

# **LES COMPARAISONS CLES DU CONCEPT A LA PRATIQUE [KEY COMPARISON: FROM CONCEPT TOWARDS PRACTICE]**

**Soraya Amarouche  
Laboratoire National d'Essais BNM-LNE  
1, rue Gaston Boissier 75724 Paris Cedex 15**

## **Résumé**

Le document du CIPM (Comité International des Poids et Mesures), "Reconnaissance mutuelle des étalons nationaux de mesure et des certificats d'étalonnage et de mesurage émis par les laboratoires nationaux de métrologie" (MRA) [1], demande de quantifier le degré d'équivalence des laboratoires nationaux au travers de comparaisons.

Des difficultés sont très vite apparues dans la quantification des paramètres qui servent à reconnaître l'équivalence des laboratoires de métrologie nationaux.

Les contraintes techniques, ainsi que des conditions spécifiques pour l'utilisation de certains outils statistiques ne permettent pas de se satisfaire d'une seule méthode d'exploitation. La diversité des problèmes rencontrés lors de l'exploitation des résultats, dont cet article donne une idée, nécessite l'élaboration de guides d'exploitation. Un premier guide, qui vient d'être publié, propose des éléments de réponse.

## **Abstract**

The document of CIPM (International Committee for Weights and Measures), "Mutual recognition of national measurement standards and of calibration and measurement certificates issued by national metrology institutes" (MRA) [1], requires the quantification of the degree of equivalence of national metrology laboratories through comparisons.

Difficulties were soon encountered in the determination of the parameters used as recognition of the equivalence of national metrology laboratories.

Technical recommendations and specific conditions for the use of statistical tools do not allow to establish a unique satisfactory exploitation method. The large number of problems encountered in exploitation of results - presented in this paper - necessitates the elaboration of exploitation guides. A first guide recently being published proposes means for the assessment of key comparison data.

## Le concept au travers du MRA [1]

Les laboratoires nationaux de métrologie collaborent et effectuent des comparaisons internationales de leurs étalons nationaux de mesure depuis plus de cent ans. Cependant, la reconnaissance ad hoc des étalons qui en résultait n'était plus considérée comme suffisante.

Un document, "l'arrangement de reconnaissance mutuelle des étalons nationaux de mesure et des certificats d'étalonnage et de mesurage émis par les laboratoires nationaux de métrologie" (MRA), a été rédigé par le CIPM (Comité International des Poids et Mesures), afin d'établir le degré de confiance que chaque laboratoire participant peut avoir dans les résultats publiés par les autres.

Le point de départ de l'application de ce document est la date de sa signature, le 14 octobre 1999, par les directeurs de laboratoires de métrologie (LNM) de 38 états membres de la convention du mètre, afin de jouer un rôle majeur dans la mondialisation de la métrologie (équivalence des services de métrologie) et d'offrir un fondement technique à des accords plus larges négociés dans le cadre du commerce, du négoce et des règlements internationaux.

L'outil nécessaire pour la quantification de ce degré de confiance, est la réalisation de comparaisons internationales de mesurage, désignées par "comparaisons clés" (key comparisons). Ces comparaisons clés vont permettre d'établir le degré d'équivalence des étalons nationaux.

Les comparaisons clés sont organisées par les comités consultatifs (CC's), le Bureau International des Poids et Mesures (BIPM) et les Organisations Régionales de Métrologie (ORM) [fig1] et elles permettent :

1. de déterminer des valeurs de référence, appelées valeurs de référence des comparaisons clés, ainsi que leurs incertitudes associées.

$$\{x_{ref}, u(x_{ref})\}$$

2. de quantifier le degré d'équivalence des étalons nationaux de mesure en deux étapes par :

- a) Le degré d'équivalence de chaque étalon national de mesure qui est exprimé quantitativement par deux termes :

$$\{d_i = x_i - x_{ref}; U(d_i)\}$$

- L'écart de sa valeur à la valeur de référence de la comparaison clé.
- L'incertitude de cet écart à un niveau de confiance de 95%.

- b) Le degré d'équivalence entre deux étalons nationaux de mesure qui est exprimé par deux termes :

$$\{D_{ij} = d_i - d_j = x_i - x_j; U(D_{ij})\}$$

- La différence de leurs écarts par rapport à la valeur de référence.
- L'incertitude de cette différence à un niveau de confiance de 95%.

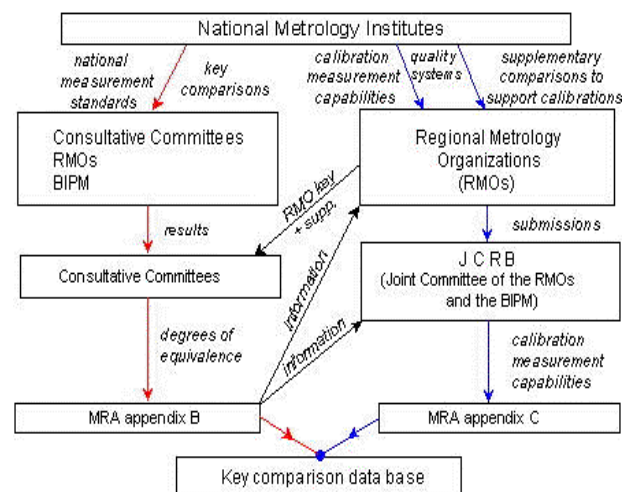


Fig.1 Les différentes comparaisons clés du MRA

La quantification de ce degré d'équivalence qui est pris au sens de conformité repose pour une grande part sur la valeur de référence et son incertitude associée. La quantification de cette valeur de référence constitue une étape essentielle dans l'analyse des résultats de mesure de la comparaison clé.

L'annexe F "Directives pour les comparaisons clés du CIPM" [2] décrit dans les grandes lignes la structure d'une comparaison clé.

Il propose même de lister les principales composantes du bilan d'incertitude que chaque participant doit évaluer et préconise d'évaluer les incertitudes selon le Guide pour l'expression des incertitudes de mesure [7]. Cette annexe est rédigée pour des étalons voyageurs ayant une bonne stabilité à court terme et au cours des transports. Il existe cependant des domaines de la métrologie où les étalons présentent des dérives malgré les précautions prises au départ.

Pour la partie de la comparaison clé, que l'on pourrait désigner par l'exploitation des résultats de mesure, le choix des méthodes d'exploitation repose beaucoup sur le pilote auquel il manque parfois des critères d'aide à la décision. En effet, si l'on se reporte au paragraphe de l'annexe F [2] "Préparation du rapport sur les comparaisons clés", il est indiqué que "Pour calculer la valeur de référence de la comparaison clé, le laboratoire pilote utilise la méthode qu'il considère comme la plus appropriée à cette comparaison clé,..." Cette liberté de choisir une méthode d'exploitation des données pourra poser des problèmes dans l'interprétation des différentes comparaisons clés.

Le MRA et ses annexes qui sont la base des comparaisons clés sont des documents qui cherchent à couvrir

l'ensemble des domaines de la métrologie et par conséquent sont des documents généralistes.

Le premier workshop sur le sujet, qui s'est tenu le 11 & 12 novembre 1999 au NPL[3] a mis l'accent sur les problèmes rencontrés lors de l'exploitation de comparaisons clés et sur lesquels aucunes solutions n'ont été proposées dans les documents officiels.

## **Les principaux problèmes d'application**

La liste n'est pas exhaustive :

1. Dérive des étalons voyageurs [4] :  
Lors d'une comparaison clé dans le domaine du dimensionnel (Euromet L – K1), certaines cales étalons ont montré une dérive significative. Cette dérive perturbe l'exploitation des données pour répondre aux objectifs du MRA, d'où la nécessité de trouver un moyen de s'affranchir de cette dérive.  
  
L'auteur propose de prendre en compte cette dérive dans la quantification de la valeur de référence au travers d'une modélisation. Si l'estimation est améliorée, tous les problèmes ne sont pas résolus car l'erreur n'étant pas systématique, elle réduit des écarts observés sur les premiers laboratoires participants non imputables à la dérive.  
  
Le sujet reste donc toujours d'actualité et nécessite une méthode consensuelle.
2. Schéma de circulation de l'étalon voyageur :  
Les différents schémas de circulations de l'étalon voyageur (Round Robin, en étoile,...) ont des implications sur la méthode d'exploitation des résultats de la comparaison clé.
3. Indépendance ou non-indépendance des résultats fournis :  
Quelques-unes ou toutes les mesures des laboratoires nationaux sont mutuellement dépendantes.
4. Valeur de référence a priori :  
La valeur de référence est déterminée à l'avance par d'autres moyens que les mesures effectuées par les laboratoires participants.

5. Etalons voyageurs multiples :  
Plusieurs étalons ont circulé avec une information plus riche mais qui nécessite de prendre des précautions dans l'exploitation des résultats.
6. Normalité :  
L'hypothèse de normalité est rejetée.
7. Le résultat de mesure est un vecteur :  
La mesure de l'étalon est composée de plusieurs informations (exemple : fréquence ; longueur d'onde).
8. L'exploitation des comparaisons bilatérales

## **Un premier guide: "Proposed guidelines for the evaluation of key comparison data [6]**

Un premier guide vient d'être rédigé par un groupe mis en place en 2001 par le directeur du BIPM et constitué de 5 laboratoires nationaux de métrologie (NPL, IMG, NIST, DFM, PTB).

Il se décompose en 2 procédures, procédure A et procédure B, pour l'évaluation des données des comparaisons clés et s'articule sur 3 conditions [fig2]:

1. L'étalon sur lequel le laboratoire doit fournir un résultat durant une circulation classique (Round robin), doit prouver une bonne stabilité à long terme ainsi que durant le transport.
2. Les laboratoires nationaux de métrologie participants doivent réaliser des mesures indépendamment des autres participants lors de la comparaison.
3. Pour chaque laboratoire participant, une distribution gaussienne peut être associée au mesurande dont il a fourni au travers de ces mesures, un estimateur.

L'application de la procédure A est soumise à la validité des 3 conditions décrites plus haut et est vivement encouragée lorsque le cas se présente. Lorsque la condition N°3 n'est pas applicable aux mesures d'un ou plusieurs laboratoires mais que les autres conditions sont validées, la procédure B doit être utilisée.

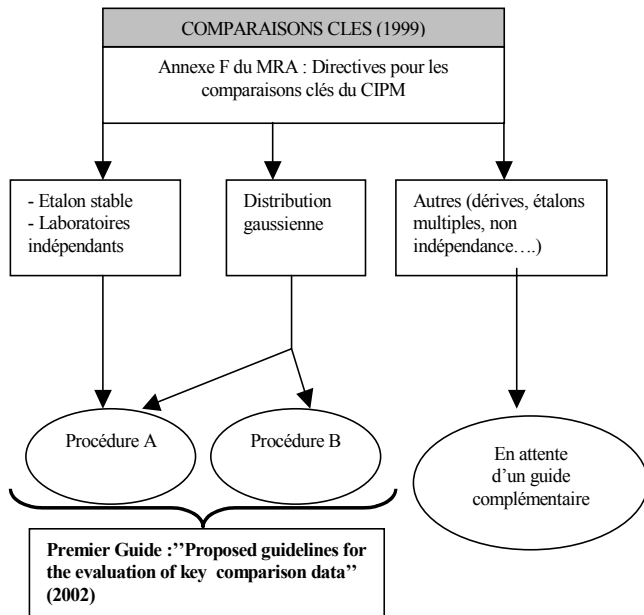


Fig.2 : Exploitation des comparaisons clés

Afin de développer les procédures A et B, le nombre de laboratoires participants est  $N$ . Chaque laboratoire fournit son résultat de mesure et l'incertitude associée, ainsi que le nombre de degré de liberté :

$$\{x_i; u(x_i); \nu_i\} \quad i=1, \dots, N$$

### Procédure A

La procédure A est basée sur l'utilisation de la moyenne pondérée, associée à une vérification de la cohérence en appliquant un test statistique d'adéquation.

- Calcul de la moyenne pondérée (weighted mean)

$$y = \frac{\sum_{i=1}^N \frac{x_i}{u^2(x_i)}}{\sum_{i=1}^N \frac{1}{u^2(x_i)}}$$

- Calcul de l'incertitude type associée à  $y$

$$\frac{1}{u^2(y)} = \frac{1}{u^2(x_1)} + \frac{1}{u^2(x_2)} + \dots + \frac{1}{u^2(x_N)}$$

- Cohérence des données : test du Khi-deux

Ce test permet de valider la condition 3, c'est à dire l'adéquation des résultats à la distribution gaussienne.

- Calcul de la statistique observée

$$\chi_{obs}^2 = \frac{(x_1 - y)^2}{u^2(x_1)} + \frac{(x_2 - y)^2}{u^2(x_2)} + \dots + \frac{(x_N - y)^2}{u^2(x_N)}$$

- Nombre de degré de liberté (ddl)

$$\nu = N - 1$$

- Le test rejette l'hypothèse de cohérence (adéquation) si :

$$P(\chi^2(\nu) > \chi_{obs}^2) < 0,05$$

- Si la cohérence est vérifiée :

- La valeur de référence de la comparaison clé  $x_{réf} = y$

- L'incertitude de la valeur de référence  $u(x_{réf}) = u(y)$

- Calcul des degrés d'équivalence :

Le degré d'équivalence de chaque étalon

$$(d_i; U(d_i)), \quad i = 1, \dots, N$$

$$d_i = x_i - x_{réf}$$

$$u^2(d_i) = u^2(x_i) - u^2(x_{réf})$$

avec,

$$U(d_i) = 2 \times u(d_i)$$

Le degré d'équivalence entre 2 laboratoires nationaux

$$(D_{ij}; U(D_{ij})) \quad i \neq j \quad i = 1, \dots, N \quad j = 1, \dots, N$$

$$D_{ij} = x_i - x_j$$

$$u^2(D_{ij}) = u^2(x_i) + u^2(x_j)$$

avec,

$$U(D_{ij}) = 2 \times u(D_{ij})$$

- Si la cohérence n'est pas vérifiée :

Identification des mesures suspectes par le test suivant au seuil de 5% :

$$|d_i| > 2 \times u(d_i)$$

S'il n'est pas possible que les laboratoires concernés modifient leurs mesures afin d'obtenir la validité de la condition N°3, l'application de la procédure B est nécessaire.

### Procédure B

Cette procédure propose une alternative à la procédure A, lorsque la normalité est rejetée, en se basant sur l'utilisation de la médiane (ou tout autre estimateur robuste plus adapté) comme valeur de référence pour l'évaluation de la comparaison clé.

$$q = (m^{(1)}, m^{(2)}, \dots, m^{(M)})$$

D'autre part, l'utilisation de la simulation via la méthode de Monte Carlo va permettre de calculer les différentes incertitudes, sur le principe de la propagation des distributions [8].

Dans ce guide, la nouveauté vis à vis des métrologues est l'intégration du principe de la propagation des distributions, alors qu'ils sont plus familiers à la propagation des incertitudes.

- Simulation numérique via Monte Carlo

L'objectif de la méthode de Monte Carlo est de reconstituer artificiellement un phénomène aléatoire. Concrètement, il s'agit de simuler un échantillon fictif de réalisations de ce phénomène à partir d'hypothèses sur les variables aléatoires. Les analyses statistiques effectuées sur un grand échantillon seront alors très proches de la réalité.

Les PC étant suffisamment puissants, le nombre de simulation peut être très grand (le guide recommande  $M = 10^6$ ). Les simulations numériques vont de  $r = 1$  à  $M$ .

Au départ, à l'instant  $r = 0$ , il y a les  $N$  mesures envoyées par les laboratoires participants auxquelles sont associées des lois de distributions.

Soit,

$$x^{(0)} = (x_1^{(0)}, \dots, x_N^{(0)})^T$$

( $T$  : transposé)

Remarque :

Il serait intéressant d'avoir plus d'informations sur ces lois de distributions. Est-ce les laboratoires participants qui doivent fournir cette information ou le laboratoire pilote qui dans son exploitation propose des lois de distributions pour chaque laboratoire ?

La simulation numérique est lancée :

Pour  $r = 1$  à  $r = M$ , on obtient une matrice  $Z_{(N, M)}$  :

$$Z = \begin{pmatrix} x_1^{(1)}, x_1^{(2)}, \dots, x_1^{(M)} \\ x_2^{(1)}, x_2^{(2)}, \dots, x_2^{(M)} \\ \vdots \\ x_N^{(1)}, x_N^{(2)}, \dots, x_N^{(M)} \end{pmatrix}$$

Les médianes :  $m^{(1)}, m^{(2)}, \dots, m^{(M)}$

Remarque :

Il peut être surprenant de ne pas voir apparaître la médiane  $m^{(0)}$ , calculée à partir des données fournies par les laboratoires. Les données qui sont exploitées par la suite sont toutes des valeurs simulées.

Calcul du vecteur médiane  $q$  à partir de la matrice  $Z$  :

On obtient ainsi la distribution empirique de la médiane, ce qui permet de calculer :

- La valeur de référence et son incertitude-type associée :

La valeur de référence est calculée à partir de la moyenne arithmétique des médianes :

$$x_{réf} = \frac{\sum_{r=1}^M m^{(r)}}{M}$$

L'incertitude associée à la valeur de référence est calculée à partir de l'écart-type des médianes :

$$u(x_{réf}) = \sqrt{\frac{\sum_{r=1}^M (m^{(r)} - x_{réf})^2}{M - 1}}$$

Pour obtenir l'intervalle de confiance à 95 %, se reporter à l'annexe B du guide [4].

- Les degrés d'équivalence :

- Le degré d'équivalence de chaque étalon

$$(d_i; U(d_i)) \quad i = 1, \dots, N$$

$$d_i = x_i - x_{réf}$$

On utilise à nouveau les résultats de la simulation via Monte Carlo au travers de la matrice  $Z_{(N, M)}$ .

Pour chaque ligne,  $i = 1, \dots, N$  :

$$(x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(M)}) - (m^{(1)}, m^{(2)}, \dots, m^{(M)})$$

Pour obtenir l'intervalle de confiance à 95 %, se reporter à l'annexe B du guide [4].

- Le degré d'équivalence entre 2 laboratoires nationaux

$$(D_{ij}; U(D_{ij})) \quad i = 1, \dots, N \quad ; \quad j = 1, \dots, N \quad (i \neq j)$$

$$D_{ij} = x_i - x_j$$

$$(x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(M)}) - (x_j^{(1)}, x_j^{(2)}, \dots, x_j^{(M)})$$

Pour obtenir l'intervalle de confiance à 95 %, se reporter à l'annexe B du guide [4].

Ce guide répond, en partie, aux souhaits des laboratoires nationaux participants aux comparaisons clés d'avoir un complément plus explicite sur la partie exploitation des résultats de l'annexe F du MRA [2]. Il est un bon début dans cette voie de rédaction de guides complémentaires.

Ce guide étant récent, il est nécessaire d'avoir un retour d'expérience afin d'évaluer son adéquation au besoin.

## **Conclusion**

Depuis la signature du document officiel, le MRA qui définit le cadre technique de la reconnaissance mutuelle des laboratoires nationaux de métrologie, d'innombrables comparaisons clés ont été exploitées avec plus ou moins de difficultés.

Très rapidement, la nécessité d'un guide est apparu, pour répondre aux interrogations des laboratoires participants.

Ce premier guide permet de répondre de façon relativement simple aux problèmes de base et amorce efficacement le travail vers d'autres guides complémentaires.

[5] Proposed guidelines for the evaluation of key comparison data. An introduction, M.G.Cox, National Physical Laboratory, UK (2002)

[6] Proposed guidelines for the evaluation of key comparison data, W.Bich, M.G.Cox, W.T. Estler, L.Nielsen, W.Woegger, BIPM Director's Advisory Group on Uncertainties (2002).

[7] Guide pour l'expression de l'incertitude de mesure, 2<sup>nd</sup> édition, Genève, Organisation Internationale de Normalisation, 1995.

[8] Guide to the Expression of Uncertainty in Measurement, Supplement 1: Numerical Methods for Propagating Probability Distributions, Technical Report, Joint Committee for Guides in Metrology, 2002. Draft.

## **Référence**

[1] Reconnaissance mutuelle des étalons nationaux de mesure et des certificats d'étalonnage et de mesurage émis par les laboratoires nationaux de métrologie, Comité International des poids et mesures (1999). (MRA)

[2] Directives pour les comparaisons clés du CIPM (Annexe F du MRA), Comité International des poids et mesures (1999).

[3] Report on the workshop held at NPL on 11 and 12 November 1999: Statistical Analysis of Interlaboratory Comparison, Maurice Cox NPL (1999)

[4] Drift problems with the circulated standard during comparisons (Workshop data analysis of Interlaboratory Comparisons Berlin 5/12/2002), Michèle Désenfant , Laboratoire National D'Essais (2002).