

# Management of reference measurement standards as part of Smart Metrology: Gauge Blocks

Jean-Michel POU<sup>1\*</sup>, Laurent LEBLOND<sup>2</sup>, Christophe DUBOIS<sup>1</sup>

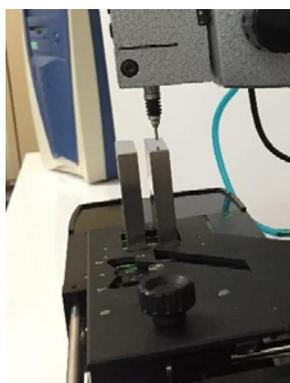
<sup>1</sup> Deltamu, 48 Rue de Sarliève, 63800 Cournon d'Auvergne

<sup>2</sup> PSA Group, C.T. de Vélizy A, 2, route de Gisy, 78943 Vélizy-Villacoublay Cedex, [laurent.leblond1@mpsa.com](mailto:laurent.leblond1@mpsa.com),

**Abstract.** The question of traceability of measurement standards used for calibration is a crucial one. Optimized or not, calendar periodicity is one of the accepted pillars of quality in measurement standard verification. Aware that this practice by itself cannot guarantee the metrological performance of measurement standards, laboratories make a point of also resorting to monitoring. In fact, calibrating independent and diverse instrumentation should of itself provide the means to detect problems with the reference gauges. Indeed, there is no logical reason why a collection of independent measurement instrumentation should, on average, have a common bias. If such a bias were observed, it would most likely point to a bias in the reference gauge used for the calibrations. In this paper, this argument is investigated with reference to a practical example of gauge block calibration.

## 1. Introduction

Calibration is defined as the comparison of a reference gauge to an entity to be calibrated, i.e. a measurement standard of the same type or measuring instrument<sup>a</sup>. The objective is to determine the deviation between the measurement standard and the entity to be calibrated, a quantity that then gives the value of the entity.



**Figure1: Calibration of a gauge block**

Because this deviation is the result of a measurement (carried out in strictly controlled conditions called 'calibration conditions'), it is affected by an uncertainty, termed calibration uncertainty. This uncertainty includes all the imperfections of the various factors affecting the calibration process (repeatability, environment, etc.) but

<sup>a</sup> The new definition of the term 'calibration' by VIM3 goes further than simply revealing the existence of a deviation. Figure shows the

also, obviously, the uncertainty of the measurement standard itself. The calibration uncertainty also includes uncertainty as to a possible change in the value of the measurement standard, in other words a drift with time.

The reference gauges of laboratories carrying out calibration for customers must also be calibrated and these calibrations are crucial. They on the one hand guarantee traceability of the results produced by the laboratory to the International System of Units, and on the other hand determine the uncertainty associated with these results. Accreditation bodies are particularly strict about this question of traceability to the reference standard. If changes in a reference gauge, when monotonic, can be estimated by analysing the periodic calibration results produced by it, the effects of accidents, such as mechanical shock, are more difficult to predict. So even if laboratories carry out intermediate monitoring, it will not necessarily detect behaviour due to accidents.

In this paper we present a dynamic way of monitoring the state of reference gauges, and even of revealing their value without having to compare them with superior reference gauges. Our demonstration explores a practical example of the calibration of gauge blocks.

calibration of a gauge block by comparison with a reference gauge block of the same nominal value.

\* Corresponding author: [jmpou@deltamu.com](mailto:jmpou@deltamu.com)

## 2 Principle of gauge block calibration via mechanical comparison

In industrial laboratories, calibration of gauge blocks is carried out by mechanical comparison, at a specialized calibration test bench (cf Figure). The method uses two inductive probes mounted opposite each other<sup>b</sup> to measure the deviation between the laboratory reference gauge block and the gauge being calibrated. The laboratory reference gauge of an accredited laboratory is calibrated by direct interferometry in a higher level laboratory (in France by the National Calibration Laboratory (*LNE: Laboratoire national de métrologie et d'essais*)). This method makes it possible to obtain very low uncertainties, of just a few tens of nanometres, whereas standard laboratory calibration gives uncertainties of a few hundred nanometres.

For the measurement, the test bench is set to zero relative to the reference gauge. Then the pair of probes measures the deviation between the length of the reference gauge and the length of the gauge to be calibrated. The comparison is direct, in other words of gauges of the same nominal value. For example, a gauge to be calibrated with a nominal value of 10 mm is compared with a reference gauge also with a nominal value of 10 mm. This is the case for all gauges, which means that the laboratory must have all possible reference gauges<sup>c</sup>.

The international standard for gauge blocks (NF EN ISO 3650 [1]) states that deviation should be measured at the centre of the gauges and that variation in length is determined by measuring the gauges at five points (Figure) and taking the difference between the two extremes of these five measurements.

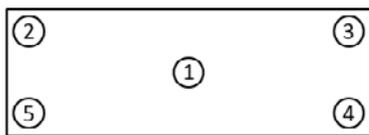


Figure2: Measurement points of a gauge block (Top view)

Variation in length is a characteristic of geometry, and thus relative, and the length of the reference gauge does not therefore enter into the calculation. On the other hand, for the length of the gauge being calibrated (defined in general as its value at the centre, i.e. at point ① of **Erreur! Source du renvoi introuvable.** and written  $Lg_{Point\textcircled{1}}$ ) the value of the reference gauge (written  $Lg_{Reference\ gauge}$ ) is indispensable, and is shown in the following relation:

$$Lg_{Point\textcircled{1}} = Lg_{Reference\ gauge} + Deviation \quad (1)$$

<sup>b</sup> Figure shows the upper probe. The opposing probe is under the gauge block and therefore not visible in the photo. The measuring force of the upper probe is slightly greater than that of the lower probe, which

## 3 Modeling the deviation

The measured deviation (written *Deviation* in (1)) is composed of different uncertainty factors, all random variables:

$$deviation = deviation_{actual} + e_{setting_0} + e_{local\ bias} + e_{\Delta temperature} + e_{repeatability} + e_{bench\ geometry} + e_{reference\ drift} \quad (2)$$

With:

$deviation_{actual}$ : actual deviation of length between reference gauge and gauge being calibrated. The  $deviation_{actual}$  will never be known perfectly, if only because the reference gauge itself cannot be perfectly known. It is this deviation that gives rise to the calibration uncertainty that needs to be assessed for any given gauge.

$e_{setting_0}$ : error due to repeatability when setting test bench to zero relative to reference gauge

$e_{local\ bias}$ : bias of measurement system composed of two opposing probes. It should be remembered that this local error is of the order of a few nanometres as the gauges are of equal nominal length (in general).

$e_{\Delta temperature}$ : error due to temperature at time of measurement (not automatically equal to 20°C). Prior to measurement, the gauge blocks are stabilized at a virtually identical temperature (by being placed next to each other for as long as necessary, typically several hours). Moreover, the gauges are usually made of the same material (stainless steel), which allows natural correction of a large part of the deviation between the temperature at time of measurement and the 20°C reference temperature, since both gauges experience the same dilatation. Occasionally, the gauges are made of ceramic or tungsten carbide. To limit effects due to component material, only results obtained for gauges of the same material as the reference gauge are used in this study.

$e_{repeatability}$ : error due to repeatability at time of measurement of gauge being calibrated.

$e_{test\ bench\ geometry}$ : test bench geometry error, notably any parallelism error between the position on the test bench of the gauges relative to the measurement probes. If this geometry problem does exist, it only effects the measured deviation, since both gauges are on the same plate.

$e_{reference\ drift}$ : actual deviation between the reference gauge value at time of calibration and the value at time of use.

ensures that the lower surface of the gauge block is firmly seated during the measurement.

<sup>c</sup> Certain gauge test benches have a measuring probe that can reduce the number of reference gauge blocks needed.

### 3 Utilization of laboratory measurements

Since the creation of the International System of Units, calibration has been the means of ensuring that a measuring instrument is related to a specific measurement unit. This relationship is established via a succession of calibrations that can be traced back to the unit as defined by the General Conference on Weights and Measures (CGPM).

When the measuring device is new, calibration is the means to ensure that the manufacturer has done a good job. For example, to produce 10 mm gauge blocks, the manufacturer must be sure that the manufacturing procedure is managed in such a way as to obtain a length of 'almost' 10 mm, the 'almost' being as low a figure as possible.

Once the measuring device is in service, calibration is the means to evaluate any drift, remembering, of course, that this drift could be either positive or negative. Gauge blocks can wear (therefore decreasing in length, even if experience shows that this is rare), but they can also increase in size as a result of crystallographic effects, such as retained austenite.

Under the hypothesis that the manufacturer of the device has endeavoured to manufacture it as perfectly as possible and that instrumentation drift is not in one direction only, it is reasonable to suppose that the average of a large collection of gauges of a nominal value of 10 mm is indeed equal to 10 mm. It is just as reasonable to suppose that the average error produced by a large number of measuring instruments will be zero. The 'average' gauge and the 'average' measurement device obviously do not physically exist, but the laboratory has access to the measured values for all the entities it has calibrated of the same type (all the 10 mm gauges of its customers) and can therefore calculate the average. Under the preceding assumption, if the average is not equal to 10 mm (i.e. a deviation of 0), the deviation observed can only be logically explained by the deviation of the reference gauge itself. As all the manufacturers are independent of each other, there is no physical reason for a bias...

The question to be asked, therefore, is this: is it possible to establish the value of the reference gauge by utilizing the large numbers of values measured in the laboratory?

### 4 Deviation uncertainty and comparison of results

The errors  $e_{rsetting0}$ ,  $e_{repeatability}$ ,  $e_{\Delta temperature}$  of the model (2) for measurement deviation are hypothetically random errors with a theoretical average of zero.

Whether local bias (i.e.  $e_{local\ bias}$ ), which represents the measurement system error over a low measurement range, is of the systematic type (defect in the linearity of the system) or purely random, it has a theoretical average of zero since any bias (constant error) is corrected by the setting up of the reference gauge.

The test bench geometry error (i.e.  $e_{test\ bench\ geometry}$ ) is considered negligible since the deviation of length between the two gauges is so low that the geometric effect would have no impact. Moreover, the type of test bench in question has been specifically engineered so as not to exhibit this type of defect.

The drift error of the reference (i.e.  $e_{Reference\ Drift}$ ) is a figure that can be obtained by analysing the calibration results obtained for each reference gauge during the respective calibrations. A drift error can be modeled by a uniform distribution of total amplitude equal to the variation in length of the reference gauge over a given period.

The terms that constitute the deviation being independent, the theoretical variance of the deviation is equal to the sum of the theoretical variances of the errors and of the theoretical variance of the actual deviation. The latter is probably much greater than the variance of all the other terms that derive from the measurement itself. Therefore, under this hypothesis, the theoretical average of the deviation, estimated via the arithmetic average of the deviation values obtained for the gauges calibrated, should be equal to the actual deviation from the true value of the reference gauge, the latter being assumed to be stable.

Of course, as the setting up is done relative to the reference gauge, if this is in reality lower than its nominal value, the deviations measured will on average be positive. Conversely, if it is in reality higher than its nominal value, the deviations will on average be negative.

Consequently, if we write  $x_i$  as the value for the deviation measured at  $i^{\text{th}}$  calibration, we get the following estimates:

$$\widehat{deviaton}_{actual} = -\frac{1}{n} \sum_{i=1}^n x_i$$

$$\hat{\sigma}_{deviation_{actual}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \frac{1}{n} \sum_{i=1}^n x_i)^2}$$

The confidence interval for confidence level  $1 - \alpha$  of  $deviation_{actual}$  is thus given when  $n$  is high under normality distribution, by:

$$\left[ \widehat{deviation}_{actual} - t_{\frac{\alpha}{2}} \frac{\hat{\sigma}_{deviation_{actual}}}{\sqrt{n}}; \widehat{deviation}_{actual} + t_{\frac{\alpha}{2}} \frac{\hat{\sigma}_{deviation_{actual}}}{\sqrt{n}} \right]$$

With  $t_{\frac{\alpha}{2}}$  the fractile of a Student law at  $n - 1$  degrees of freedom.

As we have the deviation from nominal for each reference gauge block and the associated uncertainties, we can, via a Student's t test of comparison of average, check our hypothesis. We shall consider that the averages are statistically identical if

$$\left| \frac{\widehat{deviation}_{actual} - deviation_{Nominal\ ref\ gauge}}{\sqrt{\frac{(\hat{\sigma}_{deviation_{actual}})^2}{n} + u_{calibration}^2}} \right| \leq 2$$

With:

$deviation_{Nominal\ ref\ gauge}$  : The deviation measured by the LNE for the reference gauge  $V_{measured\ Ref\ Gauge} - V_{nominal\ Ref\ Gauge}$ .

$u_{calibration}$ : Standard uncertainty of calibration, given by the LNE, for reference gauges.

### 5 Application and discussion

Our investigation involved gauges to be calibrated of different nominal values, namely 1, 5, 10, 50 and 100 mm. The data at our disposal was obtained over periods of almost a year. The calibrations were carried out by comparison with different reference gauges (two different reference gauge blocks per nominal value). The reference gauges were calibrated twice during the periods in question for the first reference, and once for the second.

Laboratory calibrated gauges returning a significant deviation from the reference were excluded graphically from the scope of observation in order to limit the investigation to a homogenous group of gauges.

The example shown in figure 3 is the histogram of values for the 5 mm gauge (reference 1). The abscissa axis has  $1/10 \mu m$  graduations. On this scale, therefore, 0.1 represents 10 nanometres. The histogram shows frequency of deviations found per class.

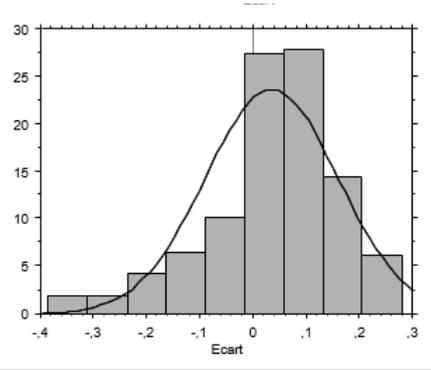


Figure 3: 5 mm gauges

For this example, we had at our disposal 120 measured deviations. We excluded 12 values. The histogram is therefore constructed from 108 values.

The LNE measured the reference gauge. They reported a value of 4.999987 mm, i.e. a deviation from nominal of -13 nanometres. The standard uncertainty for this value is 10 nanometres.

We calculated, using the 108 values retained, a measured average deviation of -35 nanometres and a standard deviation of 125 nanometres, i.e. average standard uncertainty of  $\frac{125}{\sqrt{108}} \approx 12$  nanometres

Relation (2) allows us to calculate the test statistic:

$$\left| \frac{-13 - (-35)}{\sqrt{10^2 + 12^2}} \right| \approx 1.41$$

Our conclusion was that there is statistically no significant difference between the two deviations.

The same process was followed for each nominal value.

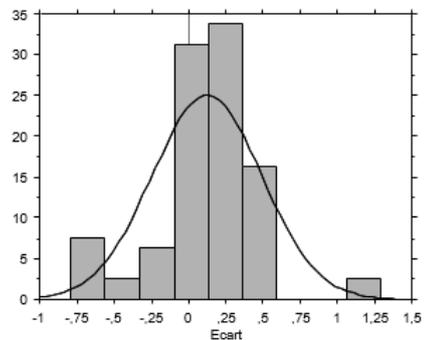


Figure 4: 100 mm gauge

For example, for the 100 mm gauge (reference 2), we had at our disposal 40 measured deviations. No values were excluded and the histogram (cf. figure 4) is therefore constructed from these 40 values. The measured average deviation is -125 nanometres and the standard deviation is

368 nanometres, hence an average standard uncertainty of  $\frac{368}{\sqrt{40}} \approx 58$  nanometres.

The LNE measured the reference gauge. They reported a value of 99.999917 mm, i.e. a deviation from nominal of -83 nanometres. The standard uncertainty for this value is 18 nanometres.

As in the previous case, relation (2) allowed us to calculate the test statistic, and led to the same conclusion:

$$\left| \frac{-83 - (-125)}{\sqrt{18^2 + 58^2}} \right| \approx 0.70$$

A summary of results is shown in the table in figure 5.

Boite étalons	Etalonnage	Nombre de cales clients retenus	Nominal	Ecart_LNE	uLNE	Ecart_moyen	Ecart_type Moyenne	Ecart (Nanomètre)	Comparaison de moyennes
Boite 1	Etalonnage 1	116	1	-0,047	0,010	-0,078	0,012	31	1,98
Boite 1	Etalonnage 2	36	1	-0,039	0,010	-0,084	0,018	45	2,18
Boite 1	Etalonnage 1	112	5	-0,021	0,010	-0,011	0,011	-10	0,66
Boite 1	Etalonnage 2	37	5	-0,005	0,010	-0,009	0,022	4	0,17
Boite 1	Etalonnage 1	110	10	0,019	0,011	0,033	0,012	-14	0,85
Boite 1	Etalonnage 2	36	10	0,025	0,011	0,035	0,023	-10	0,40
Boite 1	Etalonnage 1	121	50	0,054	0,014	0,313	0,033	-259	7,18
Boite 1	Etalonnage 2	40	50	0,067	0,014	0,045	0,071	22	0,30
Boite 1	Etalonnage 1	120	100	-0,086	0,018	0,559	0,035	-645	16,60
Boite 1	Etalonnage 2	40	100	-0,083	0,018	-0,125	0,058	42	0,69
Boite 2	Etalonnage 1	111	1	0,044	0,010	0,020	0,012	24	1,53
Boite 2	Etalonnage 1	108	5	-0,013	0,010	-0,035	0,012	22	1,39
Boite 2	Etalonnage 1	104	10	0,043	0,011	0,084	0,014	-41	2,37
Boite 2	Etalonnage 1	116	50	-0,063	0,014	-0,005	0,027	-58	1,91
Boite 2	Etalonnage 1	119	100	0,347	0,018	0,536	0,038	-189	4,50

Figure 5: Summary Table

This table summarizes 15 series of measurements of gauges of different lengths, from 1 mm to 100 mm. In twelve cases, the test of equality of averages gives a positive result, in other words, the average of the deviations measured is not statistically different from the value of the reference gauge block given by the LNE. For these 12 values, the average deviation between the average and the reference value is 5 nanometres.

It is likely that the laboratory would improve its measurement quality by knowing how to use this data on a daily basis, other than simply putting it on a calibration certificate. An alert for an incongruous piece of data with regard to other data of the same type would very likely enable the confirmation (or rejection) of a measurement.

The three cases where our test detected a significant difference raise the question of why the data from the laboratory is not consistent with that of the LNE for certain gauges, while for others it is.

For the 50 mm gauge of box 1, a significant difference in the comparison of values is apparent in the case of calibration 1, but not in the case of calibration 2. Now the value of the reference gauge has only changed by 13 nanometres (i.e. its calibration uncertainty, which does not suggest a drift of this gauge), while the average deviation is 268 nanometres!

The behaviour of the measured average deviation during the period when the data was collected is shown in Figure 6.

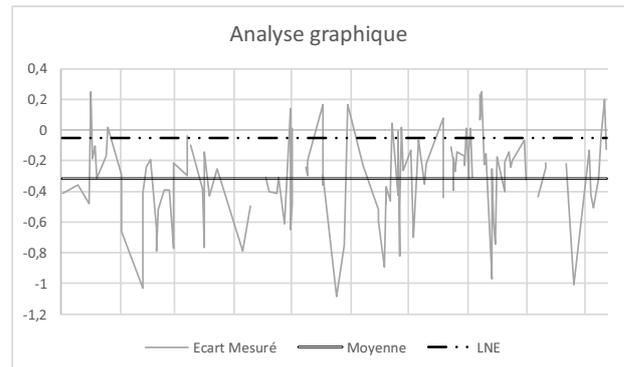


Figure 6: Behaviour of 50 mm gauge (Calibration 1)

This behaviour exhibits constancy of deviation over the period in question. For calibration 2 (Figure 7), however (fewer values), the dispersion of the measured deviations is evidently centred around an average close to the LNE value.

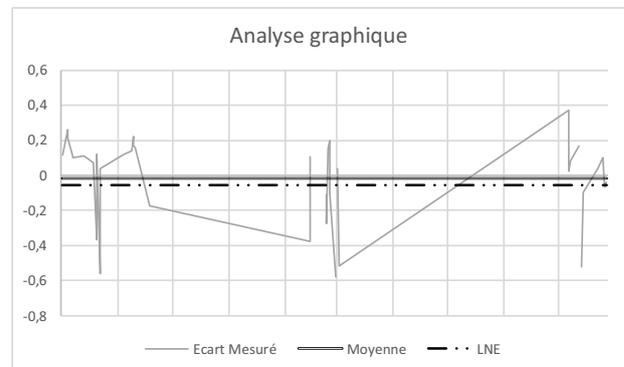
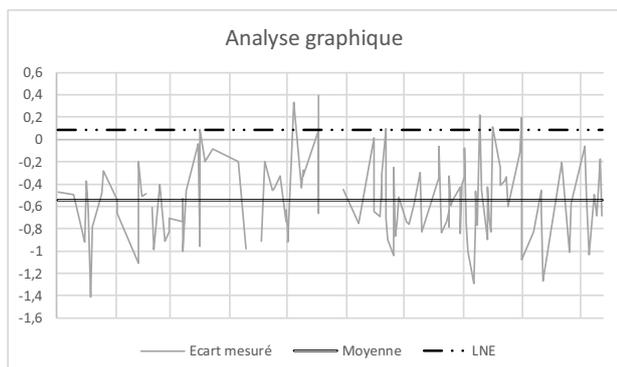


Figure 7: Behaviour of 50 mm gauge (Calibration 2)

The causes of this phenomenon are difficult to determine posteriorly. The reason cannot be measurement system bias (because this would be corrected at each setting up of the reference gauge), nor a temperature problem, nor a geometry problem (this is explained above, in paragraph 4). Bearing in mind the significant size of the deviation, it is quite likely that the cause would have been found if it had been investigated at the time. For instance (and this is something that could feasibly happen), the laboratory may have inadvertently swapped its own reference gauge with a customer gauge, only realizing this when the time came to send the laboratory gauges to the LNE for calibration. This is obviously just a hypothesis.

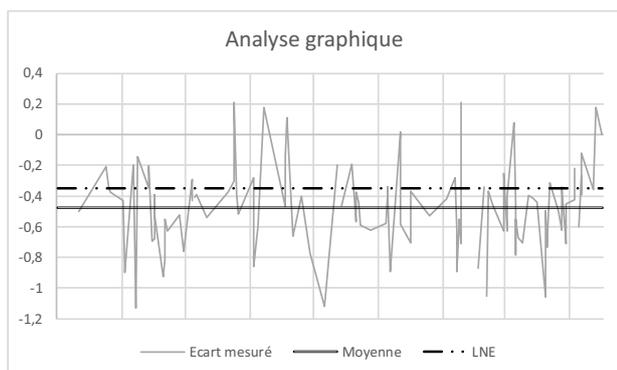
For the 100 mm gauge, a similar phenomenon can be seen. Now the value of this reference gauge has only changed by 3 nanometres (i.e. its calibration uncertainty, which does not suggest a drift of the gauge) and the average deviation is 684 nanometres!

The behaviour of the measured average deviation during the period when the data was collected is shown in Figure 8 and the conclusions are the same as those drawn previously.



**Figure 8: Behaviour of 100 mm gauge (Calibration 1)**

For the 50 mm gauge of box 2, the deviation of 189 nanometres is smaller than that found in the two preceding cases, but the comparison of averages clearly highlights that the values cannot be identical. Because we only have data from a certain period of the reference gauge calibration, we cannot know, a posteriori, the reasons for the difference. However, we can see that the change in measured deviation is, once again, somewhat constant over the period (*cf.* Figure 9).



**Figure 9: Behaviour of 100 mm gauge (box 2)**

## 6 Conclusion

It is evident that the hypothesis that the average of values measured in a laboratory (since they come from calibrated instrumentation) should reflect the 'true' value of the reference gauge is not unequivocally demonstrated by the data used in our investigation. However, it is worth noting that the hypothesis is convincing in 80% of the cases investigated. As the data used is not recent, we have not been able to do more than conjecture as to the reasons for the significant deviations encountered in the remaining 20% of cases: however, the results do clearly demonstrate the advantage of real time monitoring of calibration results in laboratories.

Our intention is to encourage those with the job of daily calibration to reflect on the broader possibilities open to them. By analysing the results of their calibrations, it would seem to be possible to detect anomalies in the reference apparatus almost immediately. IT today offers almost effort-free opportunities not only to improve measurement quality (*Why does the laboratory have a 268 nanometre deviation on its 50 mm customer gauges for the best part of a year, only for the phenomenon to mysteriously disappear?*) but also to reduce internal costs. If a laboratory can demonstrate the ability to detect an anomaly in its own measurement standards the moment one appears it will be able to lengthen its periodicities without risk, and thus take a significant step towards the optimization of costs that is so crucial in an increasingly competitive market.

## References

1. AFNOR, NF EN ISO 3650: Geometrical Product Specifications (GPS) - Length measurement standards - Gauge blocks (*Spécification géométrique des produits (GPS) - Étalons de longueurs - Cales étalons*) (AFNOR, Mars 1999)