

INCERTITUDE D'ETALONNAGE : METHODE ANALYTIQUE (G.U.M) VERSUS METHODE EXPERIMENTALE (ISO 5725).

Jean-Michel POU et Dimitri VAISSIERE

Delta Mu

25, Rue Joseph Desaynard

63000 Clermont-Ferrand

Résumé

La norme NF ENV 13005 [1] est aujourd'hui la référence dans le cadre de l'évaluation des incertitudes de mesure. Elle impose de répondre aux interrogations : Quelles causes considérer ? Quelle loi de distribution choisir pour les évaluations dites de type « B » ? Quelles covariances prendre en compte, et comment ?

En comparant, pour un même calcul, les résultats obtenus en appliquant le GUM et ceux relevant d'un essai inter laboratoires (ne nécessitant aucune hypothèse mais s'appuyant sur la réalité expérimentale), certains problèmes apparaissent ...

NF ENV 13005 (ISO Guide 99-3 [2]) is actually the reference for measurement uncertainty evaluation. This evaluation according to GUM leads to answer following questions: Which factors taking into account? Which density of probability for "B type" evaluation? What about covariances?

Considering a calculus using GUM method and experimental method like interlaboratories comparison (which needn't hypothesis because based on reality observation), some problems arise.

Introduction

Pour réaliser la comparaison des résultats obtenus par l'application du GUM et ceux obtenus par un essai inter-laboratoires organisé suivant la série des normes ISO 5725 [3], les auteurs se basent sur deux analyses :

1. Un calcul d'incertitude type obtenu par compilation de calculs présentés par des laboratoires accrédités COFRAC pour l'étalonnage de l'erreur de justesse totale des comparateurs à cadran au 1/100 de mm.
2. Les résultats d'un essai inter-laboratoires organisés en 2008 par Delta Mu en regroupant 8 laboratoires (1 laboratoire a été exclu suite aux tests de cohérence) dont un des essais reposait sur les comparateurs ci-dessus désignés.

L'erreur de justesse totale d'un comparateur à cadran

La norme française NF E 11-050 (remplacée maintenant par la norme internationale ISO 463 [4] mais toujours utilisée par les laboratoires français) prévoit, pour l'évaluation de l'erreur de justesse totale, de réaliser, dans un premier temps, une courbe d'étalonnage. Il s'agit, pour quelques points de référence (la norme prévoit un minimum de 1 point par tour d'aiguille, soit 11 points - en considérant le 0 - pour une course de 10 mm), de déterminer l'erreur de justesse (Différence entre la moyenne de 5 mesures et la valeur indiquée par le banc de mesure) en chaque point puis de tracer une courbe d'étalonnage. Les extremums de la courbe permettent de calculer l'erreur de justesse totale (e_{jt}), qui se définit donc par l'équation simple : $e_{jt} = e^+ - e^-$ (e^+ représente la plus grande erreur de justesse et e^- la plus petite, en valeur algébrique). Cette définition mathématique de e_{jt} implique de mettre en œuvre la loi de propagation.

L'approche GUM : Calcul type

Différents laboratoires COFRAC ont donc été interrogés pour savoir quelles étaient les causes qu'ils considéraient dans le calcul d'incertitude pour ladite erreur.

Le Tableau 1 ci-dessous dresse la liste des causes considérées, le moyen d'étalonnage et l'incertitude pour laquelle le laboratoire est accrédité.

Ainsi, nous pouvons constater que tous les laboratoires considèrent sensiblement les mêmes

Causes considérées	LABO 1	LABO 2	LABO 3	LABO 4
Répétabilité	O	O	O	O
Incertitude d'étalonnage du Banc	O	O	O	O
Quantification du Banc	O	O	O	N
Quantification du comparateur	O (Dans répétabilité)	N	O	O
Ecart de température entre Banc et Comp	O	O	O	O
Méconnaissance des coefficients de dilatation	O	O	O	O
Flexibilité du support	O	N	O	N
Alignement Banc/Comp	N	O	O	O
<i>Considérées en erreur systématique</i>				
Justesse du banc	O	O	O	O
Hystérésis du banc	N	O	O	O
Loi de propagation	O	O	N	N

U annoncée (pour une course de 10 mm)	± 2,9 µm	± 4 µm	± 2,8 µm	± 4,8 µm
Moyen	Banc SIP	Banc SIP	Banc Mahr	Banc Sylvac

Tableau 1 : Causes prises en compte par les laboratoires interrogés
O signifie que la laboratoire considère la cause décrite, N signifie qu'il ne la considère pas.

Quantification du comparateur	
Labo 1	Intégrée dans la répétabilité
Labo 2	Non considérée
Labo 3	Négligeable
Labo 4	Le Tiers de la quantification suivant une loi uniforme

Tableau 2 : Stratégies de prise en compte concernant l'erreur liée à la lecture sur le comparateur

De plus, lorsque nous admettons, dans le GUM, que la loi de distribution est « normale » et que l'erreur maximale est divisée par 3, on suppose connaître ladite erreur maximale avec un niveau de confiance de 99,73%. Si nous avons supposé que cette erreur maximale correspondait en fait à un niveau de confiance

causes. Au-delà, il est intéressant d'observer la stratégie au regard des erreurs liées à la quantification du comparateur (Tableau 2). On remarque que les hypothèses sont différentes alors que nous traitons du même sujet. Le lecteur pourra admettre que considérer l'erreur de lecture dans la répétabilité (Labo 1) ou la considérer égale au tiers de la quantification suivant une loi uniforme (Labo 4) conduit évidemment à un résultat différent.

de 99%, il aurait fallu diviser par 2,58 environ. Ce qui représente 26 % d'écart ($(3^2 - 2,58^2) / 3^2 = 0,26$). Ainsi, une différence de 0,73 % sur l'erreur maximale considérée conduit à une différence de 26% sur sa participation dans le calcul final, en restant dans l'hypothèse gaussienne qui peut elle aussi être contestée. On

comprend aisément, sur la base de ce simple calcul, tout le problème des estimations dites de type « B ».

Il est également intéressant de voir comment est considérée l'erreur de justesse du banc de mesure utilisé pour l'étalonnage. Le V.I.M [5] définit l'erreur de justesse comme étant systématique, ce qui conduit par cet effet « sémantique » tous les laboratoires à additionner linéairement ce paramètre à l'écart-type composé. Il suffit de regarder une seule fois un certificat d'étalonnage d'un banc de mesure pour se rendre compte que cette erreur n'a absolument rien de systématique ! Si tel était le cas, et quelque soit la valeur de l'étalon de référence, l'erreur du banc serait, soit constante et la courbe d'étalonnage serait alors strictement parallèle à l'axe des abscisses, soit en relation avec la valeur de référence (dans le cadre d'une représentation traditionnelle des résultats d'étalonnage de ce type de moyen, une droite ou tout autre modèle ...). De plus, et si c'était le cas, l'application rigoureuse de la loi de propagation permettrait d'éliminer tout ou partie de cette cause d'erreur qui participerait directement à la covariance. Elle participerait en effet de façon connue dans e^+ et dans e^- et « disparaîtrait » logiquement pour la part commune dans e_{jt} qui se présente comme la différence des deux. Cette explication « de bon sens » est confirmée par la formule :

Application de la loi de distribution

$$e_{jt} = e^+ - e^-$$

D'où

$$u_{e_{jt}}^2 = \left(\frac{\partial e_{jt}}{\partial e^+} \times u_{e^+} \right)^2 + \left(\frac{\partial e_{jt}}{\partial e^-} \times u_{e^-} \right)^2 + 2 \times \left(\frac{\partial e_{jt}}{\partial e^+} \right) \times$$

Moyenne	Unité	Sr	SL	SR	U (k = 2)
0,01537	mm	0,00060	0,00108	0,00124	0,00248
Participation	/	23,5%	76,5%		

Tableau 3 : Résultats obtenus suivant l'ISO 5725-2 [6]

$$\left(\frac{\partial e_{jt}}{\partial e^+} \right) \times Cov(u_{e^+}; u_{e^-})$$

Soit :

$$u_{e_{jt}}^2 = (1 \cdot u_{e^+})^2 + (-1 \cdot u_{e^-})^2 - 2 \cdot Cov(u_{e^+}; u_{e^-})$$

Le signe « - », qui provient de la dérivée partielle due à e^- , vient confirmer le bon sens évoqué ci-avant. Dans ce cas en effet, puisque les incertitudes sur e^+ et sur e^- évoluent dans le même sens, la covariance entre les deux est positive.

L'approche expérimentale

Les méthodes expérimentales présentent l'avantage de n'avoir aucune hypothèse à faire pour estimer les écart-types, donc les incertitudes, des méthodes étudiées. Des hypothèses portent néanmoins sur les tests d'exclusion car les valeurs seuils sont données pour des phénomènes gaussiens. Il est par ailleurs possible de revoir lesdits seuils si l'hypothèse de normalité n'est pas confirmée.

Comme indiqué en préambule, 7 laboratoires ont été considérés dans le cas qui nous intéresse. Les spécialistes noteront que 7 laboratoires représentent un faible échantillon. Il nous faut en ce cas rappeler qu'un écart-type expérimental obtenu sur peu d'individus est plus probablement minorant que majorant.

Pour être sûr d'observer le même mesurande, les points d'étalonnage du comparateur étaient fixés. Ils étaient donc les mêmes pour tous les participants.

Les résultats (

Tableau 3) obtenus sont les suivants :

Nous pouvons donc constater que l'incertitude expérimentale est plus faible que l'incertitude obtenue par application du G.U.M. Cette différence n'est pas surprenante car les approximations inévitables liées à la méthode analytique ne peuvent que conduire à des écarts. En moyenne, les laboratoires annoncent une incertitude sur cette erreur de 3,6 µm pour une incertitude expérimentale de 2,5 µm, soit un écart de $(3,6 - 2,5) / 2,5 = 44\%$!

Aucun statisticien n'est à convaincre des avantages de la méthode expérimentale ... Ici, pas de question de sémantique (sur l'erreur de justesse notamment) ou d'hypothèse (Loi normale ou autre, suivant quel niveau de confiance, ...) ou encore sur les causes à considérer (l'erreur de quantification doit-elle être prise en compte ou non ?) jusqu'aux covariances éventuelles et leur prise en compte par la loi de propagation.

Dans la réalité, les causes se sont exprimées, chez l'un et chez l'autre, et l'écart-type résultant n'est que le fruit de ces expressions. Evidemment, le nombre d'observations perturbe la qualité des écart-types expérimentaux et ce problème peut facilement être résolu par l'augmentation du nombre de laboratoires participants ...

L'approche expérimentale, complément ...

L'erreur de justesse totale du comparateur ne relève pas du hasard dans le cadre de cette comparaison entre les 2 approches. En effet, par conception, le comparateur à cadran est sujet à des erreurs importantes qu'il est possible d'observer aisément. Il est composé de plusieurs roues dentées fabriquées par emboutissage. Chaque dent de chaque roue étant différente, elles ont leur propre erreur et le choix de telle ou telle « dent » (via le point de référence) impacte directement le résultat final. Ainsi, le choix des points d'étalonnage devrait impacter sensiblement le résultat final.

Pour confirmer cette hypothèse, nous avons demandé aux laboratoires participants, après avoir étalonné le comparateur sur les points que nous avons fixés, de dérouler leur propre

procédure. Les résultats sur l'erreur de justesse totale sont présentés ci-dessous (Tableau N°4) :

Ejt	EN
0,0163	1,3
0,0094	-1,0
0,0120	-0,1
0,0085	-1,3
0,0137	0,4
0,0095	-1,0
0,0120	-0,1
0,0134	0,3
0,0173	1,6
Ecart-type	0,0031
Moyenne	0,0125

Tableau 4 : Résultats suivant les procédures d'étalonnage propre à chaque laboratoire

Le test d'exclusion proposé se limite à l'étude de l'écart normalisé des laboratoires, noté EN dans la seconde colonne du tableau 4. Aucun d'entre eux n'a un écart supérieur à 2, ils sont donc tous considérés dans l'estimation de l'écart-type des erreurs de justesse sensé représenté l'incertitude d'étalonnage ...

La valeur obtenue (0,0031 mm) est bien supérieure à celle obtenue précédemment en ayant fixé les points de référence (0,0012 mm,

Tableau 3).

La différence des deux valeurs est donc quasi uniquement due aux choix des points de référence (échantillonnage). Par soustraction de variances, il est possible de déterminer la valeur, et donc le poids, de cette cause considérée par aucun des laboratoires ayant soumis leur calcul théorique :

$$0,0031^2 = \text{Cause liée aux choix des points}(CDP)^2 + 0,0012^2$$

Soit :

$$CDP = \sqrt{0,0031^2 - 0,0012^2} = 0,0028$$

Soit le poids de cette cause dans l'incertitude finale :

$$Poids_{CDP} = \frac{0,0028^2}{0,0031^2} \approx 85\% !$$

Conclusion

L'approche expérimentale a donc permis de mettre en évidence les limites de l'approche analytique (avec ses inévitables approximations). Elle ouvre également un champ de réflexion sur le choix du nombre de points d'étalonnage.

Comme nous l'avons vu, et pouvons aisément l'admettre, le choix et le nombre des valeurs de référence impactent le résultat final sur le comparateur, objet de cette étude. Il semble évident que tous les autres moyens de mesure sont concernés par cette remarque ...

Cette question relève probablement, en tout cas pour partie, de la science statistique car il s'agit bien d'un problème lié à l'échantillonnage. Evidemment, il faut également tenir compte, et peut être avant tout, de la technologie de l'instrument concerné car c'est bien elle qui détermine les problèmes potentiels. Il est donc essentiel de la connaître pour choisir.

La nouvelle définition du VIM [5] concernant l'étalonnage va imposer aux « étalonneurs » d'avancer sur ce sujet. En effet, il ne s'agit plus maintenant de simplement dresser un tableau de valeurs indiquant la valeur de référence et la valeur mesurée correspondante par le moyen, en précisant évidemment les incertitudes associées. Il faudra, dans le futur, mettre en relation ces 2 valeurs pour permettre à l'utilisateur de corriger, le cas échéant, les indications de son instrument.

Au-delà de la correction, il faudra bien sur être en mesure de fournir :

1. L'incertitude sur cette correction qui ne pourra pas être parfaite ...
2. Le résiduel de justesse, c'est-à-dire la part aléatoire due au fait que chaque « dent » - chaque graduation - a sa propre petite erreur et ne pourra probablement jamais être corrigée.

Références

1. *NF ENV 13005*, in *Guide pour l'expression de l'incertitude de mesure*. 1999.
2. *ISO/CEI Guide 98-3*, in *Incertainde de mesure - Partie 3 : guide pour l'expression de l'incertitude de mesure (GUM:1995)*. 2008.
3. *ISO 5725*, in *Exactitude (justesse et fidélité) des résultats et méthodes de mesure. Parties 1 à 6*. 1994.
4. *ISO 463*, in *Spécification géométrique des produits (GPS) - Instruments de mesurage dimensionnel : comparateurs mécaniques à cadran - Caractéristiques de conception et caractéristiques métrologiques*. 2006.
5. *VIM, 3ème édition, JCGM 200*, in *Vocabulaire international de métrologie. Concepts fondamentaux et généraux et termes associés*. 2008.
6. *ISO 5725-2*, in *Exactitude (justesse et fidélité) des résultats et méthodes de mesure. Partie 2: méthode de base pour la détermination de la répétabilité et de la reproductibilité d'une méthode de mesure normalisée*. 1994.