

Evaluate and quantify the drift of a measuring instrument

Jean-Michel Pou^{1*}, and Laurent Leblond²

¹Deltamu, Président, 63800 Courmon d'Auvergne, France

²Goupe PSA, Expert statistician, Centre Technique de Vélizy A, 78943 Vélizy-Villacoublay, France,

Abstract. The question of drift in measuring instruments is an essential one in metrology. Calibration establishes the "metrological state" of the instrument on the date it is carried out, but what does it tell us about tomorrow, a month from now, or at the end of the calibration validity period? If, with certain instrumentation (in particular "fixed" instruments, for instance "physical" standards and limit gauges) the situation is relatively straightforward, it is not so simple with instruments that take measurements. The FD X 07-014[1] booklet, moreover, distinguishes between the two types of device, those where drift can be observed (since calibration uncertainties tend to be minor, which is generally true of size standards) and the others (measuring instruments). This article deals with the specific case of measuring.

1 Measuring instruments

Unlike physical standards with fixed dimensions (gauge blocks, calibration weights, plain ring gauges, etc.), "measuring instruments" force us to differentiate between two eventualities:

- "systematic" drift, where the model that describes the relationship between the measured value (denoted by y) and the "true" value (denoted by x) changes with time
- random drift, in other words, residual bias, which come to light as deviations between the model and the values experienced during calibration (residual errors). It is important to point out that these deviations do not originate solely from intrinsic imperfections of the device. Even if we know the model exactly, it will never be perfect. Residual error also includes the modeling error and the dispersal that is manifest during calibration (repeatability of the calibration process, notably).

2 Calibration modelling

* Corresponding author: author@email.org

With the advent of VIM 2008[2], calibration ceased to be more or less confined to a table of deviations between measured values and standard values. Common practice since that date has been to establish the relationship between y and x so as to investigate the instrument over its whole range. This relationship, termed the calibration model, often takes the form of a polynomial of suitable degree n (usually 1 or 2):

$$y = b_0 + b_1x + b_2x^2 + \dots + b_nx^n.$$

NB: for the rest of this article, the convention used is to designate coefficients of this type of calibration model by the letter b , the index being associated with the power of x .

The French College of Metrology (C.F.M.) has published a technical guide [3] on this subject. In fact, it offers a free download of M-CARE, an Excel application that enables the recommendations in the guide to be implemented. Yet few laboratories, to our knowledge, currently offer this service. It is to be hoped that industry will become more exacting in this matter, and set about creating these types of model with the care and attention they deserve.

It is already possible to acquire very useful information using Excel. The LINEST matrix function gives an estimate of the b_i coefficients of the polynomial model. As well as the values themselves, it also gives their associated measurement uncertainties. These coefficients and their uncertainties are calculated via the so-called Ordinary Least Square method (often abbreviated to OLS). It should be remembered, however, that this coefficient estimation method is unsuitable for the vast majority of calibration situations. In fact, the estimations are only meaningful if the following three conditions are satisfied:

1. There is no uncertainty associated with x (which is not the case during calibration, since there are always measurement standard uncertainties).
2. The measurement uncertainty is constant across the full range of measurements (a condition that is rarely met in metrology, as uncertainties are often proportional to the measured value).
3. There is no covariance between either the (x_i) , the (y_i) , or the x_i and y_i , whereas the latter variances are frequent. (We shall not develop this point further in this article: See [4]).

Nevertheless, these "infringements" of the correct use of the coefficient estimation method are not detrimental to the approach that we propose for evaluating and taking into account the drift of measuring instruments, if there really is drift...

3 Does the calibration model actually drift?

It is not easy to demonstrate the drift of a model obtained from calibration. The coefficients of the model will have uncertainties and these uncertainties are correlated. As with any statistical estimation, the reality is "hidden" behind the estimates, which makes it difficult to determine whether or not a drift between two calibrations really exists just from a simple algebraic comparison of the estimates obtained from the latest and from previous parameters. Are the inevitable deviations due to drift, or simply the effect of uncertainties?

To elucidate this vital question, we propose a simple and effective graphical analysis. In order to simplify the demonstration, we shall restrict our investigation to a linear model, but it is

perfectly feasible to deal with higher order models in the same way. We shall therefore model the calibration results on a linear relationship between y and x in the form $y = b_0 + b_1x$, bearing in mind that theoretically b_0 is close to 0 and b_1 close to 1.

In order to determine whether or not this calibration model, characterized by its b_0 and b_1 coefficients, changes over time or not, we shall construct a graph with the b_0 values on the abscissa, and on the ordinate axis the b_1 values obtained from each calibration. The points are plotted in chronological order of calibration and joined by a line.

3.1 Case where drift is not demonstrated

If segments between two consecutive points keep crossing, it is safe to conclude that the apparent differences between coefficients are in reality the result of estimation uncertainties. This situation allows us conclude without too high a risk that the instrument is not exhibiting drift, a hypothesis that should be either proved or disproved during further calibrations by simply adding the new points. The "drift" factor in the analysis of the causes of uncertainty of the process concerned may be deemed "negligible".

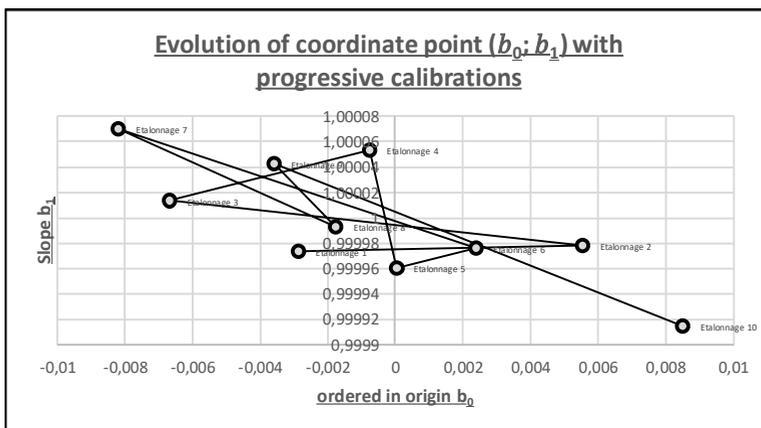


Figure1: Case where drift is not ascertained

3.2 Case where drift is probable

Conversely, if segments do not cross (or only cross rarely), there is very good reason to suppose that the instrument is drifting.

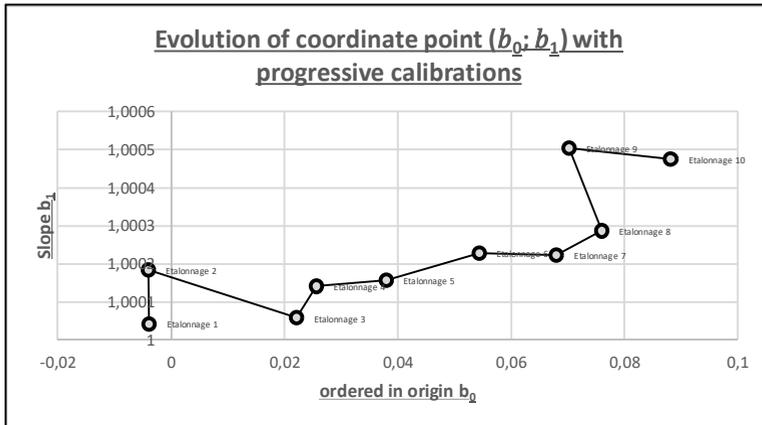


Figure 2: Case where drift is probable

In this instance, it is advisable to model the behavior of the b_0 and b_1 parameters as a function of time. Obviously this modeling will not be totally predictive, but it can allow us to estimate the portion of uncertainty that is due to drift in an uncertainty calculation.

3.2.1 Modeling the behavior of b_0 and b_1 with time

The calibration model makes it possible to establish the relationship between measured values (y) and true values (x). This relationship examines coefficients (b_0 and b_1), which, we shall presume (Figure 2), change with time. We now, therefore, want to model the temporal evolution of these coefficients.

The regression of each of the coefficients with time can be carried out using the OLS method (for example with Excel), with the temporal uncertainties t (days) on the abscissa being negligible, and the uncertainties on the ordinate being independent and of the same amplitude for all calibrations.

To carry out this regression, one marks the calibration dates on the abscissa and the appropriate parameter (b_0 or b_1) on the ordinate. The "LINEST" function will calculate the coefficients by converting dates into "number of days". The parameters are thus given in "variation per day".

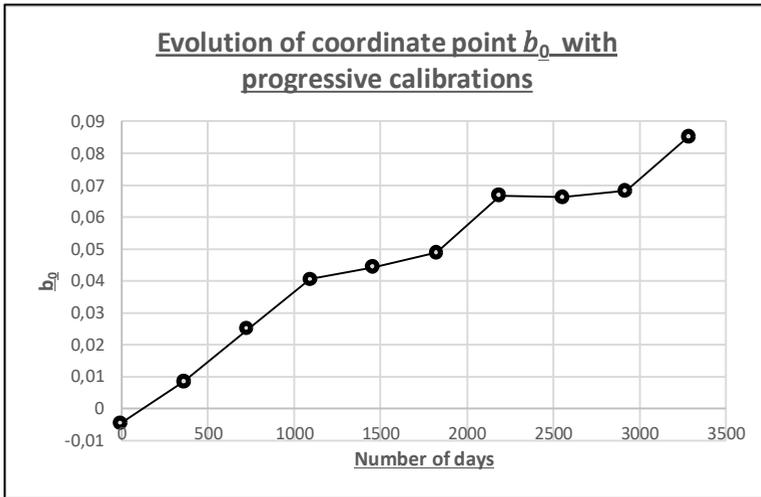


Figure 3: Slope on ordinate that gives (b_0) of instrument calibration model

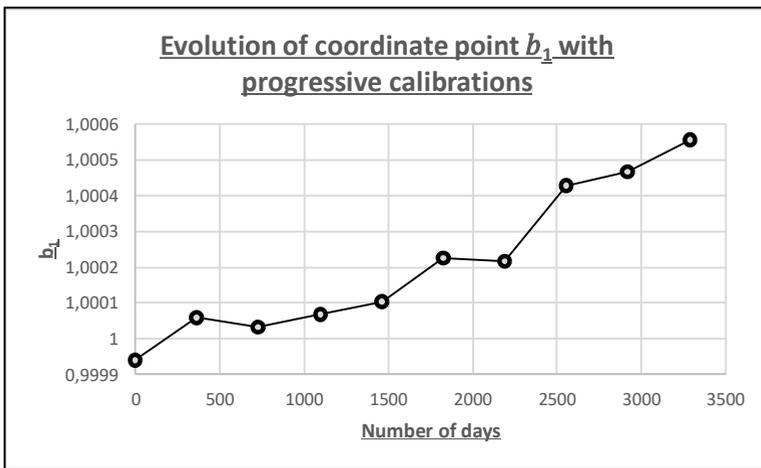


Figure 4: Slope on ordinate that gives (b_1) of instrument calibration model

In this example, we have ten calibrations and therefore ten experimental points with which to describe the behavior of each of the calibration model coefficients. Figure 3 and Figure 4 show that we can choose to model both behaviors with straight lines. If observation revealed that either of the progressions was not linear, then a more suitable model would have to be chosen, and the calculations shown below carried out with the chosen models.

The coefficients of these linear models for the evolution of coefficients (b_0 or b_1) are annotated with the letter a as follows, where t represents time in days:

$$b_0(t) = a_{00} + a_{01} \times t$$

$$b_1(t) = a_{10} + a_{11} \times t$$

As always, the fewer calibrations we have, the more imprecise will be our knowledge of the coefficients of these models. This knowledge will develop as more calibrations are carried out. Nevertheless, the uncertainties of these coefficients, uncertainties whose significance is proportionate to the lack of information, will be taken into account.

So, using traditional statistical notation[†], under the assumption of normality (i.e. $B_k(t_i) \sim \text{Normal}(a_{k0} + a_{k1} \times t_i, \sigma_k)$ where t_i designates the n^{th} calibration of I and σ_k^2 **residual variance** for $k = 0$ or $k = 1$, k referring to the index b), the estimators A_{k0} and A_{k1} of coefficients a_{k0} and a_{k1} of a linear regression using OLS check:

$$\begin{pmatrix} A_{k0} \\ A_{k1} \end{pmatrix} \sim \text{Normale} \left(\begin{pmatrix} a_{k0} \\ a_{k1} \end{pmatrix}, \Sigma_k \right) \quad (1)$$

with $\bar{t} = \frac{1}{I} \sum_{i=1}^I t_i$ and $\Sigma_k = \frac{\sigma_k^2}{\sum_{i=1}^I (t_i - \bar{t})^2} \begin{pmatrix} \frac{1}{I} \sum_{i=1}^I t_i^2 & -\bar{t} \\ -\bar{t} & 1 \end{pmatrix}$ variance-covariance matrix.

The covariance between A_{k0} and A_{k1} is given by:

$$\text{Cov}(A_{k0}; A_{k1}) = -\frac{\bar{t}}{\sum_{i=1}^I (t_i - \bar{t})^2} \sigma_k^2$$

Parameter estimates are obtained from the data $(t_i, b_{k,i})_{i=1 \dots I}$ as follows:

$$\hat{a}_{k1} = \frac{\sum_{i=1}^I (b_{k,i} - \bar{b}_k)(t_i - \bar{t})}{\sum_{i=1}^I (t_i - \bar{t})^2}, \quad \hat{a}_{k0} = \bar{b}_k - \hat{a}_{k1} \bar{t}$$

$$\widehat{\sigma}_k^2 = \frac{1}{I - 2} \sum_{i=1}^I (b_{k,i} - \hat{a}_{k0} - \hat{a}_{k1} t_i)^2 \quad \text{with } \bar{b}_k = \frac{1}{I} \sum_{i=1}^I b_{k,i}$$

Estimate $\widehat{\Sigma}_k$ of the variance-covariance matrix is achieved by substituting σ_k^2 by its estimate $\widehat{\sigma}_k^2$ in the expression for Σ_k .

3.2.2 "Drift" term estimate for uncertainty calculation

When a linear calibration model drifts, the true model of the instrument has, on date t , moved away from that obtained on the last calibration date. As the days pass it gets further from the known model being utilized. Any correction applied to a model that has become obsolete only adds more uncertainties. We shall call the last known model "Model at t_0 ", and "Model at t " is the "provisional" model t days later – the date of the next scheduled calibration, for instance. The coefficients of the "Model at t_0 " will therefore be represented by b_0 and b_1 . The coefficients of the "Models at t " will be represented by $b_0(t)$ and $b_1(t)$.

[†] A capital "Z" is a random variable and a lower-case "z" is a number. When a lower-case has a hat (\hat{z}), it represents the estimate of the corresponding unknown quantity z , i.e. the realization of the random variable Z .

It is easy to imagine two straight lines, one for the "Model at t_0 ", the other for the "Model at t " moving away from the original position due to the drift. The more time goes by, the wider the deviation between the two models becomes...

As the two models are probably not parallel (unless b_0 alone changes), the deviation between the two models, "Model at t_0 " and model presumed to be "Model at t ", depends on the time elapsed and the value measured across the nominal measurement range of the instrument. Figure5 shows the deviations between the models extrapolated at 100, 500 and 1000 days with regard to the "Model at t_0 " across the measurement range of the instrument.

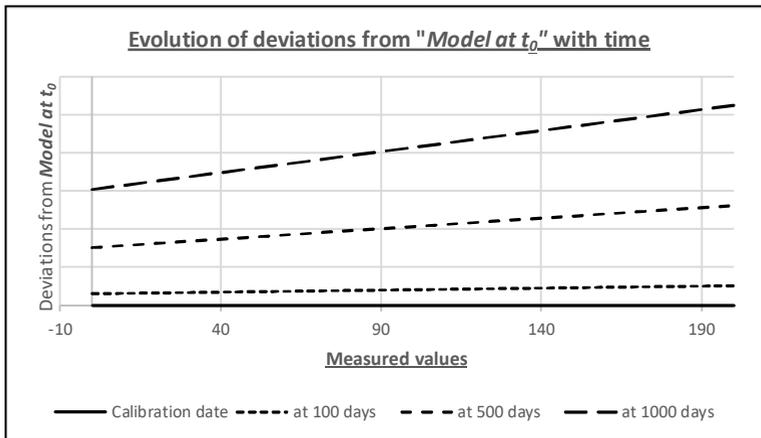


Figure5: Evolution of calibration model

What we clearly have here is a systematic phenomenon. There is nothing random about the *deviation between the "Model at t_0 " and the "Model at t ".* It is a function of the measured value and the number of days that have elapsed since the last calibration.

An interesting parallel...

Let us take the example of a laboratory: temperature is considered to be *random* and, *over time*, is considered to be comprised between two limits. In fact, at any given time, it is what it is. We could make a point of knowing what that is and taking it into account so as to make appropriate corrections. Or use the accepted approach, which is to treat the temperature of the laboratory as "random", since we need to be able to calculate the uncertainty that qualifies a result **regardless of the day on which the measurement is made**. It is therefore the moment at which the measurement is made that gives the temperature its random nature, and not the temperature itself. In the case currently under discussion, the situation is much the same.

Literally, the deviation will be whatever it is, with a measurement field value of x and a measurement day of j . As we are proceeding with an uncertainty calculation whatever the value of x in the measurement field, and whatever day j is in the period between now and the next calibration, one can consider that the "whatever x is" and "whatever j is" give a random nature to the inevitable deviation between the correction applied and the "true" correction, which nobody knows.

Therefore, to evaluate the term for the drift of an instrument, we suggest simply calculating the deviations between the two corrections ("Model at t_0 " and "Model at t ") at points dispersed across the **whole measurement field of the instrument**, then to calculate the mean and standard deviation of the deviations obtained. These two terms, mean and standard deviation, are the contributions made by the drift of the instrument to the uncertainty of the measurement process. One could say that they represent the error one risks by not using an appropriate calibration model.

The deviation mean varies according to the number of days j between now and the last calibration. Because it varies over the period, we suggest including it in the uncertainty calculation and assuming, for example, that it has a uniform distribution between 0, the date of the last calibration, and a value determined by calculation based on the periodicity in question. As for the standard deviation, this will be directly taken into consideration in the analysis, as a measurement can be made at any time and any point of the nominal measurement range of the instrument.

Table 1 shows the results obtained for an instrument with a nominal range of [0;200], the behavior models $b_0(t)$ and $b_1(t)$ being the results of 10 successive calibrations. The method used is to calculate, for the various x values possible across the nominal measurement range of the instrument, correction values derived from "Model at t_0 " and "Model at t ", with t varying between 100, 500 and 1000 days, then to calculate the deviations between the two.

Coefficients used		Coefficients modeled					
Last calibration		at 100 days		at 500 days		at 1000 days	
b_0	b_1	b_0	b_1	b_0	b_1	b_0	b_1
0,09407	1,00043	0,09708	1,00044	0,10912	1,00048	0,12416	1,00054
X	Y	Y(100)	Ecart	Y(500)	Ecart	Y(1000)	Ecart
2	2,09492	2,09796	0,00303	2,11008	0,01516	2,12524	0,03031
4	4,09578	4,09883	0,00305	4,11105	0,01527	4,12632	0,03054
6	6,09663	6,09971	0,00308	6,11201	0,01538	6,12739	0,03076
...
196	196,17778	196,18298	0,00521	196,20381	0,02603	196,22983	0,05205
198	198,17863	198,18386	0,00523	198,20477	0,02614	198,23091	0,05228
200	200,17949	200,18474	0,00525	200,20574	0,02625	200,23199	0,05250

Table 1: Calculation of correction deviations for full measurement range [0,200], according to number of days

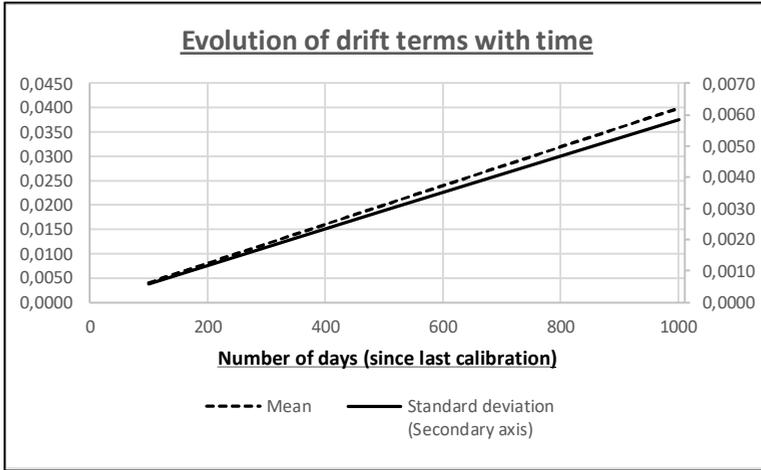


Figure 6: Evolution of "mean" and "standard deviation" terms for drift with time

Impact at (No. of days)	100	500	1000
Mean	0,0041	0,0207	0,0415
Standard deviation	0,0008	0,0041	0,0083

Table 2: Results of analysis of correction deviations between "Model at t_0 " and "Model at t "

3.2.3 How should the uncertainty of our model based on the evolution of coefficients be used in our calculations?

However many calibrations we have at our disposal to evaluate whether an instrument is suffering from drift and, where relevant, however the calibration model coefficients behave over time, no model will be free of uncertainty.

Taking into account the uncertainties of model coefficients is not a trivial matter. Digital simulation is one possible solution, but the context here is very specific. The aim of the exercise is to simulate the possible coefficients to give us $b_0(t)$ and $b_1(t)$, then to produce calculations for the mean and standard deviation of the deviations between the "Model at t_0 " and the "Model at t ".

Brief reflection on digital simulation

While a unique random variable simulation can be achieved simply using spreadsheets, the need here is to simulate **correlated** random variables. Whatever the evolution model, the model coefficients are not independent of each other. In other words, since we are talking here of the evolution of coefficients b_0 and b_1 of the calibration model, which have themselves been modeled by coefficient straight lines (a_{00}, a_{01}) (respectively (a_{10}, a_{11})), a random value of a_{00} is not necessarily compatible with a random value of a_{01} . One can, on the other hand, assume that coefficients $(a_{00}$ and $a_{01})$ and $(a_{10}$ and $a_{11})$ are independent of each other, on the basis that the evolution of b_0 and b_1 of the instrument do not have the same technical origin. It is worth putting the hypothesis to the test, and if it is disproved, the terms of the covariance between a_{00}, a_{01}, a_{10} and a_{11} should be evaluated and used.

On this subject, it is a pity that the "LINEST" function of Excel does not give the covariance between coefficient estimates, and must therefore be manually calculated using the formula given in 2.2.1.

The term $\widehat{\sigma}_k^2$, the variance of residual errors, is given in the results matrix of the "LINEST" function (Table 3). It can also be directly calculated, but in this case attention must be paid to the number of degrees of freedom (-2 versus $n - 1$ for a linear model).

LINEST (y;x>true>true) - Straight line case	
b_1	b_0
u_{b1}	u_{b0}
r^2	Standard error of extrapolation
F (tells us whether the model makes sense)	Degree of freedom
Regression sum of squares	Residual error variance

Table 3: Results matrix for "LINEST" function with a linear model

Supplement 1 to GUM [5] provides information on a method to generate correlated random numbers, to which the reader might usefully refer.

There are therefore two steps to the calculations.

Step 1: Simulate values (a_{00}, a_{01}) (respectively (a_{10}, a_{11})) according to the normal distribution (1) (cf. 2.2.1) where a_{k0}, a_{k1} and σ_k^2 are replaced by their estimates.

Using these four simulated values, calculate $b_0(t)$ and $b_1(t)$ coefficients for "Models at t " as many times t as required. Then calculate the mean and standard deviation for the model deviations (cf Table 1 for $t = 100, t = 500$ and $t = 1000$). In this way, for each t , one can obtain the same quantity of means and standard deviations as simulations.

Step 2: Construct histograms of the means (Figure 7) and standard deviations (Figure 8) obtained through simulations (step 1), then determine, at chosen confidence level α (as a rule 95%), the maximized values for the "Mean" and "Standard deviation" parameters (Table 4). This is done by searching in the simulated values for a value (mean and standard deviation) that is only exceeded in the simulations by $(1 - \alpha)\%$ (as a rule 5%) of the values.

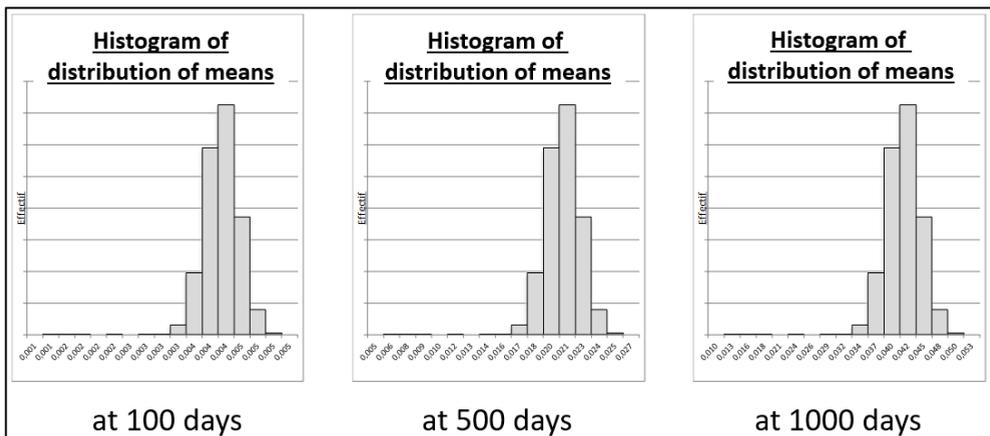


Figure 7: Histograms of means obtained via simulations

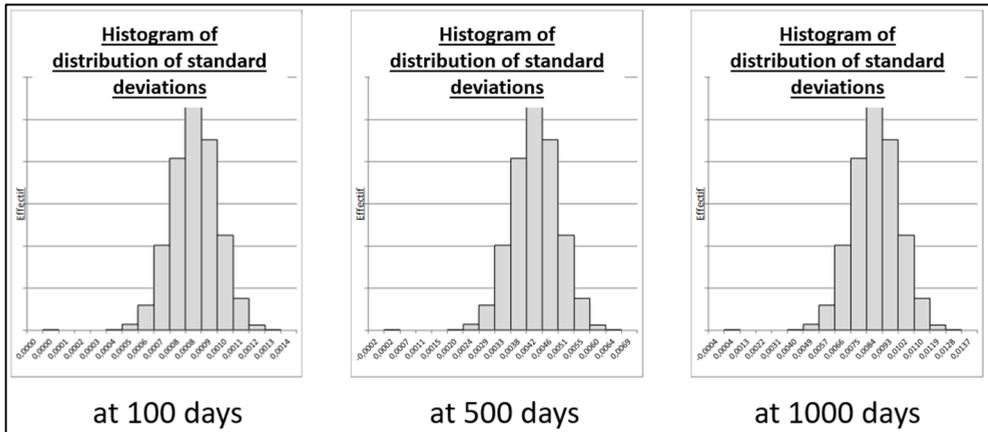


Figure 8: Histograms of standard deviations obtained via simulations

	Maximization of parameters (at 95 %)		
Number of days	100	500	1000
Mean	0,0047	0,0233	0,0466
Standard deviation	0,0011	0,0053	0,0107

Table 4: Maximization of Mean and Standard deviation parameters

NB: As can be seen from our simulations, the maximized values (*cf* Figure 9), like the mean values (*cf* Figure 6), seem to follow a linear model based on the number of days. This observation, which should be verified for each particular case, can be very useful when it comes to fixing periodicity, as the choice can now be based on the significance of drift in the instrument versus the other causes of uncertainty of the process...

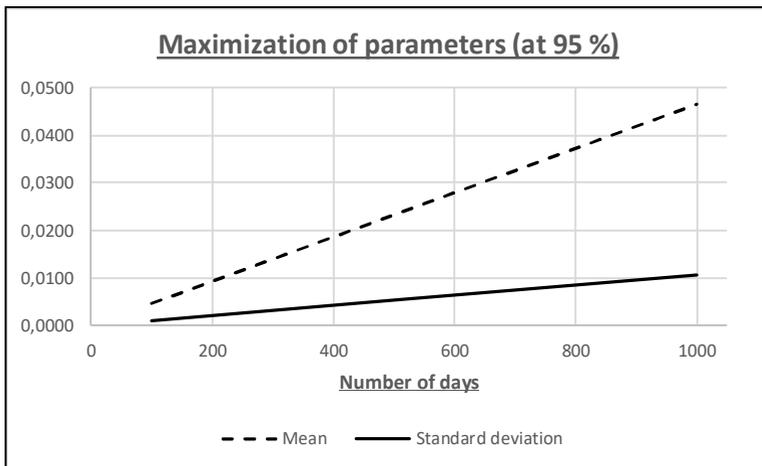


Figure 9: Evolution of maximized parameters with time

Depending on the criticality of the measurements, one can take either the average values (Table 2) or the maximized values (Table 4) of the instrument drift parameters that contribute to the overall uncertainty of the measurement process. It is nevertheless worth bearing in mind that it is rare, in uncertainty calculation, to maximize components. Who, for example, would use the upper bound of the confidence interval of a standard deviation derived from a chi-square distribution?

4 Does residual error in a calibration model actually drift?

For the residual error of the calibration model, the evolution of standard deviation with time needs to be looked at. For example, the classic Fisher-Snedecor test to compare two variances can establish whether or not the standard deviation changes between two calibrations (<http://wikistat.fr/> Inference – Statistical Tests).

The simple F test shows what we are looking for:

$$F = \frac{\max(S_{Residual\ error\ Current\ calibration}^2; S_{Residual\ error\ Prior}^2)}{\min(S_{Residual\ error\ Current\ calibration}^2; S_{Residual\ error\ Prior}^2)}$$

Once the value for F has been calculated, it is a simple matter of checking that the value is lower than the quantile of the Fisher-Snedecor distribution, determined for a given α confidence level, so we can judge if the residual error variance of the new calibration is consistent with the known a priori variance:

$$F \leq INVERSE.LAW.F(\alpha; d_{numerator}; d_{denominator}) = Limit$$

With

- $d_{numerator}$: Number of variance degrees of freedom for the numerator
- $d_{denominator}$: Number of variance degrees of freedom for the denominator

If the test gives statistically identical variances ($F \leq Limit$), we can learn more about the variance of remaining errors by calculating a weighted average of the two variances. The value thus obtained, which is more reliable than taking only one of the two values, becomes the new known value (a priori) for future calibration tests. The procedure is cumulative, and thus becomes increasingly effective in detecting changes in residual error as an increasing number of calibrations are conducted.

If the test gives statistically different variances ($F > Limit$), it will be necessary to model the evolution of residual error variance in order to take into account any deterioration over the days/months/years since the last calibration. The modeling principle of this temporal evolution is the same as that already detailed for calibration model coefficients. It becomes the third component in the contribution of instrument drift to the overall uncertainty of a measurement process.

5 Conclusion

Drift in measuring instruments is a subject that is dealt with relatively little in the literature. Some, rather summary, methods suggest calculating, calibration point by calibration point, the difference between measured values from calibration at date $t - 1$ and calibration at date t , taking the higher as the absolute value in the uncertainty calculations, and applying a (typically uniform) probability distribution.

These methods, seemingly simple, do not actually show us drift. The subtractions do not reveal any actual evolution of the instrument between two dates, as each value of the subtraction contains calibration error. The deviation obtained might therefore only be the result of two different calibration errors, without the instrument having in any way altered. Moreover, the method forces us to analyze the same measurement points at each calibration. The number of points during calibration will always be limited for the obvious reason of cost, and the lower the number of points, the greater the calibration model uncertainty. By

changing the points used at each calibration, as our method allows, more information about the instrument can be gained with each calibration. So when there is no drift in the instrument, the various calibration results can be used to estimate the model parameters, thereby reducing uncertainties related to the reference device.

To conclude, these calculations allow us to set calibration periodicity rationally, using our knowledge of the contribution made by drift to the overall uncertainty of the process in which the instrument is used.

References

- [1]: *FD X 07-014: «Métrologie - Optimisation des intervalles de confirmation métrologique des équipements de mesure »* (AFNOR 2006)
- [2]: *NF ISO/IEC Guide 99: «Vocabulaire international de métrologie - Concepts fondamentaux et généraux et termes associés (VIM) »* (AFNOR 2011)
- [3]: *«Etalonnage suivant le VIM 3 - Application du nouveau concept »* (Collège Français de Métrologie, 2013)
- [4]: C. Dubois, L. Leblond, JM Pou, « *Covariance evaluation by means of uncertainty assessment* » (https://www.researchgate.net/publication/311749082_Covariance_evaluation_by_means_of_uncertainty_assment)
- [5]: *ISO/IEC GUIDE 98-3/S1 «Measurement Uncertainty - Part 3: Guide to the expression of uncertainty in measurement (GUM:1995) - Propagation of distributions using a Monte Carlo method»* (ISO 2008)