 17043:2023[11].

Laboratories can ensure the validity of their measurement results internally by use of; reference/quality control materials, alternative instrumentation, functional check(s) of measuring equipment, control charts, intermediate checks, retained items, replicate tests or calibrations using the same or different methods, correlation of results for different characteristics of an item, and intra-laboratory comparisons[1].

Intermediate check is a simple form to ensure the validity of results and maintain confidence in measuring equipment. Intermediate checks shall be carried out according to a procedure[1]. The intermediate check procedure shall identify how the check is performed, check frequency, evaluation criteria and required actions. Check frequency may be planned according to the following: calibration interval, intensive use, manufacture recommendations, maintenance and data from previous calibration certificates (drift/accuracy).

2 Materials and Methods

2.1 Load Cells

The load cell is the most known electrical force transducer used as a force-proving instrument. The

---

*Corresponding Author: E-mail: seifelnasr_nis@yahoo.com

Table 1 — Acceptance criteria according to ISO 376:2011 and ASTM E74:2018

Application	Class according to		Criteria (%)		
	ISO 376:11	ASTM E74-2018	Repeatability	Reversibility	Accuracy
Reference standard	0	AA	0.05	0.07	0.032
	0.5	AA	0.1	0.15	0.032
Working standard	1	A	0.2	0.3	0.16
	2	A	0.4	0.5	0.16

load cell is used to disseminate force from the primary level down to the working standards. Load cells can be used at all levels, ranging from serving as a transfer standard in force standard machine comparisons to being a working standard for firms.

Long-term stability plays a crucial role in determining appropriate recalibration intervals for measuring instruments. It reflects the ability of an instrument to maintain its metrological characteristics within specified limits over extended periods of operation. Instruments demonstrating high long-term stability exhibit minimal drift, allowing measurement accuracy and uncertainty to remain acceptable for longer durations, which supports the extension of recalibration intervals. Conversely, poor long-term stability leads to increased drift and uncertainty, necessitating more frequent recalibration to mitigate the risk of erroneous measurements. The evaluation of long-term stability is commonly based on historical calibration data and trend analysis, and it is often integrated into risk-based calibration strategies in accordance with standards such as ISO/IEC 17025 and ISO 376:2011. ASTM E74:2018 defines force transducers calibration intervals not to exceed 24 months, taking into consideration stability criteria for class AA is less than or equal to 0.032 % and is 0.16 % for class A[2], while ISO 376:2011 states that the calibration interval shall not exceed 26 months after demonstration of stability, supporting the adopted recalibration interval[3].

Force laboratories have to set stability criteria for load cells in service. Stability criteria could be based on accuracy, reproducibility, reversibility (if applicable) or a combination of them. For reference load cells (load cells used to calibrate others), it is recommended to use the three criteria. ISO 376:2011 states criteria for classifying load cells based on reproducibility or reversibility, while ASTM E74: 2018 sets criteria based on accuracy for load cell classification. Table 1 shows ISO 376:2011 and ASTM E74:2018 acceptance criteria for load cells checked for accuracy, reproducibility or reversibility based on their applications. However, a manufacturer may set different criteria as a kind of competition.

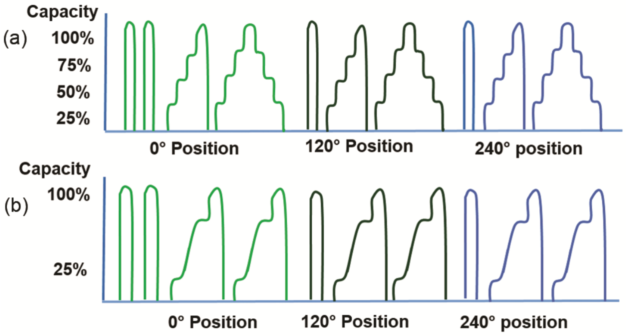

Fig. 1 — Loading scheme for load cell intermediate check

Load cell users can be categorized as users who have the capabilities to calibrate load cells and users who use load cells to perform measurements or calibrations for testing machines. Each user should perform intermediate checks; each user prefers to perform intermediate checks with his facilities.

Load cells can be checked using force standard machines (FSMs); FSM with suitable capacity to the load cell (under check) is selected, a minimum of two points (25 % and 100 % of the load cell nominal capacity) and a maximum of four points (25 % - 50 % - 75 % -100 % of the load cell nominal capacity) are determined to perform the check on them. FSM apply forces according to the loading scheme shown in Fig. 1. Figure 1 (a) shows the four-point loading scheme check for load cells used in increasing and decreasing loads. Figure 1 (b) shows the two-point loading scheme check for load cells used with increasing loads only.

The measured values may be compared analytically to the acceptance criteria mentioned in Table 1 or to other criteria set by the laboratory. In case of performing an accuracy check; the calculated average should be compared with that one on the previous calibration certificate. A control chart may be used to graphically check the load cell accuracy, where upper and lower limits are allocated based on the previous calibration certificate (upper / lower limit = mean value from latest calibration certificate ± expanded uncertainty from latest calibration certificate), and the responses of the checked load cell are plotted on the same graph (Measured Values), this graphical representation illustrates check data more clearly.

$$\overline{X}_r = \frac{\sum_{m=1}^{m=j} X_m}{j}$$

where; j = number of ascending loading series … (1)

$$b = \left| \frac{X_{\max} - X_{\min}}{\overline{X}_r} \right| \times 100 \quad \dots (2)$$

$$b' = \left| \frac{X_2 - X_1}{\overline{X}_{wr}} \right| \times 100 \quad \dots (3)$$

$$\overline{X}_{wr} = \frac{X_1 + X_2}{2} \quad \dots (4)$$

where; $\overline{X}_{wr}$ is the average value of the response without rotation.

$$v = \frac{v_a + v_b + v_c}{3} \quad \dots (5)$$

$$v_n = \left| \frac{X'_n - X_n}{X_n} \right| \times 100 \quad \dots (6)$$

where; n = a, b, c

Users who haven't FSM have to monitor their load cells performance, the widest spread load cells are based on the Wheatstone bridge circuit, which transfers the sensed force to electric signal[12]. The change in the arm ratio is directly proportional to the applied force. For a balanced circuit the resistance between points 1 and 3 must be equal to that between points 2 and 4 the change of this resistance over time may be an indication of load cell deterioration and load cell zero shift (Fig. 2).

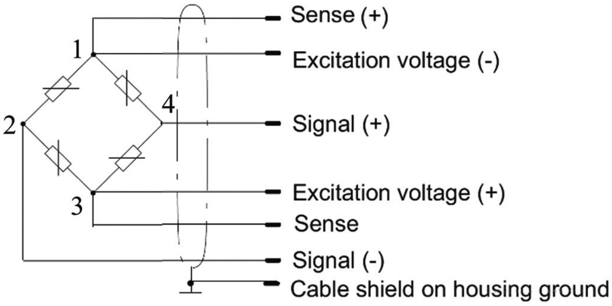

Fig. 2 — 6- Wires Wheatstone bridge circuit

As the load cell calibration certificate is issued; the resistance between the load cell terminals should be measured. A multimeter may be used to measure the resistance of the circuit point however it is a four or six-wire circuit, and measurements are performed at housing plug or wire terminals according to the load cell wiring. At check intervals; the resistance of the same points should be re-measured and compared to the original values. Resistance is measured to check strain gauge circumstances and change in resistance is used as an indicator for undesirable effects, as it is worthy known that change in load cell zero signal indicates overload and undesirable permanent deformation.

Table 2 — Testing machines acceptance criteria according to ISO 7500-1:2015

Class ISO 7500:1	Criteria (%)		
	Accuracy	Repeatability	Reversibility
0.5	0.5	0.5	0.75
1	1	1	1.5
2	2	2	3
3	3	3	4.5

2.2 Tension/Compression Testing Machines

Tension and compression testing machines are widely used by testing laboratories for material and final product testing. ISO 7500–1:2018 recommends verifying the force system at intervals not exceeding 12 months[4,13]. The time between two verifications depends on the type of testing machine, the standard of maintenance and workload. The force testing machine shall be verified if it is moved to a new location necessitating dismantling or if it is subject to major repairs or adjustments[4]. ISO 7500-1:2018 states criteria for classifying testing machines based on indication error (accuracy), repeatability and reversibility, while ASTM E4 - 2024 set 1 % as compliance criteria for accuracy and repeatability[5]. Table 2 shows acceptance criteria for force testing machines checked for accuracy, reproducibility or reversibility. The quality of results obtained from testing machines may be performed using force-proving instruments or by testing items with known values (reference material or retained items).

In the case of using load cells; five points are distributed between 20 % -100 % of the testing machine working range using one or more force-proving instruments with overlapping checks (Fig. 3). Figure 3 (a) shows a loading scheme for machines

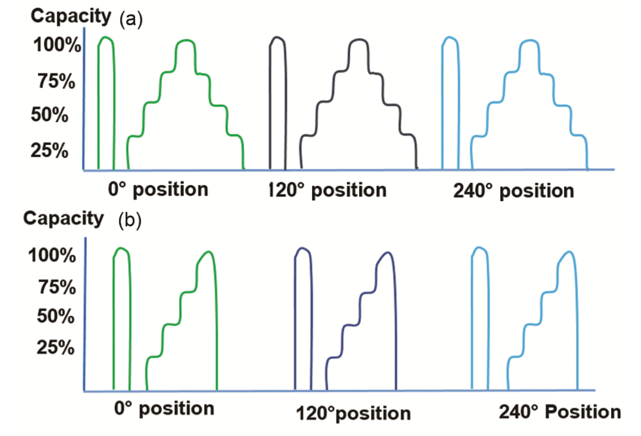

Fig. 3 — Loading scheme for testing machine intermediate check using load cell

used in increasing and decreasing loads, the scheme may be simplified as shown in Fig. 3 (b) for machines used for increasing loads only. Check results may be evaluated analytically or graphically to the acceptance criteria mentioned in Table 2 or others. For accuracy check; the calculated average should be compared with that one on the previous calibration certificate and uncertainty values on the previous calibration certificate may be used to set upper and lower acceptance limits.

The reference materials may be used routinely before daily tests or as a tool to ensure the validity of testing machine results. Reference materials are those materials produced by a producer complying with ISO 17034[14]. These materials are produced with certified certificates that comply with ISO Guide 31[15] after performing intensive measurements for homogeneity and stability for defining the nominal value and the associated uncertainty values for characterizing the produced patch. When test samples are used to assess the performance of force testing machines, the reference value is determined from measurements performed using a calibrated reference force-measuring system with traceability to the SI units. The reference value is typically taken as the mean of repeated force measurements obtained under controlled environmental conditions to minimize systematic and random effects. The associated uncertainty of the reference value is evaluated in accordance with the Guide to the Expression of Uncertainty in Measurement (GUM) and includes contributions from the calibration uncertainty of the reference system, repeatability of the measurements, resolution of the measuring devices, environmental influences, and any corrections applied to account for alignment, loading rate, or creep effects[16]. At least three pieces of reference material with certified value that comply with the testing machine capacity should be used to allow statistical analysis. Results are evaluated using the Zeta-score test (ζ)[17], Eq. (7) shows how (ζ) is calculated. If (ζ) ≤ 2.0, this indicates satisfactory results; if| $\zeta \geq 3.0$, this indicates unsatisfactory results and immediate action must be taken, starting with investigation and may end by re-check or recalibration. If $|\zeta|$ value is between 2.0 and 3.0, this indicates questionable results and the need to evaluate the initiated risk, which may lead to shortening the check period.

$$\zeta = \frac{x_i - X_{ref}}{\sqrt{u_i^2 + u_{ref}^2}} \qquad \ldots (7)$$

where;
x is the average of the three samples
u_{i} standard uncertainty of the average
X_{ref} is the certified reference force value of the reference material
u_{ref} standard uncertainty of the reference material

Using reference materials to check material testing machines is expensive if compared to retained items. Retained items are a set of items stored from a previously tested set. The interested laboratory prepares a set of identical samples and split them into groups with equal numbers of samples at least three samples per group.

One group should be tested on the testing machine after the machine is calibrated, the mean value of the tested group is determined and considered as the assigned value of the retained items, the rest groups are retained for frequent checks of the machine. Interested laboratories should determine the number of checks according to the nature and the workload on the machine; at least one check should be performed quarterly per year.

Interested laboratories may evaluate results using the Z-score test. Equation (8) shows how Z value is calculated[17]. If $Z \leq 2.0$, this indicates satisfactory results; if| $Z \geq 3.0$ this indicates unsatisfactory results and immediate action must be taken, starting with investigation and may end by re-check or recalibration if Z value is between 2.0 and 3.0, this indicates questionable results and the need to evaluate the initiated risk which may lead to shortening the check period.

Table 3 — Force standard machines

Force standard machine	Deadweight machines		Hydraulic amplification machines	Lever amplification machines	Comparator standard machines	
	Individual loading system	Subsequent loading system			Single transducer system	Individual loading system
Principle of operation	Generated force is calculated from the mathematical model of the force generation system $F = m_m g \left(1 - \frac{\rho_a}{\rho_m}\right)$ Where: ρm is the density of the object ρa is the air density where the object is suspended		Hydraulic system with a piston-cylinder with different effective areas amplifies deadweight force	Mechanical lever system amplifies deadweight force	These machines are based on force transducers, calibrated in a force standard machine. The generated force is calculated as the sum of the forces being measured by the individual transducers. Multiple transducers are loaded in parallel	
Typical CMCs	$5 \cdot 10^{-5} - 1 \cdot 10^{-4}$		$1 \cdot 10^{-4}$ to $5 \cdot 10^{-4}$	$1 \cdot 10^{-4}$ to $5 \cdot 10^{-4}$	$5 \cdot 10^{-4}$ to $5 \cdot 10^{-3}$	

$$Z = \frac{\bar{x}_i - X_{ref}}{\hat{\sigma}} \qquad \ldots (8)$$

where; $\bar{x}_i$ is the average of the three samples.

X_{ref} is the average of the pre-tested samples

$\hat{\sigma}$ is the standard deviation of the reference value

2.3 Force Standard Machines

Force Standard Machines (FSM) are machines used to calibrate the force-proving instruments, where known force values are applied to the force-proving instruments according to valid calibration procedure to determine the metrological characteristics of the force-proving instrument. Different force standard machines are available based on different principles[18]. Deadweight machines (DWM) present the head of the traceability pyramid in force measurements, as they generate forces using masses at a known gravitational field, hydraulic and lever amplification machines amplify the force generated by masses to generate accurate forces at a known gravitational force, while comparators standard machines use a reference force-proving instrument to determine the generated force[19]. Table 3 shows the characteristics of force standard machines.

Deadweight machines are the primary standard of force measurements. Deadweight machine consists of: a rigid frame, load load-generating system (Individually adjusted traceable masses), and a control system. Load load-generating system in deadweight machines depends on using weights of known masses, the mass values are adjusted so that, at a specific location, they generate particular forces[20].

Deadweight machines are compared to each other through interlaboratory comparisons, either bilateral or through Regional Metrology Organizations (RMOs). Organizing an interlaboratory comparison may take more than seven years to complete[21]. Force laboratories must implement a procedure over this period to ensure the quality of results. Bilateral comparisons and intra-laboratory comparisons may be used in cases where laboratories have more than one Deadweight machine. Laboratories having only one deadweight machine may ensure the quality of results via reference load cells by implementing one of the following two procedures.

Procedure (A) – Normalized error: this procedure utilizes at least two class (00) load cells (transfer standards according to ISO 376 classification), there must be at least 25 % overlapping measurements in the nominal range of the reference load cells, the overlapped range will permit repeatable measurement for one nominal force on the two load cells to facilitate evaluation and support decisions. When dealing with overlapping measurements in the assessment of force testing machines, a quantitative acceptance criterion is required to distinguish meaningful agreement from differences attributable to measurement uncertainty. A commonly applied criterion is based on comparing the absolute difference between measurement results (or between a measurement result and a reference value) with a fraction of the associated expanded uncertainty. The choice of 25 % represents a conservative margin that ensures a high level of confidence in the consistency of results while avoiding excessive sensitivity to normal measurement variability. This threshold is widely used in calibration, comparison, and proficiency testing practices because it provides a practical balance between discrimination power and robustness, allowing early detection of systematic effects without penalizing results that are in reasonable agreement within their stated uncertainties. The interested laboratory calibrates the reference load cells annually. The performance of the DWM is evaluated by measuring whether the results of the second calibration are within the uncertainty of the first calibration using (E_n) value as evaluation criteria. Results are evaluated using Eq. (9) which shows how the (E_n) is calculated. If $E_n \leq 1$ this indicates satisfactory results, and if $E_n > 1$ this indicates unsatisfactory results and different actions shall be taken as following,

$$E_n = \frac{\overline{X}_{r.1} - \overline{X}_{r.2}}{\sqrt{U^2_{\exp,1} + U^2_{\exp,2}}} \qquad \ldots (9)$$

where;

$\overline{X}_{r.1}$ average with rotation from the first calibration

$U_{\exp,1}$ Expanded uncertainty from the first calibration

$\overline{X}_{r.2}$ average with rotation from second calibration

$U_{\exp,2}$ Expanded uncertainty from the second calibration

In case of the E_n value was greater than 1 at the overlapping forces, the laboratory should halt using the DWM because evaluation indicates unsatisfactory results, which may be an indication of a major change in the DWM characteristics. The proper action is to perform an immediate investigation and re-evaluation, followed by comparison with trusted NMI.

In the case of one load cell indicating satisfactory results and some points on the other load cell indicate unsatisfactory results, this may be due to the change in the load cell metrological characteristics. However, measurements must be repeated on both load cells, and a third one may replace the unsatisfactory one to ensure the results before making decisions. If the E_n value for the third load cell is less than or equal 1, this means satisfactory results, and the replaced load cell must be recalibrated, if the E_n value is still higher than unity, this means more investigation is required on the DWM behaviour.

Procedure (B) – Accuracy in this procedure, the deadweight machine results are ensured through using one or more class (00) load cells according to ISO 376. The load cell is used to check the deadweight machine-generated forces where ten subsequent forces (not less than 50 % of the load cell nominal capacity with the same magnitude are applied to the load cell with the same loading conditions. Results are evaluated graphically by plotting the load cell response under the applied forces at each checkpoint. The upper and lower limits of this graph are the verified or claimed calibration measurement capabilities (CMC) of the deadweight machine. Acceptance criteria depend on the allocated measured response; If the checkpoints are between the upper and the lower limits, this ensures the DWM results. If the first and the second points are out of limits; this may be because of neglect to perform the preload process and the results of the machine are still valid, but if four points or more are out of limits, an action must be taken first by repeating the check using another load cell then proper action must be taken.

3 Experimental Results

Measurements were performed as per the pre-proposed procedures to investigate the metrological performance of load cells, tension testing machines, and force standard machines.

3.1 Load cells

The 500 kN load cell was checked using a 500 kN deadweight machine according to the loading scheme shown in Fig. 1 (b). Results for the 25 % and 100 % of the load cell capacity were plotted in Figs. 4 (a) and (b). The average of the measurements was plotted

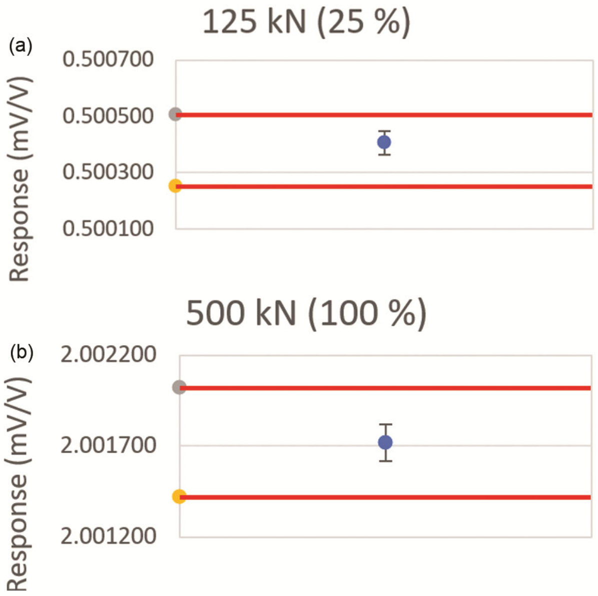

Fig. 4 — Graphical representation of load cell accuracy using uncertainty limits, (a) 25 % of the 500 kN load cell, and (b) 100 % of the 500 kN load cell

Table 4 — Results of resistance measurements

Load Cell		Resistance measurements (kΩ)			Maximum change (Δ)		
Capacity		Signal	Excitation	Sense	$\Delta_{1,2}$ %	$\Delta_{2,3}$ %	$\Delta_{3,4}$ %
100 kN	1st	1.374	0.960	0.960	0.208		
	2nd	1.371	0.962	0.962		0.104	
	3rd	1.372	0.961	0.961			0.00
	4th	1.372	0.961	0.961			
10 kN	1st	1.190	0.825	0.825	0.121		
	2nd	1.189	0.824	0825		0.121	
	3rd	1.189	0.824	0.824			0.084
	4th	1.189	0.824	0.824			

with uncertainty limits to be compared with the value in the latest calibration certificate with acceptance and refusal limits based on the uncertainty value in the latest calibration certificate. Results show no intersection between measured values and the limits which permit using the load cell under test until the next calibration If an intersection was observed, correction action should be taken according to the laboratory's internal procedure, considering associated risks.

The performance of two load cells with capacities of 10 kN and 100 kN was checked with force standard machines but with monitoring and measuring the resistances between terminals. Arm resistance (resistance between the load cell terminals) for the two load cells were monitored and recorded using a multimeter over a year on a three-month basis. The results of resistance change were as shown in Table 4.

Comparing the maximum change with the accuracy criteria stated in Table 1 (0.032 % for reference standard and 0.16 % for working standard) shows that the 100 kN and the 10 kN load cells under investigation are accepted if they are classified as working standard and rejected if they are classified as reference standard.

The resultant change in the strain gauge resistance over time may be an indication of load cell deterioration and load cell zero shift. A load cell zero signal change may be due to overload or mishandling of the load cell.

3.2 Force Testing Machines

A check for the 400 kN tensile testing machine was performed by load cells using the proposed procedure mentioned in part (2.2). Four testing loads distributed between 20 % - 100 % of the testing machine working range using a 500 kN load cell class 00. The load is applied according to the loading scheme shown in Fig. 2 (b). Figures 5 (a) and (b) show the results for measurements at 25 % and 100 % of the tensile testing machine capacity. The average of the measurements was plotted with uncertainty limits to be compared with the value in the latest calibration certificate with acceptance and refusal limits based on the uncertainty value in the latest calibration certificate. The results show that the average value of the measured point is between the acceptance limits; however, there is an intersection between the upper limit and the measured value if the uncertainty of measurements is added, this intersection indicates a warning sign for the operator, and he should act as per the laboratory/s internal procedure, considering associated risks.

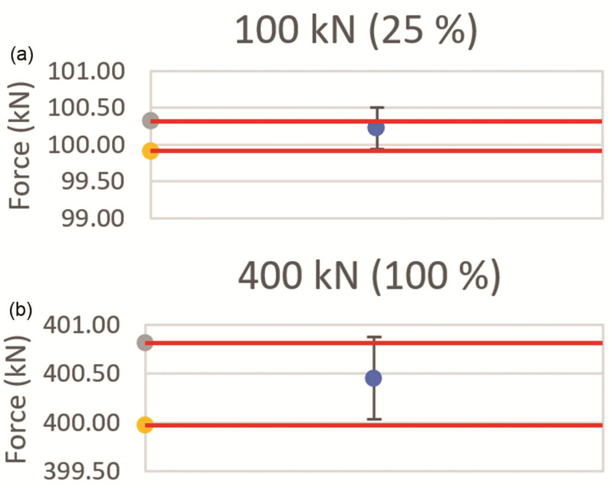

Fig. 5 — Graphical representation of tensile testing machine accuracy using uncertainty limits, (a) 25 % of the 400 kN tensile testing machine, and (b) 100 % of the 400 kN load cell

A 100 kN tensile testing machine was checked using two groups (group (A) and group (B)) of retained items. The retained items were stored from a previously tested set. The nominal values for the retained items are 14.87 kN and 84.43 kN for group

Table 5 — Results evaluated by Zeta-score test (ζ)

No.	Test measurements	Mean value (x_i)	Standard uncertainty of measurements ($u(x_i)$)	Assigned value (x_{pt})	Standard uncertainty of the assigned value ($u(x_{pt})$)	Zeta- score test (ζ)	Status
	kN	kN	kN	kN	kN		
1	13.2	13.98	0.48	14.87	0.34	1.51	Satisfactory
2	13.5						
3	13.9						
4	14.5						
5	14.8						

Table 6 — Results evaluated by Z-score test (Z)

No.	Test measurements	Mean value (x_i)	Standard uncertainty of measurements ($u(x_i)$)	Assigned value (x_{pt})	Standard uncertainty of the assigned value ($u(x_{pt})$)	Z-score test (Z)	Status
	kN	kN	kN	kN	kN		
1	83.5	84.78	0.48	84.43	0.34	0.49	Satisfactory
2	85.2						
3	83.9						
4	84.5						
5	86.8						

(A) and group (B), respectively. Five samples from each group were tested, results for group (A) were evaluated using the Zeta-score test (ζ) as shown in Table 5, results for group (B) were evaluated using the Z-score test (Z) as shown in Table 6.

Results show that the average value of the measured point is between the acceptance limits, however there is intersection between the upper limit and the measured value if the uncertainty of measurements are added, this intersection indicates warning sign for the operator and he should act as per the laboratory's internal procedure, considering associated risks.

When assessing a load cell using a deadweight force machine, risk levels can be specified by relating the expanded measurement uncertainty to the applicable tolerance, class of the load cell, or maximum permissible error (MPE) of the load cell. A common risk-based classification distinguishes between low, medium, and high risk. Low risk is associated with situations where the expanded uncertainty is small compared with the tolerance (typically ≤ 25 %), indicating high confidence that the conformity decision is reliable and the probability of false acceptance or rejection is negligible. Medium risk corresponds to uncertainty levels between approximately 25 % and 50 % of the tolerance, where conformity decisions remain acceptable but require increased attention to stability, environmental control, and monitoring of drift. High risk arises when the expanded uncertainty exceeds about 50 % of the tolerance, as the probability of incorrect conformity decisions becomes significant; in such cases, corrective actions such as reducing the uncertainty, tightening recalibration intervals, or improving the reference force realization are necessary. This uncertainty-based risk classification supports objective decision-making and aligns with metrological best practices and conformity assessment principles outlined in standards such as ISO/IEC 17025, ISO 7500-1, and ISO 376.

3.3 Force standard machines

A 100 kN class (00) load cell was used to check the status of a 100 kN dead weight machine, the load cell was calibrated using the same procedure for the latest calibration, calibration results were compared to the results in the latest calibration certificate using the normalized error (En). Table 7 shows the results for evaluation using (En), all En values were less than 1 which indicates satisfactory results which reflects the stability for both the dead weight machine and the load cell. If the results were unsatisfactory this reflects a change in the behaviour of the interaction between the deadweight machine and load cell which requires more investigation to define the causes.

Table 7 — Results evaluated by (E*n*)

No.	Calibration step	Mean value (x_1)	Standard uncertainty ($u(x_1)$)	Reference value (x_{pt})	Expanded uncertainty of the reference value ($U(x_{pt})$)	|*En*-value|	Status
	kN	mV/V	mV/V	mV/V	mV/V		
1	10	0.198365	0.000040	0.198395	0.000046	0.5	Satisfactory
2	20	0.398736	0.000080	0.398786	0.000092	0.4	Satisfactory
3	30	0.597191	0.000119	0.597211	0.000137	0.1	Satisfactory
4	40	0.795566	0.000159	0.795596	0.000183	0.1	Satisfactory
5	50	0.993931	0.000199	0.993989	0.000229	0.2	Satisfactory
6	60	1.192396	0.000238	1.192436	0.000274	0.1	Satisfactory
7	70	1.391561	0.000278	1.391597	0.000320	0.1	Satisfactory
8	80	1.587626	0.000318	1.587716	0.000365	0.2	Satisfactory
9	90	1.789791	0.000358	1.789821	0.000412	0.1	Satisfactory
10	100	1.988656	0.000398	1.988708	0.000457	0.1	Satisfactory

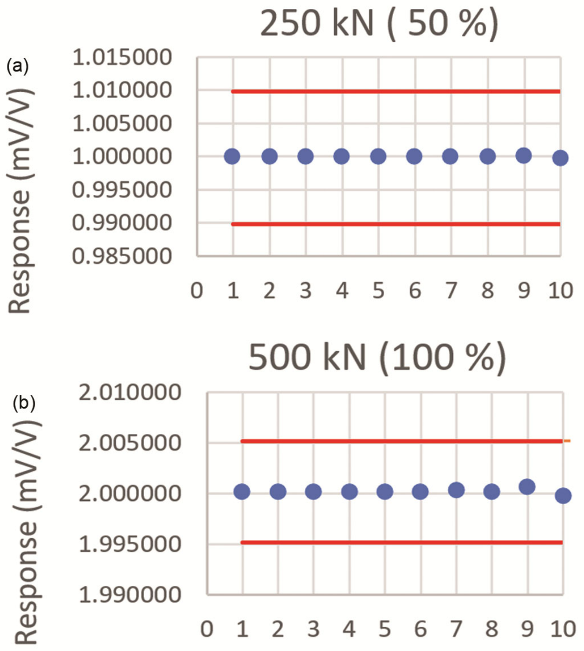

Fig. 6 — Graphical representation 500 kN force standard machine checks (a) 50 % of the 500 kN force standard machine, and (b) 100 % of the 500 kN force standard machine

A 500 kN class (00) load cell was used to check the status of a 500 kN dead weight machine. A 250 kN and 500 kN loads are applied on the load cell, each load is applied for ten times subsequently under the same loading conditions. Figures 6 (a) and (b) show the graphical representation for the measurements results at 50 % and 100 % of the force standard machine capacity; the upper and the lower limits of the graphs are the average of the results plus /minus the claimed calibration measurement capabilities (CMC) of the deadweight machine (0.001 %). The checkpoints are between the upper and the lower limits and this ensures the DWM results. But if the first and the second points are out of limits; this may be because of neglecting performing preload process and the results of the machine are still valid if other eight points are within the limits, but if four points or more are out of limits an action must be taken first by repeating the check using another load cell then proper action must be taken.

4 Conclusion

Ensuring the validity of measurement results are vital in force measurements and main requirement as per ISO/IEC 17025:2017 which can be ensured internally or externally. Different internal proposals were used to check the performance in force measurements on the three main levels; primary, secondary and working levels. Each proposal defines the measurement procedure, data analysis techniques, evaluation process, and acceptance criteria taking into considerations the associated risks. The proposed validation procedures use different analysis techniques as normalized error E_n, Z-score test and the Zeta-score test (ζ) in addition to graphical data representation with control limits. The proposed techniques may be considered as a start to set a general guide in ensuring the quality and validity of force measurements results.

References

1 ISO / IEC 17025 – 2017 “General requirements for the competence of testing and calibration”.
2 ASTM E74-18; “Standard Practices for Calibration and Verification for Force-Measuring Instruments "https://www.astm.org/Standards/E74.htm.
3 ISO 376:2011; “Metallic materials-Calibration of force proving instruments used for the verification of uniaxial testing machines”.

4 ISO 7500-1:2018; Metallic materials-Calibration and Verification of static uniaxial testing machines- Part 1: Tension/compression testing machines – Calibration and verification of the force-measuring system" https://www.iso.org/standard/72572.html.

5 ASTM E4-24; "Standard Practices for Force Verification of Testing Machines" https://www.astm.org/search/fullsite-search.html?query=E4&.

6 OIML R65:2006 ; "Force measuring system of uniaxial material testing machines" https://www.iso.org/standard/44661.html.

7 EURAMET Cg-4; "Uncertainty of Force Measurements" Calibration guide, Version 2.0 (03/2011) https://www.euramet.org/Media/docs/Publications/calguides/EURAMET_cg 4__v_2.0_Uncertainty_ of_Force_ Measurements.pdf.

8 CIPM MRA-D-05, Measurement comparisons in the CIPM MRA, Version 1.6, March 2016, WWW.BIPM.org.

9 Yon-Kyu Park, *et al.*, *Metrologia*, 56 (2019) 07003.

10 Cox M G, *Metrologia*, 39 (2002) 589.

11 ISO/IEC 17043: 2023, Conformity assessment—General requirements for proficiency testing providers.

12 Ştefănescu D M, "Handbook of Force Transducers – Principles and Components," Springer-Verlag Berlin and Heidelberg Gmb H & Co K, February 2011, ISBN 978-3-642-18295-2.

13 Osman S M & Khaled K M, *Measurement*, 146 (2019) 643.

14 ISO 17034:2016; "General requirements for the competence of reference material producers" https://www.iso.org/obp/ui/#iso:std:29357:en.

15 ISO guide 31: 2015;""Reference materials, Contents of certificates, labels and accompanying documentation" https://www.iso.org/obp/ui/#iso:std:iso:guide:31:ed-3:v1:en.

16 Khaled K M & Mahmoud G M, *Measurement*, 114 (2018) 436.

17 ISO 13528-2015, Statistical methods for use in proficiency testing by interlaboratory comparison.

18 Wu G, Shao J, Xu Y, He Z & Liu S, *Micromachine*, 16 (9) (2025) 1069.

19 Osman S M, Hasan E H & El-Hakeem H M, *MAPAN*, 25 (2010) 245.

20 E H Hasan, *et al.*, *J Braz Soc Mech Sci Eng*, 38 (2026) 955.

21 Kazunaga Ueda, *et al.*, *Metrologia*, 49 (2012) 07008.