

DE LA VALIDATION DES METHODES D'ANALYSE A L'EVALUATION DE L'INCERTITUDE DES RESULTATS DE MESURE

Michèle Désenfant – Marc Priel – Cédric Rivier
Laboratoire National d'Essais BNM-LNE
1, rue Gaston Boissier 75724 Paris Cedex 15

Résumé

La norme ISO/CEI 17025 demande de valider les méthodes lorsque les laboratoires utilisent des méthodes non normalisées ou hors du domaine d'application de la norme. Ce référentiel demande également d'évaluer l'incertitude des résultats fournis. Les laboratoires d'analyses chimiques ont une longue expérience et une longue tradition de validation de leurs méthodes d'analyse mais éprouvent parfois des difficultés pour évaluer l'incertitude de leurs résultats.

Dans cet article, les auteurs souhaitent faire le point sur cette notion de validation.

Les performances d'une méthode peuvent s'exprimer à l'aide de caractéristiques telles que : la sélectivité, la spécificité, la justesse, la linéarité, la répétabilité, la reproductibilité, la robustesse, les limites de détection.... Ces caractéristiques s'évaluent grâce à des travaux expérimentaux, réalisés soit en interne, soit de manière collective en impliquant plusieurs laboratoires.

La connaissance des caractéristiques des méthodes est une information tout à fait pertinente pour l'évaluation de l'incertitude des résultats d'analyse et les auteurs présentent leur approche pour assurer le passage des données de caractérisation à l'évaluation de l'incertitude du résultat.

Abstract

The ISO/IEC 17025 standard requires to validate the testing and analytical methods when they are not standardized or when standardized methods are used outside of their scope. A measurement uncertainty must also be associated with the results. Analytical laboratories have a long tradition in the process of methods validation, but some difficulties occur for the evaluation of their uncertainties.

In this article, the authors want to clarify the concept of method validation and its relationship with the uncertainty of measurement.

The performances of a method may be described with several characteristics as : selectivity, specificity, trueness, linearity, repeatability, reproducibility, ruggedness, limit of detection... These characteristics are evaluated through experiments. The experimental works can be managed either in an internal manner in the laboratory or in an

external manner with collaborative works implying several laboratories.

The knowledge of the method's characteristics is a pertinent information for the uncertainty evaluation. The authors wish to present their views and approach in order to establish a link between the characterization data and the measurement uncertainty.

Validation ou caractérisation d'une méthode ?

La validation se définit comme la « confirmation par examen et l'apport de preuves objectives du fait que les prescriptions particulières en vue d'une utilisation prévue déterminée sont remplies » ISO/CEI 17025 § 5.4.5.1

Il faut donc démontrer que la méthode mise en œuvre par le laboratoire est apte à l'emploi prévu (besoin du client).

Il faut déterminer les besoins et attentes du client, et peut-être ne pas s'arrêter à ce qu'il spécifie. Comme la norme l'indique, il faut « élucider la demande » et considérer ce dont le client aura besoin pour utiliser l'information qui lui sera remise, examiner les exigences réglementaires et légales, et rechercher les informations nécessaires à la mise en œuvre de la méthode.

Le laboratoire doit donc traduire les besoins du client en terme de performance pour les comparer ensuite avec les caractéristiques estimées de la méthode candidate.

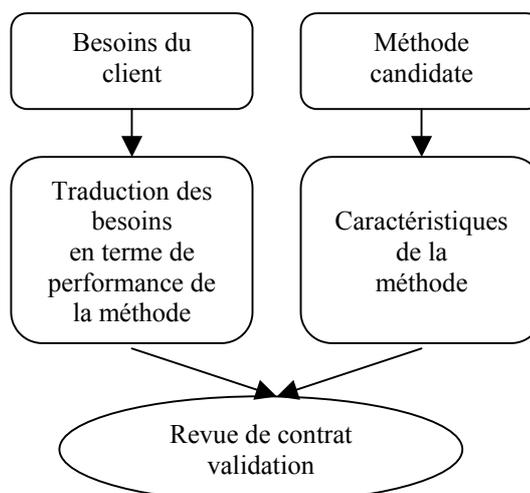


Fig.1 Relation entre caractérisation et validation d'une méthode

Comment décrire les performances d'une méthode d'analyse ?

Depuis de très nombreuses années, l'approche des analystes a été de caractériser les performances des méthodes d'analyse en évaluant au travers de travaux expérimentaux des caractéristiques.

Ces caractéristiques sont décrites dans différents ouvrages tels [1] ou documents émanant de l'association européenne Eurachem [2]

Voici la liste des caractéristiques les plus souvent citées :

- Sélectivité, spécificité
- Répétabilité
- Reproductibilité
- Linéarité
- Sensibilité
- Capacité de détection
- Robustesse
- Justesse

Comment évaluer ces caractéristiques ?

Sélectivité, spécificité

Cette caractéristique est étudiée en premier, en effet elle permet de vérifier que le signal mesuré correspond bien à l'analyte recherché.

La sélectivité peut se définir comme l'aptitude d'un élément de la méthode d'analyse (appareil de mesure, milieu de culture, etc....) à discerner un analyte donné dans un mélange complexe.

La spécificité est la propriété d'une méthode d'analyse de convenir exclusivement à la caractéristique ou à l'analyte, avec la garantie que le résultat de l'analyse ne provient que de l'analyte.

Très souvent la spécificité se fonde sur une absence d'interférence. [3]

Lorsque l'on ne peut pas évaluer l'influence de chaque constituant de la matrice, une technique commode d'emploi, permettant d'évaluer l'influence de certaines interférences, est la méthode des ajouts dosés. Elle consiste à ajouter dans l'échantillon avant, pendant ou après sa préparation des quantités connues de l'analyte. On examine alors si la réponse de l'instrument de mesure correspond au signal initial augmenté du signal correspondant à la quantité ajoutée.

Les effets liés à la sélectivité et à la spécificité sont évalués globalement lors de l'évaluation de la justesse.

Répétabilité et Reproductibilité (interne)

Pour définir la répétabilité, il faut déjà définir la fidélité qui est l'étroitesse de l'accord entre des résultats d'essais. La répétabilité est la fidélité sous des conditions de répétabilité ; c'est à dire des conditions où des résultats

d'essais indépendants sont obtenus par la même méthode sur des individus d'essais identiques dans le même laboratoire, par le même opérateur, utilisant le même équipement et pendant un court intervalle de temps.

La répétabilité représente la valeur minimale de la fidélité de la méthode lorsqu'elle est utilisée par le laboratoire, elle est souvent quantifiée par un écart-type noté S_r .

La répétabilité est évaluée à k niveaux de concentration (réalisés par des solutions étalons ou des matériaux de référence), en répétant n mesures. Après vérification de l'homogénéité des variances (test de Cochran) on évalue l'écart-type de répétabilité, la norme expérimentale [4] « Protocole d'évaluation d'une méthode alternative d'analyses physico-chimiques quantitatives par rapport à une méthode de référence », propose un nombre d'échantillons d'au moins $k=10$ et de faire au moins $n=2$ répétitions.

Pour la reproductibilité, on note qu'il coexiste deux définitions, l'une proposée par le Vocabulaire International des Termes Fondamentaux et Généraux de Métrologie (NF X 07-001) et l'autre par la norme NF ISO 3534-1 (1993). Le traitement statistique utilisé est celui décrit par la norme NF ISO 5725 [7], le facteur laboratoire S_L estimé avec p laboratoires, étant alors remplacé par le facteur étudié à p modalités (par exemple le temps). On note qu'il s'agit d'une reproductibilité qui est interne au laboratoire. L'écart-type de reproductibilité s'exprime par l'expression suivante :

$$S_R = \sqrt{S_L^2 + S_r^2}$$

Quand le nombre de répétitions est le même pour chaque modalité, cette expression se traduit par la formule suivante :

$$S_R = \sqrt{\frac{1}{p-1} \sum_{j=1}^p (\bar{x}_i - \bar{\bar{x}})^2 + S_r^2 \left(1 - \frac{1}{n}\right)}$$

On constate dans cette expression qu'un terme provient de la répétabilité et que l'autre exprime la dispersion des résultats aux p modalités (\bar{x}_i) par rapport à la moyenne générale ($\bar{\bar{x}}$).

Linéarité, sensibilité

La linéarité et la sensibilité sont deux caractéristiques qui sont évaluées simultanément. Des expérimentations proches de ce qui a été décrit pour l'évaluation de l'écart-type de répétabilité peuvent être mises en œuvre. Le guide Eurachem préconise $k=6$ niveaux et $n=3$ répétitions, la norme XP T 90-210 propose $k=5$ niveaux et $n=5$ répétitions.

Le test statistique consiste à comparer les écarts entre les moyennes des points expérimentaux et la droite estimée, à l'erreur résiduelle (c'est-à-dire la répétabilité) par une analyse de variance. L'évaluation de cette caractéristique,

permet de préciser le domaine d'utilisation de la méthode en fonction de « l'erreur de linéarité » que l'on est prêt à accepter.

Capacité de détection

Les caractéristiques, limite de détection et limite de quantification, dont les définitions ne font pas encore l'objet d'un consensus international servent essentiellement à définir le domaine d'utilisation de la méthode d'analyse. Elles n'ont pas d'incidence directe sur l'évaluation de l'incertitude. Néanmoins, elles ont une importance sur la façon de rapporter les résultats, notamment lorsque les valeurs estimées du mesurande sont inférieures à la limite de détection.

Robustesse

La norme FD V 01-000 [3] propose pour définition : l'aptitude d'une méthode d'analyse à fournir de faibles variations du résultat lorsqu'elle est soumise à des modifications contrôlées des conditions d'application (exemple : température ambiante, lumière, pression atmosphérique, humidité, réactifs, appareillage, etc....). La technique de la planification d'expérience permet d'analyser l'influence des différents facteurs qui sont supposés avoir une action sur la variable de réponse (le résultat de l'analyse). Les facteurs peuvent être quantitatifs (on fixe les différents niveaux lors de la préparation de l'expérience) ou qualitatifs (par exemple l'opérateur et on parle alors de modalité).

On fait alors l'hypothèse que le phénomène peut être décrit sous la forme d'un polynôme où la variable de réponse (le résultat de l'analyse) s'exprime en fonction des facteurs susceptibles d'agir sur le processus d'analyse.

$$y = a_1x_1 + a_2x_2 + \dots + a_jx_j + \dots + a_ix_i$$

Le modèle présenté ici ne tient pas compte des interactions entre facteurs. Les méthodes statistiques de traitement et d'analyse des résultats d'essais permettent un calcul des coefficients du polynôme. L'analyse de variance permet de vérifier que les effets définis comme influents pour la réponse étudiée sont réellement associés au facteur étudié et qu'ils ne sont pas les résultats d'une variabilité naturelle du phénomène étudié. On peut résumer le test statistique en disant que c'est un test d'hypothèse dans lequel les variations induites par un facteur A sur la réponse, notée Var(A) sont comparées aux variations aléatoires de la réponse (Var résiduelle ou variance de répétabilité).

En fonction des résultats de ces tests, si un facteur est déclaré comme influent, on peut soit modifier ou bloquer les niveaux ou les modalités pour rendre la méthode insensible, soit inclure dans l'incertitude l'effet du ou des facteurs influents. Il est certain que l'étude de la robustesse de la méthode se situe entre les phases de développement et de caractérisation de la méthode.

Justesse

Si l'on adopte la définition de la norme NF ISO 3534, la justesse est définie comme l'écart entre la valeur moyenne obtenue à partir d'une large série de résultats d'essais et une valeur de référence acceptée.

Pour évaluer la justesse, il faut donc disposer de références. C'est au travers de cette caractéristique, que l'on assure la traçabilité aux unités du Système International (SI). Les valeurs de référence acceptées peuvent provenir : d'un matériau de référence, d'une valeur fournie par une méthode de référence, d'une valeur provenant d'un essai inter-laboratoire (essai d'aptitude) et éventuellement de l'utilisation de la méthode des ajouts dosés. Bien évidemment la « qualité » de la valeur de référence est primordiale.

Le principe du test statistique consiste à calculer un écart normalisé E_N

$$E_N = \frac{|\bar{x}_i - x_{Ref}|}{\sqrt{u_i^2 + u_{Ref}^2}}$$

Si E_N est inférieur à 2, on considère que l'écart n'est pas significatif. Il est aussi possible de comparer l'écart $|\bar{x}_i - x_{Ref}|$ à une tolérance L .

Si $|\bar{x}_i - x_{Ref}| \leq L$, alors l'erreur de justesse est considérée comme non significative.

L'attitude d'un laboratoire est généralement d'essayer d'améliorer la méthode, jusqu'à ce que cet écart devienne faible par rapport aux incertitudes sur la mesure et sur la valeur de référence. En analyse chimique, il est rare de faire des corrections.

Caractéristiques de la méthode et composantes d'incertitude

Toutes les caractéristiques qui servent à décrire les performances d'une méthode d'analyse n'ont pas une incidence directe sur l'évaluation de l'incertitude du résultat d'analyse. Le tableau ci-dessous résume l'impact des différentes caractéristiques sur l'évaluation de l'incertitude.

Caractéristique	Impact sur l'incertitude
Sélectivité, spécificité	Les effets liés à la sélectivité et à la spécificité sont évalués globalement lors de l'évaluation de la composante d'incertitude liée à la justesse.
Répétabilité, Reproductibilité	Ces variances figurent parmi les éléments les plus significatifs pour l'évaluation de l'incertitude
Linéarité, Sensibilité	1 ^{er} Cas : Si le test statistique montre que le modèle est linéaire, pas d'incidence sur l'incertitude 2 ^{ème} Cas : Si le test montre que le modèle n'est pas linéaire, on peut

	alors soit réduire le domaine de mesure, soit ajouter une composante (de non linéarité) dans l'évaluation de l'incertitude Cette composante pouvant être l'écart maximal au modèle
Capacité de détection	Cette caractéristique n'a pas d'incidence sur l'évaluation de l'incertitude, elle sert pour définir le domaine d'utilisation de la méthode.
Robustesse	La technique des plans d'expérience permet de quantifier l'effet d'un facteur, pour un facteur dont le domaine de variation est limité à $\pm a$, alors l'incertitude correspondante est $u_a = c \frac{a}{\sqrt{3}}$ ou c est le coefficient de sensibilité de la réponse aux variations du facteur a .
Justesse	En analyse chimique, d'une manière générale, on n'applique pas de corrections de justesse. L'objectif lors de la mise au point de la méthode est de vérifier qu'il n'y a pas d'erreur systématique significative. Comme nous l'avons présenté au paragraphe traitant de la justesse : Si $E_n \leq 2 \quad u_{Justesse} = u_{Ref}$ Ou bien si $ \bar{x}_i - x_{Ref} \leq L \quad u_{Justesse} = \frac{L}{\sqrt{3}}$

Quelles stratégies pour l'évaluation de l'incertitude et la prise en compte des caractéristiques de la méthode ?

Le laboratoire d'analyse face à la question de la validation des méthodes et à l'estimation de l'incertitude des résultats se trouve face à deux situations possibles : soit l'approche intra-laboratoire avec une démarche analytique qui lui permet d'évaluer l'incertitude du résultat de mesure en combinant toutes les sources d'incertitude, soit une approche collective (inter-laboratoire), si la méthode de mesure est aussi pratiquée par d'autres collèges.

L'approche intra-laboratoire

Cette approche se décompose en plusieurs possibilités.

Si un modèle mathématique complet du processus de mesure est disponible, une évaluation analytique de type

GUM est possible. On rappelle que la procédure décrite dans le GUM comporte 4 étapes principales :

- Identification des facteurs d'influence qui ont une action sur le résultat de l'analyse et prise en compte de ces facteurs pour la construction du modèle mathématique
- Evaluation de l'incertitude des différentes grandeurs d'entrée du modèle (facteurs d'influence) à l'aide de méthodes de type A et de méthodes de type B
- Application de la loi de propagation de l'incertitude pour évaluer l'incertitude sur le résultat en fonction de l'incertitude des différentes grandeurs d'entrée du modèle.
- Evaluation d'une incertitude élargie.

Note : le laboratoire peut avec profit utiliser toutes les informations dont il dispose et notamment celles citées ci-après car elles correspondent bien souvent à l'effet cumulé de plusieurs grandeurs d'entrée du modèle.

Si le modèle n'est pas disponible, ou incomplet, ou que les efforts pour l'établir sont trop importants par rapport aux objectifs, le laboratoire doit utiliser toutes les informations disponibles :

- Cartes de contrôle pour évaluer la répétabilité ou la reproductibilité dans le temps
- Organiser des essais en faisant varier le maximum de conditions et de facteurs ayant une influence a priori sur le résultat, observer la dispersion des résultats et calculer la variance correspondante

Le laboratoire doit ensuite combiner la variance obtenue avec les autres causes d'incertitude (elles-mêmes exprimées sous forme de variance) qui n'ont pas pu s'exprimer au cours des répétitions (par exemple la justesse).

Le modèle statistique peut s'écrire :

$$y = m + C_{Jus} + C_{Lin} + \sum_i c_i x_i + e$$

où

y : résultat de la mesure

m : valeur vraie

C_{Jus} : correction de justesse de la méthode

C_{Lin} : correction de non-linéarité

$c_i x_i$: termes correctifs d'effets tels que la robustesse

e : erreur résiduelle

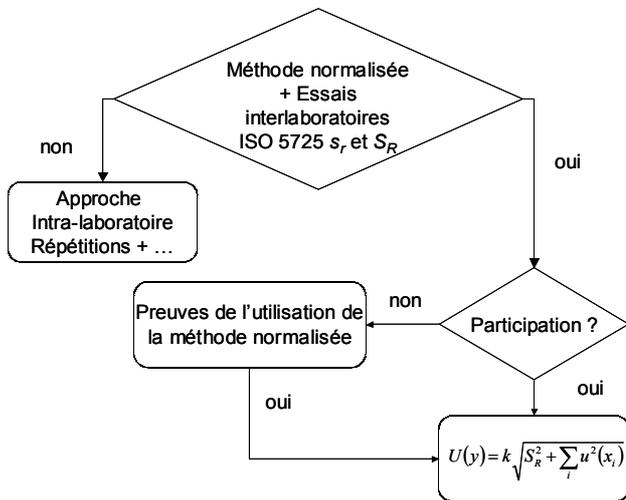
La variance de y s'écrit donc :

$$u^2(y) = u^2(C_{Jus}) + u^2(C_{Lin}) + \sum_i c_i^2 u^2(x_i) + S_r^2$$

Ces variances sont évaluées lors de la caractérisation de la méthode développée au paragraphe précédent. Selon les données disponibles, ces caractéristiques peuvent être évaluées une à une ou de manière groupée.

L'approche inter-laboratoire

Dans les approches collectives, on utilise les résultats d'essais inter-laboratoires. En fait il existe plusieurs types d'essais inter-laboratoires : les essais qui contribuent à l'évaluation d'un laboratoire (essais d'aptitude), qui sont organisés en se référant au guide ISO 43-1 [6] et les essais qui servent à déterminer les performances d'une méthode NF ISO 5725 [7].



Comme le propose le document ISO/DTS 21748, à la reproductibilité qui est évaluée collectivement via l'essai inter-laboratoires, on peut ajouter un terme de justesse (noté dans ce document $\hat{\mu}$) plus un terme qui comprend toutes les autres contributions à l'incertitude qui n'ont pu être mises en oeuvre par l'essai inter-laboratoire.

Le modèle statistique s'écrit :

$$y = m + \delta + B + \sum_i c_i x_i + e$$

où

y : résultat de mesure

m : valeur vraie

δ : justesse de la méthode

B : effet laboratoire

$c_i x_i$: termes correctifs d'effets non pris en compte lors de l'essai interlaboratoire

e : erreur résiduelle

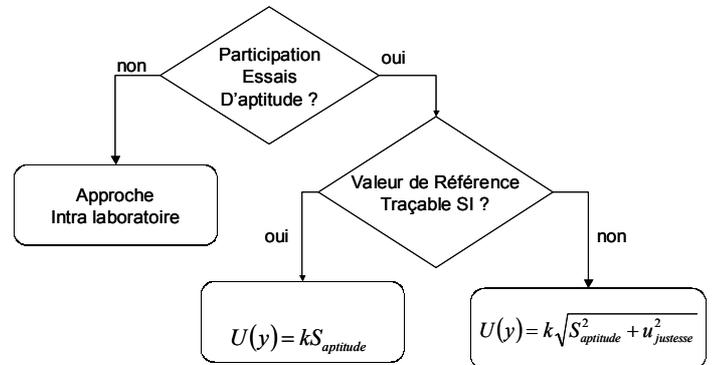
En introduisant la variance de reproductibilité

$$S_R^2 = u^2(B) + u^2(e)$$

La variance de y s'écrit alors :

$$u^2(y) = u^2(\hat{\mu}) + s_R^2 + \sum_i c_i^2 u^2(x_i)$$

L'utilisation de résultats provenant d'essais d'aptitude mérite que l'on se préoccupe de la traçabilité de la valeur de référence. En effet dans de très nombreux essais d'aptitude, la valeur de référence par rapport à laquelle sont évalués les écarts des laboratoires est bien souvent la moyenne (arithmétique ou pondérée) des réponses des laboratoires et la traçabilité de cette valeur aux unités du Système International (SI) n'est pas toujours prouvée.



Pour évaluer l'incertitude, il faut ajouter à l'écart-type provenant de la dispersion des résultats des laboratoires, une composante d'incertitude liée à la justesse de la méthode.

Conclusion

Dans les approches intra-laboratoire et inter-laboratoire, le principe pour évaluer l'incertitude est le même : combiner les sources d'incertitude les plus significatives en faisant la somme de leurs variances.

En fonction des données disponibles, qui sont différentes suivant l'approche choisie, on compose, dans la démarche intra-laboratoire : justesse, répétabilité, non-linéarité, robustesse et d'autres facteurs si nécessaire ; dans l'approche inter-laboratoire : au minimum justesse et reproductibilité inter-laboratoire.

Références

- [1] Max Feinberg, La validation des méthodes d'analyse, Masson, (1996)
- [2] Eurachem Guide : The fitness for Purpose of Analytical Methods – A laboratory guide to method validation and related topics (1998)
- [3] FD V 01-000 (décembre 1999) Analyse des produits agricoles et alimentaires – Terminologie
- [4] X PT 90-210 (1999) Qualité de l'eau - Protocole d'évaluation d'une méthode alternative d'analyses physico-chimique quantitative par rapport à une méthode de référence
- [5] ISO/DTS 21748 Application of statistics - Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation

[6] ISO/CEI Guide 43-1 Essais d'aptitude des laboratoires par intercomparaison – partie 1 Développement et mise en oeuvre de systèmes d'essais d'aptitude

[7] NF ISO 5725 Application de la statistique -Exactitude (justesse et fidélité) des résultats et méthodes de mesure