



HAL
open science

Validation des méthodes d'analyse quantitative par le profil d'exactitude

Max Feinberg, Michel Laurentie, Annie Guerin, Marie Huyez-Levrat

► To cite this version:

Max Feinberg, Michel Laurentie, Annie Guerin, Marie Huyez-Levrat. Validation des méthodes d'analyse quantitative par le profil d'exactitude. Cahier des Techniques de l'INRA, 2010, N° Spécial: Validation des méthodes. hal-04792231

HAL Id: hal-04792231

<https://hal.inrae.fr/hal-04792231v1>

Submitted on 20 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

Le Cahier des Techniques de l'Inra

Validation des méthodes
d'analyse quantitative
par le profil d'exactitude



Délégation au Partenariat avec les Entreprises

Institut national de la recherche agronomique - 147 rue de l'Université - F-75338 Paris Cedex 07

Validation des méthodes d'analyse quantitative par le profil d'exactitude

Numéro spécial 2010

*Collectif*¹

Responsables scientifiques : *Max Feinberg*² et *Michel Laurentie*³
Animatrice : *Annie Guerin*⁴ - Coordonnatrice : *Marie Huyez-Levrat*⁵

Sommaire

Introduction	p. 03
Témoignages sur les écoles techniques	p. 05
Chapitre I : Validation des méthodes	p. 11
Chapitre II : Application dans les laboratoires à l'Inra	p. 71
Conclusion	p. 135
Index des auteurs	p. 137
Index des articles	p. 139

Cet ouvrage est téléchargeable à partir du site de l'Inra

https://intranet.inra.fr/cahier_des_techniques

¹ Le *Cahier des Techniques de l'Inra* est une publication destinée à échanger des méthodes et des savoir-faire. Ce n'est pas une revue à Comité de lecture. Les auteurs sont seuls responsables de leurs textes.

² UR1204 Met@risk - Méthodologie d'analyse de risque alimentaire)- INRA – F- 75231 Paris cedex 05.

☎ 01 44 08 16 52 ✉ max.feinberg@paris.inra.fr

³ Unité de Pharmacocinétique-Pharmacodynamie - AFSSA – F-35302 Fougères cedex, m.laurentie@afssa.fr

⁴ US0010 LAS, INRA – F- 62000 Arras – www.lille.inra.fr/las

⁵ UAR1185- Délégation au partenariat avec les entreprises DV/DPE - INRA - F-75338 Paris Cedex 07

☎ 01 42 75 91 65 ✉ marie.huyez@paris.inra.fr



Introduction

Max Feinberg

Pour celle ou celui qui a entre les mains ce numéro spécial du *Cahier des Techniques de l'Inra*, deux questions peuvent se poser :

- *Pourquoi devrais-je valider les méthodes d'analyse que j'utilise ?*
- *Si je me décide, comment dois-je m'y prendre ?*

Réponse à la question : Pourquoi valider ?

Depuis toujours les analystes « valident » leurs méthodes en effectuant des séries de mesures plus ou moins bien organisées pour essayer de démontrer qu'elles conviennent à leurs objectifs. Bien sûr, celles et ceux qui sont dans des laboratoires accrédités savent que c'est insuffisant et qu'une validation doit être conduite selon une procédure expérimentale plus stricte où on parle de justesse, de reproductibilité, etc. La validation est devenue un élément majeur pour la démonstration de la compétence de ces laboratoires accrédités.

Cela signifie-t-il que lorsqu'on n'est pas accrédité, on n'a pas à se plier à cette démarche ? Si la réponse pouvait être un « oui » timide il y a encore deux ou trois ans, ce n'est plus le cas aujourd'hui. D'une part, la notion de qualité a gagné tous les niveaux de l'activité technique et scientifique. Ainsi, les industriels avec lesquels les unités INRA peuvent être amenées à collaborer ont des systèmes d'assurance qualité qui, par ricochet, les obligent à ne travailler qu'avec des contractants qui ont eux-mêmes un système qualité. D'autre part, les divers exemples présentés dans ce numéro illustrent bien les raisons pour lesquelles nombre de nos collègues ont été amenés à valider de façon formalisée leurs méthodes.

L'approche proposée s'appelle le « profil d'exactitude ». C'est une approche relativement nouvelle qui, cependant, est déjà « standardisée » dans la norme Afnor NF V03-110 qui vient d'être publiée. Elle a l'avantage de proposer une méthode d'interprétation graphique très simple et visuelle qui ne s'embarrasse pas de tests statistiques toujours délicats à décrypter. Son objectif est de servir les analystes plutôt que de les transformer en statisticiens... Elle est universelle et s'applique à toutes les méthodes de mesure, comme le prouve une littérature maintenant abondante.

Réponse à la question : Comment démarrer ?

Pour démarrer, vous avez toute une série de possibilités :

- télécharger et lire les documents disponibles sur l'Intranet de la Délégation à la qualité :
https://intranet.inra.fr/mission_qualite/acces_thematiques/formations_et_ecoles/ecole_des_techniques_validation_methodes_2008 sous le lien « Le recueil des présentations de l'école 2008 » (vous devez avoir un login LDAP pour y accéder) ;
- pour faire les calculs, télécharger des fichiers modèles Excel® (qui marchent aussi sous Open Office®) à partir de :
http://www.paris.inra.fr/metarisk/downloads/software_programs/excel_templates ;
- vous renseigner auprès de nos collègues qui ont déjà utilisé les profils d'exactitude et dont vous avez les adresses électroniques ;
- interroger la Délégation à la qualité, comme l'indique dans son article François Jacquet, le délégué à la qualité de l'Inra.

Notre but est de vous convaincre que c'est « beaucoup plus simple qu'on ne veut le faire croire ». Merci de nous aider à atteindre cet objectif.

Témoignages sur les écoles techniques de validation des méthodes d'analyse

*Max Feinberg*¹

À deux reprises, le service de la Formation permanente nationale (FPN) de l'Inra, sous l'impulsion de Marie-Andrée Piédallu, a organisé une « École des techniques » sur la validation des méthodes d'analyses quantitatives. La première a eu lieu du 3 au 5 décembre 2007 près de Strasbourg et la seconde du 10 au 13 Juin 2008 à Batz-sur-Mer. En ces deux occasions, l'assistance a été nombreuse, jusqu'à 60 participants, et assidue. De plus, le « taux de satisfaction » des participants a été très... satisfaisant pour les conférenciers ! Ce qui les a amenés à préparer ce numéro spécial du *Cahier des techniques de l'Inra* pour que celles et ceux qui n'ont pas pu participer puissent se faire une opinion sur cette question fondamentale pour l'Inra qu'est la validation des méthodes d'analyse.

En outre, il nous a semblé intéressant d'illustrer, de façon moins aride qu'un tableau chiffré², ce que j'ai appelé le taux de satisfaction des participants et les bénéfices qu'ils ont pu retirer de ces deux écoles. Vous trouverez donc ci-après une compilation de témoignages classés selon le numéro d'unité, émanant des

1. UR1204 Met@risk - Méthodologie d'analyse de risque alimentaire – INRA - F-75231 Paris cedex 05.
☎ 01 44 08 16 52 ✉ max.feinberg@paris.inra.fr

2 Cependant, le lecteur intéressé peut retrouver ces statistiques pour l'école de 2008 en téléchargeant le fichier https://intranet.inra.fr/mission_qualite/content/download/3881/36715/version/1/file/Synth%E8se+%E9valuation+ETVM+physicochimiques+2008.pdf

personnes qui ont accepté de faire part de leurs impressions. Je voudrais les en remercier car il est toujours un peu difficile de savoir si une formation a servi à quelque chose. De plus, elles m'évitent d'avoir à décliner une liste d'arguments en faveur de la validation des méthodes ; tout y est, à travers la diversité des exemples pratiques qui y sont décrits et des commentaires effectués.

De fait, s'il fallait trouver une autre preuve de satisfaction, plusieurs études de validation ont été réalisées (et réussies) à l'issue de ces deux écoles. Certaines font l'objet d'articles dans ce numéro spécial, d'autres ont été soumises à des revues internationales en vue d'une publication, enfin quelques-unes dont j'ai pu être informé ne sont pas publiées mais rendent déjà des services dans les unités.

Rappelons que l'ensemble des présentations et fichiers modèles Excel® sont sur le site :

https://intranet.inra.fr/mission_qualite/acces_thematiques/formations_et_ecoles/ecole_des_tec_hniques_validation_methodes_2008/recueil_presentations_edt_validation_methodes_2008



UR 50 - LBE - Laboratoire de biotechnologie de l'environnement – INRA – F-11100 Narbonne

Voici la réponse groupée que nous souhaitons faire au nom du LBE à Narbonne... Sans afficher de nom précis car la démarche qualité telle que nous l'entendons s'inscrit dans notre collectif. Nous partageons cette vision et nous aimerions la renforcer à l'avenir (d'où la demande ci dessous) sans pour autant pénaliser nos forces vives liées à la recherche, l'innovation et le « fourmillement » d'idées et de concepts qui ne sont pas forcément formalisés dans des procédures établies et validées... Courant 2008, une personne du laboratoire a suivi cette formation, son avis est le suivant :

« Cette formation est de bon niveau et donne de façon simple et claire les outils et la méthodologie à mettre en place pour conduire la validation de méthodes ».

Cela l'a motivée à mettre en place ces outils et cette méthodologie sur les instruments de mesure dont elle a la charge, à faire partager les connaissances acquises au cours d'une présentation synthétique et à sensibiliser ses collègues à cette pratique nécessaire pour la fiabilisation des résultats de recherche. Cette pratique est d'ailleurs appliquée dans l'unité mais sans réellement mettre derrière le vocabulaire adéquat, le formalisme qualité et la systématisation de ces activités. Une demande a été formulée quant à la systématisation de ces formations sur site en local car tout comme la formation dédiée aux incertitudes de mesure, nous devrions tous et, tous corps confondus, nous sentir concernés et investis par la validation des méthodes et la gestion des incertitudes de mesure, garantie de la fiabilité des résultats de recherche.

UR 251 - PESSAC - Physico-chimie et écotoxicologie des sols d'agrosystèmes contaminés – INRA – F-78000 Versailles

Christian Mougin, Directeur de l'unité

Les activités de recherche s'appuient sur l'interprétation de résultats expérimentaux. Il est indispensable que les scientifiques bénéficient de résultats dignes de confiance, issus de mesures et d'analyses qui doivent être justes et fiables. L'exercice de validation des méthodes physicochimiques d'analyse quantitative relève donc des mêmes préoccupations que la mise en place de l'AQR (assurance qualité en recherche) dans les unités. Il vise pour un collectif à s'approprier des concepts, des bases théoriques et des outils et à mettre en place des modes de pensée et de fonctionnement communs et partagés. Si la validation des méthodes me paraît profondément ancrée dans certaines disciplines comme par exemple la chimie analytique et concerne, à mon sens, une population encore trop réduite, elle doit être étendue à toutes les disciplines, mettant en œuvre des mesures, d'une unité de recherche. Ainsi, dans des unités fonctionnant sur la base de plateaux techniques disciplinaires, la validation est un moyen d'interaction fort entre les agents, et de mise en place d'une homogénéité de fonctionnement. Par exemple, la communauté des biologistes devrait rapidement intégrer les bases théoriques de la validation des méthodes, et effectuer la transition de méthodes optimisées à des méthodes validées, de l'empirique au raisonné. La validation doit aussi permettre de définir le cadre et les limites d'utilisation des méthodes.

En conclusion, la maîtrise des méthodes relève de la valorisation des « cœurs de métiers » des ITA, avec des retombées pour l'ensemble des agents : en effet, les données une fois validées sont ensuite destinées à répondre aux questions scientifiques soulevées par les chercheurs. Par rapport à nos activités, c'est une indispensable source d'échange et de partage, un outil d'organisation des activités et une valorisation des métiers

UE 999 – Unité expérimentale de Pech-Rouge- INRA- F-11430 Gruissan

Jean-Louis Escudier, Directeur de l'unité

L'Inra dispose d'un unique réseau d'unités expérimentales, lieux pérennes d'observations de multi données dans leurs contextes. En faire l'acquisition avec les méthodes adaptées, les valider, les enregistrer, les conserver, les rendre accessibles durablement est un des socles de ce que l'on construit pour demain. Ceux qui nous succéderont dans ces unités utiliseront certaines de ces données pour des objets de recherche non connus à ce jour. Nous avons pour eux aussi un devoir de rigueur. La validation des méthodes est un préalable

L'unité expérimentale de Pech Rouge collecte les données disponibles depuis la création du domaine, en 1956, à ce jour en termes de relevés météo, d'écophysiologie de la vigne, de qualité du raisin et du vin. C'est un exercice difficile au niveau de l'accès aux données anciennes, pour les saisir dès lors que les méthodes sont validées.

UMR 1091 – EGC - Environnement et grandes cultures - INRA- F-8850 Thiverval-Grignon

Enrique Barriuso, Directeur de l'unité

Les moyens analytiques pour la quantification des contaminants organiques de l'environnement (sols, eaux, air) sont de plus en plus performants. Leur automatisation et la simplification des interfaces de contrôle et de traitement permettent de concevoir l'acquisition de données à des flux très importants et avec des précisions de plus en plus élevées. C'est le cas, par exemple, pour l'analyse chromatographique de pesticides au niveau des traces et d'ultra-traces. Face à cette réalité, des angoisses perpétuelles des scientifiques se manifestent : Que valent ces données ? Reflètent-elles une réalité ? Pour des conditions expérimentales données, allons-nous trouver les mêmes résultats ? Quelle est la contribution analytique à la variabilité de nos répétitions ? Comment allons-nous certifier la qualité de données ? C'est dans ce contexte que la validation des méthodes vient à réduire nos angoisses, introduisant de la rigueur dans nos activités analytiques et apportant une démarche permettant de préciser et de faire valoir vis-à-vis de nos partenaires la pertinence de nos méthodes analytiques, méthodes développées ou adaptées de méthodes existantes, la qualité de nos données, leur justesse, leur reproductibilité, etc.

UR1119 URGSE - Géochimie des sols et des eaux- INRA -F-13545 Aix-en-Provence

Fabienne Trolard, Directrice de l'unité

Dans le domaine de la géochimie des sols et des eaux, nous sommes absolument convaincus et depuis longtemps de la nécessité de mettre en place des méthodes robustes de validation de méthodes analytiques. En effet, la difficulté principale que nous rencontrons, en particulier pour les analyses des eaux, est que nous ignorons les niveaux de concentrations des anions et cations qui peuvent être très variables avec la saison. Seules des mesures justes et fidèles constituent des acquis analytiques fiables dans le temps et permettent de développer des modèles géochimiques. Grâce à ces deux de formation écoles nous avons pu initier dans l'unité et mener jusqu'au bout, la validation des mesures des anions dans les eaux par chromatographie ionique. Ceci a conduit à la rédaction d'un article par Monique Mayor et Guilhem Bourrié.

UR 1213 -URH - Unité de recherche sur les Herbivores – INRA-F-63122 Saint-Genès-Champanelle

Jean-François Hocquette, Directeur de l'unité

L'unité de Recherches sur les Herbivores, constituée le 1er janvier 1999, regroupe en 2009 174 agents dont 62 chercheurs et ingénieurs. Elle est organisée en 8 équipes de recherche, une équipe logistique, une installation expérimentale et dispose d'un abattoir expérimental. De plus, elle accueille chaque année de 80 à 100 étudiants (dont 15 à 20 thésards et post-doc). La mission de l'URH est de contribuer au développement des systèmes durables d'élevage des herbivores en accord avec les attentes sociétales, c'est-à-dire économique et développement rural, respect de l'environnement, qualité des produits et bien-être animal. Dans ce contexte, les activités de l'URH sont majoritairement orientées vers la mesure des caractères qui déterminent les phénotypes des animaux, principalement les herbivores.

Afin d'atteindre les deux objectifs prioritaires « fiabilité des résultats mesurables » et « traçabilité des travaux de recherche », le groupe qualité a identifié différentes actions à mettre en place qui ont permis en particulier de :

- Fiabiliser les résultats grâce à la gestion des instruments de mesures et la mise en place d'actions de métrologie sur différents équipements tels que les balances, les pipettes, les spectrophotomètres, les pH-mètres et le suivi des températures dans les enceintes climatiques.
- Harmoniser les pratiques lors de la mise en place des expérimentations animales par la mise à disposition de procédure et de trame-type de protocole expérimentaux.
- Pérenniser les savoir-faire par la rédaction d'environ 400 modes opératoires décrivant les différentes méthodes d'analyses utilisées.

Ce travail a permis à l'unité d'atteindre environ 90% du niveau 1 du Référentiel INRA en 2008 (Auto-évaluation 2008). Cependant, il nous est apparu important de compléter ces actions par la validation de nos méthodes de laboratoires les plus « stratégiques ». En effet, l'Unité dispose d'une large gamme de méthodes analytiques (quantitatives, semi-quantitatives et qualitatives) relevant de la chimie, de la biochimie, de la physiologie et de la biologie cellulaire et moléculaire, ces méthodes originales n'étant pour la plupart que partiellement, voire pas du tout, validées au sens « qualité » du terme (domaine de validité, linéarité, spécificité, fidélité...). Cette étape devrait permettre d'apporter une aide au scientifique comme l'optimisation des protocoles expérimentaux (nombre d'animaux à utiliser, nombre de prélèvement à réaliser, méthodes à utiliser en fonction des niveaux attendus...). Cela participera à consolider l'obtention de résultats fiables et interprétables en vue de leur valorisation lors des nombreuses communications orales et publications scientifiques de l'unité (environ 330 publications annuelles dans le cas de l'URH).

De plus, la mise en place de cette approche de validation de méthode devrait renforcer le dialogue entre le chercheur et l'analyste, notamment au moment de la validation d'une méthode d'analyse pour s'assurer que les variables mesurées correspondront bien à l'utilisation faite ultérieurement des résultats obtenus. Plus généralement, la validation des méthodes est un levier d'action important parmi d'autres pour mieux préciser les caractères mesurés et ainsi les définir. A ce titre, la démarche engagée va contribuer au projet ATO pour « Animal Trait Ontology » (ou Ontologie des caractères des animaux) dont l'objectif majeur est la définition des caractères phénotypiques des animaux. Les définitions et les mesures de ces caractères doivent être standardisées pour être comparables partout dans le monde et d'une espèce à

l'autre. Le projet ATO a été lancé en 2009 par le département Phase (Responsable : Pierre-Yves Le Bail), en étroite collaboration avec James Reecy de l'université de l'Iowa.

Enfin, les caractéristiques obtenues lors de la validation de nos méthodes d'analyse permettront de renforcer la lisibilité de notre savoir faire et des prestations proposées auprès de nos différents partenaires « clients » (environ 50 contrats en cours chaque année) dans nos 4 grands secteurs d'activités : la Recherche et Développement (ANR (14 projets ANR en cours pour 2009) contrat Européen (9 Projets Européens en cours en 2009), industrie, instituts techniques, centre de recherche français ou étranger, ...), la formation (École doctorale des Sciences de la Vie et de la Santé de l'Université d'Auvergne, l'ENITA de Clermont-Ferrand, ...) et l'expertise (Agence Française de Sécurité Sanitaire des Aliments, ...).

UR 1268 BIA - Biopolymères, interactions assemblages- INRA – F-44316 Nantes

Estelle Bonnin, ingénieur de recherche, équipe PVPP, Responsable Qualité de l'unité

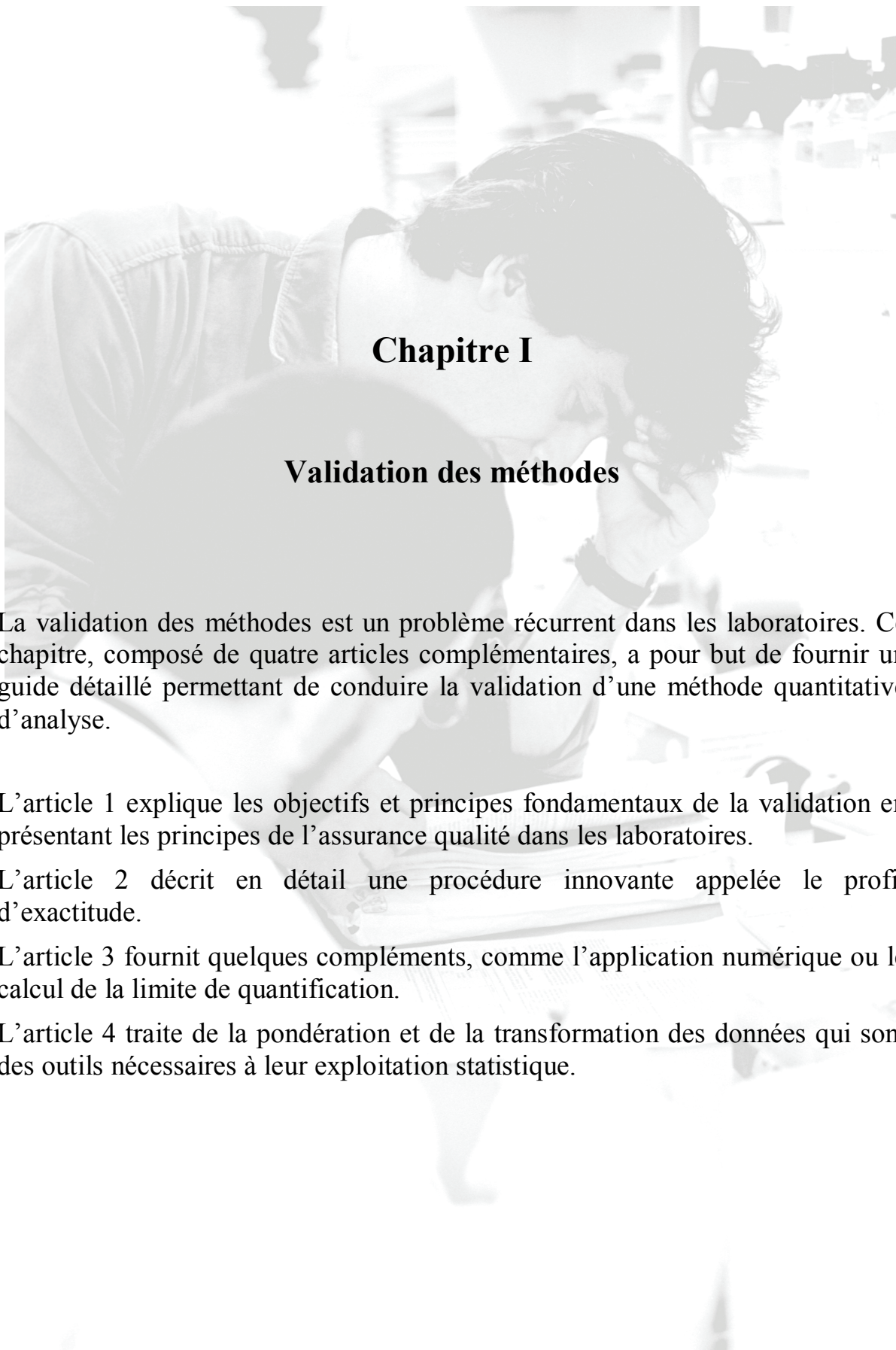
Cette formation s'applique en particulier aux méthodes utilisées en routine et par différents opérateurs (par exemple, chez nous la CPG ou les dosages d'activités enzymatiques). Cela ne peut pas être séparé du niveau 1 de la Qualité INRA puisque cela fait partie intégrante de la fiabilité des résultats. Cela signifie que maintenant que ce type de formation est mis à notre disposition, l'étape de validation devrait être systématique lors de la mise en place d'une nouvelle méthode, entre son développement et son utilisation courante. Cela nécessite donc qu'un nombre suffisant de personnes ait suivi cette formation.

Jean-Charles Gaudin, chargé de recherche, équipe ALL

L'École de validation des méthodes s'adresse surtout à des gens développant ou utilisant une méthode de dosage. L'étape de validation est indispensable pour des gens réalisant des dosages de manière très répétitive, devant en particulier garantir la validité du résultat à un tiers (prestation de service). Elle est aussi particulièrement intéressante pour tout développement de méthode car elle va permettre de déterminer si la méthode choisie est apte donner un résultat avec l'incertitude nécessaire ou demandée (elle permet en outre de mettre l'accent sur les étapes à améliorer). A mon sens, cette formation devrait être largement suivie par toute personne réalisant des dosages, particulièrement si ces dosages son fréquemment répétés. Je n'ai pas encore eu l'opportunité de mettre en application ce que j'ai appris lors de la formation, cependant elle m'a convaincu de la nécessité de mettre en place une procédure de validation pour toute nouvelle méthode utilisée ou développée à des fins de quantification.

Lucie Ribourg, technicien de la recherche, équipe ISD

Effectivement, j'ai suivi cette formation en juin l'année dernière. Je suis contente d'y avoir participé ; j'y ai appris ou revu beaucoup de choses. En tout cas, ça met l'accent sur des erreurs que l'on commet souvent et ça rappelle l'importance de la démarche qualité de l'INRA. Cette formation concerne un large public, du technicien au chercheur. C'est parfois un peu poussé pour un technicien mais ceci dit très intéressant. J'ai fait une présentation à l'équipe en rentrant sous forme de PowerPoint® et j'ai appliqué la validation à une méthode au laboratoire bien que ça ne soit pas aussi simple que sur le papier.



Chapitre I

Validation des méthodes

La validation des méthodes est un problème récurrent dans les laboratoires. Ce chapitre, composé de quatre articles complémentaires, a pour but de fournir un guide détaillé permettant de conduire la validation d'une méthode quantitative d'analyse.

L'article 1 explique les objectifs et principes fondamentaux de la validation en présentant les principes de l'assurance qualité dans les laboratoires.

L'article 2 décrit en détail une procédure innovante appelée le profil d'exactitude.

L'article 3 fournit quelques compléments, comme l'application numérique ou le calcul de la limite de quantification.

L'article 4 traite de la pondération et de la transformation des données qui sont des outils nécessaires à leur exploitation statistique.

Principes et vocabulaire pour la validation des méthodes

Max Feinberg¹

Résumé : *La validation des méthodes est un problème récurrent dans les laboratoires. Cet article a pour but d'expliquer les objectifs et principes fondamentaux de la validation en présentant les principes de l'assurance qualité dans les laboratoires.*

Mots-clés : Validation, profil d'exactitude, justesse, fidélité, incertitude, méthodes d'analyse quantitative

1. Place de la validation dans la qualité

1.1. Assurance qualité au laboratoire

Parmi les diverses préoccupations des analystes, la place prise par la validation des méthodes, doit être mise en relation avec le développement de l'assurance qualité dans les laboratoires. Les principes de l'assurance qualité sont l'objet d'une série de normes publiées par l'International Standardisation Organisation (ISO), comme la série des normes ISO 9000. S'y trouvent définies la qualité et toutes les opérations qui y concourent. Même si ces normes ne sont pas, en général, applicables en l'état aux activités des laboratoires, leur lecture n'est pas inutile. Elles insistent sur le fait que la qualité des produits ou des services doit être conçue en vue de satisfaire les **besoins d'un client**. Souvent on résume les objectifs de la qualité d'un produit ou d'un service sous la forme d'une liste comme celle-ci :

- répondre à un besoin, un usage ou un objectif bien défini ;
- satisfaire les attentes des consommateurs ;
- être conforme aux normes, spécifications, exigences réglementaires et autres ;
- avoir à un prix compétitif ;
- être fourni à un coût qui génère un profit.

Par ailleurs, pour remplir son rôle, un système d'assurance de la qualité doit présenter trois caractéristiques :

- être conçu pour le but qu'on lui assigne ;
- être correctement mis en œuvre ;
- être effectivement contrôlé.

Tous ces principes peuvent sembler bien abstraits mais en les appliquant à l'activité d'un laboratoire d'analyse, l'assurance de la qualité est tout ce qui peut individuellement ou collectivement influencer le résultat d'une analyse.

Afin de prendre en compte la spécificité des activités de laboratoire et la diversité de leurs domaines d'application, il existe au moins quatre référentiels principaux qui expliquent comment organiser l'assurance de la qualité au laboratoire :

- les **Bonnes pratiques de laboratoire** (BPL) s'appliquent à l'étude et au développement

1. UR1204 Met@risk - Méthodologie d'analyse de risque alimentaire - INRA – F- 75231 Paris cedex 05.

☎ 01 44 08 16 52 ✉ max.feinberg@paris.inra.fr

de molécules chimiques, médicaments, des pesticides et, en général, à l'étude de toute substance nouvelle ;

- le **Guide de bonne exécution des analyses** (GBEA) précise comment doit fonctionner le système d'assurance de la qualité d'un laboratoire d'analyse de biologie médicale ;
- l'**accréditation**, qui suit les principes de la norme ISO 17025 et permet d'assurer la compétence d'un laboratoire qui applique un ensemble de recommandations NF EN ISO/CEI 17025:2005, (NF X 50-061) ;
- la **certification de service** s'appuie sur les normes ISO 9000 et s'applique à une entreprise dans son ensemble et, éventuellement, au laboratoire qui s'y rattache.

Dans les années 1970, mettre en place l'assurance qualité dans un laboratoire, c'était surtout mieux prévoir son organisation afin d'améliorer la transparence de son fonctionnement. La validation se concevait alors surtout pour les méthodes normalisées qui servaient aux échanges commerciaux : c'était le rôle de la norme ISO 5725 de fixer ces critères d'inter-comparaison. Ensuite, les normes sur l'assurance qualité se sont diversifiées mais n'ont pas changé beaucoup dans leurs principes. Les textes qui définissaient des procédures de validation interne ne sont apparus que vers la fin des années 1990. La question de savoir comment seront interprétés les résultats d'analyse par les demandeurs ne s'est vraiment posée que vers 2005. L'introduction de l'incertitude et le développement de l'approche critères sont liés à cet objectif. On peut considérer qu'il s'agit de mieux prendre en compte les besoins des utilisateurs ; ce qui semble logique si on se réfère à la définition de la qualité.

De nombreux guides, recommandations et guidances plus ou moins officiels ont été publiés par des structures réglementaires, des organismes de normalisation ou des groupes d'experts. On trouvera en section 4 une liste de références pour ces normes et guides. Ces documents sont souvent disponibles sur Internet et peuvent être téléchargés. On peut ainsi constater qu'ils se recopient largement les uns les autres. En outre, ils contiennent de nombreuses définitions dont certaines sont reprises à la section 2.

Ces guides permettent cependant de comprendre la philosophie générale qui a été adoptée jusqu'à présent pour conduire une validation. Dans la plupart des cas, cela consiste en une suite d'études expérimentales qui permettent de calculer des critères de validation. Par exemple, le document 1133 du Comité français d'accréditation (Cofrac) sur la « *Validation interne des méthodes d'essais* » indique que les « *critères de performance à établir, lors de la validation d'une méthode d'essai, dépendent des caractéristiques mesurées, qualitatives ou quantitatives, des principes physiques, chimiques, mécaniques ou biologiques sur lesquels portent l'essai ou l'analyse et du type d'équipement utilisé* ». Vient ensuite la liste de ces critères, citée à titre « d'exemple » : *la justesse (obtenue par comparaison avec une méthode de référence, par l'emploi de matériaux de référence ou de matériaux dont les propriétés sont connues ou par essai inter-laboratoires) ; la fidélité (exprimée en terme de répétabilité, de reproductibilité et éventuellement d'hystérésis) ; la limite de détection et la limite de quantification ; la sensibilité ; la sélectivité ou la spécificité ; le pouvoir de résolution ou de séparation ou de discrimination ; et la robustesse.*

De même, la norme NF V03-110:1994, propose de faire des tests pour vérifier la conformité d'une méthode alternative par rapport à une méthode de référence en comparant les écarts-types de répétabilité, les moyennes et les limites de détection de deux méthodes. Divers autres tests viennent compléter ce qui est alors appelé étude de validation : *vérification de l'homogénéité de la variance dans le domaine d'application ; adéquation du modèle linéaire au cas d'étalonnage linéaire avec une variance constante ; calcul de la sensibilité ; contrôle*

de la spécificité par la méthode des ajouts dosés et étalonnage linéaire par les moindres carrés.

Mais rien n'est dit sur le niveau de garantie à atteindre pour l'estimation de ces critères, par exemple en termes de nombre de mesures, ni ce qu'on doit en faire par la suite pour effectivement valider. En particulier, comment doit-on faire pour vérifier qu'ils permettent d'atteindre un niveau de qualité donné ? Le principe du traitement statistique des données développé dans ces guides de validation consiste à réaliser un certain nombre de tests dit de signification de l'hypothèse nulle. Ils ont le gros inconvénient de favoriser les méthodes les plus dispersives, c'est-à-dire dont la fidélité est mauvaise. La démonstration de cet inconvénient des guides classiques requiert des développements importants qui alourdiraient ce texte. Le lecteur intéressé pourra trouver dans Feinberg M. (2009), Boulanger *et al.* (2003) et Feinberg M. *et al.*, (2004), tous les éléments qui appuient cette affirmation. C'est pourquoi, nous avons développé une approche mieux adaptée qui débouche sur un autre outil statistique appelé le **profil d'exactitude**.

1.2. Organiser la validation d'une méthode

Pour organiser la validation d'une méthode d'analyse, il faut d'abord l'avoir correctement mise au point. Bien souvent les analystes sont trop impatientes et vont essayer de valider une méthode dont le domaine d'application n'est encore bien défini et pour laquelle il existe des effets de matrice inconnus. C'est une perte de temps et d'argent. Une étude de validation doit démarrer lorsqu'on dispose d'un mode opératoire complètement rédigé.

Bien sûr l'organisation pratique des essais dépend ensuite du guide de validation choisi. Dans Feinberg (2010b) nous proposons une séquence de dix opérations qui débouchent sur la construction d'un profil d'exactitude. On peut déjà en définir des principes généraux :

- définir les objectifs à atteindre ;
- prévoir les plans d'expérience à réaliser ;
- prévoir les quantités et la qualité des matériaux qui seront utilisés ;
- organiser le calendrier des études et celui du personnel qui en sera chargé.

En fin d'étude, il ne faut pas sous-estimer le temps nécessaire pour interpréter les résultats. C'est pourquoi, il importe de prévoir dès le début le mode d'enregistrement des mesures, en fonction des outils de calcul dont dispose le laboratoire.

De toute façon, si le laboratoire fonctionne avec un système qualité, il se doit de rédiger une procédure et tous les documents annexes qui précisent comment l'étude sera conduite. On peut proposer cette série d'étapes qui constituera le squelette de cette procédure :

- nommer un responsable qui coordonnera l'étude de validation ;
- rédiger la version préliminaire du mode opératoire de la méthode en suivant une procédure de présentation d'un mode opératoire normalisé.
- ouvrir une fiche de validation ;
- réunir un groupe de travail qui définira le domaine d'application et les spécifications, en particulier l'objectif attendu de la méthode.
- rédiger les plans d'expérience qui devront être effectués. Prévoir la logistique, en termes de volumes d'échantillon de validation, de réactifs et de disponibilité du personnel.
- réaliser les études prévues.
- interpréter les résultats et calculer les divers critères de performance requis.
- rédiger la version finale du mode opératoire de la méthode avec ses critères de performance.

Il est alors possible de définir un certain nombre de types de **dossiers de validation** qui se solderont par des documents présentables à un organisme de contrôle. Le dossier de validation, une fois constitué, permettra aussi d'établir des valeurs seuils de qualification du matériel de mesure qu'il faudra vérifier, par exemple, avant chaque utilisation. Ces tests de qualification font l'objet de **fiches de qualification** spécifiques de chaque instrument employé. Une vérification d'aptitude doit aussi avoir lieu si la méthode est transférée sur un autre matériel. Il ne faut donc pas confondre la validation de la méthode avec un test de qualification de l'instrumentation.

2. Vocabulaire et définitions

Le vocabulaire relatif à la validation est très abondant, souvent foisonnant et souvent confus. Nous avons décidé de choisir les définitions proposées en 2007 dans la dernière version du Vocabulaire International de Métrologie (VIM) (Guide ISO/CEI 99:2007) car elles représentent l'effort le plus récent de normalisation dans ce domaine.

Terme	Référence	Définition
Mesurage, mesure	VIM 2.1	Processus consistant à obtenir expérimentalement une ou plusieurs valeurs que l'on peut raisonnablement attribuer à une grandeur.
Mesurande	VIM 2.3	Grandeur que l'on veut mesurer. Analyte, employé en chimie, est un synonyme de mesurande, d'usage plus général.
Justesse de mesure, justesse	VIM 2.15	Étroitesse de l'accord entre la moyenne d'un nombre infini de valeurs mesurées répétées et une valeur de référence.
Fidélité de mesure, fidélité	VIM 2.16	Étroitesse de l'accord entre les valeurs mesurées obtenues par des mesurages répétés du même objet ou d'objets similaires dans des conditions spécifiées.
Condition de répétabilité	VIM 2.21	Condition de mesurage dans un ensemble de conditions qui comprennent la même procédure opératoire, les mêmes opérateurs, le même système de mesure, les mêmes conditions de fonctionnement et le même lieu, ainsi que des mesurages répétés sur le même objet ou des objets similaires pendant une courte période de temps.
Répétabilité de mesure, répétabilité	VIM 2.22	Fidélité de mesure selon un ensemble de conditions de répétabilité.
Condition de fidélité intermédiaire	VIM 2.23	Condition de mesurage dans un ensemble de conditions qui comprennent la même procédure opératoire, le même lieu et des mesurages répétés sur le même objet ou des objets similaires pendant une période de temps étendue, mais peuvent comprendre d'autres conditions que l'on fait varier.
Fidélité intermédiaire de mesure, fidélité intermédiaire	VIM 2.24	Fidélité de mesure selon un ensemble de conditions de fidélité intermédiaire.

Exactitude de mesure, exactitude	VIM 2.13 et 2.14	a) Étroitesse de l'accord entre une valeur mesurée et une valeur vraie du mesurande. b) Étroitesse de l'accord entre les valeurs mesurées qui sont attribuées au mesurande.
Incertitude de mesure, incertitude	VIM 2.27	Paramètre qui caractérise la dispersion des valeurs attribuées à un mesurande, à partir des informations utilisées. L'incertitude de mesure comprend des composantes provenant d'effets systématiques, telles que les composantes associées aux corrections et aux valeurs assignées des étalons, ainsi que l'incertitude définitionnelle. Parfois, les effets systématiques connus ne sont pas corrigés mais traités comme des composantes de l'incertitude. Le paramètre qui sert à exprimer l'incertitude peut être, par exemple, un écart-type appelé incertitude-type (ou un de ses multiples) ou la demi étendue d'un Intervalle ayant un niveau de confiance déterminé.
Limite de détection	VIM 4.18	Valeur mesurée, obtenue par une procédure opératoire donnée, pour laquelle la probabilité de déclarer faussement l'absence d'un constituant dans un matériau est β , étant donnée la probabilité α de déclarer faussement sa présence. La notation β utilisée dans cette définition représente un risque d'erreur (dit de seconde espèce) et ne doit pas être confondue avec la proportion β de la définition de l'intervalle de tolérance.
Sélectivité	VIM 4.13	Aptitude d'un système de mesure, utilisant une procédure opératoire spécifiée, à fournir des résultats de mesure pour un ou plusieurs mesurandes, qui ne dépendent pas les uns des autres ou de toute autre grandeur existant dans le système en cours de mesurage.
Spécificité	UIPAC	Stade ultime de la sélectivité.
Valeur de référence	VIM 5.18	Valeur d'une grandeur, dont l'incertitude de mesure associée est considérée communément comme suffisamment petite pour que la valeur puisse servir de base de comparaison pour les valeurs de grandeurs de même nature.
Validation	VIM 2.45	Vérification que des exigences spécifiées sont adéquates pour un usage déterminé. Le terme caractérisation s'applique à la méthode, alors que le terme vérification s'applique au résultat rendu. La validation de la méthode consiste donc à vérifier si les résultats sont adéquats en vue d'un usage défini.
Vérification	VIM 2.44	Fourniture de preuves tangibles qu'une entité donnée satisfait à des exigences spécifiées, en prenant en compte toute incertitude de mesure.
Domaine d'application <de la méthode>		Ensemble des types de matrice auquel s'applique la méthode, en tenant compte de la gamme de concentrations sur laquelle doit porter la validation.

Domaine de validation	Ensemble des types de matrice auquel s'applique la méthode et gamme de concentrations sur laquelle a porté la validation.
Domaine de validité	Ensemble des types de matrice auquel s'applique la méthode et gamme de concentrations sur laquelle a porté la validation et pour lequel les futurs résultats fournis par la méthode sont jugés valides.
Méthode quantitative	Méthode d'analyse qui détermine la quantité ou la fraction pondérale d'un analyte de manière à pouvoir l'exprimer sous forme de valeur numérique dans les unités appropriées.
Méthode de référence	Méthode d'analyse reconnue par des experts ou prise comme référence par accord entre les parties, qui donne ou est supposée donner la valeur de référence acceptée du mesurande.
Matrice	Ensemble des propriétés de l'échantillon et de ses constituants autres que l'analyte. L'effet de matrice traduit l'influence possible que ces propriétés ou constituants peuvent avoir sur la réponse instrumentale. Comme l'effet de matrice peut varier aux différents stades de l'analyse (par exemple avant ou après minéralisation), pratiquement un type de matrice sera défini comme un groupe de matériaux ou de produits reconnus par l'analyste comme ayant un comportement homogène vis-à-vis de la méthode d'analyse utilisée.
Série	Ensemble de mesurages réalisés sur un même échantillon selon des conditions de connues. Par exemple, une série réunira des mesurages effectués un même jour et/ou par un même opérateur et/ou un même appareil...
Profil d'exactitude	Combinaison, sous une forme graphique, d'un ou plusieurs intervalles de tolérance d'espérance β calculés à différents niveaux de concentration et d'une ou plusieurs intervalles d'acceptabilité.
Intervalle de tolérance d'espérance β (Intervalle de tolérance)	Intervalle qui contient en moyenne une proportion $\beta\%$ définie de futurs mesurages obtenus selon un mode opératoire donné et pour une concentration donnée. NOTE. L'intervalle de tolérance, revêt une grande importance dans la suite du texte. Ses limites sont obtenues par le calcul, à partir de mesurages répétés. La valeur de β est, généralement, choisie entre 80 % et 95 %. En fait, elle n'intervient que pour situer la limite de quantification de la méthode comme on le verra à la partie 3.
Intervalle d'acceptabilité	Spécification de la performance exigée pour la méthode, exprimée comme un écart acceptable autour de la valeur de référence. Les limites de l'intervalle sont fixées par le client ou par une obligation réglementaire, parfois en fonction du niveau de concentration. Elles sont notées $\pm \lambda$ en valeurs absolues et dans l'unité du mesurande ou $(1 \pm \lambda) \times 100$ en valeurs relatives.

Échantillon de validation (échantillon témoin)		Matériau dont la valeur de référence peut être assignée, soit parce qu'il s'agit d'un matériau de référence (certifié ou non), soit parce qu'on a procédé à un ajout dosé ou un marquage de la molécule à doser.
Limite de quantification (1)		Plus petite et/ou plus grande concentration de l'analyte pouvant être quantifiée, dans les conditions expérimentales décrites de la méthode.
Limite de quantification (2)	SFSTP	Plus petite et/ou plus grande quantité de l'analyte dans un échantillon pouvant être dosée dans les conditions expérimentales décrites avec une exactitude définie. Elle correspond à la concentration la plus petite et/ou la plus grande du domaine de validité.
Linéarité <de la méthode>		Établissement qu'il existe une relation linéaire entre les quantités retrouvées (ou quantifiées) dans des échantillons et leurs valeurs de référence. La linéarité de la méthode ne doit pas être confondue avec la linéarité de la fonction de réponse de l'appareillage de mesure qui caractérise la seule réponse instrumentale lors de l'étalonnage et qui n'est pas indispensable pour obtenir une quantification exacte.

3. Principes à la base du profil d'exactitude

3.1. Rappel sur les normes de la série ISO 5725

Le profil d'exactitude est basé sur une application directe des principes décrits dans les normes de la série ISO 5725 (1, 2, 3, 4 ; 1994)]. On y propose un modèle statistique pour estimer l'exactitude (justesse et fidélité) d'une méthode ou de résultats. Ce modèle décrit un mesurage z d'un mesurande Z sous la forme :

$$z = m + B + e$$

Où m est la moyenne générale de l'échantillon homogène, B la composante du biais du laboratoire sous condition de répétabilité et e l'erreur aléatoire survenant dans chaque mesurage, sous condition de répétabilité. Dans le cas d'une étude inter-laboratoires, la composante du biais B provient des laboratoires, mais elle peut aussi provenir de toute autre source d'incertitude, comme le jour, l'opérateur, l'instrument... dans le cas d'une étude intra-laboratoire. Dans le vocabulaire statistique, un tel modèle est dit « **modèle d'analyse de la variance à un seul facteur à effets aléatoires** » ou aussi « **modèle des composantes de la variance** » (en anglais, One-Way ANOVA Random Effect Model). Il est très répandu dans la recherche agronomique.

Outre les méthodes statistiques pour calculer les critères d'exactitude, les normes ISO 5725 précisent aussi en détail l'organisation de la collecte et les précautions à respecter.

3.2. Diverses conditions d'estimation de la fidélité

D'après sa définition, la fidélité peut être estimée sous plusieurs conditions, comme la **répétabilité**, la **fidélité intermédiaire** ou la **reproductibilité**. Dans tous les cas, la fidélité est quantifiée par un écart-type, qu'il soit de répétabilité et noté s_r , de fidélité intermédiaire s_{FI} ou de reproductibilité s_R . On peut établir une échelle de complexité entre ces différents écarts-

types, en fonction du nombre de sources d'incertitude. La Figure 1 illustre cette graduation depuis les conditions de répétabilité où il n'existe aucun facteur de variation identifié et/ou contrôlé pour calculer l'écart entre des répétitions, jusqu'aux conditions de reproductibilité où le nombre de sources est inconnu. Ceci, en passant par les différentes possibilités d'estimation de la fidélité intermédiaire. La norme ISO 5725-3 décrit ces divers contextes d'application, dits de fidélité intermédiaire, et les modèles à utiliser lorsque d'autres facteurs peuvent varier, comme le jour, l'opérateur, l'instrument...

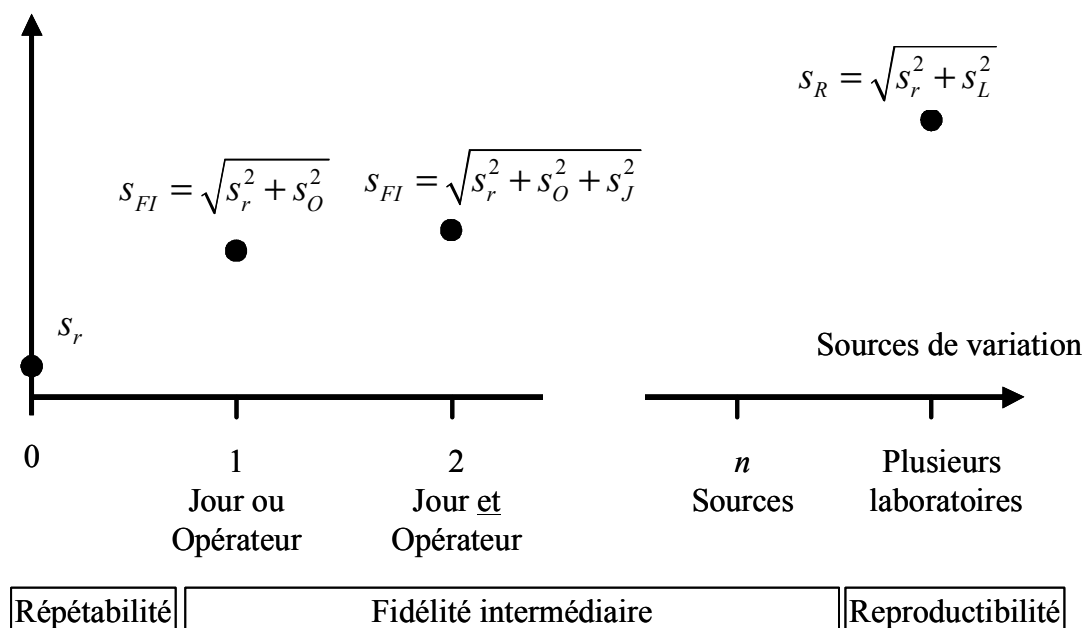


Figure 1 : les différentes estimations de la fidélité d'une méthode en fonction des sources de variation impliquées

L'application la plus classique de la norme ISO 5725-2 est le calcul de la répétabilité et de la reproductibilité, dans le cadre d'une étude inter-laboratoires qui débouche sur deux écarts-types. Mais elle décrit aussi les calculs sous des conditions intermédiaires de fidélité. Dans le cadre de la méthode du profil d'exactitude, on peut se situer aussi bien en condition de fidélité intermédiaire et la validation est alors intra-laboratoire, que de reproductibilité et la validation est inter-laboratoires.

Pour simplifier la présentation, la notion de **série** sert à désigner un ensemble de répétitions effectuées sous des conditions de répétabilité : une série regroupe tous les mesurages effectués sur un même échantillon sous les mêmes conditions, par exemple le même jour, le même opérateur, un délai court. Les plans d'expériences décrits dans la Partie 2 consistent à réaliser des essais en les répartissant en plusieurs séries (par exemple, au moins 3 jours). Pour certaines méthodes qui s'appliquent à des échantillons très instables dans le temps il faudra choisir un effet série qui ne soit pas le jour mais plutôt les opérateurs : une série regroupe alors les répétitions faites par un même opérateur. Par exemple, la détermination de l'extrait sec d'un fromage peut exiger d'utiliser différents opérateurs travaillant un même jour, si la méthode est destinée, en routine, à rendre un résultat en moins de 24 heures.

Des schémas plus compliqués peuvent être utilisés, comme dans l'exemple suivant où se combinent différents jours, opérateurs et instruments pour créer 4 séries sous la forme d'un plan factoriel fractionné à 3 facteurs.

Séries	Jours	Opérateurs	Instruments
1	Jour 1 (-)	Opérateur 1 (-)	Instrument 1 (+)
2	Jour 2 (+)	Opérateur 1 (-)	Instrument 2 (-)
3	Jour 1 (-)	Opérateur 2 (+)	Instrument 2 (-)
4	Jour 2 (+)	Opérateur 2 (+)	Instrument 1 (+)

D'une façon générale, le choix des sources de variation pour organiser les séries de mesure doit refléter au mieux les composantes de l'incertitude qui risquent d'apparaître lors de l'application en routine de la méthode à valider.

Dans la méthode du profil d'exactitude, il est indispensable de collecter les données en plusieurs séries et en contrôlant les sources de variation. Sinon, il est impossible de construire un profil d'exactitude.

3.3. Profil d'exactitude

À partir de l'écart-type de fidélité intermédiaire ou de reproductibilité, calculé selon les recommandations de la norme ISO 5725-2 (voir Partie 3), on construit un intervalle de tolérance d'espérance β qui contient une proportion $\beta\%$ de futurs résultats. Cette procédure a été établie par un groupe de travail de la Société française des sciences et techniques pharmaceutiques (Commission SFSTP 2003, 2006).

L'intervalle de tolérance d'espérance β est différent de l'intervalle de confiance. L'intervalle de confiance caractérise un paramètre statistique. Par exemple l'intervalle de confiance de la moyenne permet d'établir avec un niveau de confiance donné un intervalle autour de la moyenne observée dans lequel la moyenne vraie a de grandes chances de se trouver. Il s'exprime sous la forme Moyenne $\pm k_{IC}$ (écart-type de la moyenne), où k_{IC} est le facteur de couverture de l'intervalle de confiance. L'intervalle de tolérance caractérise, d'un côté, la dispersion des résultats déjà obtenus mais permet aussi de prévoir ceux qui seront obtenus lors de futurs mesurages. Il s'exprime aussi sous la forme Moyenne $\pm k_{IT}$ (écart-type de fidélité) mais où k_{IT} est un autre facteur de couverture, calculé de façon complètement différente de k_{IC} .

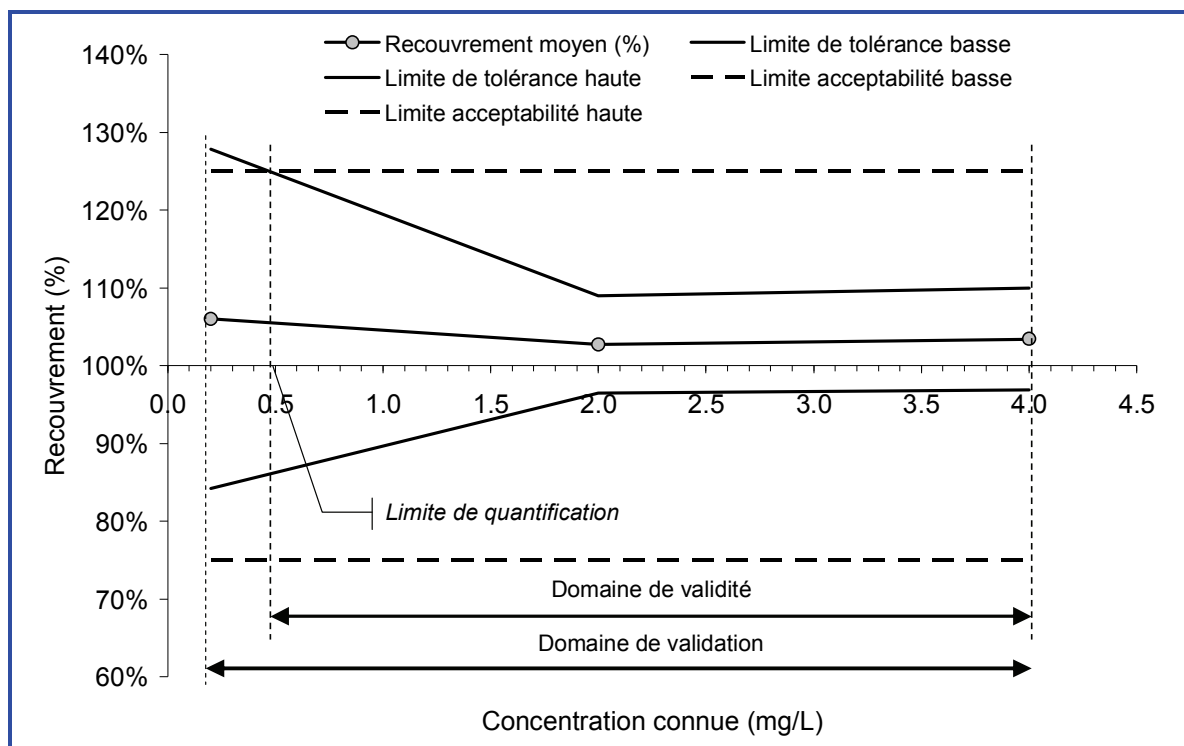


Figure 2 : profil d'exactitude établi avec 3 niveaux de concentration.

Chaque cercle gris représente la concentration moyenne retrouvée pour le niveau et quantifie la justesse. Les lignes pointillées délimitent l'intervalle d'acceptabilité et les lignes continues l'intervalle de tolérance calculé à partir des écarts-types de fidélité intermédiaire pour chaque niveau. Les traits verticaux délimitent le domaine de validité, dans lequel la méthode est capable de produire une proportion élevée et connue de résultats acceptables

Tous les calculs sont effectués séparément pour chaque niveau de concentration k , permettant d'obtenir k écarts-types de fidélité puis k intervalles de tolérance qui une fois réunis servent à construire le profil. La **figure 2** illustre un exemple de profil d'exactitude construit avec 3 niveaux de concentration – 0,4 ; 2,0 et 4,0 mg/l – qui délimitent le domaine de validation (ou d'application) de la méthode à valider.

Les éléments graphiques entrant dans le profil d'exactitude sont les suivants :

- sur l'axe horizontal, la concentration théorique des niveaux (soit les valeurs de référence assignées aux niveaux) ;
- sur l'axe vertical, simultanément :
 - les limites des intervalles de tolérance d'espérance β calculés sur les concentrations retrouvées et exprimées en pourcentages (sous la forme d'un taux de recouvrement ou d'une exactitude relative) ;
 - les intervalles d'acceptabilité, définis en fonction de l'objectif de la méthode, exprimés de la même façon que les intervalles de tolérance.

La stratégie d'interprétation de ce graphique est décrite en détail par Max Feinberg (2010 c). Cependant, d'une part, les limites d'acceptabilité, d'autre part, la proportion β qui sert à calculer les intervalles de tolérance dépendent strictement du contexte d'utilisation de la méthode et doivent être adaptés à chaque cas. Sur la figure 2, dans le domaine de validité, délimité par les traits verticaux discontinus, la méthode est capable de produire une

proportion moyenne attendue de résultats égale à β comprise entre les limites d'acceptabilité : la méthode est dite valide dans ce domaine. Le domaine d'application représente le domaine initialement choisi pour conduire la validation.

D'une façon générale, pour être en cohérence avec les modèles statistiques utilisés, il faut disposer de répétitions effectuées sous condition de fidélité intermédiaire ou de reproductibilité sur des échantillons homogènes.

Max Feinberg présente de façon détaillée les 10 étapes de la méthode du profil d'exactitude (2010 b) et des exemples, des compléments et des feuilles de calcul (2010 c).

4. Guides et normes de validation

4.1. Textes réglementaires

Commission des Communautés Européennes, *Décision 2002/658/EC : modalités d'application de la directive 96/23/CE du Conseil en ce qui concerne les performances des méthodes d'analyse et l'interprétation des résultats* (2002) Off J. Eur. Commun. L221/8

Food and Drug Administration, *Guidance for Industry: Analytical Procedures and Methods Validation, Draft guidance* (2000) US Department of Health and Human Services, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER)

Food and Drug Administration, *Guidance for Industry: Bioanalytical Methods Validation for Human Studies, Draft guidance* (1998) US Department of Health and Human Services, Center for Drug Evaluation and Research (CDER)

Food and Drug Administration, International Conference on Harmonization: *Guideline on validation of analytical procedures: definitions and terminology* (1995) Federal Register 60, 11260-11262

Food and Drug Administration, International Conference on Harmonization: *Guideline on validation of analytical procedures: methodology* (1997) Federal Register 62, 27463-27467

GBEA Arrêté du 2 novembre relatif à la bonne exécution des analyses de biologie médicale (1995) *Revue Française des Laboratoires* 272, 55-65

Health and Welfare Canada, Drugs Directorate Guidelines, *Acceptable Methods, published by Authority of the Minister of National Health and Welfare* (1992) Health Protection Branch

United States Pharmacopeia, *Current concepts for the validation of compendia assays, Pharmacopeia Forum* (1986) The United States Pharmacopeia Inc., Rockville, MD, USA, 1241-1252

United States Pharmacopeia, XXVI, general information <1225>, *Validation of compendia methods* (2003) The United States Pharmacopeia Inc., Rockville, MD, USA, p. 2439-2442

4.2. Normes et textes de consensus

EURACHEM, *The Fitness for Purpose of Analytical Methods, A Laboratory Guide to Method Validation and Related Topics* (1998) 1st Edition

EURACHEM/CITAC *Guide, Quantifying Uncertainty in Analytical Measurement* (2000) Second Edition, Final draft, April

International Conference on Harmonization (ICH) *Q2(R1) Validation of Analytical Procedures: Text and Methodology* (2005) <http://www.ich.org/cache/compo/276-254-1.html>

International Union of Pure and Applied Chemistry (IUPAC) Harmonized guidelines for single-laboratory validation of methods of analysis (2002) *Pure Appl. Chem.* 74(5) 835-855

NF V03-110:2010 *Analyse des produits agricoles et alimentaires - Protocole de caractérisation en vue de la validation d'une méthode d'analyse quantitative par construction du profil d'exactitude* (2010) Afnor, Paris

NF V03-111 *Protocole d'évaluation intralaboratoire d'une méthode alternative qualitative par rapport à une méthode de référence* (1995) Afnor Paris

Nordic Committee on Food Analysis (NMKL) *Validation (Proc. 4)* <http://www.nmkl.org>

Organisation internationale de la vigne et vin (OIV) OENO 10-2005 Guide pratique pour la validation, le contrôle qualité, et l'estimation de l'incertitude d'une méthode d'analyse œnologique alternative (2005) OIV, Paris <http://www.oiv.int>

Vassault A., alt. Protocole de Validation des Techniques. Commission SFBC (1986) *Annales de Biologie Clinique* 44, 679-745

4.3. Méthodes microbiologiques

ISO 13843 (NF T90-460) *Qualité de l'eau, Lignes directrices pour la validation des méthodes microbiologiques* (2001) ISO, Genève

ISO 16140 *Microbiologie des aliments, Protocole pour la validation des méthodes alternatives* (2003) ISO, Genève

ISO 17994 *Qualité de l'eau, Critères pour établir l'équivalence entre les méthodes microbiologiques* (2004) ISO, Genève

Bibliographie

AFNOR (2005) Norme NF EN ISO/CEI 17025 (NF X 50-061) *Exigences générales concernant la compétence des laboratoires d'étalonnages et d'essais*. Afnor, Paris

AFNOR (1994) Norme NF V03-110 *Protocole d'évaluation d'une méthode alternative quantitative par rapport à une méthode de référence* Afnor, Paris

AFNOR (1994a) Norme NF ISO 5725-1 : *Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 1 : Principes généraux et définitions*. ISO, Genève

AFNOR (1994b) Norme NF ISO 5725-2 : *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*. ISO, Genève

AFNOR (1994c) Norme NF ISO 5725-3 : *Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 3 : Mesures intermédiaires de la fidélité d'une méthode de mesure normalisée*. ISO, Genève

AFNOR (1994d) Norme NF ISO 5725-4 : *Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 4 : Méthodes de base pour la détermination de la justesse d'une méthode de mesure normalisée*. ISO, Genève

- Boulanger B., Chiap P., Dewé W., Crommen J., Hubert Ph. (2003) An analysis of the SFSTP guide on validation of chromatographic bioanalytical methods: progresses and limitations. *Journal of Pharmaceutical and Biomedical Analysis* 32, 753-765
- Commission SFSTP, Hubert P Nguyen-Huu JJ, Boulanger B, Chapuzet E, Cohen N, Compagnon PA, Dewé W, Feinberg M, Laurentie M, Mercier N, Muzard G, Valat L, (2003) Validation des procédures analytiques quantitatives, Harmonisation des démarches : Partie I *STP Pharma Pratiques* 13(3) 101-138
- Commission SFSTP, Hubert P Nguyen-Huu JJ, Boulanger B, Chapuzet E, Cohen N, Compagnon PA, Dewé W, Feinberg M, Laurentie M, Mercier N, Muzard G, Valat L, (2006) Validation des procédures analytiques quantitatives, Harmonisation des démarches : Partie II Statistiques *STP Pharma Pratiques* 16(1) 30-60
- Commission SFSTP, Hubert P Nguyen-Huu JJ, Boulanger B, Chapuzet E, Cohen N, Compagnon PA, Dewé W, Feinberg M, Laurentie M, Mercier N, Muzard G, Valat L, (2006) Validation des procédures analytiques quantitatives, Harmonisation des démarches : Partie III Exemples d'application *STP Pharma Pratiques* 16(2) 87-121
- Feinberg M., Boulanger B., Dewé W., Hubert P. (2004) New advances in method validation and measurement uncertainty aimed at improving the quality of chemical data. *Analytical Bioanalytical Chemistry* 380: 502-514
- Feinberg M., (2009) *Labo-Stat : Guide de validation des méthodes d'analyse* Lavoisier Tec&Doc, Paris
- Feinberg M. (2010b) Mise en œuvre du profil d'exactitude. In Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 27-44
- Feinberg M. (2010c) Interprétation du profil d'exactitude. In Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 45-59
- ISO (2007) Guide ISO/CEI 99 *Vocabulaire international de métrologie – Concepts fondamentaux et généraux et termes associés (VIM)*, ISO, Genève

Mise en œuvre du profil d'exactitude

Max Feinberg¹

Résumé : Cet article a pour objet d'expliquer comment la procédure de construction du profil d'exactitude peut être utilisée pour valider une méthode d'analyse quantitative. Le profil d'exactitude est applicable aussi bien à la validation interne dans un laboratoire d'essais qu'à l'interprétation de données collectées dans le cadre d'une étude inter-laboratoires conçue selon les recommandations de la norme ISO 5725-2. Elle peut aussi s'appliquer aux essais basés sur des énumérations, comme les méthodes microbiologiques, au prix d'un certain nombre de modifications qui ne seront pas présentées ici. L'approche proposée doit permettre une validation interne des méthodes dans les domaines suivants :

- les méthodes non normalisées ;
- les méthodes conçues/développées par le laboratoire ;
- les méthodes normalisées employées en dehors de leur domaine d'application prévu ;
- les amplifications ou modifications de méthodes normalisées ;
- et aussi une méthode alternative par rapport à une méthode de référence.

Mots-clés : Validation, profil d'exactitude, justesse, fidélité, incertitude, méthodes d'analyse quantitatives

Introduction

Le profil d'exactitude ne s'applique qu'aux méthodes complètement développées et mises au point. En particulier, la sélectivité/spécificité doit avoir été étudiée ainsi que le domaine d'application de la méthode à valider, en termes de types de matrice et de niveaux de concentrations. Le domaine d'application du profil d'exactitude s'inspire directement de la norme ISO 17025:2005 qui propose que

« Le laboratoire doit valider les méthodes non normalisées, les méthodes conçues/développées par le laboratoire, les méthodes normalisées employées en dehors de leur domaine d'application prévu, ainsi que les amplifications ou les modifications de méthodes normalisées, afin de confirmer que les méthodes sont aptes à l'emploi prévu. La validation doit être aussi étendue que l'impose la réponse aux besoins pour l'application ou le domaine d'application donné ».

Bien sûr, de nombreux développements statistiques ne sont pas présentés mais sont disponibles à la référence (Feinberg M., 2009).

¹ UR1204 Met@risk - Méthodologie d'analyse de risque alimentaire- INRA – F-75231 Paris cedex 05

☎ 01 44 08 16 52 ✉ max.feinberg@paris.inra.fr

1. Profil d'exactitude en 10 étapes

1.1 Définir la quantité mesurée

À partir du mode opératoire de la méthode, on définit la quantité mesurée en précisant les formules d'expression du résultat final et la procédure pour établir ce résultat. En particulier, il faut faire la distinction entre deux types de méthodes.

- Les méthodes « indirectes ou rationnelles » qui exigent un étalonnage préalable pour calculer la concentration des échantillons inconnus. Pour ces méthodes, on procède en deux étapes : d'abord construire la courbe d'étalonnage en utilisant le même principe physico-chimique que celui qui sert pour les échantillons ; ensuite faire des mesures sur les échantillons inconnus et calculer leurs concentrations à l'aide d'un modèle d'étalonnage ;
- Les méthodes « directes ou empiriques » dont l'analyte est défini par la méthode elle-même. On mesure alors la concentration des échantillons inconnus à l'aide d'une autre méthode de mesure, comme une pesée ou une titrimétrie. Parmi ces méthodes, dites critères ou de type I dans le Codex Alimentarius, on trouve l'azote organique total de Kjeldahl, l'extrait sec, les lipides totaux, etc.

Il est fondamental que la quantité mesurée lors de la validation soit bien la même que celle qui sera mesurée en routine. En particulier, si le mode opératoire prévoit d'exprimer le résultat final à partir d'un seul mesurage, il ne faut pas exprimer le résultat d'un essai de validation comme la moyenne de plusieurs répétitions.

1.2 Préciser des objectifs de la validation

Choisir le domaine de validation

Définir le domaine de validation de la méthode, sous la forme d'une gamme de concentrations absolues ou relatives, par exemple, entre 1 g et 60 g ou entre 1 et 500 µg/l. Pratiquement, c'est à travers les choix effectués pour construire les plans d'expériences, en termes de gamme de concentrations pour un ou plusieurs types de matrice, qu'on démontre dans quel domaine la méthode est effectivement valide, et si elle est capable de fournir des résultats acceptables. Le domaine où cette démonstration est faite est appelé **domaine de validité** et peut être plus petit que le **domaine d'application** défini *a priori*.

Le choix du domaine d'application peut correspondre à une obligation légale. Par exemple, dans le domaine pharmaceutique, il faut prendre en compte les domaines minimums suivants :

- pour les substances pures ou les spécialités pharmaceutiques, le domaine doit au moins couvrir entre 80 et 120 % de leurs valeur nominales ;
- pour une vérification de concentration continue, le domaine doit s'étendre entre 70 et 130 % de la concentration attendue.

Si, lors de l'utilisation en routine de la méthode on rencontre des échantillons dont la concentration n'est pas incluse dans le domaine de validité, il est interdit d'extrapoler. Il faut alors procéder à une étude complémentaire pour étendre ce domaine ou procéder à une dilution dans la mesure où on a démontré qu'elle n'entraîne pas d'effet. Une mauvaise définition du domaine d'application peut avoir des conséquences importantes sur la validité, surtout en termes de spécificité, d'interférences ou de réactions croisées

Les limites d'acceptabilité sont notées $\pm\lambda$, lorsqu'elles sont exprimées de façon absolue dans la même unité que le mesurande, ou $(1 \pm \lambda) \times 100$ de façon relative. Dans la mesure du possible, on va définir les limites d'acceptabilité, en se référant à un document, une pratique de la profession ou une exigence scientifique. La liste des approches possibles qui sont présentées ci-dessous n'est pas exhaustive et n'a qu'une valeur d'exemple. Cependant, il n'entre pas dans le champ d'application de ce texte de fixer des valeurs pour λ .

- **Législation.** Plusieurs textes réglementaires fournissent des valeurs qui peuvent servir de base pour fixer λ . Par exemple, la Décision de la commission 2002/657 indique dans son tableau 2 la justesse minimale des méthodes quantitatives pour les médicaments et résidus de médicaments vétérinaires dans les produits d'origine animale (Décision de la commission 2002/657 (CE) du 12 août 2002). D'autres documents publiés par la Commission du Codex Alimentarius définissent une « approche critère » pour différents analytes et matrices (FAO, 1997 ; CX/MAS 09/30/7) qui devrait clarifier la situation au niveau international.
- **Pratiques de la profession.** Certaines professions ont défini des limites d'acceptabilité sous la forme d'écart maximum acceptable (EMA), comme cela se pratique pour le contrôle de qualité ou la métrologie légale. C'est une approche classique en biologie médicale pour les paramètres du sang ou dans l'industrie pharmaceutique pour le contrôle des spécialités pharmaceutiques.
- **Modèle empirique.** Il existe des modèles mathématiques qui relient des performances analytiques à la concentration. Le plus connu est celui développé par Horwitz vers 1980 ; il est repris dans de nombreux textes et permet de prédire la valeur attendue de l'écart-type de reproductibilité s_R en fonction d'un niveau de concentration X (en particulier : décision de la commission 2002/657 (CE) du 12 août 2002 ; CX/MAS 09/30/7). Pour une concentration inférieure à 10^{-7} kg.kg⁻¹, il s'exprime sous la forme,

$$s_R = 0,02 \cdot X^{0,85} \quad (1)$$

- **À partir de performances connues.** Par ailleurs, si on considère que la reproductibilité limite représente un écart maximal acceptable (EMA) entre 2 mesures faites en condition de reproductibilité, pour un niveau de confiance de 95%, celui-ci vaut $|EMA| \leq 2,83s_R$ et il vaudra $|EMA| \leq 4,24s_R$ pour un niveau de 99%. En combinant cette définition à l'équation (1) on peut définir la limite d'acceptabilité à partir de l'EMA et du modèle Horwitz :

$$|EMA| \leq 0,057 \times X^{0,85} \text{ ou } |EMA| \leq 0,085 \times X^{0,85}$$

ou exprimé en %.

$$EMA\% \leq \frac{0,057 \times X^{0,85}}{X} \times 100 \text{ ou } EMA\% \leq \frac{0,085 \times X^{0,85}}{X} \times 100$$

Dans le cas où on dispose d'une estimation expérimentale de l'écart-type de reproductibilité – par exemple si la méthode a été soumise à une étude inter-laboratoires – on peut affiner les limites d'acceptabilité en utilisant la valeur observée de s_R .

- **À partir de la limite de quantification.** Lorsque l'objectif de la méthode est l'application d'une norme sanitaire, il est possible de déduire des limites d'acceptabilité, à partir de la limite de quantification calculée avec le profil d'exactitude (Feinberg M., 2010c). La décision de validation se fera alors en fonction du risque encouru, obtenu *a posteriori*.

Il est souvent plus commode que les limites d'acceptabilité soient exprimées sous la forme d'un pourcentage mais ce n'est pas une obligation. De même, pour simplifier, on choisit une valeur constante pour tout le domaine d'application. Mais, si le domaine d'application est large, on peut utiliser des valeurs variant en fonction du niveau de concentration.

Dans le cas où les valeurs de référence sont connues avec une incertitude élevée, le choix des limites d'acceptabilité devra être compatible avec cette incertitude.

1.3 Sélectionner les échantillons de validation

Choisir la ou les matrices

Les essais qui servent à calculer les éléments du profil d'exactitude doivent être réalisés sur des échantillons clairement identifiés, appelés **échantillons de validation**. Ce sont des matériaux qui doivent être le plus représentatif possible du domaine d'application de la méthode. Ils peuvent ainsi inclure différentes matrices pour les méthodes multi-matrices. Mais surtout on doit être capable de leur attribuer une **valeur de référence** qui représente la teneur attendue. Penser à choisir des matériaux stables et homogènes en quantité suffisante pour réaliser l'ensemble des essais définis par le plan d'expérience de validation.

Établir les valeurs de référence des échantillons de validation

Pour estimer la justesse de la méthode, il faut disposer d'échantillons de validation dont la concentration est connue le plus exactement possible avec une incertitude connue. Cette concentration correspond à la **valeur de référence** assignée à l'échantillon de validation et doit être fixée indépendamment de la méthode à valider, on la note X . Il existe plusieurs approches possibles pour établir la valeur de référence assignée à un échantillon de validation, parmi lesquelles :

- utiliser des matériaux de référence certifiés (MRC), externes (MRE) ou internes (MRI) ; la traçabilité de ces matériaux décroît en fonction de leur nature ;
- réaliser des ajouts dosés à partir d'une molécule étalon de pureté connue ;
- préparer des échantillons dopés ou synthétiques avec la matrice choisie ;
- utiliser une méthode de référence, appliquée en interne ou en externe par un laboratoire pair. Si la valeur de référence est estimée par des répétitions de la méthode de référence, il conviendra de sélectionner la valeur la plus représentative, comme la moyenne ou la médiane. Cette approche doit être préférentiellement appliquée dans le cas de la validation d'une méthode alternative par rapport à une méthode de référence ;
- utiliser de la dilution isotopique lorsqu'on dispose de l'analyte sous une forme marquée.

Pour la préparation d'échantillons de validation par la méthode des ajouts dosés de niveaux variés, veiller à réduire le plus possible la dépendance entre les différentes préparations. Les ajouts dosés peuvent être effectués sous une forme liquide (par ajout d'une solution étalon) ou solide (par pesée). Si on procède par des dilutions successives (1/2, 1/4, 1/8...) à partir d'une même solution mère, on crée une corrélation entre les mesures et on augmente les conséquences d'une erreur de justesse ; il est donc déconseillé de procéder de cette façon. Cette remarque s'applique aussi à la préparation des solutions étalons.

Le choix raisonné des valeurs de référence revêt un aspect fondamental pour l'application du profil d'exactitude car il conditionne la justesse de la méthode. En fonction de la technique analytique choisie, toute approche visant à cet objectif est acceptable, aussi longtemps qu'une justification scientifique fondée est apportée.

Lorsqu'on est dans l'incapacité d'assigner une valeur de référence à un échantillon de validation, il est possible de la remplacer par la moyenne ou la médiane des concentrations obtenues directement sur cet échantillon. La valeur de référence ainsi obtenue n'est plus indépendante de la méthode en cours de validation et interdit l'évaluation du biais de justesse. Cette approche très restrictive doit être clairement justifiée dans le dossier de validation.

1.4 Planifier les essais de validation

Organiser les essais de validation

Le plan d'expérience de validation sert à estimer, dans les conditions où le mode opératoire sera appliqué en routine, quelles seront les performances en routine de la méthode. Dans ce contexte, un essai consiste à faire un mesurage sur un échantillon de validation, de préférence dont la valeur de référence a été établie avec une incertitude connue. Les notations sont rappelées à la Partie 3.

Pour réaliser ce plan, prévoir :

- I séries de mesures ($1 \leq i \leq I$) ;
- pour chaque série, effectuer J répétitions ($1 \leq j \leq J$) ;
- K niveaux de concentration ($1 \leq k \leq K$) couvrant le domaine d'application de la méthode. Dans ce contexte, on entend par un niveau une valeur de référence.

Construire un tableau qui rassemblera les données brutes selon le modèle du **tableau 1**. Les valeurs de référence X_k des échantillons de validation sont exprimées en quantités absolues (mg, g, log,...) ou relatives (mg/kg, µg/kg, mg/l,...). Pour les méthodes indirectes, la réponse Y est notée dans une unité correspondant à celle de la méthode instrumentale (surface du pic, hauteur de pic, absorbance,...). Pour les méthodes directes, elle est directement exprimée dans la même unité que celle des valeurs de référence.

Tableau 1 : organisation des mesures du plan de validation

Niveaux	Séries (Jours)	Concentrations des échantillons de validation (valeurs de référence) X	Mesurages (Réponses instrumentales) Y			
			1	2	...	J
1	1	x_{11}	y_{111}	y_{121}	...	y_{1J1}
			
	I		...			
2	1					
	...					
	I					
...				
K	1					
				
	I	x_{IK}				y_{IJK}

Choisir le nombre de séries, de répétitions et de niveaux de validation

L'influence du nombre de séries I est discutée au § 6.3.2 de la norme ISO 5725-1 à propos du choix du nombre de laboratoires pour une étude inter-laboratoires (NF ISO 5725-2:1994). D'ailleurs, à partir de ces données, l'International Union for Pure and Applied Chemistry (IUPAC) a décidé que I doit être supérieur ou égal à 8 pour qu'une étude inter-laboratoires soit reconnue comme recevable au plan scientifique. Dans le cas d'une validation interne, il semble irréaliste pour des raisons économiques de maintenir une telle exigence. C'est pourquoi, on peut choisir les exigences minimales suivantes :

- le nombre de séries I égal ou supérieur à 3. Une série peut être représentée par un jour mais aussi par une combinaison de diverses sources d'incertitude, comme plusieurs appareils, plusieurs opérateurs et plusieurs jours.
- le nombre **constant** de répétitions par série et par niveau J égal ou supérieur à 2.
- le nombre de niveaux de concentration K égal ou supérieur à 3. Il est indispensable d'avoir $K \geq 3$ car on pourra alors vérifier la linéarité entre les concentrations des valeurs de référence et les concentrations retrouvées : on a besoin de 3 niveaux pour conduire cette vérification. Mais, lorsqu'il est nécessaire de valider la méthode à proximité de sa limite de quantification LQ , il est conseillé de choisir K égal ou supérieur à 4 (Feinberg M., 2010c).

Si on a choisi $I = 3$, $J = 2$ et $K = 3$, le plan de validation comporte 18 essais. À part ces limites minimales, une grande flexibilité est laissée dans le choix du nombre d'essais.

Mais il faut se rappeler que plus le nombre d'essais est élevé, meilleures sont les estimations des critères de validation, les intervalles de tolérance resserrés et la confiance qu'on peut avoir dans la validation augmentée. Cependant, pour améliorer les choses, il est généralement préférable d'augmenter le nombre de séries plutôt que le nombre de répétitions ou de niveaux.

1.5 Planifier l'étalonnage (pour les méthodes indirectes)

Organiser les essais d'étalonnage

Le plan d'étalonnage a pour objectif de permettre l'estimation des coefficients du modèle de la courbe d'étalonnage². Il n'est obligatoire que pour les méthodes indirectes qui requièrent un étalonnage pour réaliser la quantification. Dans le cas des méthodes directes (comme la cellulose brute) la quantification se fait sans étalonnage, souvent par une pesée, l'analyte étant défini par la méthode. Le dispositif expérimental utilisé (nombre de niveaux d'étalonnage, de répétitions par niveau et l'étendue du domaine d'étalonnage) doit refléter le plus exactement possible ce qui sera fait en routine. L'étendue de la gamme d'étalonnage n'est pas obligatoirement confondue avec le domaine d'application de la méthode. Pour réaliser ce plan, on va choisir les paramètres minimums suivants :

- prévoir I séries ($1 \leq i \leq I$). Ce nombre doit absolument être le même que pour le plan de validation. Par exemple, si le jour se confond avec la série, les mesurages d'étalonnage doivent être effectués les mêmes jours que les mesurages de validation. On calculera ainsi un modèle d'étalonnage pour chaque jour et non pas un modèle global ;
- pour chaque niveau, effectuer J' mesurages répétés ($1 \leq j \leq J'$). Le nombre de répétitions J' peut être différent du nombre J choisi pour le plan de validation ;

² Sous l'influence de la littérature anglo-saxonne, de nombreux analystes français utilisent le terme « calibration » à la place d'étalonnage.

- pour chaque série prévoir K' niveaux ($1 \leq k \leq K'$) étalons de concentrations connues, couvrant la gamme d'étalonnage. Le nombre de niveaux K' peut être différent du nombre K choisis pour le plan de validation. Par exemple, on peut choisir $K' = 1$ si on est sûr que la courbe d'étalonnage est une droite qui passe par le 0.

Construire le **tableau 2** qui résume l'ensemble des essais et mesurages à effectuer. Dans ce tableau les concentrations des étalons sont exprimées en quantités absolues (mg, g,...) ou relatives (mg/kg, µg/kg, mg/l,...). Après réalisation, les réponses sont notées dans les unités de la méthode instrumentale (surface du pic, hauteur de pic, absorbance,...).

Choisir le nombre de séries, de répétitions et de niveaux d'étalonnage

La valeur souhaitable pour K' dépend du type de fonction de réponse instrumentale attendue, comme l'indique le **tableau 3**. La règle préconisée consiste à prendre au moins un niveau de plus qu'il n'y a de coefficients dans le modèle d'étalonnage. Cette contrainte est liée à l'utilisation d'une méthode de régression aux moindres carrés pour estimer les coefficients du modèle (Azais J.M., Bardet J.M., 2006 - Huet S. *et al.*, 2004). Cependant, on peut aussi utiliser la valeur minimale qui figure dans le même tableau. Les K' étalons doivent permettre une bonne estimation statistique des paramètres de la fonction de réponse. La méthode des moindres carrés, qui sert aux calculs, est très sensible au dispositif expérimental et il faut veiller à ce point.

Tableau 2 : organisation des essais du plan d'étalonnage

Niveaux des étalons	Séries (Jours)	Concentrations des échantillons de validation (valeurs de référence) X	Mesurages (Réponses instrumentales) Y			
			1	2	...	J'
1	1	x_{11}	y_{111}	y_{121}	...	$y_{1J'1}$
			
	I		...			
2	1					
	...					
	I					
...				
K'	1					
				
	I	$x_{IK'}$				$y_{IJ'K'}$

Le nombre de répétitions J' pour chaque étalon dans chaque série est laissé au choix. Cependant, il est recommandé de prendre $J' \geq 2$. Lors de l'utilisation en routine, on pourra modifier le mode opératoire normalisé de la méthode en fonction des conditions démontrées comme étant capables de fournir des résultats acceptables.

Si la méthode nécessite un étalonnage, il faut veiller à ce qu'aucune des réponses y obtenues sur les échantillons de validation ne dépasse la réponse la plus élevée ou la plus petite, recueillie dans le tableau 1. En d'autres termes, il ne faut pas extrapoler le modèle en-dehors des concentrations où il a été construit.

1.6 Réaliser les essais

Il est impératif de réaliser les mesures pour le plan de validation **en appliquant la méthode telle qu'elle sera utilisée en routine**. En particulier, appliquer au mieux les conditions d'étalonnage même si celles-ci pourront être modifiées au vu des résultats obtenus par le profil d'exactitude. Par contre, le nombre de « répétitions » utilisées pour exprimer un résultat final doit être fidèlement respecté. Par exemple, si chaque résultat final est exprimé comme la moyenne de 2 répétitions, chaque essai doit être fait selon cette méthode.

Dans le cas où des effets de matrice ont été mis en évidence lors du développement de la méthode, on peut être amené à réaliser plusieurs plans de validation pour calculer un facteur de correction qui sera appliqué en routine (sauf interdiction réglementaire). Il faut alors effectuer deux plans de validation sur deux échantillons indépendants et utiliser les résultats du premier pour calculer le facteur de correction qui sera appliqué au second plan (Feinberg M., 2010c)

Le rôle du profil d'exactitude est d'estimer, à partir des résultats obtenus lors de la validation, quelle garantie aura l'utilisateur que la méthode utilisée en routine fournira des résultats acceptables. C'est pourquoi, la réalisation des plans doit respecter les conditions suivantes :

- **Exigence d'être en condition de fidélité intermédiaire ou de reproductibilité.** Les mesurages doivent être réalisés en condition de fidélité intermédiaire ou de reproductibilité, avec pour objectif de prendre en compte le plus de sources d'incertitude possible. Comme ces sources d'incertitude dépendent de la méthode et des matrices analysées, il est impossible de définir une stratégie unique. Il est recommandé de construire un diagramme causes/effets pour choisir les conditions les plus globales. Dans la plupart des cas, on pourra décider de faire les mesurages sur plusieurs jours dans la mesure où on prend ainsi en compte les incertitudes dues à l'appareillage (réglages, et préparation des réactifs), à l'étalonnage, au personnel, à la conservation de l'échantillon... Mais pour des échantillons très instables, on pourra réaliser les mesurages sur un seul jour avec plusieurs opérateurs ou instruments.
- **Exigence de couvrir le domaine d'application.** Une étude de validation doit couvrir l'ensemble du domaine d'application de la méthode. Pour satisfaire ces deux contraintes, plusieurs choix sont possibles et, dans tous les cas, avant de commencer il faut disposer de quantités suffisantes d'échantillons pour faire tous les mesurages :
 - utiliser un seul échantillon dont on sait faire varier le niveau de concentration, par exemple au moyen d'ajouts dosés (approche de choix pour une méthode indirecte) ;
 - utiliser plusieurs échantillons (ou même plusieurs matrices proches) présentant des niveaux de concentration variés (généralement le cas pour une méthode directe) ;
 - préparer des matrices synthétiques (matières premières ou produits chimiques) ;
 - utiliser des matériaux de référence externes ou certifiés...
- **Exigence de synchroniser les plans.** Dans le cas, où deux plans sont nécessaires – étalonnage et validation – il est indispensable de réaliser les essais dans les mêmes conditions de fidélité intermédiaire. Par exemple, si le jour est la source de variation choisie, il faut réaliser des mesurages d'étalonnage et de validation pour un même jour donné. Les données d'étalonnage de ce jour serviront à prédire les concentrations à partir des données de validation faites le même jour.

1.7 Calculer les concentrations prédites inverses [méthodes indirectes]

Avant tout traitement numérique, il est recommandé de procéder à une illustration graphique préalable et à un examen visuel des données afin de détecter une erreur flagrante, comme une donnée mal enregistrée. Ces graphiques peuvent être mis en annexe du dossier d'évaluation. La présence d'une donnée anormale amène à suspecter une mauvaise mise au point de la méthode et peut invalider la procédure de validation.

Calculer les modèles d'étalonnage

Pour les méthodes directes, cette étape de calcul n'existe pas.

Pour les méthodes indirectes, il est nécessaire d'exprimer la réponse instrumentale Y en fonction des concentrations x des étalons, à l'aide d'un modèle mathématique f de la forme :

$$Y = f(x) \quad (2)$$

Les fonctions f classiquement utilisées sont regroupées au **tableau 3** mais cette liste n'est pas limitative. Les paramètres a_1, a_2, \dots sont appelés les paramètres du modèle.

Calculer les paramètres du modèle d'étalonnage à partir des données recueillies pour chaque série k , de façon à obtenir k ensembles de valeurs des paramètres. Le même type de modèle doit être utilisé pour l'ensemble des données, quelle que soit la série, mais les valeurs des paramètres peuvent être différentes d'une série à l'autre. Cette approche permet de prendre en compte les variations inter-séries (inter-jours) observées.

Le calcul des estimations des coefficients du modèle d'étalonnage peut faire appel aux diverses techniques statistiques classiques détaillées dans la littérature :

1. Régression par la méthode des moindres carrés (AFNOR, 1996)
2. Régression pondérée (Azaïs J.M., Bardet J.M., 2006)
3. Régression non linéaire (Huet S., et al., 2004)

Tableau 3 : principales fonctions de réponse utilisables pour un étalonnage

Type	Équation	Souhaitable pour K'	Minimum pour K'
Droite passant par l'origine	$Y = a_1x$	2	1
Droite	$Y = a_0 + a_1x$	3	2
Fonction quadratique	$Y = a_0 + a_1x + a_2x^2$	4	3
Fonction logistique à 4 paramètres	$Y = a_0 + \frac{a_3 - a_0}{1 + \left(\frac{a_2}{x}\right)^{a_1}}$	5	5

La méthode la plus accessible est celle des moindres carrés, car elle est disponible sur les tableurs. La régression pondérée s'applique dans les cas où les variances des réponses ne sont pas homogènes entre niveaux. Les estimations des paramètres du modèle pour chaque série sont regroupées dans le **tableau 4**.

Si le type de modèle d'étalonnage a été clairement établi lors du développement de la méthode, on pourra choisir le nombre minimum pour K' . Cependant, on aura intérêt à prendre le nombre souhaitable ou plus afin de tester si d'autres modèles plus complexes ne sont pas

mieux adaptés. En effet, à partir des données du plan d'étalonnage, il est possible d'estimer plusieurs modèles d'étalonnage puis de construire ainsi plusieurs profils d'exactitude. On pourra alors retenir le modèle qui fournit le profil le plus favorable. Il conviendra alors de modifier le mode opératoire en fonction du modèle choisi, sans pour autant modifier les limites d'acceptabilité λ .

Tableau 4 : estimations des paramètres du modèle en fonction de la série

Série	1	...	I
a_0			
a_1			
...			

Calculer les concentrations retrouvées par prédiction inverse

Pour les méthodes indirectes, les modèles d'étalonnage servent à calculer les concentrations retrouvées, à partir des données au plan de validation (Tableau 2), en utilisant la fonction inverse du modèle d'étalonnage, selon le modèle mathématique suivant :

$$\hat{x} = z = f^{-1}(Y) \quad (3)$$

La fonction inverse est appelée équation de prédiction inverse et les valeurs z ainsi obtenues sont appelées concentrations retrouvées. Le **tableau 5** fournit ces équations selon la fonction de réponse choisie au tableau 3.

Tableau 5 : équations de prédiction inverse selon le type de la fonction de réponse

Fonction de réponse	Fonction inverse de calcul de la concentration retrouvée
Droite passant par l'origine	$z = \frac{Y}{a_1}$
Droite	$z = \frac{Y - a_0}{a_1}$
Fonction quadratique	$z = \frac{-a_1 + \sqrt{a_1^2 - 4a_2(a_0 - Y)}}{2a_2}$
Logistique à 4 paramètres	$z = \frac{a_2}{\left(\frac{a_3 - a_0}{Y - a_0} - 1\right)^{\frac{1}{a_1}}}$

Utiliser les données du plan de validation et les paramètres des modèles d'étalonnage estimés à partir du plan d'étalonnage pour calculer les concentrations retrouvées dans les échantillons de validation. Effectuer ces calculs par série, avec le modèle de la série correspondante. Ce mode de calcul souligne le fait qu'il est indispensable que les essais des plans d'étalonnage et de validation soient synchronisés c'est-à-dire réalisés le même jour et/ou par le même opérateur et/ou par le même laboratoire.

1.8 Calculer les critères de validation

Calculer la justesse par série

Rassembler l'ensemble des données dans un tableau selon le modèle du Tableau 6. Au cas où différents modèles d'étalonnage sont testés, construire autant de tableaux qu'il y a de modèles. Utiliser les valeurs de z pour calculer les écarts-types de répétabilité, inter-séries et de fidélité intermédiaire (ou de reproductibilité). Ce calcul s'effectue indépendamment pour chaque niveau de concentration k (avec $1 \leq k \leq K$) selon le principe de la norme ISO 5725-2, comme il est décrit dans Feinberg M. (2010c). Pour certaines mesures exprimées en logarithmes, comme les comptages microbiologiques ou unité formant colonie (ufc), exprimés en \log_{10} (ufc), il n'est pas utile de calculer les valeurs relatives.

Tableau 6 : concentrations retrouvées par prédiction inverse obtenues à partir des données du plan de validation

Niveau	Série	Répétition	Concentration théorique X	Concentration retrouvée Z	Biais absolu	Biais relatif
1	1	1	x_{111}	z_{111}	b_{ijk}	$b_{ijk}\%$
1	1	2				
1		...				
...		J				
1	I	1				
1	I	2				
1	I	...				
1	I	J				
K	I	J	x_{IJK}	z_{IJK}	b_{IJK}	

Dans le **tableau 6**, il est possible de calculer un ou plusieurs critères exprimant la justesse (sous la forme d'un biais). Ces critères sont les suivants :

$$\text{Biais absolu : } b_{ijk} = z_{ijk} - x_{ijk} \quad (4)$$

$$\text{Biais relatif : } b_{ijk}\% = \frac{z_{ijk} - x_{ijk}}{x_{ijk}} \times 100 \quad (5)$$

$$\text{Recouvrement} = \frac{z_{ijk}}{\bar{x}_k} \times 100 \quad (6)$$

Calculer la justesse et la fidélité par niveau

Rassembler avec les critères de justesse, calculés pour un niveau et non plus pour un mesurage individuel, sous la forme d'un tableau, comme le **tableau 7**.

Dans les cas très particuliers où la justesse n'est pas exigée, la valeur de référence moyenne peut être fixée comme étant égale à la concentration retrouvée moyenne. Il est alors inutile de calculer un biais moyen ou un taux de recouvrement moyen.

Tableau 7 : critères de fidélité et de justesse par niveau

Critères	Symbole	Niveaux		
		1	...	K
Valeur de référence moyenne	$\bar{\bar{x}}_k$			
Concentration retrouvée moyenne	$\bar{\bar{z}}_k$			
Écart-type de répétabilité	s_{kr}			
Écart-type inter-séries	s_{kB}			
Écart-type de fidélité intermédiaire	$s_{kFI} = \sqrt{s_{kr}^2 + s_{kB}^2}$			
Coefficient de variation de la fidélité intermédiaire	$\frac{s_{kFI}}{\bar{\bar{x}}_k} \times 100$			
Biais moyen absolu	$\bar{\bar{z}}_k - \bar{\bar{x}}_k$			
Biais moyen relatif	$\left(\frac{\bar{\bar{z}}_k - \bar{\bar{x}}_k}{\bar{\bar{x}}_k} \right) \times 100$			
Taux de recouvrement moyen	$\frac{\bar{\bar{z}}_k}{\bar{\bar{x}}_k} \times 100$			

Calculer les intervalles de tolérance

La méthode de calcul proposée par Mee (1984) est celle qui a été choisie pour cette procédure. Elle a aussi été adoptée par une commission de la Société des sciences et techniques pharmaceutiques (SFSTP - 2003, 2006). Le calcul se fait à partir des données du **tableau 7**, indépendamment pour chaque niveau de concentration k . Pour des raisons de simplification des formules, l'indice k est omis dans les formules suivantes.

Exprimer l'intervalle de tolérance comme un intervalle symétrique autour de la concentration retrouvée moyenne $\bar{\bar{z}}$ du niveau :

$$\bar{\bar{z}} \pm k_{tol} \times s_{IT} \quad (7)$$

Calculer l'écart-type de l'intervalle de tolérance s_{IT} selon la formule (8) ; il servira aussi à l'estimation de l'incertitude de mesure. Il est obtenu grâce aux formules suivantes :

$$s_{IT} = s_{FI} \left(\sqrt{1 + \frac{1}{I \times J \times B^2}} \right) \quad (8)$$

$$B = \sqrt{\frac{R+1}{J \times R+1}} \quad (9)$$

$$R = \frac{s_B^2}{s_r^2} \quad (10)$$

La quantité k_{tol} est appelé **facteur de couverture de l'intervalle de tolérance** et vaut :

$$k_{tol} = t_{v; \frac{1+\beta}{2}} \quad (11)$$

$t_{v; \frac{1+\beta}{2}}$ est le quantile de la distribution t de Student pour v degrés de liberté et β la probabilité du contenu de l'intervalle de tolérance. Le nombre de degrés de liberté v est calculé selon méthode d'approximation et donne :

$$v = \frac{(R+1)^2}{\left(R + \frac{1}{J}\right)^2 + \frac{1 - \frac{1}{J}}{IJ}} \quad (12)$$

Pour chaque niveau k , calculer l'écart-type de l'intervalle de tolérance s_{IT} , le facteur de couverture k_{tol} et les limites basse et haute de l'intervalle de tolérance. La valeur choisie pour β doit être au moins de 80%. L'ensemble des calculs est rassemblé dans un tableau de la forme du **tableau 8**.

Tableau 8 : résumé des différents éléments qui serviront au profil d'exactitude, incluant les limites des intervalles de tolérance par niveau

Critères	Symbole	Niveaux		
		1	...	K
Valeur de référence moyenne	$\bar{\bar{x}}$			
Limite de tolérance basse	$\bar{\bar{z}} - k_{tol} \times s_{IT}$			
Limite de tolérance haute	$\bar{\bar{z}} + k_{tol} \times s_{IT}$			
Limite de tolérance basse relative	$\left(\frac{\bar{\bar{z}} - k_{tol} \times s_{IT}}{\bar{\bar{x}}}\right) \times 100$			
Limite de tolérance haute relative	$\left(\frac{\bar{\bar{z}} + k_{tol} \times s_{IT}}{\bar{\bar{x}}}\right) \times 100$			
Limite d'acceptabilité basse relative	$(1 - \lambda) \times 100$			
Limite d'acceptabilité haute relative	$(1 + \lambda) \times 100$			

Le nombre v (Eq.(12)) est rarement un nombre entier et il convient d'utiliser des tables de la loi de Student qui acceptent des nombres de degrés de liberté non entiers. Cependant, il est possible d'approcher la valeur du quantile par interpolation linéaire entre les 2 nombres de degré de liberté entiers qui encadrent v . Ce calcul approché est illustré dans la feuille de calcul (dans les cellules B30 à B32) présentée par Max Feinberg (2010c).

Le rapport $R = \frac{s_B^2}{s_r^2}$ entre la variance inter-séries et la variance de répétabilité intervient dans

plusieurs formules. Il traduit l'importance relative de l'effet de la série. Par exemple, dans le cas où l'effet jour est l'effet série, si la méthode est très stable d'un jour à l'autre et est capable de fournir des résultats très proches pour un même échantillon, R est proche de 1 et ne jouera qu'un rôle secondaire. Par contre, si ce rapport augmente le nombre de degrés de

liberté est diminué et, plus le nombre de degrés de liberté est petit, plus le quantile de la distribution t de Student est élevé et élargit l'intervalle de tolérance.

1.9 Construire le profil d'exactitude

Le profil d'exactitude peut être construit de différentes façons, en fonction du type de données traité. La méthode la plus classique, lorsqu'on a à faire à des concentrations relatives, est celle présentée à la **figure 1** où les performances sont exprimées de façon relative, par un taux de recouvrement. Mais si les données sont des comptages exprimés en logarithmes, il vaut mieux exprimer les performances, comme des différences entre 2 logarithmes, ce qui est équivalent à un rapport.

Pour construire le profil d'exactitude, sélectionner dans le **tableau 8**, les lignes suivantes :

- à reporter sur l'axe horizontal
 - 1) Les valeurs de référence moyennes.
- à reporter sur l'axe vertical
 - 2) Les limites de tolérance basses relatives ;
 - 3) Les limites de tolérance hautes relatives ;
 - 4) Les taux de recouvrement moyens ;
 - 5) Les limites d'acceptabilité basses relatives ;
 - 6) Les limites d'acceptabilité hautes relatives.

Reporter ces données sur un graphique en utilisant les valeurs de référence moyennes pour dessiner l'axe des abscisses, comme l'illustre la **figure 1**.

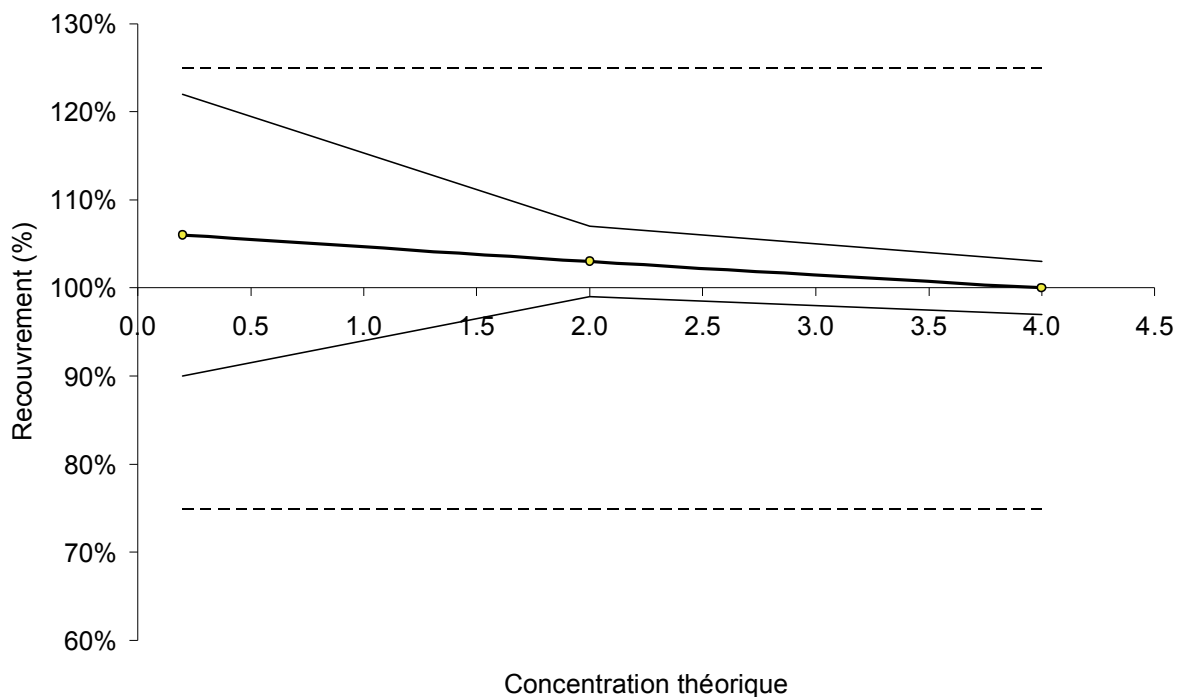


Figure 1 : profil d'exactitude exprimé par le taux de recouvrement

1.10 Interpréter le profil d'exactitude pour la validation

Pour utiliser le profil d'exactitude en vue de valider une méthode, il faut avoir fixé les deux critères de décision suivants :

- **les limites d'acceptabilité $\pm\lambda$.** Elles servent à traduire les objectifs pratiques des utilisateurs. Elles délimitent un intervalle autour de la valeur de référence. Le plus souvent, ces limites sont réglementaires ou issues de la réglementation. Mais dans le cas où il n'existe pas de référence établie, il convient de prendre en compte les attentes des utilisateurs finaux, comme une LQ donnée (voir § 1.2) ;
- **la proportion β .** Elle représente la proportion de futurs résultats qui seront en moyenne compris dans les intervalles de tolérance. La valeur choisie pour β dépend largement du champ d'application (contrôle sanitaire, contrôle de fabrication, etc.). Il est évident que plus β est petit, par exemple 70 %, plus la méthode risque de produire des résultats qui ne correspondent pas aux spécifications annoncées. C'est pourquoi, dans la méthode du profil d'exactitude cette proportion a été fixée à 80 %, au moins.

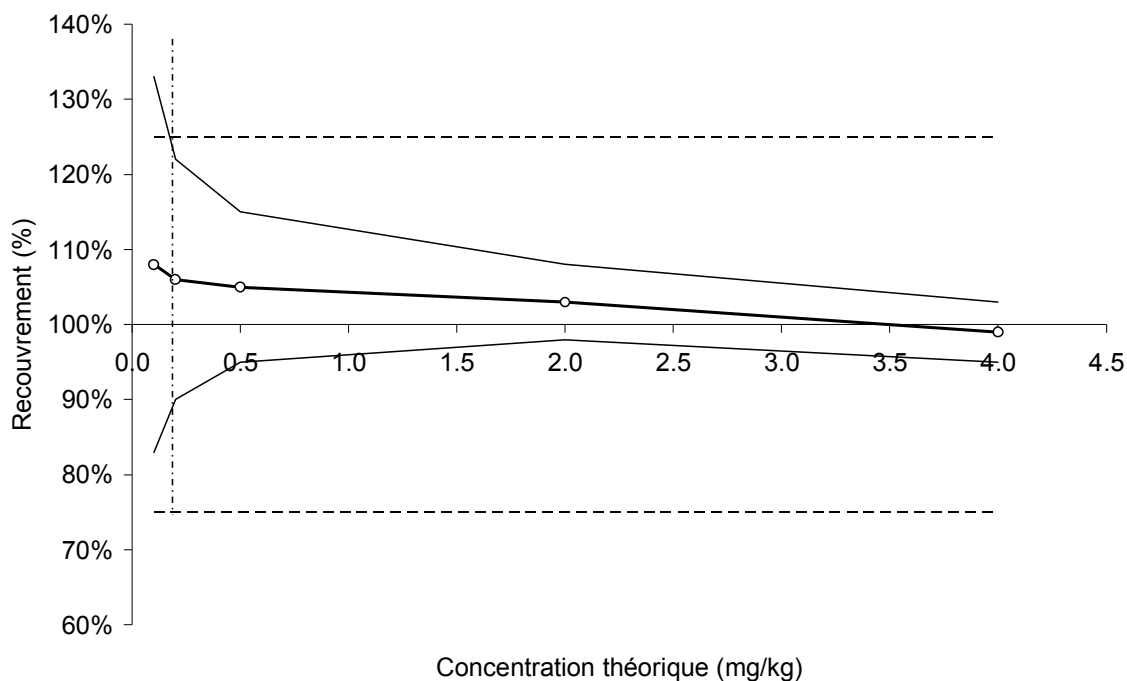


Figure 2 : exemple d'une méthode validée sur une zone réduite du domaine de validation

Règles de décision

Reporter les limites d'acceptabilité sur le graphique du profil d'exactitude pour permettre une interprétation visuelle directe des résultats. La **figure 2** illustre une situation classique où la majeure partie des intervalles de tolérance est comprise dans l'intervalle d'acceptabilité. Dès que l'intervalle de tolérance sort de l'intervalle d'acceptabilité, on peut conclure que la méthode n'est plus capable de fournir suffisamment de résultats acceptables, en fonction des choix faits au départ de l'étude. Par exemple, pour $\beta = 80\%$ et une limite d'acceptabilité de $\pm 25\%$, on peut conclure que la méthode est valide entre environ 0,2 mg/kg et 4,0 mg/kg. En outre, ce graphique fournit d'autres indications. En particulier, sur la **figure 2**, il apparaît que

la justesse varie avec la concentration. Le taux de recouvrement qui traduit la justesse est d'environ 108 % aux basses concentrations pour être proche de 100 % aux concentrations élevées. À l'aide des résultats statistiques regroupés dans le tableau construit sur le modèle du tableau 7, on peut trouver différents éléments qui expliquent le comportement de ce biais de la méthode.

Dans le cas illustré par la **figure 3**, la méthode présente un biais systématique élevé : le taux de recouvrement est seulement d'environ 70 % ; dû par exemple à un effet de matrice non contrôlé. Il faut alors conclure que la méthode n'est pas valide dans le domaine étudié. Si les pratiques de la profession ou la réglementation le permettent, il est alors possible d'appliquer un facteur de correction qui viendra corriger ce biais. La procédure à appliquer pour calculer et valider ce facteur de correction est décrite par Max Feinberg (2010c).

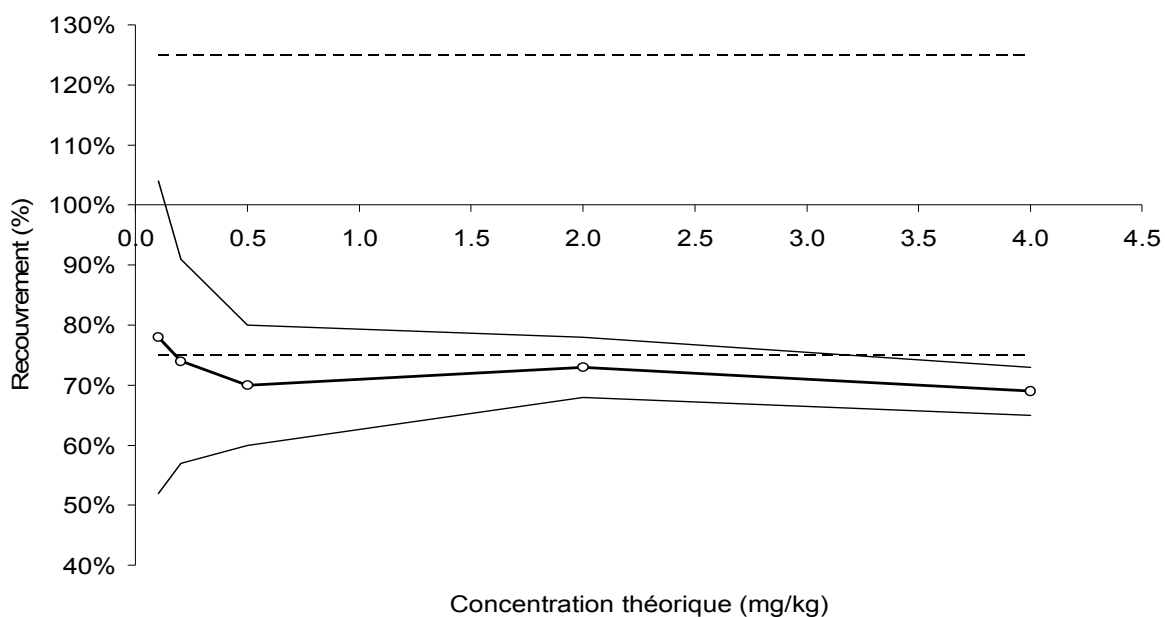


Figure 3 : exemple d'un profil d'exactitude pour une méthode présentant un biais systématique important, caractérisé par un taux de recouvrement d'environ 70 %

Définir du domaine de validité

Le domaine de validité est déterminé par la zone du domaine de validation dans laquelle la méthode fournit une proportion de résultats acceptables au moins égale à β . Il est limité par une borne supérieure qui équivaut à la limite de quantification inférieure et une limite supérieure qui correspond à la limite de quantification supérieure.

Choisir la procédure d'étalonnage pour la routine

Dans la mesure où avec les méthodes indirectes on dispose de données d'étalonnage, il est possible de construire plusieurs profils d'exactitude avec le même jeu de données mais en utilisant différents modèles d'étalonnage, par exemple un modèle passant par le zéro et construit avec une seule solution étalon qui est comparé à un modèle linéaire complet calculé à partir de 3 solutions étalons. On peut alors sélectionner le profil le plus favorable ou la procédure d'étalonnage la plus simple qui, malgré tout, permet d'atteindre l'objectif fixé.

Identifier les points atypiques

La méthode de calcul proposée dans la méthode du profil d'exactitude exige que le plan d'expérience de validation soit équilibré, c'est-à-dire que le nombre de répétitions I par série soit constant. C'est pourquoi, elle ne prévoit pas de méthodes de détection et/ou de rejet de données atypiques (aberrantes).

Toutefois, si le responsable de la validation juge que le rejet de données atypiques est nécessaire après un examen visuel d'une représentation graphique, il a toute latitude pour mettre en œuvre la procédure qui lui semblera la mieux adaptée, par exemple, en s'appuyant sur les recommandations de la norme ISO 5725-2. De plus, il est toujours possible de recommencer les mesures pour un niveau de concentration et une série donnés.

Un trop grand nombre de points aberrants souligne la mise en œuvre médiocre de la méthode d'analyse et met en doute le bien-fondé de sa validation.

2. Transférer la méthode en routine

Finalement, lors de l'application en routine de la méthode ainsi validée, il convient de vérifier de façon régulière si les résultats obtenus restent acceptables, par exemple au moyen d'une carte de contrôle, en particulier, si on utilise un facteur de correction (Feinberg M., 2010c).

Des actions correctives inefficaces peuvent ainsi conduire à la revalidation de la méthode. La revalidation peut alors être complète ou partielle. Une revalidation partielle consistant à faire un plan d'expérience à un seul niveau de concentration pour vérifier si l'intervalle de tolérance obtenu reste bien entre les limites d'acceptabilité.

Bibliographie

AFNOR (1994) Norme NF ISO 5725-2, *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.*

AFNOR (1996) Norme NF ISO 11095 *Étalonnage linéaire utilisant des matériaux de référence*, Afnor, Paris

Azaïs J.M., Bardet J.M.; (2006) *Le modèle linéaire par l'exemple, Régression, analyse de la variance et plans d'expérience illustrés avec R, SAS et Splus*, Dunod, Paris

Commission européenne (2002) Décision de la commission 2002/657 du 12 août 2002 portant modalités d'application de la directive 96/23/CE du Conseil en ce qui concerne les performances des méthodes d'analyse et l'interprétation des résultats

Commission SFSTP, P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat ; Validation des procédures analytiques quantitatives : harmonisation des démarches.

Partie I. : *STP Pharma Pratiques* (2003) 13:3, 101-138

Partie II. : Statistiques, *STP Pharma Pratiques* (2006) 16:1, 30-60

Partie III : Exemples d'application, *STP Pharma Pratiques* (2006) 16:2, 87-121

FAO (1997) Manuel sur le contrôle de la qualité des produits alimentaires. 14 : Assurance de la qualité dans le laboratoire d'analyse chimique des aliments, Rome, <http://www.fao.org/docrep/T0845F/t0845f00.HTM>

- Feinberg M., (2009) *Labo-Stat : Guide de validation des méthodes d'analyse* Lavoisier tec&Doc, Paris
- Feinberg M. (2010c) Interprétation du profil d'exactitude in Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 45-59
- Huet S., Bouvier A., Poursat M.A., Jolivet E. (2004) *Statistical Tools for Nonlinear Regression* Springer Verlag, New-York, ISBN 0-387-40081-8
- Mee R.W. (1984) β -Expectation and β -Content Tolerance Limits for Balanced One-Way ANOVA Random Model *Technometrics* 26:3, 251-254
- Commission du Codex Alimentarius (2005) Manuel de procédure de la 15^{ème} édition, http://www.codexalimentarius.net/web/procedural_manual.jsp)
- CX/MAS 09/30/7, Directives pour établir des critères pour l'identification des méthodes d'analyse adéquates (Conversion des méthodes pour les éléments traces en critères) <ftp://ftp.fao.org/codex/>)

Interprétation du profil d'exactitude

Max Feinberg¹

Résumé : À partir des données collectées pour construire le profil d'exactitude, il est possible de calculer divers critères complémentaires souvent exigés pour les dossiers de validation, comme la limite de quantification ou l'incertitude. S'appuyant sur les concepts décrits dans d'autres articles consacrés à la validation des méthodes : principe et mise en œuvre du profil d'exactitude cet article présente différents compléments et, en particulier, le mode de calcul de la répétabilité et de la fidélité.

Mots-clés : Validation, profil d'exactitude, justesse, fidélité, incertitude, méthodes d'analyse quantitatives

Notations et symboles

Une série de mesurages i , (i variant de 1 à I) contient des niveaux de concentration k (k variant de 1 à K) pour lesquels des répétitions ont été faites j (j variant de 1 à J) les indices seront repris dans l'ordre i, j, k . Les variables aléatoires sont notées en majuscules et les valeurs qu'elles prennent en minuscules.

Symbole	Description
x_{ijk}	Valeur de référence assignée à un étalon pour la série i ($1 \leq i \leq I$), la répétition j ($1 \leq j \leq J$) et le niveau k ($1 \leq k \leq K$) ou Valeur de référence assignée à un échantillon de validation pour la série i , la répétition j et le niveau k .
y_{ijk}	Mesure de la réponse instrumentale ou expérimentale observée pour un étalon ou un témoin de validation pour la série i , la répétition j et le niveau k .
z_{ijk}	Valeur retrouvée dans l'échantillon de validation de la série i , pour la répétition j et le niveau k , obtenue soit par prédiction inverse à l'aide d'un modèle d'étalonnage, soit par une mesure directe.
b_{ijk}	Biais exprimant l'erreur de justesse pour un témoin de validation entre sa valeur retrouvée et sa valeur de référence $b_{ijk} = z_{ijk} - x_{ijk}$

1. Calcul des écarts-types de répétabilité, fidélité intermédiaire et reproductibilité

Les calculs sont conduits selon le principe décrit dans la norme ISO 5725-2 ou de sa version Afnor NF X 06-041. Ils doivent être effectués indépendamment niveau par niveau k , sur les concentrations retrouvées par prédiction inverse dans le cas des méthodes indirectes ou sur les concentrations retrouvées dans le cas des méthodes directes. Dans les deux cas, les

¹ UR1204 Met@risk - Méthodologie d'analyse de risque alimentaire – INRA – F-75231 Paris cedex 05

☎ 01 44 08 16 52 ✉ max.feinberg@paris.inra.fr

concentrations retrouvées sont notées z_{ijk} . Par convention, l'indice i ($1 \leq i \leq I$) représente le numéro de la série et j ($1 \leq j \leq J$) celui de la répétition dans la série. Toutes les séries doivent avoir le même nombre de répétitions, c'est-à-dire que J est le même quelque soit i : on dit alors que le **plan d'expérience est équilibré**. Pour simplifier les notations, l'indice k ($1 \leq k \leq K$) du niveau est supprimé dans les formules qui suivent.

Calculer pour chaque niveau k :

- l'écart-type de répétabilité, noté s_r ;
- l'écart-type inter-séries, noté s_B ;
- l'écart-type de fidélité intermédiaire $s_{FI} = \sqrt{s_r^2 + s_B^2}$.

1.1 Cas d'un plan équilibré

La norme ISO 5725 ne prend en compte que les plans d'expérience où le nombre de répétitions I par série est le même pour toutes les séries. Le nombre total de répétitions doit être égal à IJ . Par contre, ce total peut être différent pour chaque niveau.

Calculer la moyenne générale d'un niveau :

$$\bar{Z} = \frac{\sum_{i=1}^I \sum_{j=1}^J z_{ij}}{I \times J} \quad \text{Concentration moyenne du niveau (1)}$$

Calculer les écarts-types de répétabilité, inter-séries et de fidélité intermédiaire en effectuant l'analyse de la variance (ANOVA) d'un modèle à effet aléatoire, selon les principes décrits de la norme ISO 5725, à savoir une décomposition de la somme des carrés des écarts totale en deux sommes des carrés d'écarts :

$$\underbrace{\sum_{i=1}^I \sum_{j=1}^J (z_{ij} - \bar{Z})^2}_{SCE_t} = \underbrace{\sum_{i=1}^I \sum_{j=1}^J (z_{ij} - \bar{Z}_i)^2}_{SCE_r} + \underbrace{\sum_{i=1}^I J \times (\bar{Z}_i - \bar{Z})^2}_{SCE_B} \quad (2)$$

Cette équation est traditionnellement écrite sous une forme abrégée, faisant appel à trois sommes de carrés d'écarts (*SCE*).

$$SCE_t = SCE_B + SCE_r \quad \text{Équation générale de l'analyse de variance}$$

Où chacune des sommes est définie comme suit pour faciliter l'interprétation :

SCE_t	Somme totale des écarts à la moyenne générale du niveau
SCE_B	Somme des écarts inter-séries
SCE_r	Somme des écarts intra-série

Comme le propose la norme ISO 5725, il n'est pas nécessaire de développer les trois sommes de carrés pour effectuer les calculs ; le calcul de SCE_B se fait par différence. Cette méthode peut poser des problèmes si le résultat est négatif : forcer alors la valeur de SCE_B à 0.

$$SCE_B = SCE_t - SCE_r \text{ si } SCE_B > 0$$

$$SCE_B = 0 \text{ si } SCE_B \leq 0$$

Calculer la variance de répétabilité du niveau, à partir des répétitions z_{ij} :

$$s_r^2 = \frac{SCE_r}{I(J-1)} \quad \text{Variance intra-série (3)}$$

Calculer la variance inter-séries notée s_B^2 , comme suit :

$$s_B^2 = \frac{\frac{SCE_B}{I-1} - s_r^2}{J} \quad \text{Variance inter-séries (4)}$$

Finalement, calculer l'écart-type de fidélité intermédiaire du niveau :

$$s_{FI} = \sqrt{s_B^2 + s_r^2} \quad \text{Écart-type de fidélité intermédiaire (5)}$$

Voir le § 2.2 « Utilisation d'un logiciel » pour un exemple de calcul sous Excel.

1.2 Cas d'un plan déséquilibré

Dans la mesure où on ne prévoit pas de tests de rejet des valeurs atypiques mais qu'on laisse à l'utilisateur la liberté de supprimer des valeurs jugées atypiques, il est possible de déboucher sur un plan d'expérience déséquilibré. Cette situation ne correspond plus aux recommandations de la norme ISO 5725. Cependant, un algorithme alternatif est présenté pour éviter des pertes de données.

Si le nombre n_i de répétitions n'est pas le même pour toutes les séries d'un même niveau, utiliser la formule (6) au lieu de la (4) pour le calcul la variance inter-séries :

$$s_B^2 = \frac{(I-1) \left(\frac{SCE_B}{I-1} - s_r^2 \right)}{N^*} \quad (6)$$

$$\text{Avec } N^* = N - \frac{\sum_{i=1}^I n_i^2}{N} \quad \text{et } N = \sum_{i=1}^I n_i$$

Utiliser les formules classiques pour les autres critères de fidélité.

Une autre alternative à cette méthode de calcul consiste à utiliser un algorithme d'estimation des variances utilisant la méthode du maximum de vraisemblance restreint (REML) ; elle disponible dans la plupart des logiciels statistiques spécialisés, comme R.

2. Conduite des calculs

2.1 Exemple de construction d'un profil d'exactitude

Différents exemples de calcul sont disponibles dans la littérature (Feinberg M., 2009 - Commission SFSTP 2003, 2006). L'exemple présenté porte sur une méthode de dosage du nicotinamide (vitamine B3) dans le lait par HPLC avec une détection fluorimétrique. C'est une méthode indirecte qui nécessite un étalonnage pour réaliser la quantification. Les 10 étapes de la procédure se déroulent ainsi.

1. **Choix du domaine et des limites d'acceptabilité.** Le domaine de validation va de 0,4 à 4 mg/l. Les limites d'acceptabilité sont fixées $\pm 10\%$ autour de la valeur cible et choisies après discussion avec l'utilisateur. La proportion β est égale à 80 % et choisie de la même façon.
2. **Matériau de validation et fixation des valeurs cible.** Le matériau de validation choisi est un lait UHT qui ne contient pas la vitamine recherchée afin de pouvoir fixer les valeurs cible par dopage avec des ajouts dosés.
3. **Plan de validation (tableau 1).** Le plan de validation est formé de $I = 3$ séries (jours) avec $K = 3$ niveaux de concentration et $J = 3$ répétitions par jour et par niveau, ce qui fait un total de $3 \times 3 \times 3 = 27$ essais. Les concentrations des valeurs de référence sont exprimées en mg/L et les réponses en surfaces de pic. Les niveaux varient de 1 à K et sont codés A, B et C et les séries de 1 à I et sont codées, Jour1, Jour 2 et Jour 3.

Tableau 1 : plan de validation (mesures exprimées en unités de surface de pic)

Niveau	Valeur de référence	Série	Répétition 1	Répétition 2	Répétition 3
A	0.40	Jour 1	22.6	22.1	22.4
		Jour 2	23.3	24.1	23.9
		Jour 3	23.8	23.6	23.5
B	2.0	Jour 1	135.0	135.1	129.9
		Jour 2	137.6	135.2	138.8
		Jour 3	136.3	135.1	134.4
C	4.0	Jour 1	275.2	276.9	261.3
		Jour 2	268.1	269.7	276.9
		Jour 3	271.6	273.3	275.0

4. **Plan d'étalonnage (tableau 2).** Le plan d'étalonnage comporte $I = 3$ séries (jours) pour $K' = 2$ niveaux de concentration et $J' = 2$ répétitions par jour et par niveau, soit un total de 12 essais. Les deux niveaux sont codés Bas et Haut.

Tableau 2 : plan d'étalonnage (mesures exprimées en unités de surface de pic)

Niveau	Série	Concentration	Répétition 1	Répétition 2
Bas	Jour 1	0.4	22.7	23.1
	Jour 2	0.4	22.9	23.2
	Jour 3	0.4	21.9	22.1
Haut	Jour 1	4.0	281.6	275.3
	Jour 2	4.0	275.3	274.6
	Jour 3	4.0	272.0	273.0

5. **Estimation des coefficients des modèles d'étalonnage (tableau 3).** Le modèle choisi est une droite du type $Y = a_0 + a_1X + e$, où a_0 est le blanc, a_1 la sensibilité et e une erreur aléatoire. La méthode de calcul est la méthode de régression aux moindres carrés. Un modèle est calculé pour chaque jour.

Tableau 3 : estimations des paramètres des modèles en fonction des séries

Série	Sensibilité a_1	Blanc a_0
Jour 1	70.986	-5.494
Jour 2	69.972	-4.939
Jour 3	69.583	-5.833

6. **Calcul des concentrations prédites inverses et calcul de la justesse (tableau 4).** À partir des coefficients du tableau 3 et des mesures des surfaces de pic des échantillons de validation rassemblées au tableau 1, calculer les concentrations prédites inverses de chaque répétition en prenant soin d'utiliser les modèles obtenus pour chaque série.

Tableau 4. Concentrations retrouvées par prédiction inverse à partir des données du plan de validation (exprimées en mg/l)

Série	Niveau	Concentration X	Réponse Y	Retrouvée Z	Biais absolu	Biais relatif
Jour 1	A	0.40	22.6	0.40	-0.004	-1.06%
Jour 1	A	0.40	22.1	0.39	-0.011	-2.82%
Jour 1	A	0.40	22.4	0.39	-0.007	-1.76%
Jour 2	A	0.40	23.3	0.40	0.004	0.89%
Jour 2	A	0.40	24.1	0.42	0.015	3.75%
Jour 2	A	0.40	23.9	0.41	0.012	3.04%
Jour 3	A	0.40	23.8	0.43	0.026	6.47%
Jour 3	A	0.40	23.6	0.42	0.023	5.75%
Jour 3	A	0.40	23.5	0.42	0.022	5.39%
Jour 1	B	2.00	135.0	1.98	-0.021	-1.04%
Jour 1	B	2.00	135.1	1.98	-0.019	-0.97%
Jour 1	B	2.00	129.9	1.91	-0.093	-4.63%
Jour 2	B	2.00	137.6	2.04	0.037	1.85%
Jour 2	B	2.00	135.2	2.00	0.003	0.14%
Jour 2	B	2.00	138.8	2.05	0.054	2.71%
Jour 3	B	2.00	136.3	2.04	0.043	2.13%
Jour 3	B	2.00	135.1	2.03	0.025	1.27%
Jour 3	B	2.00	134.4	2.02	0.015	0.77%
Jour 1	C	4.00	275.2	3.95	-0.046	-1.14%
Jour 1	C	4.00	276.9	3.98	-0.022	-0.55%
Jour 1	C	4.00	261.3	3.76	-0.242	-6.04%
Jour 2	C	4.00	268.1	3.90	-0.098	-2.45%
Jour 2	C	4.00	269.7	3.92	-0.075	-1.88%
Jour 2	C	4.00	276.9	4.03	0.028	0.70%
Jour 3	C	4.00	271.6	3.99	-0.013	-0.32%
Jour 3	C	4.00	273.3	4.01	0.011	0.29%
Jour 3	C	4.00	275.0	4.04	0.036	0.90%

7. **Calculer les biais individuels.** On observe que les biais individuels varient entre -6.04% et 6.47% et permettent d'identifier les mesures éventuellement atypiques. Aucune donnée atypique n'a été identifiée.

8. **Calcul des données de fidélité et des intervalles de tolérance.** Calculer les critères de fidélité, comme décrit dans le § 1 et les limites des intervalles de tolérance selon la méthode décrite dans Feinberg M. (2010b). Rassembler les valeurs trouvées dans un tableau récapitulatif organisé comme le tableau 5.

Tableau 5 : tableau récapitulatif des critères de validation et des éléments du profil d'exactitude

Probabilité tolérance (β)	80%		
Limite d'acceptabilité (λ)	10%		
Niveaux	Niveau A	Niveau B	Niveau C
Concentration moyenne théorique	0.4	2.0	4.0
Concentration moyenne retrouvée	0.409	2.005	3.953
Fidélité			
Écart-type de répétabilité (s_r)	0.004	0.030	0.081
Écart-type inter-séries (s_B)	0.015	0.039	0.033
Écart-type de fidélité (s_{FI})	0.016	0.049	0.087
Coefficient de variation de fidélité	3.90%	2.45%	2.21%
Justesse			
Biais (%)	2.18%	0.25%	-1.17%
Intervalle de tolérance			
Nombre de degrés liberté	2.197	3.374	6.826
Facteur de couverture (k_{IT})	1.837	1.599	1.419
Écart-type de tolérance (s_{IT})	0.018	0.055	0.093
Limite intervalle tolérance basse	0.375	1.917	3.821
Limite intervalle tolérance haute	0.442	2.093	4.086
Limite d'acceptabilité basse	0.3600	1.8000	3.6000
Limite d'acceptabilité haute	0.4400	2.2000	4.4000
Profil d'exactitude			
Recouvrement %	102.2%	100.2%	98.8%
Limite intervalle tolérance basse en %	93.8%	95.9%	95.5%
Limite intervalle tolérance haute en %	110.6%	104.6%	102.2%
Limite d'acceptabilité basse en %	90.0%	90.0%	90.0%
Limite d'acceptabilité haute en %	110.0%	110.0%	110.0%

9. **Construction du profil d'exactitude.** Construire le graphique correspondant au profil en sélectionnant les valeurs de la ligne « Concentration moyenne théorique » en abscisse et les valeurs des 5 lignes sous le titre « Profil d'exactitude » en ordonnées. On obtient ainsi la **figure 1**.
10. **Interprétation.** On observe que les limites de tolérance à 80 % sont comprises entre les limites d'acceptabilité dans un domaine de validité s'étend de 0,5 mg/l environ à 4,0 mg/., la méthode est valide dans ce domaine. La limite inférieure (limite de quantification) sera calculée, si nécessaire, de façon plus exacte comme décrit dans la section 3. Par ailleurs, on constate que, d'une part la fidélité varie en fonction de la concentration puisque son coefficient de variation passe de 3,9 à 2,2 %, d'autre part, la justesse varie aussi puisque le biais de justesse varie de + 2,2 à - 1,2 %. Cette remarque souligne l'intérêt de procéder aux calculs niveau par niveau. Cependant, ce biais n'a pas d'influence importante sur la validité de la méthode.

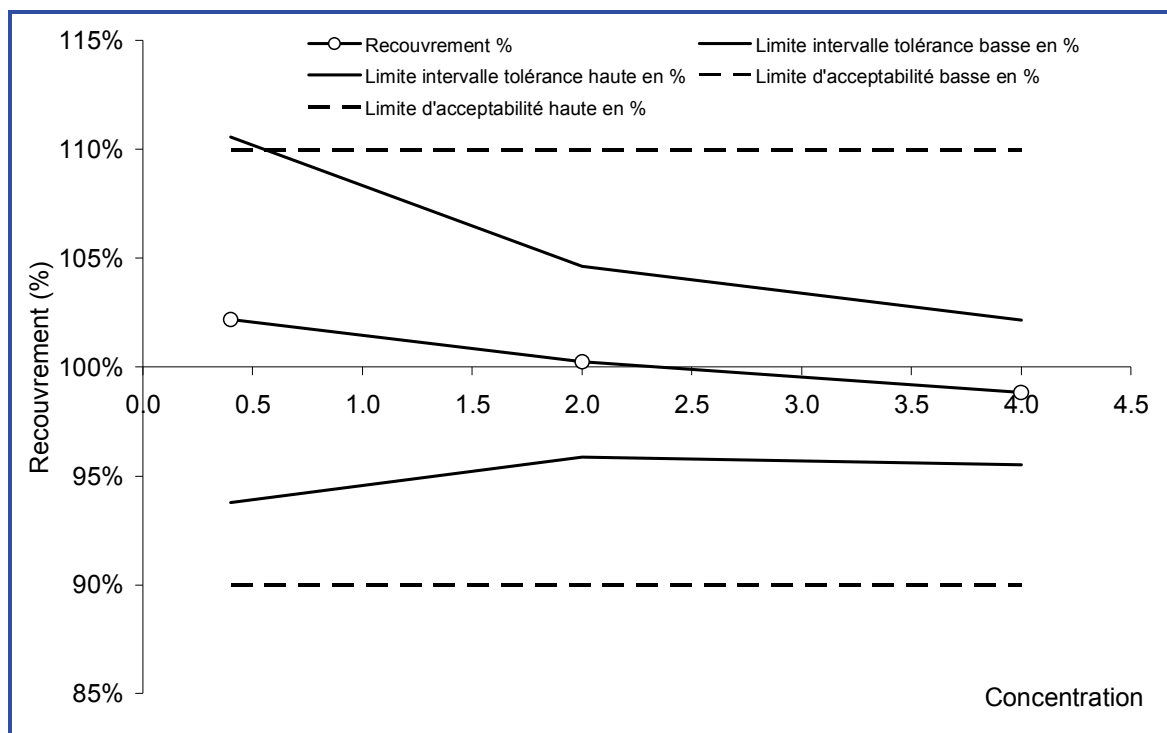


Figure 1 : profil d'exactitude pour le dosage de la vitamine B3 dans le lait par HPLC

11. **Conclusion.** On peut donc conclure à la capacité de la méthode à quantifier exactement le nicotinamide dans le lait sur le domaine de validité ainsi défini.

2.2 Utilisation d'un logiciel

Les calculs peuvent être réalisés en utilisant différents logiciels, comme Excel, OpenOffice ou R. Excel® est une marque commerciale de Microsoft. OpenOffice.org est un projet Open Source qui développe une suite bureautique libre de droits, compatible avec les autres standards du marché et qui est téléchargeable à partir de <http://www.openoffice.org/>. R est un environnement de calcul statistique disponible sous la forme d'un logiciel libre téléchargeable à partir de <http://cran.cict.fr/index.html>.

La feuille de calcul suivante permet de calculer les critères de fidélité, de justesse et les limites de tolérance pour un niveau dans un tableur. Elle est extraite d'un fichier modèle qui peut être téléchargé à partir :

http://www.paris.inra.fr/metarisk/downloads/software_programs/excel_templates.

Elle est utilisable sous Excel ou OpenOffice. Les données utilisées sont celles du niveau A de l'exemple.

	A	B	C	D	E	F	G
1	Niveau A						
2	Concentration moyenne théorique	0.4000					
3	Séries	Répétition 1	Répétition 2	Répétition 3		ni	SCEi
4	Jour (Série) 1	0.396	0.389	0.393		3	0.00003
5	Jour (Série) 2	0.404	0.415	0.412		3	0.00007
6	Jour (Série) 3	0.426	0.423	0.422		3	0.00001
7							
8	Nombre de séries (I)	3.0000	=NB(B4:B6)				
9	Nombre de mesures (IJ)	9.0000	=NB(B4:D6)				
10	Nombre de répétitions (J)	3.0000	=SI(F4*B8<>B9;"Calcul impossible";F4)				
11	SCE résiduelle	0.00011	=SOMME(G4:G8)				
12	SCE totale	0.00156	=SOMME.CARRES.ECARTS(B4:D6)				
13	SCE inter-séries	0.00145	=B12-B11				
14	Calcul intermédiaire de s ² B	0.00024	=((B13/(B8-1))-B15)/B10				
15	Variance de répétabilité (s ² r)	0.00002	=B11/(B9-B8)				
16	Variance inter-séries (s ² B)	0.00024	=SI(B14<0;0;B14)				
17	Variance de fidélité (s ² FI)	0.00025	=SOMME(B15:B16)				
18	Fidélité						
19	Concentration moyenne retrouvée	0.40873	=MOYENNE(B4:D6)				
20	Ecart-type de répétabilité (sr)	0.00419	=RACINE(B15)				
21	Ecart-type inter-séries (sB)	0.01536	=RACINE(B16)				
22	Ecart-type de fidélité (sFI)	0.01592	=RACINE(B17)				
23	Justesse						
24	Biais (%)	2.18%	=B19/B2-1				
25	Facteur de couverture de l'intervalle de tolérance						
26	Rapport des variances (R)	13.40469	=B16/B15				
27	Coefficient B ²	0.34951	=(B26+1)/(B10*B26+1)				
28	Nombre de degrés liberté	2.19709	=(B26+1)^2/((B26+1/B10)^2/(B8-1)+(1-1/B10)/B9)				
29	Probabilité tolérance (bêta)	80%					
30	t Student bas	1.88562	=LOI.STUDENT.INVERSE(1-B29;ARRONDI.INF(B28;0))				
31	t Student haut	1.63774	=LOI.STUDENT.INVERSE(1-B29;ARRONDI.SUP(B28;0))				
32	Facteur de couverture (kIT)	1.83676	=B30-(B30-B31)*(B28-ARRONDI.INF(B28;0))				
33	Ecart-type de tolérance (sIT)	0.01828	=B22*RACINE(1+1/(B9*B27))				
34	Intervalle de tolérance						
35	Limite intervalle tolérance basse	0.37516	=B19-B33*B32				
36	Limite intervalle tolérance haute	0.44230	=B19+B33*B32				

Afin de faciliter la mise en œuvre, les principales formules, telles que saisies, sont indiquées dans la colonne C et se réfèrent au contenu de la colonne B.

En ce qui concerne l'étalonnage et le calcul des modèles, il suffit d'utiliser la fonction de régression DROITEREG pour trouver les paramètres. L'organisation de la feuille doit tenir compte du fait qu'on calcule un modèle par série.

3. Critères de validation complémentaires

3.1 Limite de quantification

Principe

Sur la **figure 1** on note qu'il existe une concentration à partir de laquelle au moins une des bornes de l'intervalle de tolérance coupe une des limites d'acceptabilité. Au-dessus de ce seuil, l'analyste peut garantir que la méthode produira une proportion de résultats au moins égale à β . Selon les règles de décision proposées, la méthode est donc valide puisqu'elle permet de garantir les objectifs fixés. En revanche, en dessous de ce seuil, il n'est plus

possible de conclure à la validité. Tout naturellement, on peut proposer de définir ce seuil comme la limite de quantification (LQ) de la méthode. Si on se réfère à l'exemple de la **figure 1**, la LQ se situerait entre les niveaux A et B.

Lorsqu'il est nécessaire d'avoir une très bonne estimation de la limite de quantification, il est judicieux d'utiliser un plan de validation avec au minimum $K = 4$ niveaux. Les deux niveaux les plus bas (notés ici A et B) doivent alors être choisis de façon à encadrer au plus près la LQ supposée. L'interpolation sera alors d'autant plus exacte que cet encadrement sera étroit.

La LQ peut être obtenue exactement en calculant l'abscisse du point d'intersection. Mais les calculs doivent être réalisés sur les valeurs absolues des limites d'acceptabilité et de tolérance car le fait de passer aux valeurs relatives, utilisées en pratique pour la représentation graphique du profil, introduit une interpolation hyperbolique. En effet, le calcul des coordonnées du point d'intersection de deux droites est un problème d'algèbre bien connu. Pour calculer la LQ entre deux niveaux notés A et B, les notations suivantes sont utilisées :

- Z les valeurs absolues lues sur l'axe des ordonnées ;
- X les valeurs de référence fournies par l'axe des abscisses ;
- x_A et x_B les abscisses des deux niveaux A et B ;
- x_{At} et x_{Bt} les ordonnées sur la limite supérieure (ou inférieure) de l'intervalle de tolérance ;
- x_{Aa} et x_{Ba} les ordonnées sur la limite supérieure (ou inférieure) d'acceptabilité ;
- x_{LQ} l'abscisse du point d'intersection, avec $LQ = x_{LQ}$.

Les deux droites peuvent être représentées à l'aide d'un système de deux équations, dont la première traduit la limite de l'intervalle de tolérance (paramètres t_0 l'ordonnée à l'origine et t_1 la pente) et la seconde la limite d'acceptabilité (paramètres a_0 et a_1) :

$$\left. \begin{array}{l} Z = t_0 + t_1 X \\ Z = a_0 + a_1 X \end{array} \right\} \quad (7)$$

On peut noter que :

- a_0 doit être égal à 0 puisque la limite de l'intervalle d'acceptabilité passe toujours par le zéro.
- Si on cherche l'intersection avec la limite supérieure d'acceptabilité $z_{At} = x_A(1 + \lambda)$; il faudra prendre $1 - \lambda$ si on recherche l'intersection avec la limite inférieure.

Pour obtenir la pente t_1 , on utilise une règle de trois :

$$t_1 = \frac{z_{Bt} - z_{At}}{x_B - x_A} \quad \text{Pente (8)}$$

Pour l'ordonnée à l'origine t_0 , on l'obtient à partir de la première équation du système (7) :

$$t_0 = z_{At} - x_A \times t_1 \quad \text{Origine (9)}$$

De façon symétrique :

$$a_1 = \frac{z_{Ba} - z_{Aa}}{x_B - x_A}$$

$$a_0 = z_{Aa} - x_A \times a_1$$

Pour résoudre ce système d'équations, poser :

$$t_0 + t_1 X = a_0 + a_1 X$$

$$X(t_1 - a_1) = a_0 - t_0$$

On obtient ainsi la valeur de x_{LQ} :

$$x_{LQ} = \frac{a_0 - t_0}{t_1 - a_1} \quad (10)$$

Exemple de calcul avec un tableur

Pour illustrer la méthode, on va prendre en compte les résultats de l'exemple obtenus pour les niveaux A (0,4 mg/l) et B (2,0 mg/l), puisque c'est entre ces niveaux que la borne supérieure de l'IT coupe la limite supérieure d'acceptabilité. Ce sont les valeurs absolues qui vont servir. La **figure 2** décrit le mode de calcul sur les valeurs absolues et comment cette opération correspond à une interpolation hyperbolique sur les données relatives. Pour illustrer le calcul dans un tableur, les lignes utiles ont été sélectionnées dans le **tableau 5**.

Ligne	Niveau A	Niveau B
Concentration moyenne théorique	0,40	2,00
Limite intervalle tolérance haute	0,442	2,093
Limite d'acceptabilité haute	0,440	2,200

La feuille de calcul suivante donne les formules de calcul dans un tableur.

	A	B	C	D	E
2		Niveau A	Niveau B		
3	Concentration moyenne théorique	0.40	2.00		
4	Limite intervalle tolérance haute	0.442	2.093		
5	Limite d'acceptabilité haute	0.440	2.200		
6				Formules	
7	Droite	Pente	Origine	Pente	Origine
8	Intervalle tolérance (U)	1.032	0.030	=(C4-B4)/(C3-B3)	=B4-B8*B3
9	Limite d'acceptabilité haute	1.100	0.000	=(C5-B5)/(C3-B3)	=B5-B9*B3
10					
11	Limite de quantification	0.4337		=(C9-C8)/(B8-B9)	

Les formules utilisées sont indiquées dans les colonnes D et E. Finalement, la LQ vaut 0,434 mg/l.

Bien sûr, la limite de quantification dépend des valeurs choisies des limites d'acceptabilité et de la proportion β . Si ces deux critères n'ont pas pu être fixés à l'avance par manque de consignes ou de réglementation, mais si une LQ limite est attendue, par exemple pour un contrôle réglementaire, il est possible de les calculer en retour en se fixant une LQ préalable. Ce mode de calcul inverse se déduit directement des formules fournies.

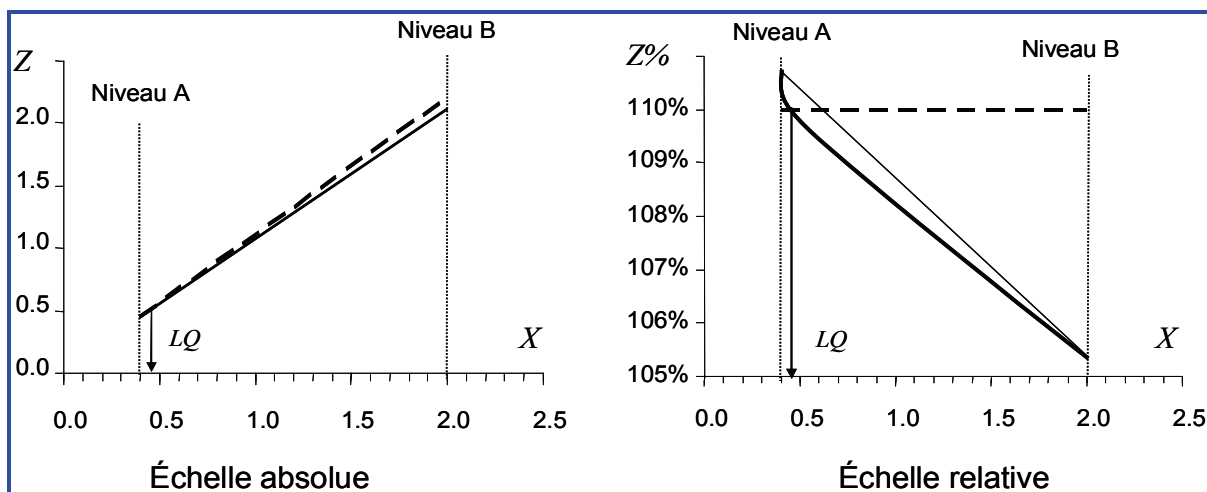


Figure 2 : illustration du mode de calcul de la limite de quantification

3.2 Limite de détection

La méthode du profil d'exactitude ne propose pas d'approche spécifique pour le calcul de la LD, par contre on peut s'appuyer sur les données recueillies pour construire le profil d'exactitude en se référant à diverses approches, dont la liste suivante n'est pas exhaustive :

- le guide EURACHEM (disponible sur le site www.lne.fr)
- la série des normes ISO 11843
- les approches proposées dans l'International conference on harmonisation (ICH), à savoir :
 - 1 - examen visuel. Applicable aux méthodes non instrumentales ou instrumentales.
 - 2 - rapport signal sur bruit. Cette approche n'est possible que lorsqu'on peut observer la ligne de base de la méthode. La détermination du rapport signal sur bruit est effectuée en comparant les signaux obtenus sur des échantillons de faibles concentrations avec ceux d'échantillons blancs. On estime qu'un rapport entre 3/1 et 2/1 est généralement satisfaisant pour estimer la limite de détection.

3.3 Calcul d'un facteur de correction

Le calcul d'un facteur de correction n'est possible que sous deux conditions :

- il n'existe pas de contrainte réglementaire qui interdise de corriger les résultats avant de les renvoyer au demandeur d'analyse. Le demandeur d'analyse est alors informé du fait que le résultat final est corrigé.
- l'analyste procède à un examen détaillé des mesures collectées lors de l'étude de validation afin de déterminer la nature du biais systématique afin de choisir un facteur de correction approprié.

Si on observe un biais systématique important dû à un taux de recouvrement moyen inférieur ou supérieur à 100 %, on peut alors proposer d'appliquer un facteur de correction sur les réponses instrumentales, pour corriger ce biais systématique. Par exemple, si on considère que le taux de recouvrement moyen dans tout le domaine est 70 %, on peut prendre un facteur $1/0,7 \approx 1,4$ et l'appliquer à tous les résultats.

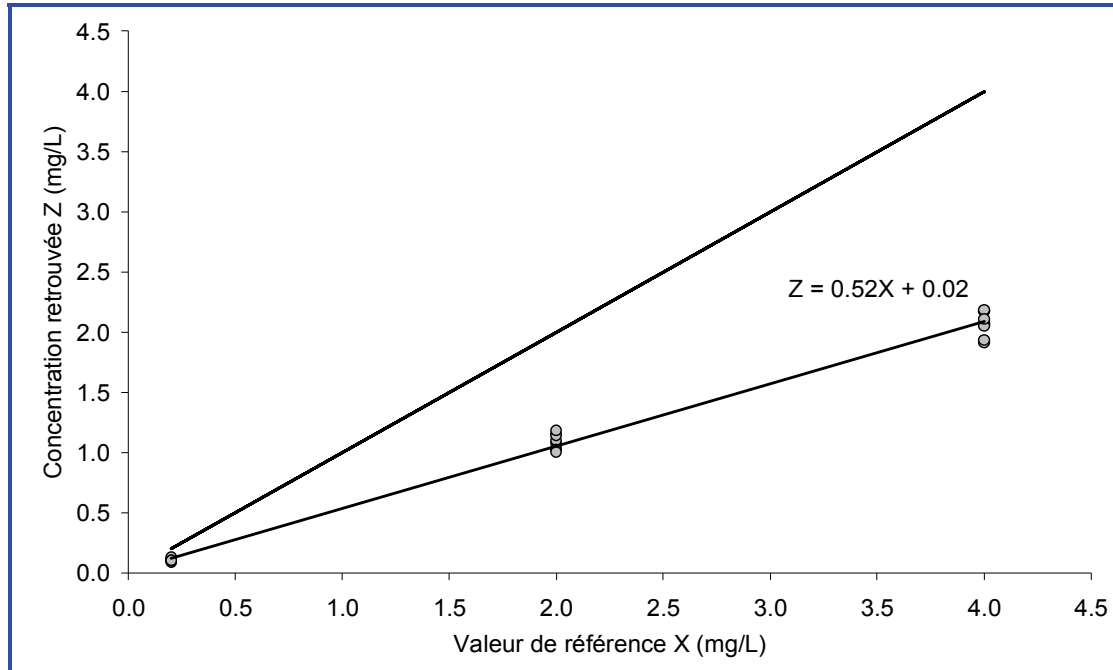


Figure 3 : représentation des concentrations retrouvées en fonction des valeurs de référence permettant de mettre en évidence un effet de matrice et de calculer une droite « moyenne » de recouvrement

Une méthode plus exacte consiste à reporter sur un graphique les concentrations retrouvées en fonction des valeurs de référence. La **figure 3** montre comment mettre en évidence un effet de matrice. On peut alors calculer, sous la forme d'une droite de régression au moindres carrés, l'équation d'une droite dont la pente représente un taux moyen de recouvrement et l'ordonnée à l'origine un biais systématique. Elle permet ainsi de définir deux facteurs de correction possibles :

- l'inverse de la pente permet de corriger un taux de recouvrement non satisfaisant. Dans l'exemple on prendra $1/0,52 \approx 2,0$;
- l'ordonnée à l'origine permet de corriger un biais systématique, par exemple généré par l'intégration d'un pic mal résolu en chromatographie. Dans l'exemple ce type de biais n'est pas important et ne sera pas pris en compte.

Le facteur de correction choisi sera alors inclus dans le mode opératoire de la méthode, par exemple, dans la partie relative à l'expression du résultat. Afin de vérifier la pertinence du facteur de correction choisi, il suffit de recalculer le profil d'exactitude avec les données corrigées. Cependant, pour assurer l'indépendance des mesures, il est préférable d'utiliser deux échantillons de validation et de réaliser l'ensemble des plans d'expérience en parallèle.

La **figure 4** illustre l'effet du facteur de correction 2,0. L'échantillon A dont les mesures n'ont pas été corrigées produit un profil d'exactitude qui mène à conclure que la méthode n'est pas valide. En revanche, les mesures faites sur un deuxième échantillon B puis corrigées par 2,0, conduisent à la conclusion que la méthode est valide si les résultats sont corrigés.

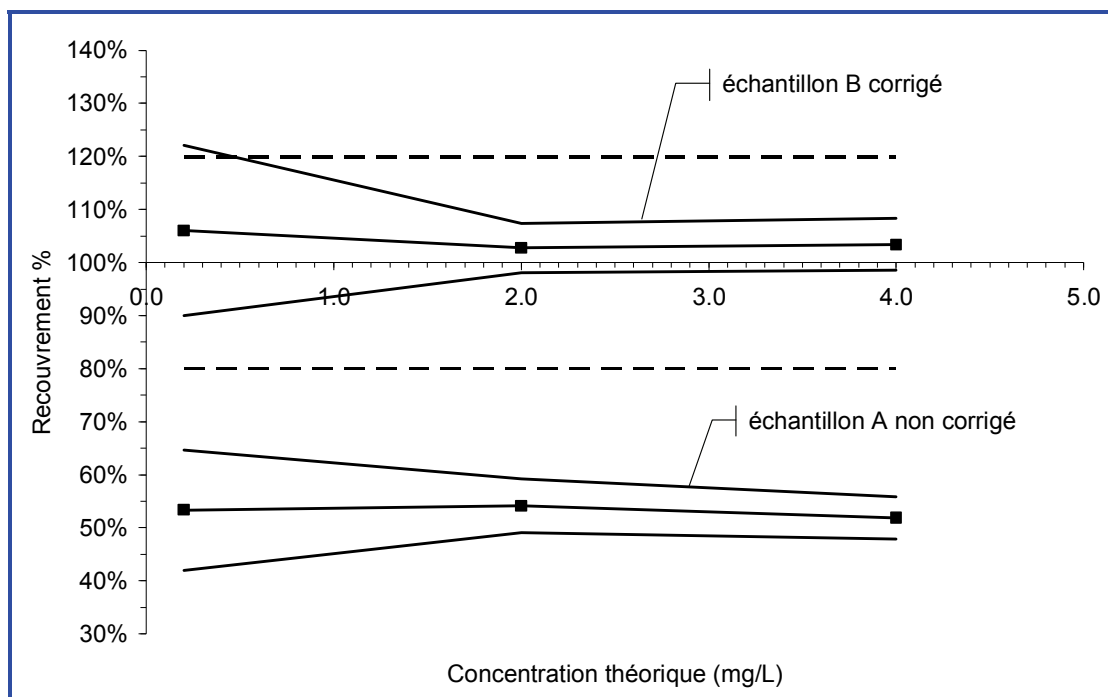


Figure 4 : illustration de l'intérêt du facteur de correction pour valider une méthode

Cette approche est volontairement empirique et ne fait pas appel à des tests de signification de la pente ou de l'ordonnée à l'origine.

3.4 Contribution au calcul de l'incertitude

Cette section propose une façon d'estimer une part, souvent importante, de l'incertitude-type composée. Elle est obtenue à partir des données de validation mais ne remplace pas l'application d'une procédure exhaustive d'estimation de l'incertitude, telle que décrite dans différents fascicules de documentation (Afnor, FD X07-021:1999 et FD V03-116:2001). En particulier, il reste indispensable d'établir la liste des principales sources d'incertitude et de l'organiser sous la forme d'un diagramme causes/effets. Le principe de l'approche proposée consiste à s'appuyer sur le modèle décrit dans le fascicule de documentation ISO/TS 21748. Ainsi, pour un niveau de concentration donné on va écrire que l'incertitude-type élevée au carré vaut :

$$u^2(Z) = u^2(D) + s_{FI}^2$$

Avec Z la valeur du mesurande et D le biais absolu moyen. Si on néglige les incertitudes sur la valeur de référence et d'autres sources d'incertitude éventuelles, on démontre que (pour plus de détails, voir (Feinberg M., 2009) :

$$u(D) = \sqrt{\frac{s_{FI}^2 - \left(1 - \frac{1}{I}\right) s_r^2}{J}}$$

En réorganisant les calculs, on trouve :

$$u(Z) = s_{FI} \sqrt{1 + \frac{1}{J B^2}} \quad (11)$$

Avec

$$B = \sqrt{\frac{R+1}{J \times R + 1}} \quad \text{et} \quad R = \frac{s_B^2}{s_r^2}$$

On remarquera alors que la formule (11) est identique à celle de l'écart-type de l'intervalle de tolérance fourni dans Feinberg (2010b). On peut ainsi conclure que l'écart-type de l'intervalle de tolérance s_{IT} permet d'estimer les composantes de l'incertitude liées, par exemple :

- au modèle d'étalonnage ;
- à l'environnement du laboratoire ;
- aux opérateurs ;
- à la manipulation et au stockage des échantillons lors de leur séjour dans le laboratoire, etc.

Cette liste n'est pas exhaustive et dépend beaucoup des conditions de fidélité intermédiaire choisies pour réaliser les essais. Cependant, elle ne prend pas en compte l'incertitude sur les valeurs de référence qui correspond à ce qu'on appelle la traçabilité des mesurages.

Bibliographie

- AFNOR Norme NF ISO 5725-2:1994, 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. Afnor, Paris
- AFNOR Norme NF X 06-041, Application de la statistique — Fidélité des méthodes d'essai — Détermination de la répétabilité et de la reproductibilité d'une méthode d'essai normalisée par essais inter-laboratoires. Afnor, Paris
- AFNOR Norme FD X07-021:1999 Normes fondamentales - Métrologie et applications de la statistique - Aide à la démarche pour l'estimation et l'utilisation de l'incertitude des mesures et des résultats d'essais, Afnor, Paris
- AFNOR Norme FD V03-116:2001 Analyse des produits agricoles et alimentaires - Guide d'application des données métrologiques, Afnor, Paris
- FD ISO/TS 21748 :2004. Lignes directrices relatives à l'utilisation d'estimations de la répétabilité, de la reproductibilité et de la justesse dans l'évaluation de l'incertitude de mesure, ISO, Genève
- Feinberg M., (2009) *Labo-Stat : Guide de validation des méthodes d'analyse* Lavoisier Tec&Doc, Paris
- Feinberg M. (2010a) Principes et vocabulaire in Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 13-26

- Feinberg M. (2010b) Mise en œuvre du profil d'exactitude *in* Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 27-44
- Commission SFSTP, P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat ; Validation des procédures analytiques quantitatives : Harmonisation des démarches.
- Partie I. : *STP Pharma Pratiques* (2003) 13:3, 101-138
- Partie II. : Statistiques, *STP Pharma Pratiques* (2006) 16:1, 30-60
- Partie III : Exemples d'application, *STP Pharma Pratiques* (2006) 16:2, 87-121
- ISO 11843-2 :2004 Capabilité de détection. Cas du modèle d'étalonnage linéaire, ISO Genève

Compléments sur la quantification des résultats pondération et transformation

Michel Laurentie¹

Résumé : *La pondération ou la transformation des données sont des outils nécessaires à leur exploitation statistique. Pondérer ou transformer des données repose sur la nécessité de stabiliser la variance. En effet, un principe théorique pour l'utilisation de la régression selon le critère des moindres carrés impose que les variances soient homogènes, c'est à dire quelles soient indépendantes du niveau des concentrations. Cet article décrit ces outils et montre au travers un exemple leur impact sur la validation des méthodes.*

Mots-Clés : pondération, transformation, validation, profil d'exactitude, méthodes d'analyse quantitatives

Problématique

De nombreux logiciels de chromatographie permettent de calculer les critères de performance d'une méthode analytique. En particulier, ils proposent des modèles d'étalonnage (fonction de réponse) permettant de lier la réponse aux différents niveaux de concentration. Bien souvent ces logiciels proposent systématiquement un facteur de pondération.

La pondération est un des moyens de respecter une règle fondamentale dans l'analyse statistique de données : l'homogénéité des variances.

Une autre approche pour respecter cette règle est celle de la transformation des données qui est plus rarement utilisée par les analystes car elle est souvent assimilée à une modification des données qui ne semble pas justifiée. À l'inverse dans d'autres disciplines telles que la microbiologie cette approche est systématique (transformation logarithmique pour analyser les comptages bactériens).

La question est alors la suivante : quand faut-il pondérer ou utiliser des données transformées ?

1. Quand faut-il agir ?

Dans une gamme de concentration, la variance des réponses peut varier en fonction de la concentration, d'autant plus que l'intervalle de gamme est important (Mac Taggart D.L. et Farwell S.O., 1992). On dit alors que la variance n'est pas homogène.

Généralement la variance est plus large aux extrémités du domaine qu'en son milieu. La Figure 1 illustre cette observation.

¹ Unité de pharmacocinétique-pharmacodynamie - AFSSA – F-35302 Fougères cedex- m.laurentie@afssa.fr

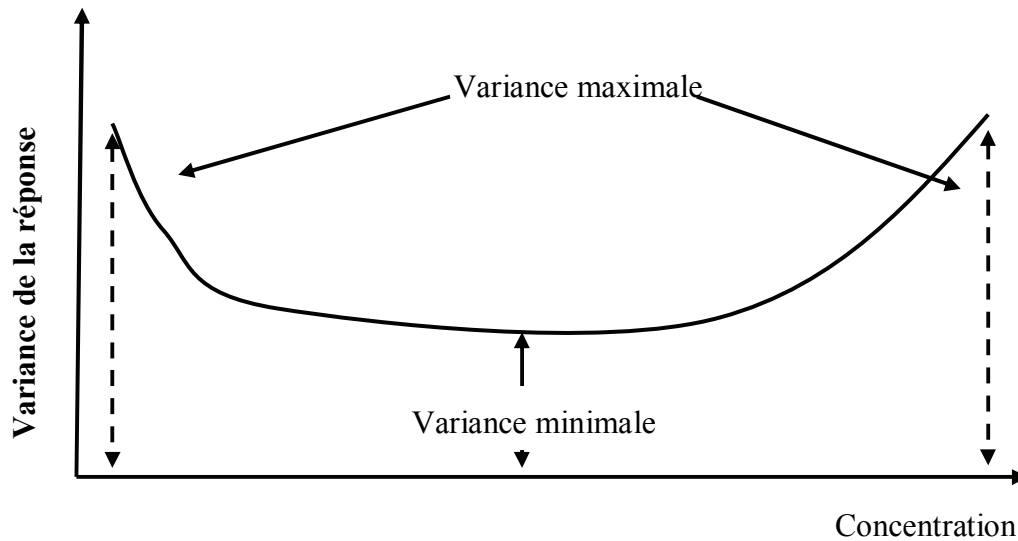


Figure 1 : représentation de l'évolution de la variance des réponses en fonction du niveau de concentration

La représentation graphique de la variance de la réponse en fonction de la concentration n'est pas fournie dans les logiciels de chromatographie, mais cette observation peut être également visualisée sur une droite d'étalonnage (figure 2).

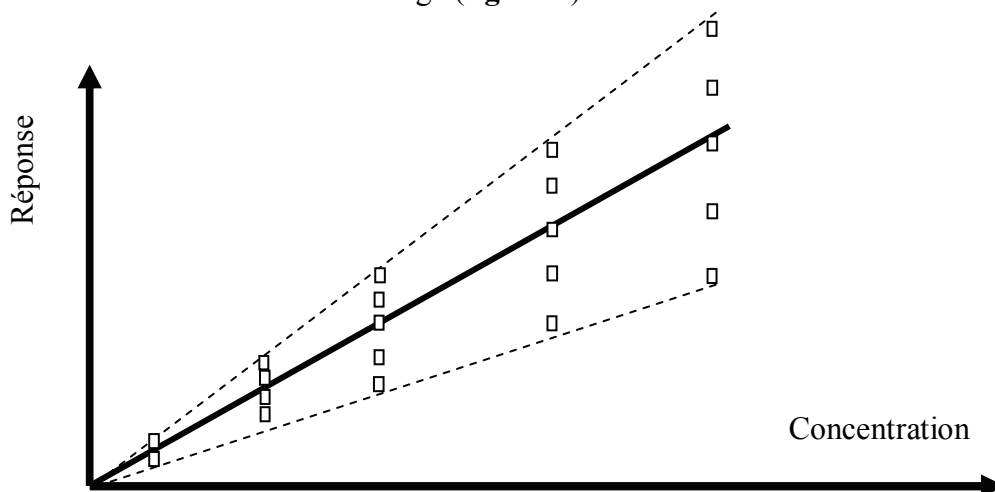


Figure 2 : évolution des réponses en fonction du niveau des concentrations, représentée par rapport à une courbe d'étalonnage

Une autre possibilité pour évaluer l'impact de la variance sur le modèle d'étalonnage est l'utilisation du graphique des résidus. Le résidu est la différence entre la valeur observée et la valeur prédite (Équation 1)

$$\text{résidu} = Y_{obs} - \hat{Y}_{cal} \quad \text{Équation 1}$$

Avec Y_{obs} la valeur observée, \hat{Y}_{cal} la valeur calculée par le modèle d'étalonnage e.g. un modèle linéaire.

Si la répartition des résidus est homogène dans l'intervalle des concentrations, il est conclu à la validité du modèle et à l'homogénéité des variances par niveau des concentrations. Si au

contraire les résidus paraissent structurés, *i.e.* que distribution suit une courbe particulière (droite par exemple) ou si la distribution des résidus montre une variabilité importante ou une forme particulière, par exemple, en trompette, il y a hétérogénéité des variances et par conséquent il est conclut à la non validité du modèle. Ceci est illustré sur les figures 3A et 3B.

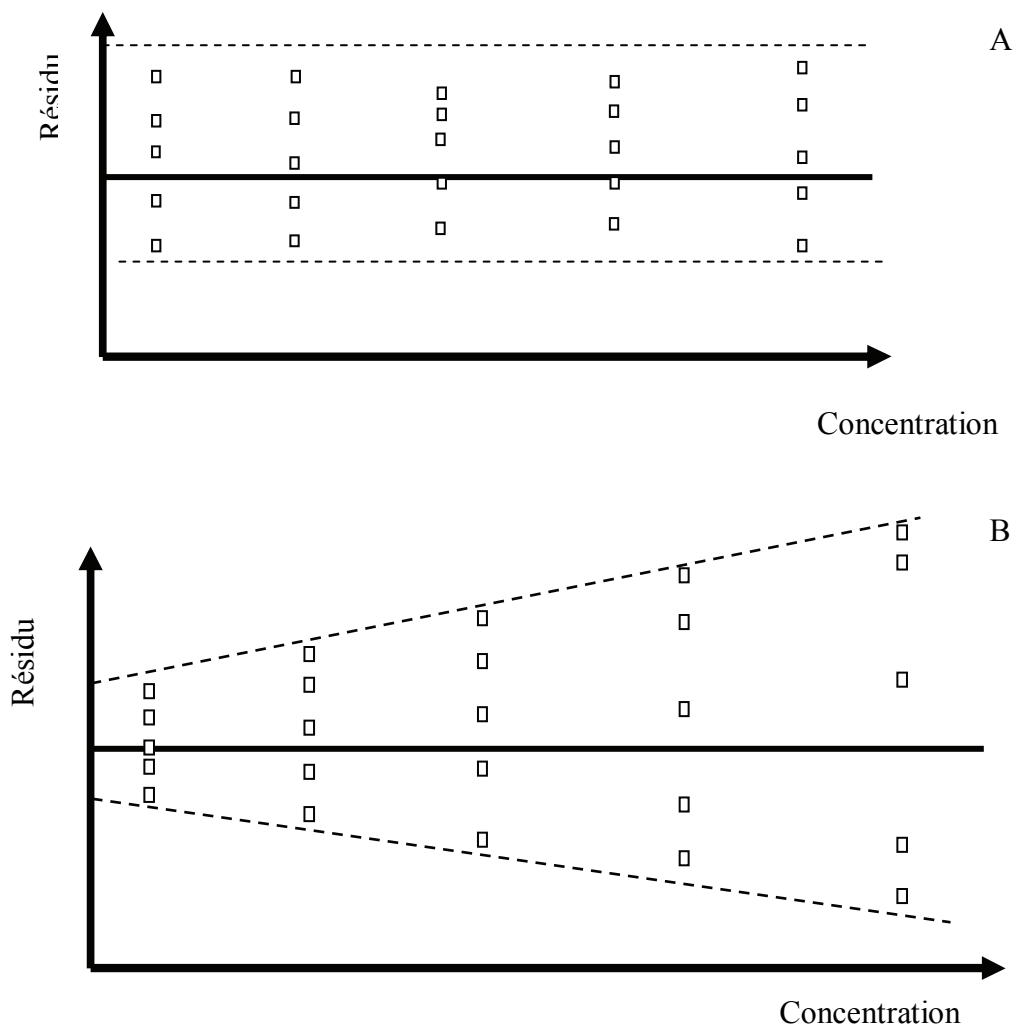


Figure 3 : (A) répartition des résidus pour un modèle d'étalonnage pour lequel la variance des réponses est homogène.

(B) répartition des résidus pour un modèle d'étalonnage pour lequel les variances des réponses sont non homogènes (forme dite en trompette).

La connaissance de la variance, à chaque niveau de concentration, est donc primordiale pour établir correctement le modèle d'étalonnage ; ceci implique que le schéma expérimental doit être rigoureux, avec en particulier, un nombre suffisant de répétition pour évaluer la variance aux extrémités du domaine de dosage.

Cette analyse graphique peut-être complétée par une analyse statistique basée sur l'utilisation de tests tel que les tests de Cochran ou de Levène (Hubert Ph. *et al.*, 1999). Ces test sont largement décrits dans la littérature et ne seront pas décrits dans cet article. Toutefois il

convient de préciser que ces tests sont des drapeaux pour indiquer l'hétérogénéité des variances et non des tests pour éliminer des niveaux de concentration qui seraient dits aberrants.

2. La stabilisation des variances

Si une hétérogénéité des variances est observée, il convient de la prendre en compte. Deux approches sont possibles : la pondération ou la transformation des données.

2.1 La transformation des données

La transformation des données permet de stabiliser la variance. Les plus utilisées sont les logarithmes ou les racines carrées.

Dans le cas de l'utilisation de la transformation, logarithmique le modèle utilisé s'écrit :

$$\log Y = a \log C + b \quad \text{Équation 2}$$

Où le $\log Y$ est le logarithme népérien de la réponse Y et $\log C$ le logarithme népérien de la concentration.

Les techniques de régression sont alors utilisées comme pour une régression classique, en particulier l'utilisation des moindres carrés. L'observation des résidus doit montrer une stabilisation des variances. Dans ce cas le modèle de fonction de réponse utilisée en routine implique de transformer les réponses avant de calculer les concentrations en retour.

La deuxième approche est l'utilisation de la pondération.

2.2 La pondération

Le principe de la pondération est simple. Pour chaque niveau de concentration d'un intervalle de dosage, la variance n'est pas proportionnelle à la concentration. Ceci implique de donner un poids à chaque niveau de concentration pour corriger le manque de proportionnalité.

Prenons le cas d'une régression linéaire de type (Tomassone R., Lesquoy E., Miller C., 1983) :

$$Y = aC + b \quad \text{Équation 3}$$

Avec Y la réponse, a et b les coefficients de la droite de régression.

Lors d'une estimation des paramètres a et b nous cherchons à minimiser la différence entre les valeurs calculées et les valeurs estimées ce qui se traduit par l'équation

$$SCE = \sum_{i=1}^n (Y_i - \hat{Y})^2 \quad \text{Équation 4}$$

Avec SCE : somme des carrés des écarts entre la i ème observation Y_i et l'estimation \hat{Y} par la droite de régression.

La prise en compte du poids (w_i) se fera par l'introduction d'un facteur dans l'équation 4.

$$SCE = \sum_{i=1}^n w_i (Y_i - \hat{Y})^2 \quad \text{Équation 5}$$

La sélection de la valeur du poids peut se faire graphiquement. Sur la **figure 4**, les différentes relations entre la variance des réponses et les niveaux de concentration sont représentées.

Plusieurs types de relation existent entre la variance et la concentration [4]. A chaque type de structure il est possible de définir un poids (w_i pour weighting factor) qui permettra de prendre en compte l'influence de la variance. Quatre cas sont généralement observés :

- 1 - Il n'y a pas de relation entre la concentration et la variance (le cas idéal) : le poids $w_i = 1$. En d'autres il n'y a pas de nécessité de pondérer
- 2 - la variance varie avec une relation de type $1/X$ i.e. inversement proportionnelle à la concentration, le poids qui doit être pris en compte est de la forme $w_i = 1/C$.
- 3 - la variance varie selon une relation de type $1/X^2$ i.e. inversement proportionnelle au carré de la concentration, le poids à utiliser est de la forme $w_i = 1/C^2$
- 4 - la relation n'est pas une relation simple entre la concentration et la variance, chaque variance variant indépendamment du niveau de concentration, le poids $w_i = 1/S^2$ avec S^2 la variance observée à chaque niveau de concentration.

Les cas 2 et 3 sont les plus fréquemment rencontrés mais le cas 4, cas le plus difficile, peut également s'observer.

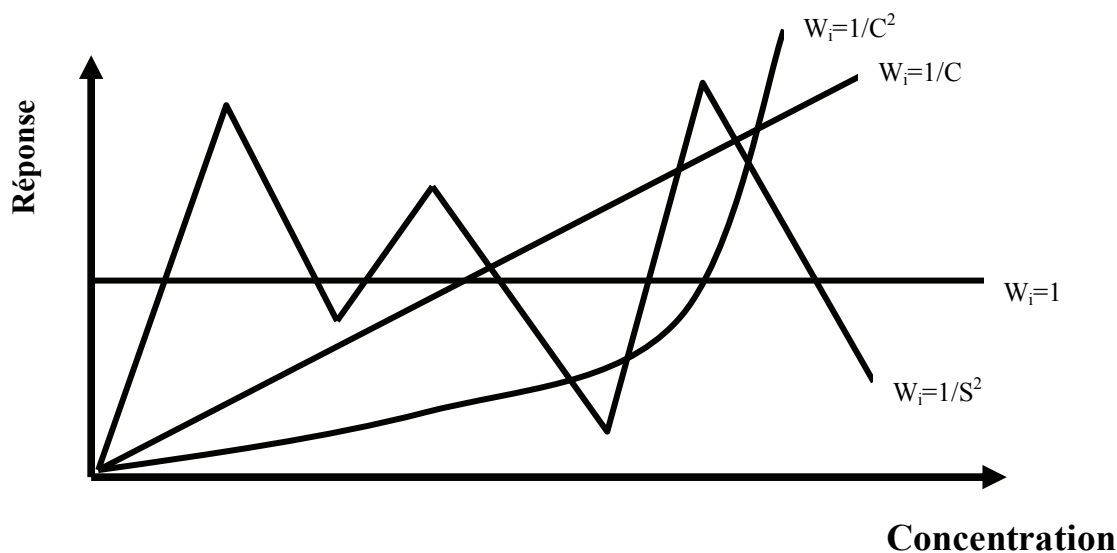


Figure 4 : types de relation fréquemment observées entre la variance de la réponse et les niveaux de concentration. Les quatre courbes présentées correspondent aux quatre cas décrits dans le texte

Ce type de graphique peut être facilement réalisé sur une feuille Excel, à partir des données obtenues lors de l'étude de validation ou de pré-validation. Il devrait être un préalable avant toute analyse statistique car ce graphique permet de visualiser s'il faut appliquer une pondération.

La pondération est applicable quel que soit le modèle choisi et l'utilisation d'un facteur permet de minimiser la *SCE* et donc d'avoir des meilleurs estimateurs pour les coefficients de la droite de régression.

3. Exemple

Un exemple d'utilisation de la pondération ou de la transformation des données est montré dans l'exemple du dosage de l'acrylamide.

L'acrylamide est un produit neo-formé lors de la cuisson de certains aliments. Il est issu d'une réaction de Maillard par combinaison de sucre (*i.e.* glucose) et de certains acides aminés comme l'asparagine. Si de nombreuses études documentent la teneur en acrylamide dans les aliments peu d'études ont permis d'en quantifier l'absorption après ingestion. Une étude de pharmacocinétique, réalisée chez le porc doit permettre de déterminer la biodisponibilité de l'acrylamide. Au préalable une méthode analytique est nécessaire pour quantifier dans le plasma de porc les concentrations en acrylamide. La méthode retenue pour effectuer ce dosage est une méthode HPLC associée à une détection par spectrométrie de masse (MS). La gamme étudiée varie de 10 à 5000 ng/ml, compte tenu que nous n'avons pas d'information sur les taux plasmatiques circulants après ingestion d'aliment contenant de l'acrylamide.

Deux gammes sont préparées : une gamme standard d'étalonnage et une gamme dans du plasma de porc (standard de validation). À 200 µl de plasma est ajouté 100 µl de solution saturée de ZnSO₄ puis 1000 µl d'acétonitrile et 100 µl de standard interne (acrylamide D5). Après agitation et centrifugation le surnageant est évaporé. L'éluât est repris avec 200 µl d'acétate d'ammonium 0,01M pH 6. Le volume d'injection est de 50 µl. Les conditions analytiques utilisées sont un débit chromatographique de 0,2ml/mn à travers une colonne Hypercarb (5µ 50-2 mm) et une détection par spectrométrie de masse sur l'ion moléculaire 72 de l'acrylamide. La gamme standard d'étalonnage en acrylamide varie de 10 à 5000 ng/ml et est réalisée dans une solution d'acétate d'ammonium 0,01M ramenée à pH 6 avec de l'acide formique.

Le plan d'expérience utilisé consiste en 5 jours, 6 niveaux et 2 répétitions ($5 \times 6 \times 2$) soit 60 essais pour les standards d'étalonnage et de validation.

La **figure 5** montre l'évolution de la variance des réponses en fonction des niveaux de concentration.

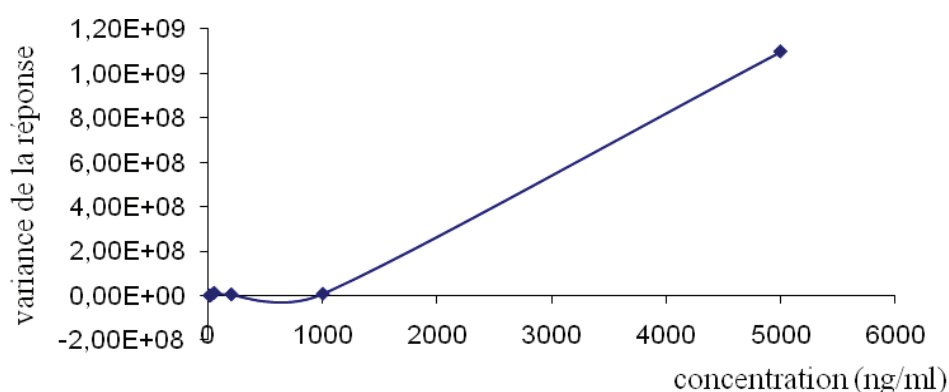


Figure 5 : évolution de la variance des réponses en fonction des niveaux de concentrations

La forme de la courbe montre que la variance n'est pas homogène sur l'intervalle de dosage, avec une très forte augmentation pour le niveau de concentration de 5000 ng/ml. Pour

stabiliser cette variance il est possible de pondérer soit par $1/C$ ou par un $1/C^2$ ou tester la transformation logarithmique des données.

La **figure 6** montre la distribution des résidus obtenue par un modèle de régression linéaire pondérée ou non ou appliqué sur des données logarithmiques.

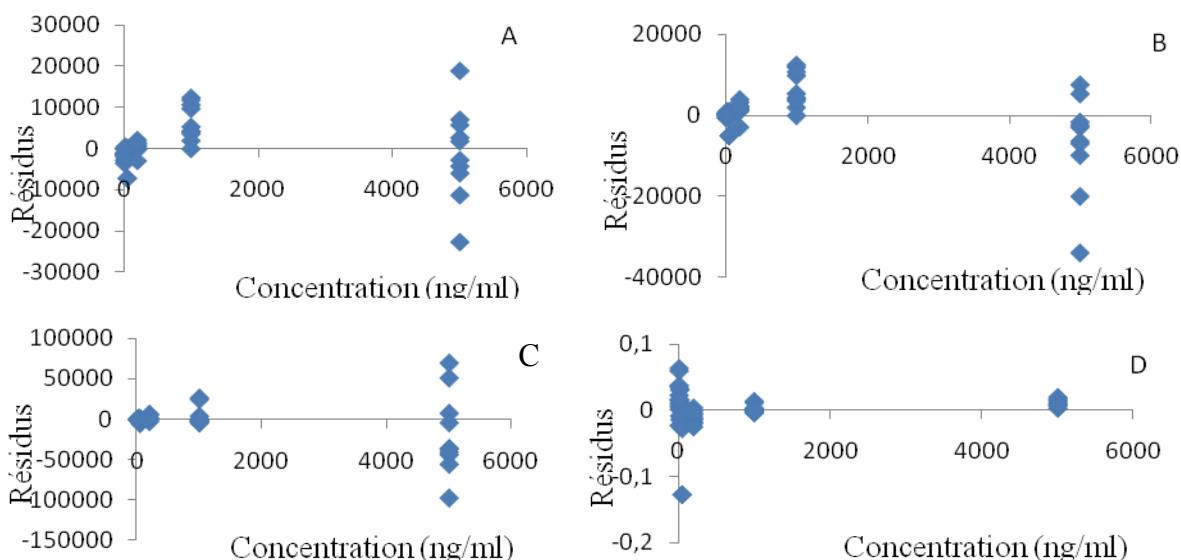
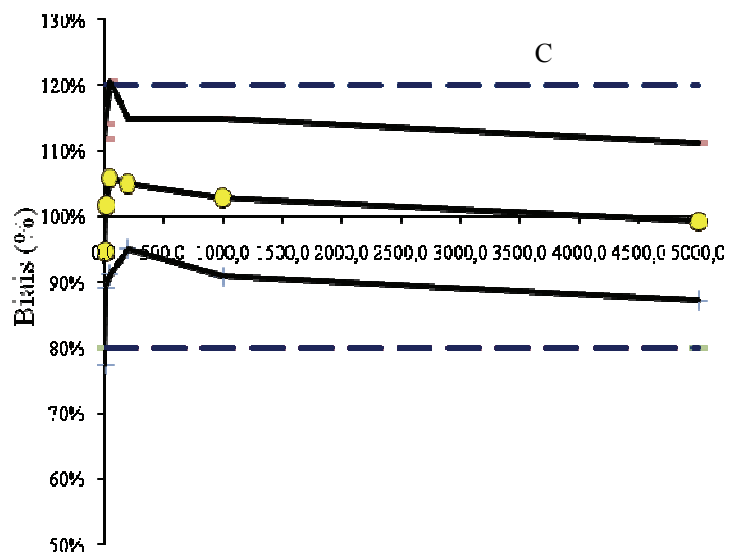
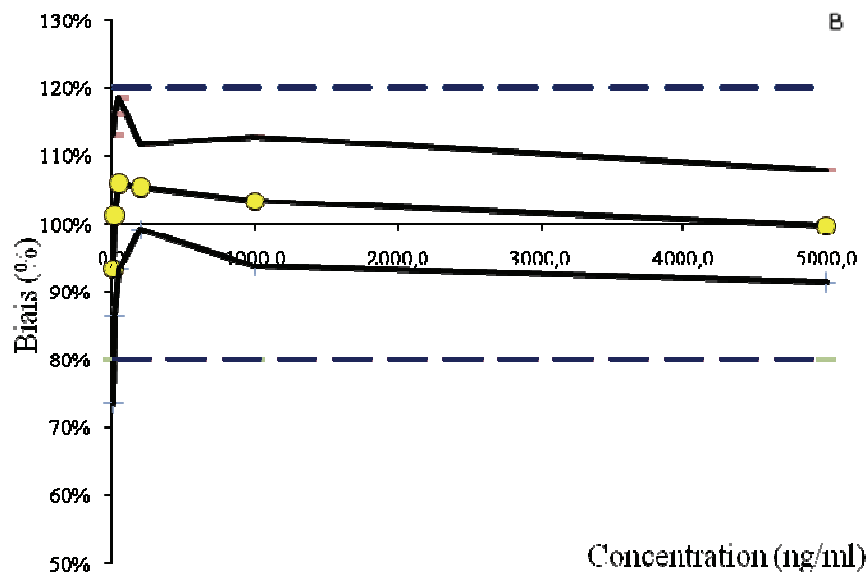
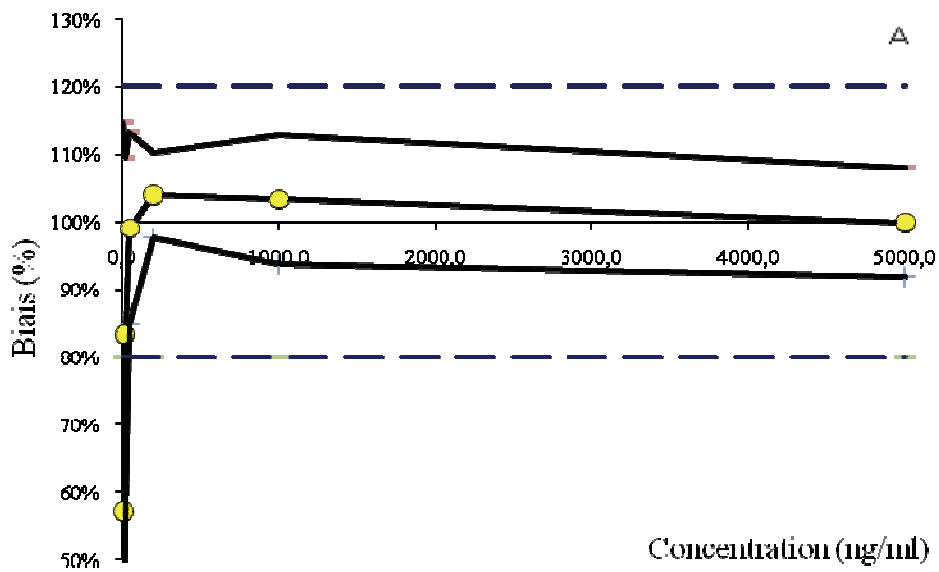


Figure 6 : évolution des résidus en fonction des concentrations (A) modèle linéaire ; (B) modèle linéaire pondéré ($1/C$) ; (C) modèle linéaire pondéré ($1/C^2$) ; (D) modèle linéaire appliqué après transformation logarithmique des données

L'observation des résidus montre que le modèle linéaire simple (**figure 6A**) conduit à une distribution des résidus en forme de trompette. Ceci implique que le modèle ne sera pas très performant. La pondération $1/C$ (**figure 6B**) diminue cet effet trompette. En revanche la pondération de type $1/C^2$ ne modifie pas la distribution des résidus (**figure 6C**). L'application du modèle linéaire sur les données transformées (**figure 6D**) montre une stabilisation de la variance. Ce modèle pourrait être retenu pour décrire les données.

La **figure 7** montre l'impact de la pondération et de la transformation sur le profil d'exactitude. En utilisant une fonction de réponse linéaire sans pondération ou transformation, les faibles concentrations présentent un biais de plus de 40 % (**figure 7A**). Les profils obtenus après pondération par $1/C$ ou par $1/C^2$ ou après transformation des données sont acceptables. La pondération $1/C$ présente une variabilité moins grande et après transformation logarithmique un moindre biais est observé. Le choix se fera sur l'objectif assigné de la méthode, c'est-à-dire privilégier le biais ou la fidélité.



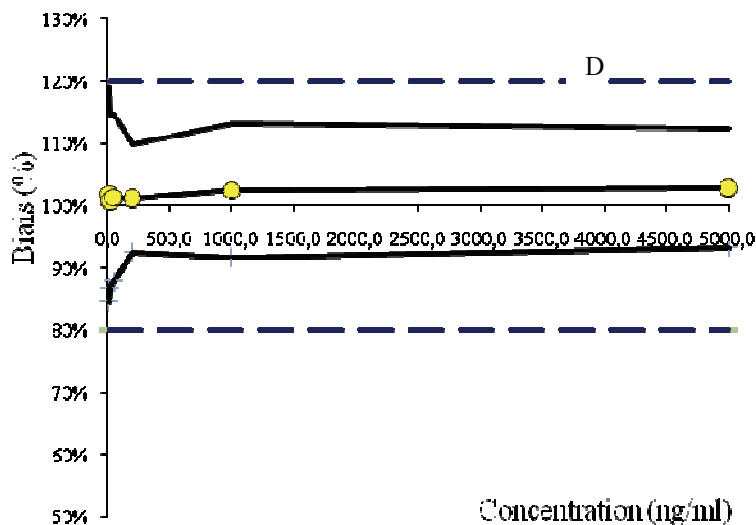


Figure 7 : profil d'exactitude obtenu avec (A) modèle linéaire, (B) modèle linéaire avec pondération ($1/C$), (C) modèle linéaire avec pondération ($1/C^2$), (D) modèle linéaire après transformation logarithmique des données

Conclusion

La pondération et / ou la transformation des données sont des outils qui permettent de stabiliser la variance. Elle n'est pas systématiquement nécessaire mais elle est un préalable à l'utilisation de modèle linéaire ou non linéaire, si la variance des réponses n'est pas homogène sur le domaine des concentrations étudiées. Le choix de la pondération ou de la transformation repose sur des aspects graphiques (résidus) ou sur la réalisation de tests statistiques appropriés. Dans certains cas difficiles, il faut avoir recours à la fois à la pondération et à la transformation des données. L'utilisation de ces outils optimise le modèle et ainsi d'obtenir les meilleurs paramètres descriptifs de la relation réponse / concentration.

Bibliographie

- Hubert Ph., Chiap P., Crommena J., Boulanger B., Chapuzet E., Mercier N., Bervoas-Martin S., Chevalier P., Grandjean D., Lagorce P., Lallier M., Laparra M.C., Laurentie M., Nivet J.C. (1999) The SFSTP guide on the validation of chromatographic methods for drug bioanalysis: from the Washington Conference to the laboratory. *Analytica Chimica Acta*, 391, 135 -148
- Mac Taggart D.L., Farwell S.O. (1992) analysis use of linear regression. Part 1: regression procedures or calibration and quantitation. *J. of the AOAC* 75, 594-614
- Tomassone R., Lesquoy E., Miller C. (1983) La régression, nouveaux regards sur une ancienne méthode statistique. Inra, Actualités scientifiques et agronomiques, p. 180



Chapitre II

Application dans les laboratoires à l'Inra

Photo © Christophe Maitre/Inra

En quoi la connaissance des performances d'une méthode permet-elle d'optimiser la mise en place d'une expérimentation animale ?

Agnès Thomas, Denys Durand et Isabelle Ortigues-Marty¹

Résumé : Parmi les programmes de recherche conduits par l'Unité de recherche sur les herbivores, la mesure des flux de nutriments énergétiques à travers le foie nécessite de déterminer avec exactitude les concentrations sanguines de glucose au niveau des vaisseaux afférents (veine porte) et efférents (veine hépatique). Les faibles différences de concentrations attendues entre vaisseaux, rendent la mesure de ces concentrations déterminante. Par ailleurs, la lourdeur en coût et en temps, des expérimentations réalisées sur des animaux cathétérisés ne permet pas d'envisager plusieurs répétitions de la même expérimentation. Aussi, nous avons souhaité déterminer les performances de notre méthode de dosage du glucose (selon la norme V03-110) avant d'établir le plan expérimental. Nous avons ainsi montré que selon les exigences que l'on se fixe (niveau de probabilité de 0,05 ou 0,01 ; dosages effectués dans la même journée ou sur plusieurs jours), le nombre d'animaux nécessaire pour mettre en évidence d'éventuelles différences significatives entre traitements variait de 4 à 10. Cet article reprendra d'après cet exemple spécifique, les étapes de caractérisation d'une méthode de dosage ainsi que l'intérêt de l'utilisation des performances de la méthode pour l'établissement du plan expérimental.

Mots clefs : caractérisation de méthode, plan expérimental, bilan hépatique, glucose,

1. Introduction

L'unité de recherche sur les herbivores (URH) du centre Inra de Theix, regroupe 172 agents permanents répartis dans 8 équipes de recherche, un abattoir expérimental, une équipe de ressources techniques et administratives et une installation expérimentale. La mission de l'URH est de contribuer au **développement des systèmes durables d'élevage** des herbivores en accord avec les attentes sociétales, c'est-à-dire conciliant **viabilité socio-économique** et **développement rural, respect de l'environnement, qualité des produits et bien-être animal**. Pour répondre à ces missions, l'URH met en place des projets de recherche faisant intervenir de plus en plus de partenaires extérieurs (publics, privés ...). Les résultats obtenus à l'issue de ces projets sont valorisés principalement au travers de publications (environ 300 par an) et d'enseignements (320 heures/an).

Outre le fait que dans le cadre du référentiel "Qualité" INRA, les exigences concernant la fiabilité des résultats peut nécessiter que nos méthodes soient validées, une autre contrainte propre à notre secteur d'activité nous a conduits à réfléchir à une approche complémentaire de la "**validation de méthode**" qu'est la "**caractérisation de méthode**". Cette approche peut complètement s'intégrer dans les dispositions déjà prises dans notre Unité pour répondre aux exigences du référentiel Inra de niveau 1 (Thomas A. et Durand D., 2009).

¹ UR1213 URH - Unité de recherche sur les herbivores - NEM (Équipe Nutriments et Métabolismes) - INRA – F-63122 Saint-Genès-Champanelle ☎ 04 73 62 45 43 ✉ agnes.thomas@clermont.inra.fr

Une grande partie des expérimentations sont conduites sur des animaux (bovins, ovins, caprins, rongeurs...) ce qui implique une approbation de nos protocoles par le comité d'éthique. La charte du comité d'éthique énoncée en 9 articles (Verschuere B., 2008), précise que le scientifique doit avoir employé toutes les stratégies débouchant sur une utilisation d'un nombre réduit et adéquat d'animaux pour obtenir des données suffisantes permettant de répondre à une question de recherche.

Nous allons montrer, à travers un exemple extrait d'une expérimentation conduite sur des moutons ce qu'apporte la caractérisation de méthodes par rapport aux exigences relatives à la démarche qualité et celles relatives à l'éthique animale. Nous aborderons successivement les différentes étapes de caractérisation d'une méthode en prenant l'exemple du dosage du glucose sanguin ; puis nous montrerons comment nous avons utilisé les résultats de cette caractérisation dans l'élaboration du plan expérimental.

2. Contexte expérimental

Les finalités de l'équipe Nutriments et Métabolismes sont d'améliorer les qualités nutritionnelles et sensorielles de la viande de ruminant, tout en préservant les performances de croissance des animaux. Afin d'atteindre ces objectifs scientifiques, il est nécessaire d'évaluer la répartition des nutriments entre les tissus et les organes de manière à déterminer les étapes limitantes de l'utilisation métabolique des nutriments.

Pour évaluer la répartition des nutriments entre organes, il est impératif de mesurer au mieux la quantité de nutriments qui entre (flux afférent) et qui sort (flux efférent) des organes d'intérêt (Majdoub L. *et al.*, 2003). Pour répondre à cette problématique, l'équipe a mis en place une approche chirurgicale permettant de prélever du sang à l'entrée et à la sortie des organes d'intérêt (pose de cathéters implantés en chronique), et de mesurer les débits sanguins des vaisseaux cathétérisés (pose de sondes débitométriques ou dilution de marqueur). Ainsi il est possible de mesurer des flux de nutriments entre organes et d'en calculer les flux nets au travers des organes (de production ou captation) au cours de la journée en fonction de différents facteurs de variation (composition de la ration, niveau de production, état physiologique de l'animal...). Parmi les différents nutriments d'intérêt, le glucose est l'un des principaux nutriments énergétiques étudié (Ortigues I. et Visseiche A.-L., 1995), c'est pourquoi nous développerons dans cet exemple la caractérisation de cette méthode.

3. Méthode de dosage du glucose

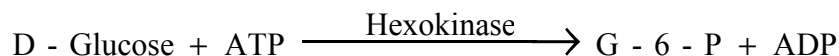
3.1 Principe du dosage

La concentration en glucose est déterminée sur un échantillon sanguin prélevé sur anticoagulant (EDTA ou Héparine de lithium). L'échantillon sanguin est déprotéinisé immédiatement pour minimiser la glycolyse qui peut conduire à une sous estimation de la concentration de glucose d'environ 10 à 20 % (Kunst *et al.*, 1984).

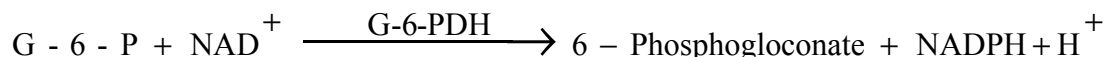
La déprotéinisation est réalisée en ajoutant 2 volumes d'acide perchlorique à 0.6 mol/l dans un volume de sang fraîchement prélevé. Le surnageant obtenu après une centrifugation de 20 mn à 4500 tours / mn est récupéré puis conservé à -20°C jusqu'à l'analyse.

La détermination du glucose sanguin est ensuite réalisée sur le surnageant en présence de deux enzymes l'hexokinase et la glucose-6-phosphate deshydrogénase (G-6-PDH) selon la méthode décrite par Bergmeyer *et al.*, (1974).

Le glucose est phosphorylé par l'adénosine triphosphate (ATP) au cours d'une réaction enzymatique catalysée par l'hexokinase, pour produire du glucose-6-phosphate (G-6-P) selon l'équation :



Le glucose-6-phosphate est ensuite oxydé en 6-phosphogluconate avec la réduction du NAD^+ en NADPH par l'action de la G-6-PDH.



La quantité de NADPH formé est proportionnelle à la concentration du glucose présent dans l'échantillon et son absorbance est mesurée à 340 nm.

La quantité de glucose présent dans les échantillons sanguins, exprimé en g/l, est prédite à partir d'une équation d'étalonnage de type $y = ax + b$, établie à partir d'une gamme étalon.

3. 2 Principe de la caractérisation

Rappel des principes de la caractérisation de méthode

De manière générale, la caractérisation d'une méthode sert à démontrer, à partir de données expérimentales, que les performances de la méthode permettent de répondre aux exigences de l'usage auquel elle est destinée (NF ISO 17025, mai 2000). La caractérisation est l'étape ultime du développement d'une nouvelle méthode analytique avant son application en routine ce qui suit la logique du cycle de vie d'une méthode d'analyse (**figure 1**). Elle doit donc être réalisée lorsque les étapes de mise au point et/ou d'optimisation ont été finalisées et que le mode opératoire ait été rédigé (Feinberg M., Septembre 2001).

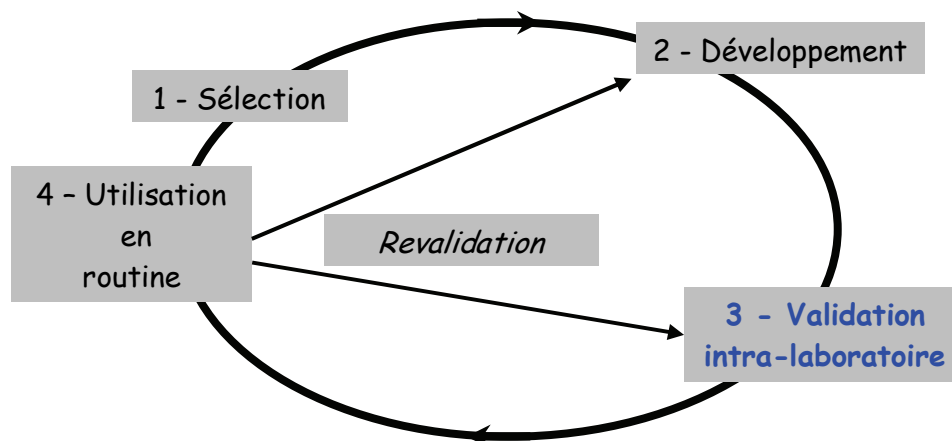


Figure 1 : cycle de vie d'une méthode d'analyse

La caractérisation d'une méthode d'analyse s'effectue selon une procédure bien définie (protocole de caractérisation) établie d'après la norme AFNOR V03-110 et les recommandations de Max Feinberg (Feinberg M., 2001).

Les performances d'une méthode sont évaluées selon plusieurs critères. Les critères les plus fréquents sont la linéarité, la spécificité, la répétabilité, la fidélité intermédiaire, la justesse, la limite de détection et la limite de quantification (Feinberg M., 2010^a).

Lorsque cette procédure de caractérisation est achevée, l'analyste doit confronter les résultats obtenus pour chacun des critères aux exigences définies par le chercheur de manière à vérifier que la méthode sera en mesure d'y répondre. C'est seulement après cette étape que la méthode d'analyse pourra être utilisée en routine.

3. 3 Caractérisation de la méthode

Nous avons choisi, dans le cas du dosage du glucose, de déterminer les performances de la méthode au travers des critères suivants : linéarité, limite de détection et de quantification et fidélité.

Les différents calculs ont été effectués à l'aide d'une application développée sous Excel (A. Thomas, M. Feinberg, Y. Anglaret et D. Durand - Mars 2000), les tableaux de résultats étant pour la plupart issus directement de l'application Excel.

3.3.1 Linéarité de la « fonction de réponse »

Objectif

Dans la méthode que nous avons développée, la quantité de glucose sanguin est obtenue à partir d'une méthode indirecte qui nécessite la mise en place d'un étalonnage. La linéarité d'une méthode est sa capacité, à l'intérieur d'un certain domaine, à fournir des résultats directement proportionnels à la concentration en analyte à doser. Le test de linéarité va permettre de vérifier la validité du modèle de régression et de vérifier le domaine d'étalonnage pour en déduire les caractéristiques d'étalonnage. Il faut donc choisir *a priori* le domaine d'étalonnage. Nous avons choisi celui-ci à partir des « exigences » inhérentes à l'expérimentation elle-même, à savoir une « plage » de concentration variant 0,40 et 0,80g/l.

Organisation

Les exigences de la norme V03-110 préconisent pour mettre en évidence une non-linéarité au risque de 1% de préparer, un minimum de 5 solutions étalons et de faire 5 répétitions de chacune d'elle. Nous avons choisi de tester 9 niveaux de concentrations analysés sur 9 journées. Les niveaux de concentrations ont été répartis régulièrement entre 0 à 1,6 g/l, donc encadrant largement la « plage » de concentration souhaitée au départ (0,40 et 0,80 g/l) de façon à pouvoir utiliser cette méthode dans d'autres situations physiologiques plus "extrêmes" (comme lors d'un stress intense, lors des phases peripartum ...).

Les niveaux de concentration ont été préparés par dilutions indépendantes les unes des autres à partir d'une solution mère de glucose à 2 g/l (Sigma). L'analyse des 9 niveaux de concentration a été répétée sur 9 jours à partir de solution fraîchement préparée de manière à éviter tout biais systématique (**tableau 1**). Les essais ont été organisés en suivant le tableau ci-dessous :

Tableau 1 : tableau récapitulatif de l'organisation des essais conduits en vue de la détermination des performances d'étalonnage du glucose

Niveaux	Séries (Jours)	Concentrations des points de gamme (y)	Réponses instrumentales (x)
1	1	0	Y11
	...	0	...
	9	0	Y19
2	1	0.2	Y12

	9	0.2	Y29
...
9	1	1.6	Y91

	9	1.6	Y99

Évaluation des caractéristiques

Après réalisation du plan expérimental complet, la première étape consiste à établir la fonction de réponse la plus appropriée à nos données. Dans notre exemple, nous avons délibérément testé le modèle linéaire, la grande majorité de nos méthodes utilisées dans l'unité obéissant à ce modèle. Le modèle de régression que nous avons testé était donc le suivant :

$$y = a_1 x + a_0$$

où :

y : la concentration

a₁ : la pente

x : la réponse instrumentale (densité optique (DO))

a₀ : l'ordonnée à l'origine

La méthode utilisée pour réaliser le test de linéarité met en œuvre un test d'erreur d'adéquation du modèle basé sur l'analyse de variance. Ce test consiste à décomposer la somme des carrés des écarts des réponses, en différentes SCE, selon l'équation ci-dessous :

$$SCE_y = SCE_e + SCE_{em} + SCE_{reg}$$

où :

SCE_{reg} ⇒ somme des carrés des écarts due à la régression

SCE_{em} ⇒ somme des carrés des écarts due à une erreur de modèle

SCE_e ⇒ somme des carrés des écarts résiduels

Le test de linéarité va permettre de vérifier 2 hypothèses :

- le modèle choisi est-il acceptable ?
- le domaine d'étalonnage est-il jugé acceptable, c'est à dire n'existe-t-il pas une courbure significative dans le domaine choisi ?

Résultats

Tableau 2 : résultats du test de linéarité du modèle d'étalonnage

Source de variation	Somme des carrés	Degrés de liberté	Variance	F calculé	F de table de fisher P 0,01	Conclusion
Linéarité (SCEI)	15,9516	1	15,95163	21670	7,001	acceptable
Non linéarité (SCEnl)	0,0014	7	0,00020	0,269	2,898	acceptable
Résiduelle (SCr)	0,0530	72	0,00074			
Totale (somme)	16,0060	80				

Au vu des résultats statistiques (**tableau 2**) entre les concentrations 0 et 1,6 g/l, le modèle de régression choisi est acceptable (F calculé > F limite déterminé à partir de la table de Fisher) et le modèle est linéaire dans tout le domaine choisi (F calculé < F critique déterminé à partir de la table de Fisher).

A partir de ces données ayant permis d'établir la fonction réponse il est possible de déterminer les limites de détection et de quantification.

3.3.2 Limite de détection et de quantification

Objectif

Il est important pour le chercheur de connaître ces limites qui pourront lui permettre de savoir à partir de quelle concentration il pourra mettre en évidence avec certitude la présence de l'analyte recherché (ici le glucose) (limite de détection) et à partir de quelle concentration il pourra garantir la fiabilité de son résultat et donc en tirer une interprétation mécanistique (limite de quantification).

Organisation

L'estimation des limites de détection et de quantification a été réalisée à partir des paramètres obtenus à partir du modèle linéaire de la fonction de réponse.

Évaluation des limites de détermination

Les mesures collectées sont utilisées pour calculer la sensibilité ou pente (a_1) et son écart-type (s_{a1}), ainsi que le blanc ou ordonnée à l'origine (a_0) et son écart-type (s_{a0}).

Les limites de détection et de quantification sont déterminées de la façon suivante :

$$\text{Limite de détection (LD)} = \frac{a_0 + 3 s_{a0}}{a_1}$$

$$\text{Limite de quantification (LQ)} = \frac{a_0 + 10 s_{a0}}{a_1}$$

Résultats

Tableau 3 : estimation des limites de quantification et de détection

Limite de détection, de quantification		
Variance expérimentale	X g/l	0,00069
Ecart-type expérimental	X g/l	0,02624
Limite de détection X_{LD}	X g/l	0,089
Limite de quantification X_{LQ}	X g/l	0,133

Comme nous pouvons le lire dans le **tableau 3**, la limite de détection est de 0,089 g/l, la valeur à partir de laquelle nous pouvons garantir la fiabilité de notre mesure ($P < 0,01$) est de 0,133 g/l.

3.3.3 Fidélité

Objectif

Ce test va permettre à l'analyste d'estimer la dispersion associée à la mesure (erreur aléatoire). Elle peut être évaluée dans plusieurs conditions, de répétabilité (intra-série ou intra-jour), de fidélité intermédiaire (inter-jours) et de reproductibilité (inter-laboratoires). Dans cette approche, nous n'avons traité que les conditions de répétabilité et de fidélité intermédiaire.

Organisation

Afin de répondre aux besoins de l'étude, ce test a été réalisé sur la matrice d'intérêt, à savoir un échantillon sanguin d'ovin situé dans des conditions physiologiques de base. Cet échantillon a été traité comme les échantillons expérimentaux c'est-à-dire qu'après prélèvement au niveau de la veine jugulaire, le sang a été déprotéinisé, aliquoté puis conservé à -20°C .

Ensuite, un aliquote d'échantillon déprotéinisé a été analysé en triple (conditions de répétabilité) et répété sur 9 jours (conditions de fidélité intermédiaire). Le nombre de répétitions intra et inter journée (ou série) a été défini en fonction des exigences de la norme française (NF ISO 5725-1, 1994) qui recommande de faire au moins 8 séries d'au moins 2 répétitions de l'échantillon.

Évaluation des caractéristiques

Ce test utilise comme modèle statistique, une analyse de variance à un seul facteur à effet aléatoire. Cette analyse permet de déterminer l'écart-type de répétabilité, de fidélité intermédiaire ainsi que les limites de répétabilité et de fidélité intermédiaire pour un risque d'erreur de 5%.

Résultats

Tableau 4 : estimation des critères de fidélité

Tableau 4 : Calculs de la répétabilité et de la Fidélité intermédiaire	
Moy générale	= 0,82
S ² r	= 0,00016
S ² L	= 0,00035
S ² R	= 0,00051
Sr	= 0,013
SR	= 0,023
Répétabilité et fidélité intermédiaire	CV de répétabilité et de fidélité intermédiaire
r _{95%} = 0,04	CVr = 1,55%
R _{95%} = 0,06	CVR = 2,75%

La méthode enzymatique que nous avons utilisée pour doser le glucose sanguin est caractérisée par des coefficients de variation (CV) de répétabilité de 1,55% et de fidélité intermédiaire de 2,75%.

3.3.4 Conclusion

La méthode sélectionnée et développée sur du sang ovin permet de déterminer des concentrations de glucose comprises entre 0,13 g/l et 1,6 g/l avec des performances qui nous apparaissent comme très satisfaisantes par rapport aux variations attendues pour la majorité des études conduites dans notre équipe.

Par ailleurs, il apparaît important que l'on puisse associer les performances ainsi déterminées aux besoins spécifiques de certaines expérimentations en particulier dans le cas d'expérimentation exigeant une très bonne limite de quantification. Nous allons dans la suite de cet article illustrer par un exemple d'expérimentation conduite dans notre équipe cette exploitation des performances d'une méthode.

4. Exemple d'utilisation des caractéristiques d'une méthode à l'établissement du plan expérimental

4.1 Objectifs

Lors de la mise en place d'une expérimentation, le chercheur formule une hypothèse à tester, pour répondre à une question scientifique élaborée à partir des connaissances disponibles. Il doit ensuite élaborer le protocole expérimental et faire des choix « opérationnels » les plus appropriés pour tester l'hypothèse émise. Pour cela, le chercheur doit sélectionner les variables qu'il voudra mesurer mais également la méthode de dosage de ces variables. Le choix des méthodes associées aux variables dépend de leur facilité de mise en œuvre, de l'équipement disponible dans l'équipe, du coût mais peut également dépendre des performances exigées. C'est sur ce dernier point que nous allons particulièrement insister.

4.2 Contraintes liées à l'expérimentation

Nous avons dans un premier temps réalisé des pré-essais sur des animaux recevant différentes rations alimentaires. Nous avons ainsi mesuré les concentrations en glucose dans plusieurs vaisseaux d'intérêt comme l'aorte, la veine porte, la veine sus-hépatique et la veine iliaque pour étudier les flux nets de glucose à travers le foie, le tube digestif et le train arrière. Nous avons déterminé les variations de concentrations en glucose au cours de la journée et ceci pour chacun des vaisseaux et en fonction des différents traitements expérimentaux appliqués aux animaux (niveaux d'énergie variables entre lots). Il est apparu clairement que la différence minimale en glucose entre vaisseaux relevée parmi ces traitements expérimentaux était voisine de 0,050 g/l soit seulement 8 % des concentrations moyennes habituellement enregistrées chez des ovins. Compte tenu de ces différences attendues et des performances de notre méthode nous avons pu apporter des éléments objectifs de choix pour l'élaboration du plan expérimental. C'est ce que nous allons présenter dans la suite de ce paragraphe.

4.3 Quel plan expérimental mettre en place ?

D'après le principe de l'intervalle de confiance, il est possible de calculer le nombre d'animaux théorique à mettre en expérimentation par lot (par traitement alimentaire) pour mesurer avec une incertitude contrôlée les différences entre vaisseaux. Le calcul (**équation 1**) fait intervenir la différence (Δ) entre traitements (ici entre vaisseaux), les écarts-types de répétabilité ou de fidélité intermédiaire de la méthode (σ) et le nombre de répétitions de la mesure (n) qui correspond ici au nombre d'animaux par traitement.

Equation 1

$$\Delta \geq 2 \times \left(t \times \frac{\sigma}{\sqrt{n}} \right)$$

Où :

Δ = différence

t = variable de Student (probabilité 0,05 ou 0,01)

σ = écart-type de répétabilité ou écart-type de fidélité

n = nombre d'animaux

Nous avons simulé plusieurs situations :

- soit les analyses de glucose peuvent-être réalisées lors d'une même série de dosage (donc le même jour), et dans ce cas nous avons utilisé l'écart type de répétabilité, soit elles sont réalisées sur plusieurs jours et dans ce cas nous avons utilisé l'écart-type de fidélité intermédiaire. Ces deux situations peuvent être rencontrées dans les études mises en œuvre par l'équipe de recherche.
- nous avons fixé 2 niveaux de confiance habituellement utilisés en biologie et qui sont 0,01 et 0,05.

Tableau 5 : nombre d'animaux à utiliser selon les conditions de dosage et précision recherchées

σ	Différences		Différences	
	Répétabilité (0,013 g/l)		Fidélité intermédiaire (0,023 g/l)	
P	0,05	0,01	0,05	0,01
n				
2	0,413	2,071	0,234	1,170
3	0,114	0,264	0,065	0,149
4	0,073	0,134	0,041	0,076
5	0,057	0,095	0,032	0,054
6	0,048	0,076	0,027	0,043
7	0,043	0,064	0,024	0,036
8	0,038	0,057	0,022	0,032
9	0,035	0,051	0,020	0,029
10	0,033	0,047	0,019	0,027
11	0,031	0,044	0,017	0,025
12	0,029	0,041	0,017	0,023

Le **tableau 5** rassemble les résultats de ces simulations et montre clairement que pour mettre en évidence une différence de concentration en glucose entre vaisseaux avec la probabilité la plus exigeante ($P < 0,01$) il serait nécessaire de conduire 10 ou 6 animaux selon que le laboratoire réalise tous les dosages dans une même série avec le même opérateur (conditions de répétabilité) ou dans des séries différentes (conditions de fidélité intermédiaire). Le nombre d'animaux se réduit à 6 (condition de répétabilité) ou 4 (condition de fidélité intermédiaire) si l'expérimentateur n'exige qu'une probabilité de 5 %. Ainsi, il apparaît clairement que le nombre d'animaux à mettre en expérimentation est variable selon les possibilités analytiques du laboratoire et les exigences du chercheur en termes de niveau statistique souhaité.

Conclusion

La caractérisation d'une méthode permet donc de connaître les performances de cette méthode et d'en déduire des indications précieuses pour l'élaboration du plan expérimental à mettre en œuvre. Elle garantit que le dosage effectué est compatible avec les « possibilités » qu'offre la méthode (mesure réalisée au dessus de la limite de quantification, dans le domaine de linéarité, avec une incertitude liée aux critères de fidélité...). En revanche, elle ne permet pas d'accompagner la valeur rendue au commanditaire avec des incertitudes connues.

Il apparaît clairement que la validation de méthode, telle qu'elle sera abordée dans les chapitres suivants, permet à l'analyste de garantir que chaque mesure réalisée avec cette méthode sera « accompagnée » d'une limite d'acceptabilité et d'un intervalle de tolérance qui définissent des incertitudes sur la valeur transmise au commanditaire. Ces deux paramètres, limite d'acceptabilité et intervalle de tolérance, sont fixés par le futur utilisateur (le plus souvent le chercheur) et/ou par l'analyste en fonction de ses besoins (Feinberg M., 2010^b).

Bibliographie

- Bergmeyer H.U., Bernt, Schmit F. and Stork H. (1974) Determination with hexokinase and glucose-6-phosphate dehydrogenase. *In* "Methods of enzymatic analysis" volume 3rd ed. H.U. Bergmeyer Verlag Chemie Academic Press, Inc. 1196-1227
- Feinberg M. (2001) Validation interne des méthodes d'analyse. Techniques de l'ingénieur. Mesures – Analyse – Qualité et sécurité au laboratoire. P224, 1-23
- Feinberg M. (2010^a) Principes et vocabulaire *in* Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 13-25
- Feinberg M. (2010^b) Mise en œuvre du profil d'exactitude *in* Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 27-44
- Kunst A. *et al.*, (1984) D-glucose. UV-methods with hexokinase and glucose-6-phosphate dehydrogenase. *In* Methods of Enzymatic Analysis. 3rd edition. H. U. Bergmeyer, editor. Verlag Chemie, Weinheim, Germany, 163-172
- Majdoub L. *et al.*, (2003) Intraruminal propionate supplementation modifies hindlimb energy metabolism without changing the splanchnic release of glucose *in* growing lambs. *British Journal of Nutrition* 89, 39-50
- Ortigue I. et Visseiche A.-L. (1995) Whole-body fuel selection in ruminants : nutrient supply and utilization by major tissues. *Proceedings of the Nutrition Society* 54, 235-251
- Thomas A. et Durand D. (2009) Démarche qualité de l'unité de recherche sur les herbivores, Inra. *In* La démarche qualité dans la recherche publique et l'enseignement supérieur. Editions QUAE. Granier C., Mas L.-Y., Finot L., Arnoux B., Pasqualini N., Dollé V. coordinateurs. 13, 121-131
- Verschuere B. (2008) Expérimentation animale : les comités d'éthique en 2008. *Bull. Acad. Vét. France*. Tome 161 N°5 <http://www.academie-veterinaire-defrance.org/> (Consulté le 5/2/2010)
- NF ISO 5725-1, (Décembre 1994) Exactitude (justesse et fidélité) des résultats et méthodes de mesure. Partie 1 : Principes généraux et définitions.
- NF V03-110, (Décembre 1998) Procédure de validation intralaboratoire d'une méthode alternative par rapport à une méthode de référence. AFNOR
- NF ISO/CEI 17025, (Mai 2000) Prescriptions générales concernant la compétence des laboratoires d'étalonnages et d'essais. AFNOR

Validation de l'analyse des hydrocarbures aromatiques polycycliques dans les eaux de pluie par la méthode du profil d'exactitude

Marjolaine Bourdat-Deschamps¹

Résumé : Une méthode d'analyse des hydrocarbures aromatiques polycycliques (HAP) contenus dans l'eau de pluie a été optimisée au laboratoire par la méthodologie des plans d'expériences. La méthode d'analyse (absorption sur barreau magnétique SBSE – analyse en chromatographie liquide haute performance avec détection de fluorescence) présente les deux caractéristiques suivantes : l'étape d'extraction ne permet pas d'analyser plus de 12 échantillons par jour et c'est une méthode multi-analytes (15 HAP). Cet article décrit sa validation par la méthode du profil d'exactitude ; les résultats pour le pyrène et le benzo(k)fluoranthène sont présentés à titre d'exemple.

Mots clés : hydrocarbures aromatiques polycycliques (HAP), eau de pluie, validation, profil d'exactitude

Introduction

Les HAP sont des molécules apolaires, issues de la combustion incomplète des matières organiques dont certaines sont cancérigènes, mutagènes et toxiques pour la reproduction. Les flux de HAP provenant des dépositions atmosphériques totales (retombées sèches et humides) sont de l'ordre de 150 µg/m²/an (zone urbaine et péri-urbaine, Azimi *et al.*, 2005). Disposant d'un pluviomètre permettant de récolter les retombées humides (pluie) et suivant par ailleurs les teneurs en HAP dans les sols dans le cadre d'un suivi expérimental de longue durée, nous avons souhaité connaître le flux des HAP *via* l'eau de pluie en regard du stock en HAP du sol. Une méthode d'analyse a été optimisée (Bourdat-Deschamps *et al.*, 2007) qui permet de concentrer les HAP par absorption sur barreau magnétique (SBSE : Stir Bar Sorptive Extraction), de les désorber dans un très faible volume d'acétonitrile et de les analyser par chromatographie liquide haute performance avec un détecteur de fluorescence. La méthode a été caractérisée (linéarité, répétabilité à deux niveaux de concentration, fidélité intermédiaire, limite de quantification). La méthode du profil d'exactitude a été employée pour valider l'analyse des HAP contenus dans les eaux de pluie. Dans une première partie, chaque étape de la méthode de validation est détaillée (cf. Feinberg M, 2010^b). Dans une deuxième partie, différents profils d'exactitude du pyrène et du benzo(k)fluoranthène sont présentés et commentés.

¹ UMR1091 EGC- Environnement et Grandes Cultures, INRA-F-78850 Thiverval-Grignon
☎ 01 30 81 53 12 ✉ marjolaine.deschamps@grignon.inra.fr

1. Étapes pour la validation (Max Feinberg, 2010^b)

1.1. Étape 0 : description du mode opératoire

Objectif de la méthode

Quantifier 15 HAP (naphtalène, acénaphène, fluorène, phénanthrène, anthracène, fluoranthène, pyrène, benzo(a)anthracène, chrysène, benzo(b)fluoranthène, benzo(k)fluoranthène, benzo(a)pyrène, dibenzo(ah)anthracène, benzo(ghi)pérylène, indéno(1,2,3-cd)pyrène) contenus dans des eaux de pluie. La méthode doit être juste (nous avons besoin des concentrations en ng/l pour calculer les flux de HAP *via* l'eau de pluie et les comparer au stock du sol, qui est déterminé par une autre méthode analytique). La méthode doit être fidèle (les analyses sont réalisées une fois par an, il n'y a pas de répétition par échantillon). Les résultats sont destinés aux scientifiques utilisant le dispositif expérimental.

Méthode analytique optimisée et mode opératoire rédigé sous trame qualité du laboratoire

Voici quelques informations concernant le mode opératoire (Bourdat-Deschamps *et al.*, 2007).

Les eaux de pluies sont collectées de façon hebdomadaire et sont congelées. Après décongélation, un échantillon moyen mensuel est réalisé. Dans un volume de 10 ml d'eau de cet échantillon, additionné de 200 µl d'acétonitrile, on place un Twister® (barreau aimanté recouvert d'un film polymérique), préalablement conditionné. L'extraction est réalisée dans un bain marie à 25°C, avec une agitation de 800 tr/mn pendant 140 mn. Le barreau est ensuite placé dans un flacon HPLC avec insert. Un volume de 100 µl d'acétonitrile est rajouté. L'étape de désorption réalisée à l'abri de la lumière dure 25 mn dans un bain marie à 50°C. Le barreau est ensuite sorti du flacon. Le flacon est rebouché, stocké à l'abri de la lumière avant analyse par HPLC-fluorescence (colonne à polarité de phases inversée spéciale HAP, gradient acétonitrile-eau). La quantification est réalisée par étalonnage externe : dopage d'eau milli-Q à 4 ou 5 niveaux de concentrations de HAP, extraction des HAP selon la procédure détaillée ci-dessus, quantification des HAP contenus dans l'eau de pluie avec la droite d'étalonnage réalisée dans l'eau milli-Q dopée.

Initialement, la méthode prévoyait deux désorptions successives dans l'acétonitrile afin de maximiser le rendement de désorption. Cependant la réalisation de deux désorptions successives limitait à 8 le nombre d'expériences qui pouvaient être réalisées dans une journée. Afin d'augmenter à 12 le nombre d'expériences réalisables par jour, le protocole a été modifié et désormais une seule désorption est réalisée.

En général, 12 échantillons d'eau de pluie sont extraits par jour. Le lendemain, ce sont les échantillons d'étalonnage (eau Milli-Q dopée) qui sont extraits. Puis, l'analyse HPLC est réalisée pour l'ensemble des échantillons d'eau de pluie et d'étalonnage.

Caractérisation réalisée : linéarité, répétabilité à deux niveaux de concentration, fidélité intermédiaire, limite de quantification (pour eau milli-Q) (Cf. Bourdat-Deschamps *et al.*, 2007).

Teneurs approximatives des échantillons à analyser : nous avons analysé des échantillons d'eau de pluie à l'aide de la méthode décrite ci-dessus. Les teneurs sont variables selon les HAP et selon les saisons (de quelques ng/l à quelques dizaines de ng/l).

1.2 Étape 1 : Définition du mesurande

Que mesure-t-on ? Chacun des 15 HAP contenu dans de l'eau de pluie (échantillon moyen mensuel)

Comment exprime-t-on le résultat final ? Un seul échantillon d'eau par mois est analysé (pas de répétition). La teneur (ng/l) est déterminée par étalonnage externe (eau milli-Q dopée).

1.3 Étape 2 : paramètres de la validation

Domaine d'application (*a priori*) : d'après l'étude de linéarité (caractérisation de la méthode) : de 1 à 50 ng/l (voire jusqu'à 100 ou 200 ng/l pour certains HAP). Les analyses préalables d'échantillons d'eau de pluie ont montré que les teneurs maximales de chaque HAP n'excédaient pas 30 ng/l. Le domaine d'application choisi a donc été de **1 à 30 ng/l**.

Limites d'acceptabilité (λ) : $\pm 20\%$, ce qui signifie que pour une teneur de 1 ng/L le résultat sera compris au maximum entre 0,8 et 1,2 ng/L et pour une teneur de 30 ng/L, il sera compris au maximum entre 24 et 36 ng/L.

Intervalle de tolérance (β) : 80 % (pas de réglementation pour cette application, en général, on choisit $\beta \geq 80\%$).

1.4 Étape 3 : échantillons de validation et définition des valeurs de référence

Quels échantillons pour l'étape de validation ? Afin de représenter au mieux la matrice « eau de pluie », un échantillon moyen a été réalisé en rassemblant des eaux de pluies récoltées tous les mois d'une année. Cet échantillon a été homogénéisé et a été stocké à l'abri de la lumière à 4°C pour la durée des expériences de validation. Afin de minimiser la modification de la matrice au cours du temps (développements microbiens), l'ensemble des expériences pour la validation a été réalisé sur 1,5 semaine.

Quelle est la valeur de référence pour chaque échantillon ? Des ajouts dosés à différentes concentrations ont été réalisés sur la matrice ainsi préparée, avec une solution étalon commerciale contenant les 15 HAP, chacun à la même concentration.

1.5 Étape 4 : essais de validation

Quatre niveaux de concentration (1,9 ; 4,7 ; 9,5 ; 28,5 ng/l) ont été choisis. Selon les saisons et les HAP, certaines teneurs risquent d'être proches des limites de quantification. Lors de la caractérisation de la méthode, ces limites avaient été évaluées (rapport signal sur bruit de l'ordre de 10) autour de 1 à 5 ng/l selon les HAP, ce qui explique notre choix pour les deux niveaux bas de concentration.

La matrice que nous avons dopée contient probablement des HAP (c'est de l'eau de pluie naturelle). Un échantillon non dopé a également été analysé.

Deux répétitions par niveau de concentration ont été réalisées car nous pouvons effectuer au maximum 12 extractions par jour. Or (4 niveaux de concentration + 1 échantillon non dopé) \times 2 répétitions = 10 extractions.

Trois séries de mesures ont été réalisées, sur trois jours.

1.6 Étape 5 : essais d'étalonnage

Nous avons choisi d'appliquer le protocole d'extraction SBSE aux échantillons d'étalonnage qui sont constitués d'eau Milli-Q dopée avec les HAP.

Comme indiqué dans le mode opératoire à l'étape 0, les extractions des échantillons de terrain sont réalisées un jour et les extractions des échantillons d'eau milli-Q dopée sont réalisées un autre jour. Nous avons donc procédé de même pour la validation de la méthode. **Pour chaque série**, les extractions SBSE des échantillons d'étalonnage ont été réalisées le lendemain des extractions d'échantillons de validation. Ensuite, les injections HPLC de tous ces échantillons de la même série ont été réalisées sur une nuit.

Cinq niveaux de concentration (de 0,9 à 27,6 ng/l) ont été choisis. Le point le plus élevé de la gamme d'étalonnage est légèrement inférieur à celui de la gamme de validation. Comme le rendement d'extraction SBSE est inférieur à 83 %, cela ne posera pas de problème car les aires obtenues pour les échantillons de validations seront comprises dans la gamme des aires des échantillons d'étalonnage.

Une analyse par niveau a été réalisée. Ce choix a été fait car dans l'organisation de la semaine de travail, il n'était pas possible de réaliser deux répétitions (le conditionnement des barreaux demandant beaucoup de temps). Par ailleurs, le coefficient de variation caractérisant la répétabilité de la méthode pour chaque HAP, déterminé notamment pour de l'eau Milli-Q dopée à 5ng/l, est inférieur à 10% (excepté pour le benzo(ghi)pérylène : 12,8 %), ce qui a conforté ce choix.

Trois séries de mesures ont été réalisées c'est-à-dire trois jours, comme pour les échantillons de validation.

1.7 Étape 6 : réalisation des essais (validation et étalonnage)

Pour les étapes 6 à 9, seuls les résultats concernant le pyrène seront présentés, la méthodologie étant la même pour les 14 autres HAP.

Essais de validation

Tableau 1 : mesurages correspondant aux essais de validation (eau de pluie dopée)

niveaux	séries	valeur ref (ng/L)	Mesurages pyrène (aires)*	
			rep 1	rep 2
1	1	1,9	36539	36785
	2		60086	35295
	3		57695	59731
2	1	4,7	102066	98495
	2		99897	93547
	3		115298	111584
3	1	9,5	188665	191294
	2		188657	198683
	3		221678	194983
4	1	28,5	595999	604704
	2		520857	501025
	3		557258	541355

* pour chaque série, les aires contenues dans les « blancs » (matrice non dopée) ont été soustraites aux aires obtenues pour chaque niveau.

Essais d'étalonnage avec l'eau milli-Q dopée subissant l'intégralité du protocole

Tableau 2 : mesures correspondant aux essais d'étalonnage (eau milli-Q dopée)

niveaux	séries	valeur ref (ng/L)	Mesurages pyrène (aires)
1	1	0,9	43083
	2		24719
	3		35684
2	1	4,6	117767
	2		112319
	3		115840
3	1	9,2	238120
	2		202957
	3		249807
4	1	18,4	450132
	2		398342
	3		440029
5	1	27,6	682393
	2		611523
	3		633748

1.8 Étape 7 : Modèle d'étalonnage pour chaque série

À partir des données du **tableau 2**, pour chaque série d'étalonnage, les coefficients du modèle d'étalonnage sont déterminés (**tableau 3**), sachant que dans notre cas, le modèle d'étalonnage est une droite. Ensuite, à partir des mesurages réalisés sur les échantillons de validation (**tableau 1**), ces coefficients permettent de calculer les concentrations retrouvées des échantillons de validation (tableau 4).

La **figure 1** est une représentation graphique des tableaux 1 et 2.

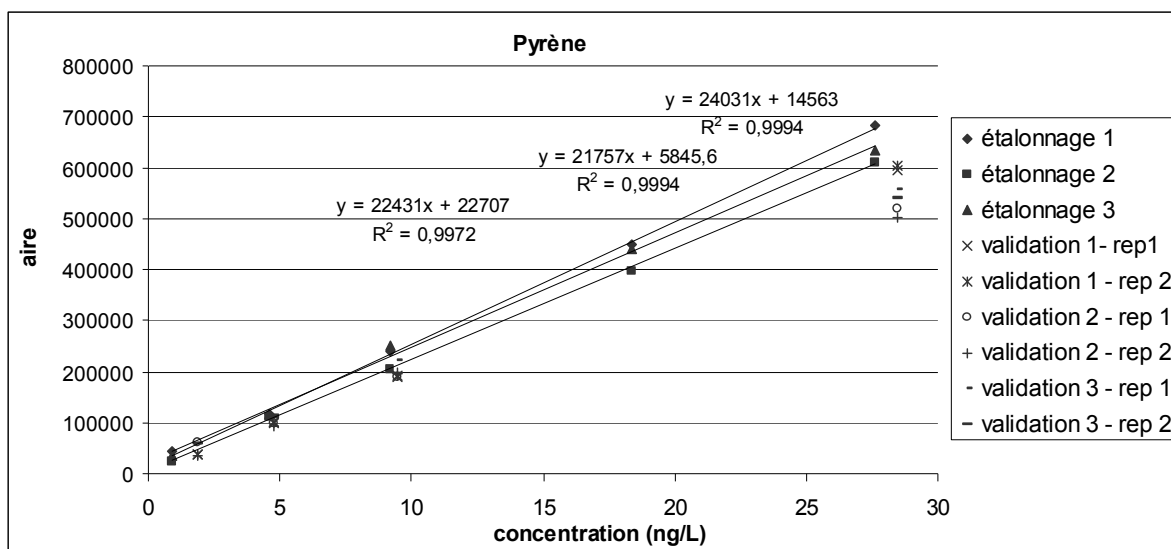
**Figure 1** : représentation graphique des aires mesurées lors des essais d'étalonnage et de validation, en fonction de la concentration

Tableau 3 : détermination des coefficients du modèle d'étalonnage pour chaque série

	a_1	a_0
étalonnage série 1	24031	14563
étalonnage série 2	21757	5845
étalonnage série 3	22431	22707

Tableau 4 : concentrations retrouvées des échantillons de validation

Niveaux	Séries	Valeur de référence (ng/l)	Concentrations pyrène (ng/l)	
			Rép. 1	Rép. 2
1	1	1,9	0,91	0,92
	2		2,49	1,35
	3		1,56	1,65
2	1	4,7	3,64	3,49
	2		4,32	4,03
	3		4,13	3,96
3	1	9,5	7,24	7,35
	2		8,40	8,86
	3		8,87	7,68
4	1	28,5	24,20	24,56
	2		23,67	22,76
	3		23,83	23,12

1.9 Étapes 8 et 9 : calcul et tracé du profil d'exactitude

Les données du **tableau 4** sont saisies dans le fichier Excel® pour réaliser les calculs et tracer le profil d'exactitude. L'exploitation des données de chaque HAP, comme présentée aux paragraphes 1.7 à 1.9 permet de tracer le profil d'exactitude pour chacun des HAP.

2. Résultats et interprétations

2.1 Résultats pour le pyrène

Le tracé du profil d'exactitude pour le pyrène est présenté sur la **figure 2**.

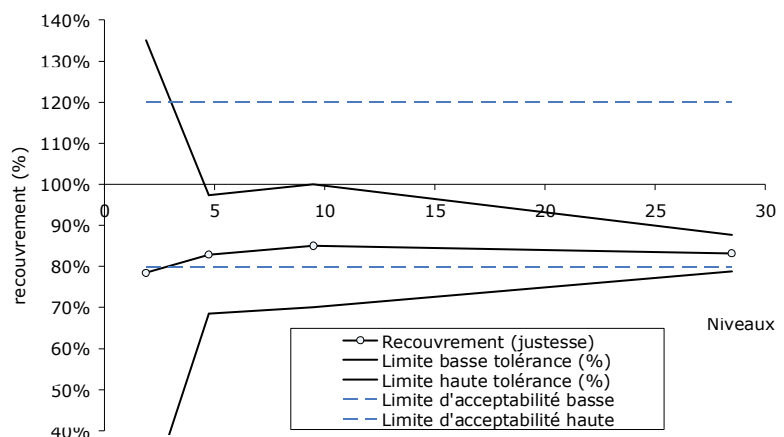


Figure 2 : profil d'exactitude pour le pyrène (d'après tableau 4), pour $\lambda=20\%$ et $\beta=80\%$

La méthode présente un **biais systématique** (la ligne de recouvrements moyens n'est pas à 100%) qui est explicable par un effet matrice. Or cela invalide la méthode puisque nous avons besoin qu'elle soit juste. En traçant les concentrations retrouvées (données du **tableau 4**) en fonction des concentrations théoriques pour les échantillons de validation (**figure 3**), il est possible de déterminer la pente moyenne des taux de recouvrement. L'inverse de cette pente moyenne donne le **facteur de correction** (1,20) que l'on peut appliquer à chacune des concentrations retrouvées afin de retracer un profil d'exactitude corrigé (**figure 4**).

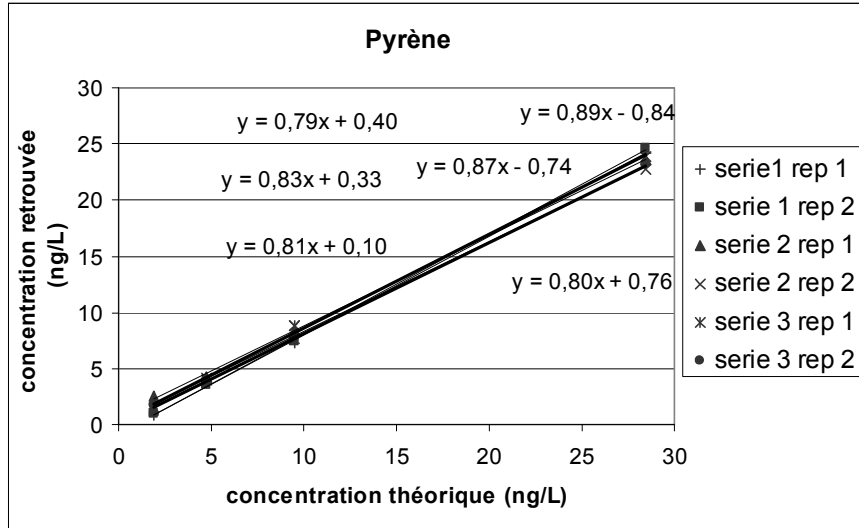


Figure 3 : concentrations de pyrène retrouvées en fonction des concentrations théoriques des échantillons de validation

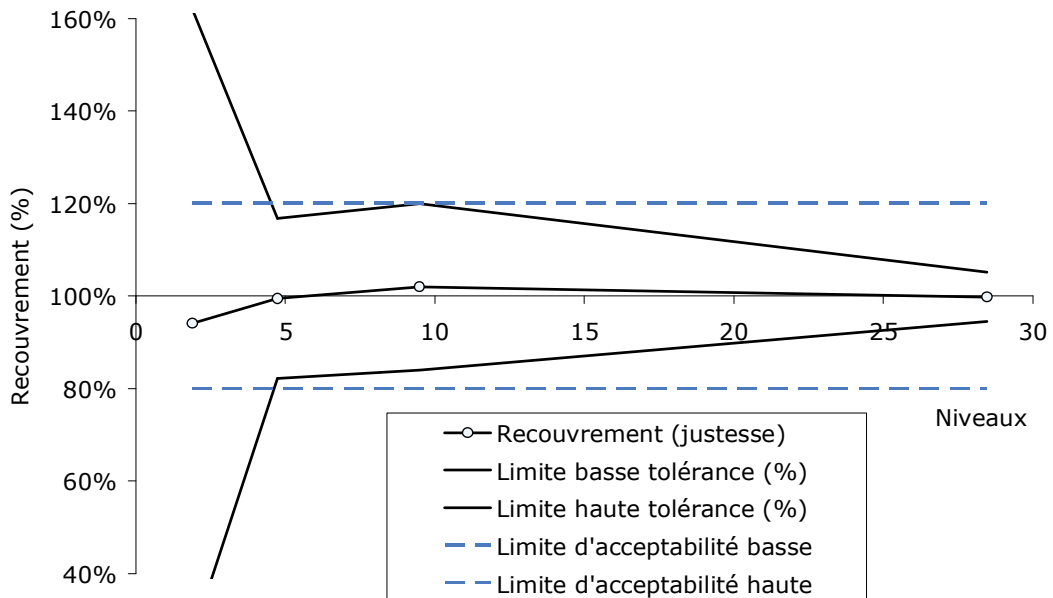


Figure 4 : profil d'exactitude pour le pyrène en appliquant un facteur de correction (1,20) aux concentrations de validation ($\lambda=20\%$ et $\beta=80\%$)

En corrigeant les valeurs retrouvées avec le facteur de correction, la **méthode devient juste** (biais inférieur à 3% pour des concentrations entre 5 et 28 ng/l).

Afin de valider le facteur de correction, une nouvelle série de mesures a été réalisée (échantillons de validation : 4 niveaux et 2 répétitions par niveau ; échantillons d'étalonnage : 5 niveaux sans répétition). Le facteur de correction déterminé auparavant a été appliqué aux concentrations retrouvées. Les teneurs corrigées se situent dans le profil d'exactitude présenté **figure 4**.

La **limite de quantification basse** est déterminée par l'intersection entre la limite d'acceptabilité et la limite de tolérance. Elle est de 4,7 ng/l. La **limite de quantification haute** correspond au point haut de gamme des échantillons de validation (28,5 ng/l).

La **méthode est donc validée entre 4,7 et 28,5 ng/l ($\lambda = 20\%$ et $\beta = 80\%$)** car dans ce domaine, l'intervalle de tolérance est compris dans l'intervalle d'acceptabilité.

En dessous de 4,7 ng/l, les **limites de tolérance haute et basse divergent très rapidement** et de façon symétrique.

Vers 10 ng/l, la **limite haute de tolérance est proche de la limite d'acceptabilité haute**, ce qui signifie que le risque moyen d'avoir en routine une proportion moyenne de résultats non acceptables est proche de 1/10 (car $\beta = 80\%$, soit une proportion moyenne de 1/5, et l'intervalle de tolérance est proche de la limite d'acceptabilité sur la partie haute du profil seulement, ce qui divise le risque par 2). Mais une proportion moyenne de 1/10 signifie qu'elle peut varier autour de cette moyenne. Comme les intervalles de tolérance se resserrent de 10 à 28 ng/l, ce risque diminue fortement.

Le profil d'exactitude peut être employé pour **choisir les teneurs des contrôles qualités (CQ)** qui seront employés lors des analyses en routine. En effet, on pourra choisir un CQ proche de 10 ng/l. Le CQ se comportant comme les échantillons, s'il ne passe pas, il faudra refaire la série d'analyse. Si on choisit d'utiliser 3 CQ de niveaux de concentrations différents, on peut choisir un CQ à 5 ng/l (proche de la limite de quantification basse), un CQ à 10 ng/l (concentration « critique » du profil) et un CQ à 25 ng/l (proche de la limite haute).

Les résultats présentés à la **figure 1** ont été obtenus en prenant une droite ($a_1 X + a_0$) comme modèle d'étalonnage. D'autres modèles d'étalonnage ont été testés (droite passant par zéro, modèle quadratique), ainsi que des pondérations ($1/X$) de ces modèles. Ces modèles d'étalonnage, pondérés ou non, ont été déterminés à partir des données du **tableau 2**, en utilisant le logiciel Statgraphics Plus 4.1. Les coefficients de ces modèles ont été appliqués aux données du **tableau 1** afin de déterminer les concentrations de pyrène retrouvées dans chaque échantillon de validation. Les profils d'exactitude (avec application d'un facteur de correction aux données de validation) correspondant à chaque modèle ont été tracés (**figure 5**).

Aucun modèle ne donne de profil plus satisfaisant (justesse, étroitesse de l'intervalle de tolérance, étendue du domaine de validité de la méthode) que le profil obtenu avec le modèle linéaire avec facteur de correction (**figure 4**).

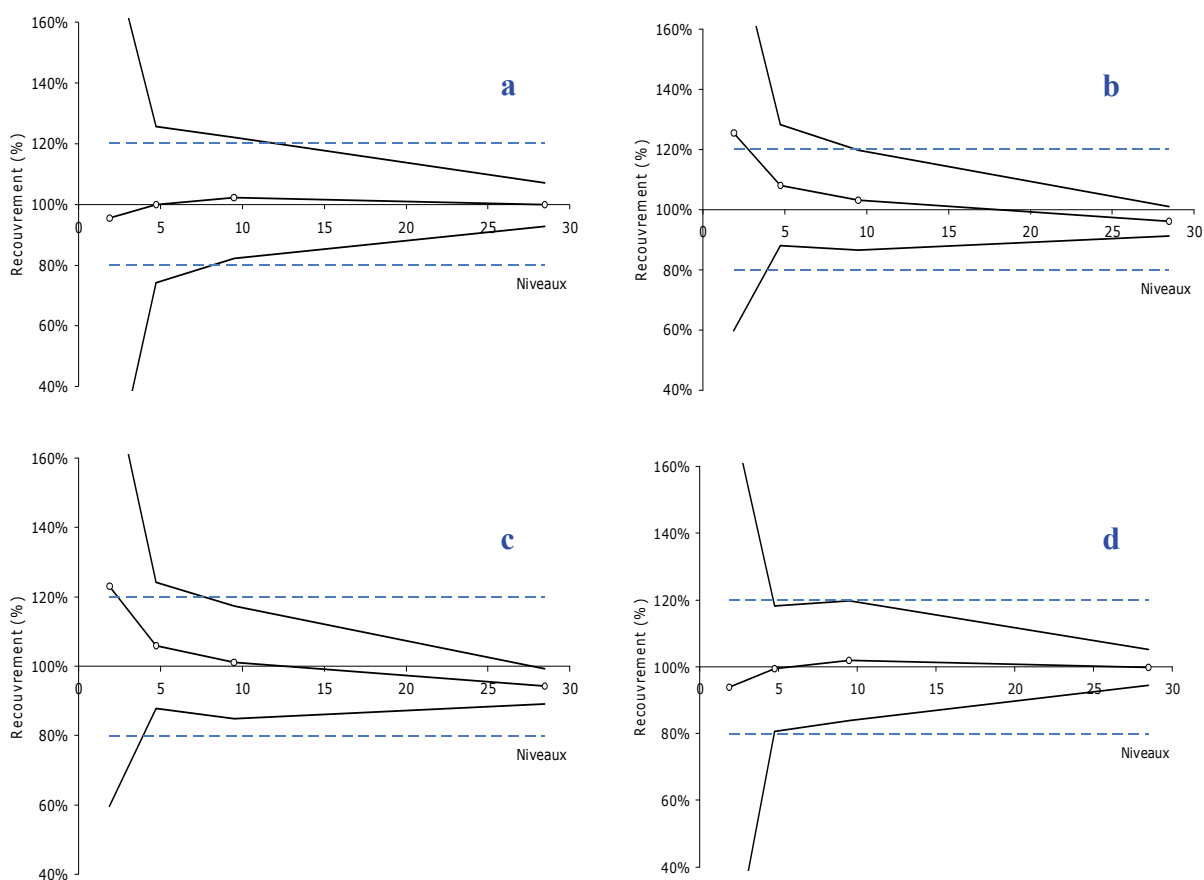


Figure 5 : profils d'exactitude pour le pyrène ($\lambda=20\%$ et $\beta=80\%$)

- (a) modèle linéaire ($a_1 X + a_0$) pondéré ($1/X$), facteur de correction (1,20)
- (b) modèle linéaire passant par zéro ($a_1 X$), facteur de correction (1,16)
- (c) modèle linéaire passant par zéro ($a_1 X$) pondéré ($1/X$), facteur de correction (1,16)
- (d) modèle quadratique ($a_2 X^2 + a_1 X + a_0$), facteur de correction (1,20)

Nous avons choisi d'appliquer aux échantillons d'étalonnage l'ensemble du protocole expérimental (extraction SBSE, désorption, analyse HPLC) (cf. étape 5). Cependant, nous pourrions choisir comme échantillons d'étalonnage des solutions dans un solvant organique (HAP dans l'acétonitrile), qui ne subiraient pas le protocole d'extraction. Cette méthode présente l'avantage de raccourcir la durée des expérimentations, mais l'inconvénient de ne pas montrer l'impact de l'extraction sur la fonction de réponse (droite dans notre cas). Le profil d'exactitude, **figure 6**, a été déterminé en réalisant les essais d'étalonnage avec des étalons de HAP dans l'acétonitrile (4 niveaux de concentrations, 2 répétitions par niveau, 3 séries, qui ont été injectés en HPLC en même temps que les échantillons de validation et les échantillons d'étalonnage subissant l'intégralité du protocole), et en appliquant un facteur de correction de 1,5. La méthode n'est juste (biais inférieur à 5 %) qu'entre 10 et 28 ng/l et la limite de quantification inférieure est assez élevée (13,4 ng/l), ce qui nous conforte dans notre choix de faire subir aux échantillons d'étalonnage l'intégralité du protocole extraction-analyse.

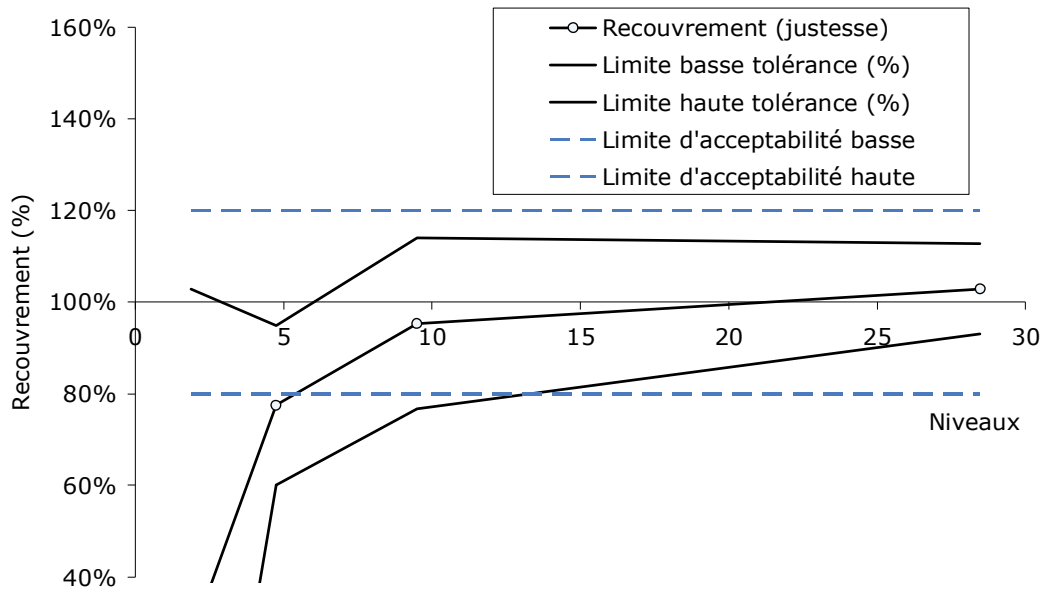


Figure 6 : profil d'exactitude pour le pyrène en appliquant un facteur de correction (1,5) aux concentrations de validation ($\lambda = 20\%$ et $\beta = 80\%$). Les échantillons d'étalonnage étaient des étalons de HAP dans l'acétonitrile et n'ont pas subi le protocole d'extraction.

2.2 Autre exemple de résultats : le benzo(k)fluoranthène

Pour le benzo(k)fluoranthène (**figure 7a**), la méthode présente un biais (effet matrice) qui est plus important que dans le cas du pyrène (car celui-ci est plus apolaire que le pyrène et présente de plus fortes interactions avec la matière organique dissoute de l'eau de pluie). Après application d'un facteur de correction (**figure 7b**), le biais est corrigé pour les teneurs entre 8 et 18 ng/L, mais il subsiste un biais qui s'accroît quand la concentration diminue. La méthode non corrigée est assez fidèle (**figure 7a**), mais après application du facteur de correction (figure 7b), les intervalles de tolérance s'élargissent considérablement. Ce qui est normal puisque le facteur de correction est supérieur à 1 et a pour effet d'augmenter les écarts-types.

Après application du facteur de correction, la limite haute de tolérance dépasse la limite haute d'acceptabilité à 17,5 ng/l, ce qui correspond à la limite basse de quantification. Sachant que les teneurs en benzo(k)fluoranthène dans les échantillons d'eau de pluie sont généralement inférieures à cette valeur, la méthode n'est donc pas validée pour réaliser les analyses de ce composé.

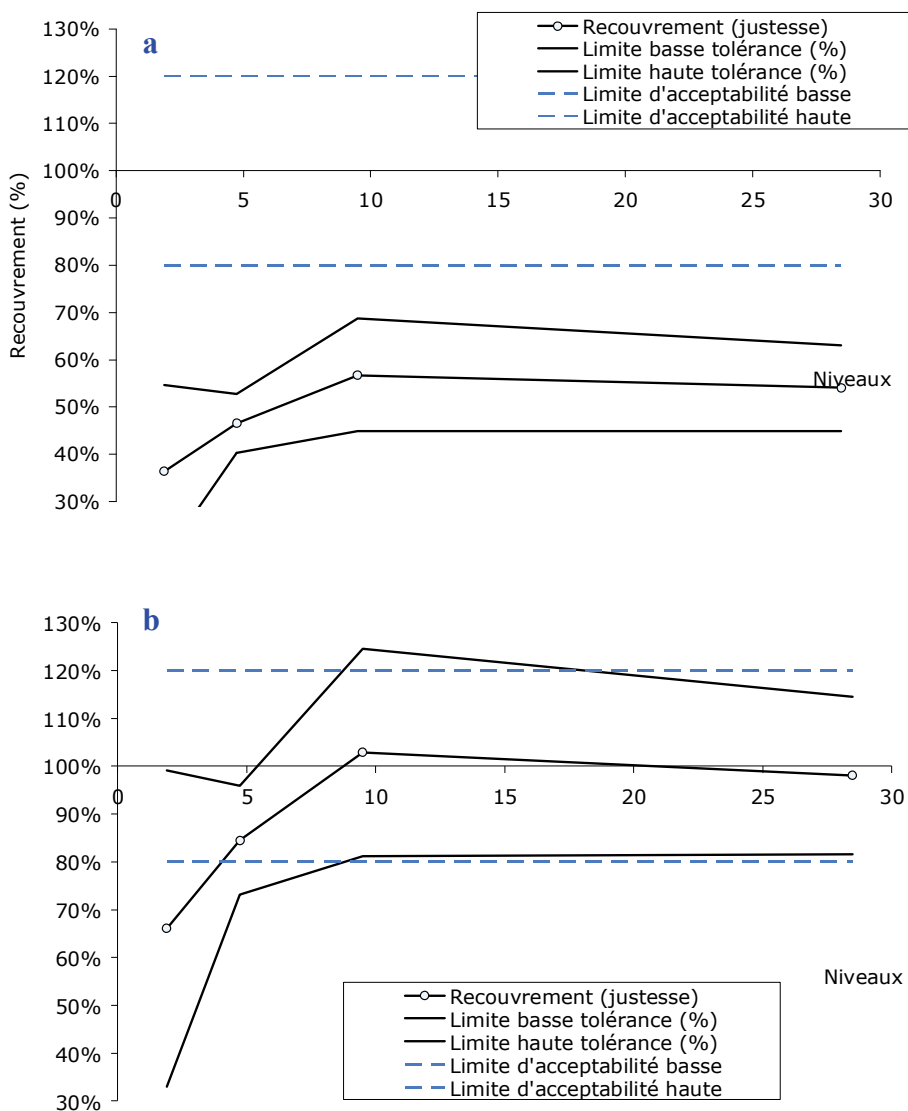


Figure 7 : profils d'exactitude pour le benzo(k)fluoranthène pour $\lambda=20\%$ et $\beta=80\%$, modèle d'étalonnage : droite
 (a) tracé à partir des valeurs des plans d'expériences d'étalonnage et de validation
 (b) après application d'un facteur de correction (1,81)

Pour que la méthode soit validée, le facteur β pourrait être diminué (70 % par exemple, **figure 8**) ; dans ce cas, le risque moyen d'avoir un résultat qui ne sera plus acceptable serait augmenté et il est peu raisonnable d'augmenter ce risque (on choisit en général un β d'au minimum 80 %).

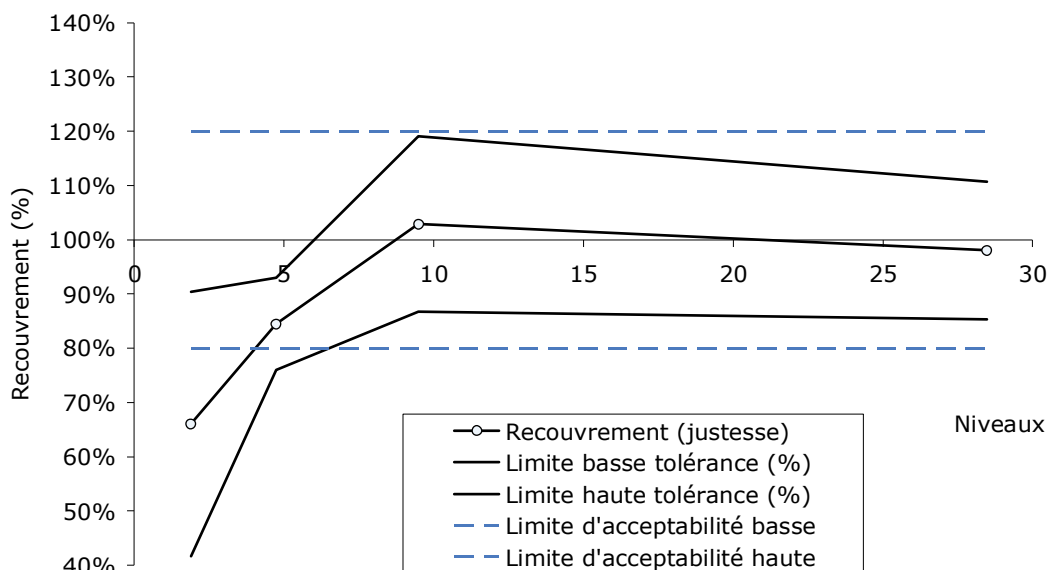


Figure 8 : profil d'exactitude pour le benzo(k)fluoranthène pour $\lambda = 20\%$ et $\beta = 70\%$, avec facteur de correction (1,81)

En revanche, **augmenter la limite d'acceptabilité λ à 25 %** (avec $\beta = 80\%$) (**figure 9**) permet de valider la méthode sur un domaine plus large (entre 6 et 28,5 ng/l), tout en conservant un risque acceptable. Augmenter λ de 20 à 25% signifie que pour une teneur de 10 ng/l, le résultat sera compris au maximum entre 7,5 et 12,5 ng/l (au lieu de 8 à 12 ng/l). Ceci reste acceptable pour l'utilisation des résultats qu'il est prévu de faire. À la limite de quantification basse (6 ng/l), le biais est de 10%. Il devient inférieur à 3% entre 8 et 28 ng/l.

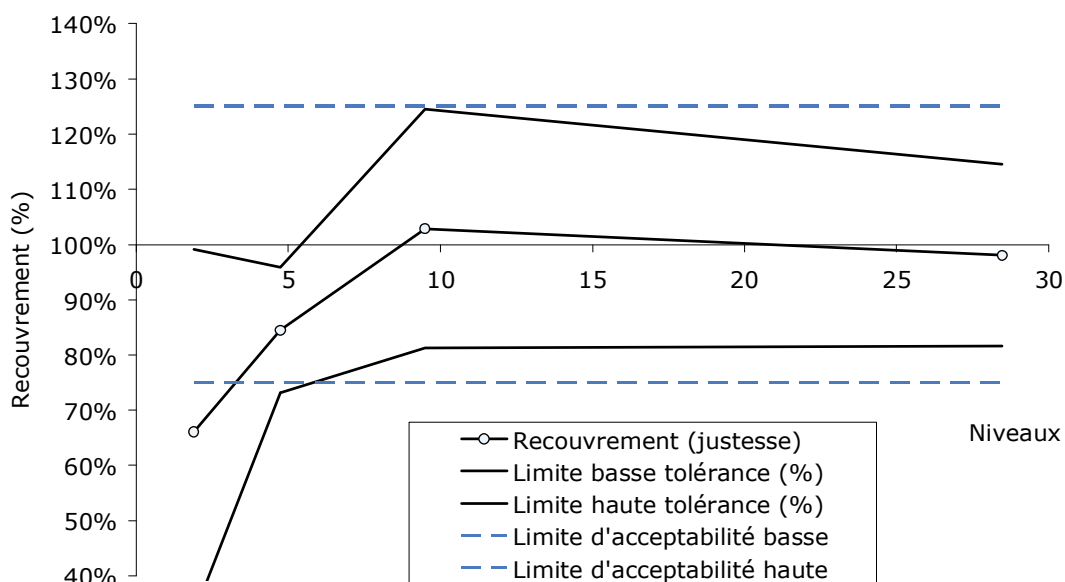


Figure 9 : profil d'exactitude pour le benzo(k)fluoranthène pour $\lambda = 25\%$ et $\beta = 80\%$, avec facteur de correction (1,81)

D'autres profils d'exactitude ont été tracés avec différents modèles d'étalonnages (pondérés ou non). Il s'avère que le modèle linéaire donne le profil (**figure 9**) le plus adapté aux objectifs de la méthode (intervalle de tolérance le plus étroit, meilleure justesse, limite de quantification la plus basse).

Comme dans le cas du pyrène, nous avons tracé le profil d'exactitude en utilisant des échantillons d'étalonnage dans l'acétonitrile (ne subissant pas le protocole d'extraction) (**figure 10**), et après application d'un facteur de correction (2,45). La figure 10 est à comparer avec la figure 7b, où les échantillons d'étalonnage ont subi l'intégralité du protocole extraction-analyse.

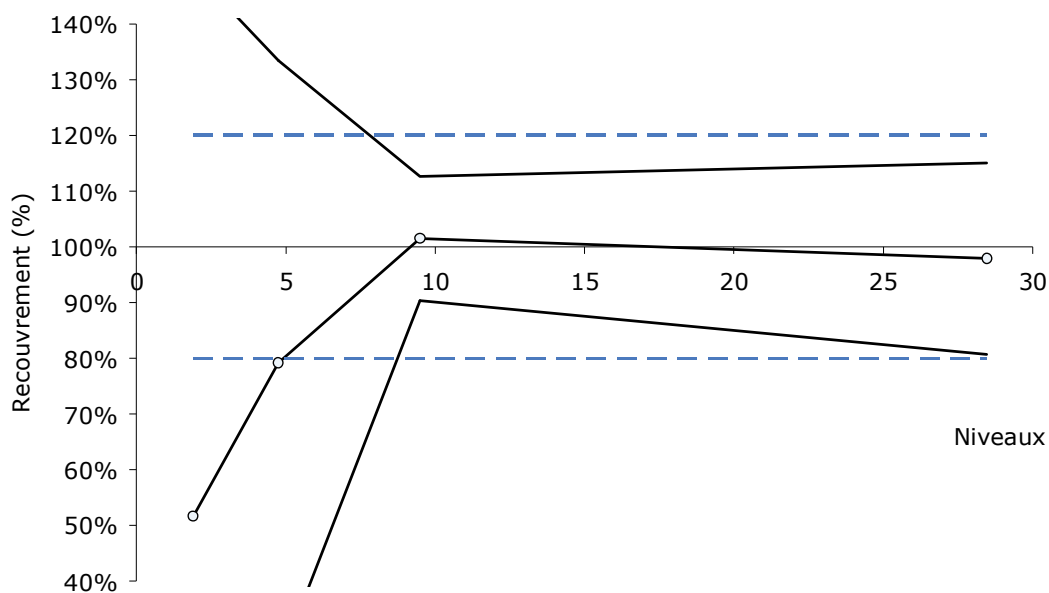


Figure 10 : profil d'exactitude pour le benzo(k)fluoranthène pour $\lambda = 20\%$ et $\beta = 80\%$, après application d'un facteur de correction (2,45)

La méthode est juste entre 8 et 28 ng/l (biais inférieur à 5 %), comme c'était le cas figure 7b. En revanche, figure 10, l'intervalle de tolérance se situe à l'intérieur des limites d'acceptabilité entre 8,8 et 28 ng/l. Le domaine de validité de la méthode ainsi défini, avec $\lambda = 20\%$ et $\beta = 80\%$, est plus étendu que celui du profil d'exactitude figure 7b. Pour valider la méthode en utilisant des étalons qui subissent le protocole d'extraction, nous avons été contraints d'augmenter λ à 25 % (**figure 9**), mais le domaine de validité était plus étendu (entre 6 et 28 ng/l). Dans le cas du benzo(k)fluoranthène, le choix de la méthode d'étalonnage est à orienter en fonction des besoins scientifiques : domaine de validité étendu et $\lambda = 25\%$ (ce que nous avons choisi) ou domaine de validité plus restreint (limite basse de quantification plus élevée) et $\lambda = 20\%$.

Conclusion

Nous avons réalisé un total de 47 essais pour essayer de valider la méthode d'analyse de 15 HAP contenus dans de l'eau de pluie selon la méthode du profil d'exactitude. Les expériences ont été réalisées avec le mélange des 15 HAP (méthode multi-analytes), mais les résultats ont été traités individuellement pour chaque HAP.

L'étendue du domaine de validation varie selon les HAP. Pour la plupart des HAP, il est nécessaire d'appliquer un facteur de correction afin de corriger le biais (effet matrice) et de valider la méthode. Pour certains HAP, nous avons choisi d'augmenter la limite d'acceptabilité λ à 25 % (au lieu de 20 %) afin de valider la méthode sur un domaine de concentration plus large (mais toujours compris entre les niveaux bas et haut des échantillons de validation, l'extrapolation hors du domaine testé étant proscrite).

Outre l'intérêt pour la validation de la méthode, le profil d'exactitude est également un outil diagnostique pour déterminer les points faibles de la méthode et également pour choisir judicieusement les contrôles qualité qui seront employés en routine. Le tracé des profils d'exactitude pour des modèles d'étalonnages variés, pondérés ou non, permet de déterminer le modèle d'étalonnage qui donnera le profil le plus en adéquation avec les objectifs scientifiques. De même, le tracé des profils d'exactitude en prenant des échantillons d'étalonnage qui subissent ou non le protocole d'extraction permet de déterminer la procédure d'étalonnage qui valide la méthode.

Bibliographie

- Azimi S., Rocher V., Muller M., Moilleron R., Thevenot D. (2005) Source, distribution and variability of hydrocarbons and metals in atmospheric deposition in an urban area (Paris, France). *Sci. Total Environ.* 337:223-239
- Bourdat-Deschamps M., Daudin J.-J., Barriuso E. (2007) An experimental design approach to optimize the determination of polycyclic aromatic hydrocarbons from rainfall water using stir bar sorptive extraction and high performance liquid chromatography – fluorescence. *J. Chromatogr. A.* 1167 :143-153
- Feinberg M. (2010^b) Mise en œuvre du profil d'exactitude *in* Validation des méthodes d'analyse quantitatives physico-chimiques par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 27-44

Validation de la méthode de dosage du pigment chlorophylle *a* dans l'eau douce par extraction et par HPLC-DAD

Jérôme Lazzarotto¹, Jennifer Maherou

Résumé : *La chlorophylle a est un pigment vert présent dans tous les végétaux, dont les algues. L'unité mixte de recherche CARRTEL de l'Inra de Thonon-les-Bains, mesure la concentration de ce pigment dans des lacs alpins afin d'étudier ces écosystèmes. Ce pigment est dosé dans l'eau par HPLC-DAD après une étape de filtration / extraction. Cette méthode qui fait l'objet d'un protocole défini, est utilisée depuis plusieurs années au laboratoire de l'unité. Afin de mieux répondre aux besoins des chercheurs et de se conformer à la démarche qualité de l'Inra, cette méthode a été validée par un profil d'exactitude. Cette validation a été effectuée en deux étapes afin d'évaluer la méthode dans son ensemble : une étape avec un matériau de référence puis à partir de cultures d'algues. Cet article décrit le plan d'expérience et la validation de cette méthode avec un profil d'exactitude dans ce contexte.*

Mots clés : limnologie, phytoplancton, chlorophylle *a*, validation, profil d'exactitude

Introduction

L'unité mixte de recherche CARRTEL de l'Inra de Thonon-les-Bains étudie les écosystèmes lacustres ; son laboratoire de chimie réalise les analyses des eaux et des sédiments lors d'études limnologiques. Depuis plusieurs années, le laboratoire de chimie est engagé dans une démarche qualité selon le Référentiel Inra. Un des objectifs du plateau technique est que toute méthode utilisée en routine soit validée.

Afin d'étudier les écosystèmes lacustres, la biomasse algale des lacs doit être évaluée. Cette biomasse peut être évaluée soit par une détermination taxonomique soit par une estimation en quantifiant les pigments présents.

Nous traiterons ici de la validation de la méthode d'analyse de la chlorophylle *a* par chromatographie liquide avec détections sur barrette de diodes (HPLC-DAD) après extraction liquide-solide avec la méthode du profil d'exactitude (Max Feinberg, 2010b).

1. Contexte d'application de la méthode

Depuis le milieu du XX^{ème} siècle, la qualité de l'eau des grands lacs s'est fortement dégradée avec le phénomène d'eutrophisation. Un apport excessif en nutriments par les effluents urbains et par les terres agricoles a provoqué une abondante prolifération phytoplanctonique. La décomposition de cette importante biomasse a entraîné une désoxygénation des eaux. Le phosphore est le facteur limitant de ce phénomène d'eutrophisation (Barroin, 2004). Il est aussi l'un des principaux indicateurs de la qualité de l'eau avec la biomasse algale et la transparence (DCE, 2000 ; Seq-EAU, 1999).

La chlorophylle *a* est le pigment le plus répandu dans les différentes classes d'algues. Sa quantification permet une estimation de la biomasse algale en présence.

¹ UMR0042 CARRTEL – Centre alpin de recherche sur les réseaux trophiques des écosystèmes mimniques – INRA - F-74200 Thonon-les-bains ☎ 04 50 26 78 21 ✉ lazzarotto@thonon.inra.fr

L'analyse de la chlorophylle *a* est réalisée en routine par le laboratoire de chimie de l'unité CARRETEL par extraction liquide-solide puis HPLC-DAD.

Ces analyses sont effectuées lors de programmes de recherche limnologique afin d'évaluer la qualité d'un milieu lacustre (Lazzarotto et Rapin, 2008). Le stade d'eutrophisation d'un milieu est estimé d'après plusieurs paramètres, dont la chlorophylle *a* : oligotrophe : Chl *a* < 35 µg/l ; mésotrophe : 35µg/l < Chl *a* < 100 µg/l ; et eutrophe : Chl *a* > 100 µg/l (Vollenweider and Kerekes, 1982). D'autre part ces études sont effectuées avec plusieurs analyses, au minimum 4 prélèvements différents dans l'année.

Les exigences des demandeurs d'analyses seront fixées selon ces fréquences d'observation et les distinctions entre les stades d'eutrophisation.

2. Matériel et méthode

L'eau est filtrée pour mesurer sa teneur en chlorophylle *a*. L'analyse de la chlorophylle *a* est réalisée par extraction liquide-solide à partir du filtre puis HPLC-DAD.

Cette méthode est établie et fixée par un mode opératoire, inspiré de la norme NF T90 116. Les pigments sont photosensibles et thermosensibles, les échantillons sont donc conservés au froid et à l'obscurité jusqu'à l'analyse.

2.1 Filtration – Extraction

L'eau brute est filtrée sur un filtre GFC. Le volume filtré est choisi entre 50 ml et 5 l selon la transparence de l'échantillon. Les algues contenant les pigments sont retenues sur ce filtre. Les filtres sont déposés dans un tube à centrifuger avec le solvant d'extraction (méthanol/acétate d'ammonium 0.5 M (98/2 ; v/v)). Ils sont soumis à des ultra-sons afin de briser les cellules algales et de libérer les pigments. Les tubes sont ensuite centrifugés.

Le surnageant est récupéré et filtré à travers un filtre seringue pour le séparer des débris

2.2 HPLC-DAD

L'extrait est ensuite analysé par HPLC-DAD, sur une chaîne Waters avec un passeur d'échantillons. 100 µl sont injectés et élués selon un gradient d'éluants à base de 3 éluants (méthanol / 0,5 M acétate d'ammonium (80/20) ; acétonitrile/eau (90/10) ; éthyle acétate) de 30 mn.

La chlorophylle *a* est identifiée par son temps de rétention ainsi que par son spectre d'absorption spécifique (Jeffrey *et al.*, 1997).

2.3. Expression du résultat

La quantification est permise grâce à un étalonnage externe à partir d'étalons certifiés de chlorophylle *a* dissous dans du méthanol. La quantité mesurée dans le volume injecté est rapportée au volume d'eau filtré. La concentration ainsi mesurée est exprimée en µg Chl *a*/l.

3. Démarche de validation avec le profil d'exactitude

Afin de valider cette méthode, nous avons utilisé la méthode du profil d'exactitude. Cet outil permet de vérifier si une méthode répond aux exigences fixées. En effectuant plusieurs

analyses sur plusieurs jours, il détermine les paramètres pris en compte pour la validation de méthode tels que la justesse, la fidélité intermédiaire et les incertitudes.

Pour cette démarche, nous avons choisi de différencier les deux étapes de la méthode : l'extraction puis la quantification. Dans un premier temps, nous établissons le profil d'exactitude pour la quantification seule *via* HPLC-DAD, avec des matériaux de référence certifiés (MRC) et différents de ceux utilisés pour effectuer l'étalonnage. Nous validons ainsi la justesse de la méthode et nous estimerons la variabilité de cette étape.

Dans un second temps, nous établissons le profil d'exactitude pour la méthode entière, soit l'extraction puis la quantification. Pour ce faire, nous utilisons plusieurs cultures d'algues.

L'étape d'extraction peut apporter un biais qu'il faut considérer. Les pigments présents dans l'échantillon initial peuvent ne pas être extraits en totalité. Nous ne pouvons pas déterminer le recouvrement réel de la méthode. Nous avons estimé en étude préliminaire le rendement maximum de la méthode. Nous avons effectué des mesures en modifiant la méthode et en utilisant une méthode d'extraction considérée comme optimale, avec une étape supplémentaire de broyage du filtre. Le recouvrement observé par rapport à la méthode considérée comme optimale est de 78 % pour des quantités entre 0.1 µg Chl *a* et 6 µg Chl *a* extraites. Ce recouvrement n'est pas un recouvrement absolu mais relatif et correspond au recouvrement maximum possible de la méthode.

D'après ces connaissances de la méthode que nous utilisons en routine, nous suivons la démarche de validation décrite par Max Feinberg (2010b)

3.1 La quantité mesurée

La quantité mesurée est la concentration en chlorophylle *a* dans le phytoplancton de l'eau. La méthode est une méthode indirecte qui nécessite un étalonnage à partir d'extraits méthanoliques étalons.

Les résultats sont exprimés en µg Chl *a* extrait, cette quantité est ensuite reportée au volume filtré. L'expression finale du résultat est en µg Chl *a*/l.

3.2 Les objectifs de la validation

3.2.1 Le domaine de validation

Les résultats habituels observés varient entre 0,2 et 100 µg Chl *a*/L, avec des volumes filtrés entre 50 ml et 5 l. La quantité extraite varie donc entre 1 et 5 µg Chl *a*.

Le domaine de validation est exprimé en quantité extraite. Il est choisi plus large qu'entre 1 et 5 µg Chl *a* extraite afin de pouvoir mesurer le pigment si le volume filtré n'est pas optimal et de ne pas avoir à refaire une filtration. Le domaine de validation choisi est de 0,4 à 8,0 µg Chl *a* extrait.

3.2.2 Les limites d'acceptabilité et intervalle de tolérance

Les exigences des demandeurs d'analyses sont fixées selon les fréquences d'observation et les distinctions entre stades d'eutrophisation explicitées en 1. Il est souhaité que au moins 3 analyses sur 4 donnent un résultat à +/- 25 % de la valeur vraie de l'échantillon, soit que 75 % des résultats obtenus aient un recouvrement de 25 %.

Ainsi les limites d'acceptabilité (□) sont de 25 % et l'intervalle de tolérance (□) est de 75 %.

3.3 Les échantillons de validation

Il existe des étalons certifiés d'extrait de chlorophylle *a* méthanolique. Mais il n'existe pas de solution d'algue avec une concentration connue en chlorophylle *a*, du fait des problèmes liés à l'évolution des algues. Il est impossible aussi de doper une matrice puisque le pigment mesuré est à l'intérieur des cellules de plancton.

Pour cette validation, nous avons choisi de différencier les deux étapes de la méthode : l'extraction puis la quantification.

Dans un premier temps, nous établirons un profil d'exactitude pour la quantification seule *via* HPLC-DAD, avec des étalons d'extrait de chlorophylle *a* méthanolique certifiés. Nous validons ainsi la justesse de la méthode et la variabilité liée à cette étape.

Dans un second temps, nous établirons deux profils d'exactitude pour la méthode entière, soit l'extraction puis la quantification. Pour ce faire, nous utilisons deux cultures d'algues, afin de contrôler les espèces d'algues en présence et l'évolution de la biomasse. Nous ne connaissons pas la valeur vraie de la concentration en chlorophylle *a* de ces cultures. Ces profils d'exactitude nous renseigneront sur la variabilité globale de la méthode.

Ces profils sont effectués pour deux algues cultivées spécifiques : diatomées et cyanobactéries. Ces cultures seront filtrées dans les mêmes conditions que les échantillons.

3.4 Les essais de validation

3.4.1 Etape HPLC-DAD

Nous avons choisi 4 niveaux de quantité de chlorophylle *a* : 0,4 ; 0,8 ; 2,0 ; 8,0 $\mu\text{g Chl } a$. Afin d'évaluer la répétabilité et la fidélité intermédiaire de cette étape, nous avons réalisé pour chacune de ces quantités 3 réplicats et cela pendant 3 jours. Soit 9 mesures par niveau, soit 36 mesures au total.

3.4.2 Analyse globale extraction – HPLC-DAD

Nous avons filtré différentes quantités des cultures d'algues afin d'approcher les quantités évaluées lors de la première étape : 1 ml ; 3 ml ; 10 ml ; 90 ml. Afin d'évaluer la répétabilité de cette étape, nous avons réalisé pour chacune de ces filtrations 3 réplicats. Etant donné que la culture d'algues évolue dans le temps, la fidélité intermédiaire ne peut être déterminée en effectuant plusieurs mesures sur différents jours, nous avons donc effectué 3 séries de filtration dans la même journée avec deux opérateurs. Soit 9 mesures par niveau, 36 mesures par culture d'algues et 72 mesures au total.

3.5 L'étalonnage

L'étalonnage est effectué avec des solutions étalons méthanoliques injectées à 3 niveaux : 0,28 $\mu\text{g Chl } a$; 4,60 $\mu\text{g Chl } a$; 9,20 $\mu\text{g Chl } a$. L'étalonnage est effectué avec une analyse par niveau.

L'étalonnage devant être effectué à chaque série du plan de validation, il sera effectué lors des 3 séries de mesures pour les étalons certifiés ainsi que pour les algues.

4. Résultats

La fonction d'étalonnage est linéaire. Les quantités extraites lors de chaque série sont calculées à partir des fonctions linéaires déterminées à chaque étalonnage.

Tableau 1 : Résultats des étalonnages obtenus lors de chaque série du plan ($Aire = a x [Chl a] + b$)

	ESSAIS Etalons certifiés			ESSAIS Culture cyanobactéries			ESSAIS Culture diatomées		
	I	II	III	I	II	III	I	II	III
a	5 851 243	5 920 726	5 955 871	5 603 378	5 738 962	5 465 850	5 451 452	5 515 275	5 645 598
b	-3 636	-6 268	1 621	-922	-5 150	-1 359	3 490	3 650	5 696
r ²	1.0000	0.9999	0.9999	1.0000	0.9999	1.0000	1.0000	0.9999	0.9997
n	4	4	4	4	4	4	4	4	4

Les résultats des essais de validation sont insérés dans le fichier Excel (https://intranet.inra.fr/mission_qualite/acces_thematiques/formations_et_ecoles/ecole_des_tech_niques_validation_methodes/profil_d_exactitude). Le fichier génère automatiquement les tableaux de calcul

4.1 Profil d'exactitude de l'étape analytique HPLC-DAD seule

Le profil d'exactitude de cette étape analytique (figure 1) nous indique que la méthode est juste et répétable dans le domaine de 0,4 à 8,0 µg Chl a. Les incertitudes relatives sont de l'ordre de 10 % et les limites de tolérance sont comprises dans les limites d'acceptabilité jusqu'à la valeur basse de 0,4 µg Chl a.

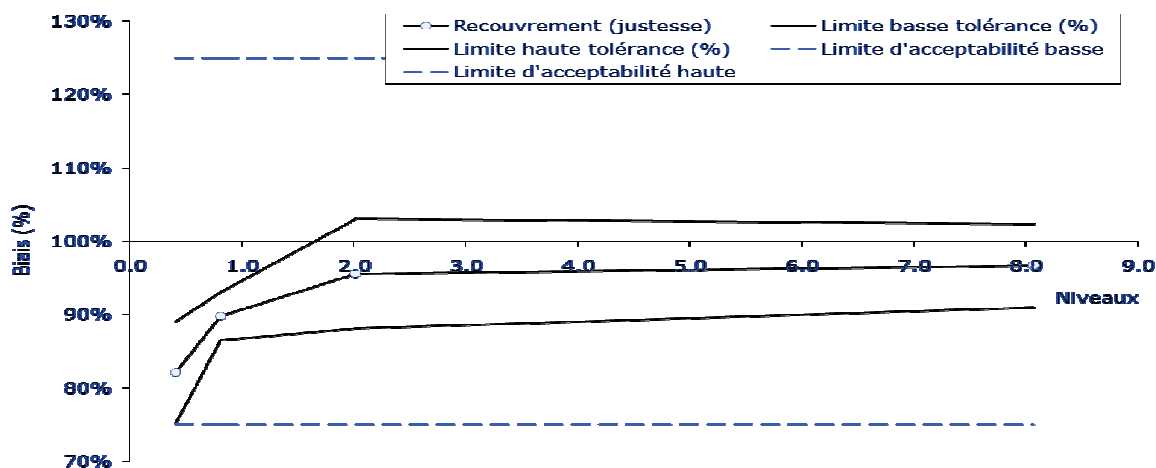


Figure 1 : profil d'exactitude de la méthode d'analyse HPLC-DAD de la chlorophylle a avec des matériaux de référence

4.2 Profils d'exactitude de la méthode globale extraction - analyse HPLC-DAD à partir de cultures d'algues

Les profils d'exactitude à partir des cultures d'algues ne renseignent pas sur la justesse de la méthode. La valeur cible correspond à la moyenne des mesures. Ces profils indiquent la variabilité de la méthode dans son ensemble : l'extraction et l'analyse HPLC-DAD.

Les figures 2 et 3 représentent ces profils.

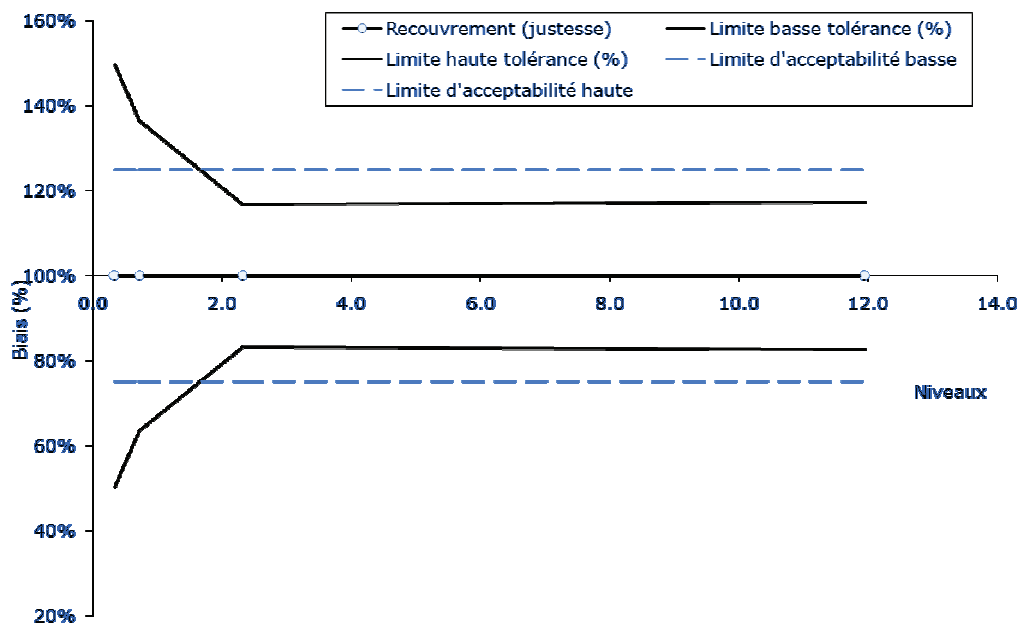


Figure 2 : profil d'exactitude de la méthode d'analyse HPLC-DAD de la chlorophylle a avec une culture de cyanobactéries

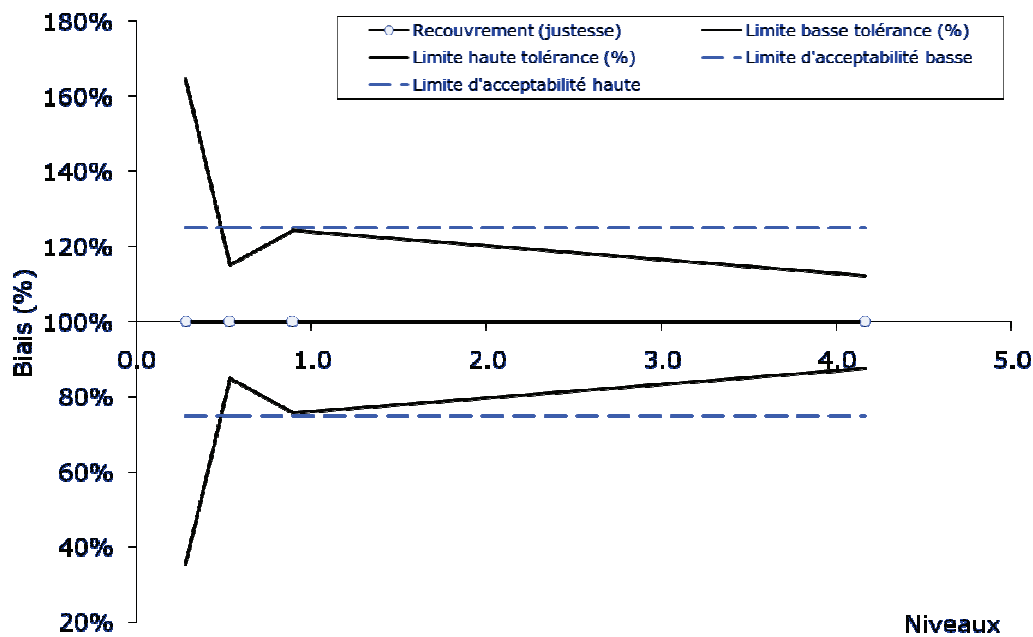


Figure 3 : profil d'exactitude de la méthode d'analyse HPLC-DAD de la chlorophylle a avec une culture de diatomées

Ces profils sont différents selon la culture d'algues. L'homogénéité des pigments dans les algues et les solutions ne sont donc pas les mêmes selon les cultures d'algues. La variabilité est acceptable de 1,8 µg Chl *a* à 12 µg Chl *a* pour la culture de diatomées et de 0,5 µg Chl *a* à 4,1 µg Chl *a* pour la culture de cyanobactéries.

Conclusion

L'étape d'analyse HPLC-DAD est validée selon les exigences spécifiées : la justesse de la méthode est validée dans le domaine de 0,4 à 8,0 µgChl *a*.

Les résultats des essais de la méthode globale extraction – HPLC-DAD indique une variabilité différente selon les algues en présence.

En prenant en considération ces résultats, la méthode est validée selon les exigences spécifiées dans le domaine de 4,1 µg Chl *a* à 8 µg Chl *a*, soit avec des volumes filtrés de 50 ml à 5 l de 0,8 à 160 µg Chl *a*/l. Le domaine validé correspond aux besoins usuels. 75 % des résultats indiqueront des valeurs autour de 25 % de la valeur vraie.

Étant donné l'importance de la variabilité due au prélèvement et aux classes d'algues en présence, il est conseillé d'apporter un soin particulier à l'échantillonnage sur le terrain, en effectuant par exemple plusieurs échantillons à pooler.

Une carte de contrôle à partir de solutions de chlorophylle *a* certifiées, est effectuée afin de garantir la stabilité de la méthode validée.

L'outil profil d'exactitude a facilité cette validation et a permis de confronter les exigences des demandeurs d'analyses avec les possibilités du laboratoire.

Bibliographie

Barroin G. (2004) Phosphore, azote, carbone... Du facteur limitant au facteur de maîtrise. *Courrier Env. INRA*, 52, 23-30.

DCE, Directive Cadre Eau (2000) Directive 2000/60/CE JO CE L 327/1 du 22.12.2000, p. 72

Feinberg Max (2010^b) Mise en œuvre du profil d'exactitude *in* Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 27-44

Lazzarotto J., Rapin F. (2008) Évolution physico-chimique des eaux du Léman. Rapports sur les études et recherches entreprises dans le bassin lémanique. Programme quinquennal 2006-2010. Campagne 2007, p. 31-55

SEQ-Eau (1999) Système d'évaluation de la qualité des cours d'eau, rapport de présentation SEQ-Eau. Agences de l'eau, les études des agences de l'eau n° 64 - 59p.

Vollenweider, R.A., and Kerekes, J. (1982) Eutrophication of waters. Monitoring, assessment and control. OECD Cooperative programme on monitoring of inland waters (Eutrophication control), Environment Directorate, OECD, Paris. P. 154

Validation d'une méthode automatisée de titration de l'alcalimétrie complète dans l'eau douce

Jérôme Lazzarotto¹

Résumé : *Le titre alcalimétrique complet (TAC) d'une eau est la mesure du pouvoir alcalin de la solution. Le TAC est principalement expliqué par le pH et la concentration en ions carbonate. Cette titration est effectuée de manière automatisée à l'unité mixte de recherche CARTEL de l'Inra de Thonon-les-Bains*

Cette méthode a été validée avec un profil d'exactitude afin de comparer les besoins des études utilisant cette mesure avec ses caractéristiques.

Cet article décrit le plan d'expérience choisi pour cette validation ainsi que les résultats.

Mots clés : limnologie, hydrologie, titre alcalimétrique complet, carbonate, validation, profil d'exactitude

Introduction

L'unité mixte de recherche CARTEL de l'Inra de Thonon-les-Bains étudie les écosystèmes lacustres et son laboratoire de chimie réalise les analyses des eaux et sédiments lors d'études limnologiques, notamment pour son observatoire de recherche environnemental : ORE Lacs. Depuis plusieurs années, le laboratoire de chimie est engagé dans une démarche qualité selon le Référentiel Inra, et l'un de ses objectifs est que toute méthode utilisée en routine soit validée. La mesure du TAC est effectuée en routine dans l'étude des lacs et des rivières des bassins versants de l'ORE Lacs.

Nous traiterons ici de la validation de la méthode automatisée de titration de l'alcalimétrie complète avec la méthode du profil d'exactitude (Feinberg M., 2010^b- 2010^c).

1. Contexte d'application de la méthode

L'ORE Lacs représente un suivi établi depuis plusieurs dizaines d'années des trois grands lacs alpins (Léman, Annecy et Bourget). Durant ce suivi, les paramètres physico-chimiques de base sont analysés dans les eaux. Le TAC fait partie de ces paramètres.

Dans l'étude d'un écosystème lacustre, la connaissance du TAC est nécessaire pour le calcul de la production primaire. Il estime le carbone inorganique présent pouvant être consommé avec les nutriments lors de la photosynthèse (Lazzarotto J. et Rapin F., 2008). D'autre part, dans un cours d'eau et un lac, le TAC est un des paramètres indiquant la minéralisation d'une eau.

Le TAC n'est pas un polluant, mais le système d'évaluation de la qualité des cours d'eau (SEQ-Eau) indique des seuils de variations pour l'utilisation en production d'eau potable et en tant qu'indicateur de la qualité de l'eau par altération (Seq-EAU, 2003). Ces seuils et les indices d'aptitude correspondants sont décrits dans le **tableau 1**.

¹ UMR0042 CARTEL – Centre alpin de recherche sur les réseaux trophiques des écosystèmes minniques – INRA - F-74200 Thonon-les-bains ☎ 04 50 26 78 21 ✉ lazzarotto@thonon.inra.fr

Tableau 1 : valeurs seuils du TAC selon le SEQ-Eau

TAC (°f)	8 à 40	5 à 58	3 à 75	0 à 100
TAC (még/l)	1,6 à 8,0	1,0 à 11,6	0,6 à 15,0	0,0 à 20,0
État du milieu	Très bon (bleu)	Bon (vert)	Mauvais (orange)	Très mauvais (rouge)

2. Matériel et méthode

La mesure de l'alcalinité complète est une titration acide-base. Au laboratoire de l'unité CARTELE, la méthode utilisée est une titration automatique avec mesure du point final déterminé à l'aide de la dérivée seconde du pH en fonction du volume d'acide. Cette méthode est inspirée de la norme NF T90-036 qui utilise une titration manuelle avec indicateur coloré. Le pH est mesuré avec une électrode de verre connectée au titrateur. Le titrateur utilisé est un Metrohm Basic Titrino 794 avec un passeur d'échantillons 730.

La prise d'essai est de 100 ml. La burette a un volume de 20 ml de HCl à 0.1 mol/l. Le TAC est calculé par équivalence de la quantité d'acide ayant réagi avec l'ensemble des espèces alcalines de la solution.

3. Démarche de validation avec le profil d'exactitude

Afin de valider cette méthode, nous avons utilisé la méthode du profil d'exactitude. Cet outil permet de vérifier si une méthode répond aux exigences fixées. En effectuant plusieurs analyses sur plusieurs jours, il détermine les paramètres pris en compte pour la validation de méthode tels que la justesse, la fidélité intermédiaire et les incertitudes.

La mesure du TAC utilise la mesure du pH. La méthode de pH a été validée, l'étalonnage est effectué lors de chaque essai avec des solutions de référence certifiées. La stabilité du système est contrôlée avec une carte de contrôle.

Pour cette validation, nous suivons la démarche de validation décrite par Max Feinberg (2010^b).

3.1 La quantité mesurée

La quantité mesurée est la quantité d'espèces alcalines présentes dans l'eau à analyser. Cette quantité est ramenée au volume de la prise d'essai. Le résultat est exprimé en még/l.

3.2 Les objectifs de la validation

Le domaine de validation

Les eaux douces des écosystèmes que nous étudions ont un TAC qui varie entre 1 et 5 még/l. Le volume de la burette permet d'atteindre des valeurs de 20 még/l.

Afin que le domaine d'utilisation soit compris dans le domaine de validation, nous choisissons un domaine de 0,5 à 10,0 még/l.

Les limites d'acceptabilité et intervalle de tolérance

Les exigences pour ces analyses sont définies en fonction de leur utilisation dans l'ORE Lacs et d'après les seuils indiqués par le SEQ-Eau. Les écarts entre les valeurs minimales (environ 1 méq/l) sont de l'ordre de 0,5 méq/l et de l'ordre de 4 méq/l pour les valeurs maximales (environ 15 méq/l), soit des variations de 36 à 50 %. La limite d'acceptabilité doit donc être inférieure à 36 % afin de pouvoir discriminer deux valeurs de seuils maximales. La limite d'acceptabilité est donc choisie à 25 %.

Par rapport à la fréquence des échantillonnages et des analyses de l'ORE, il est souhaité que au moins 4 résultats sur 5 soient dans les limites d'acceptabilité, soit 80 % des résultats avec un recouvrement de 25 %.

Ainsi les limites d'acceptabilité (\square) sont de 25 % et l'intervalle de tolérance (\square) est de 80 %.

3.3 Les échantillons de validation

Étant donné que le volume de la prise d'essai est important (100 ml) et qu'un plan d'expérience minimal nécessitera au minimum 27 mesures (avec 3 réplicats à 3 niveaux pendant 3 jours), le volume minimal d'étalon utilisé serait de 2,7 l. Le coût de cette validation serait donc trop élevé si nous utilisons des solutions de référence certifiées, par rapport à l'utilisation de la mesure.

Nous choisissons donc de réaliser des solutions étalons à base de carbonate de sodium et d'hydrogencarbonate de sodium.

Les échantillons de validation seront donc préparés à partir de poudres séchées puis pesées sur des balances vérifiées, ensuite dissoutes dans des fioles de classe A vérifiées.

3.4 Les essais de validation

Étant donné que la méthode est automatisée et que le domaine est grand, nous choisissons 5 niveaux de TAC dans le domaine de validation. Les 5 niveaux choisis sont : 0,5 ; 1,0 ; 1,5 ; 5,0 et 10,0 méq/l. Afin d'évaluer la répétabilité et la fidélité intermédiaire de cette étape, nous avons réalisé pour chacun de ces niveaux 3 réplicats et cela pendant 5 jours, soit 15 mesures par niveaux, soit 75 mesures au total.

3.5 L'étalonnage

La détermination du TAC est une méthode directe qui ne nécessite pas d'étalonnage.

4. Résultats

Les résultats des essais de validation sont insérés dans le fichier Excel (https://intranet.inra.fr/mission_qualite/acces_thematiques/formations_et_ecoles/ecole_des_tech_niques_validation_methodes/profil_d_exactitude). Le fichier génère automatiquement les tableaux de calcul et le profil d'exactitude (**figure 1**).

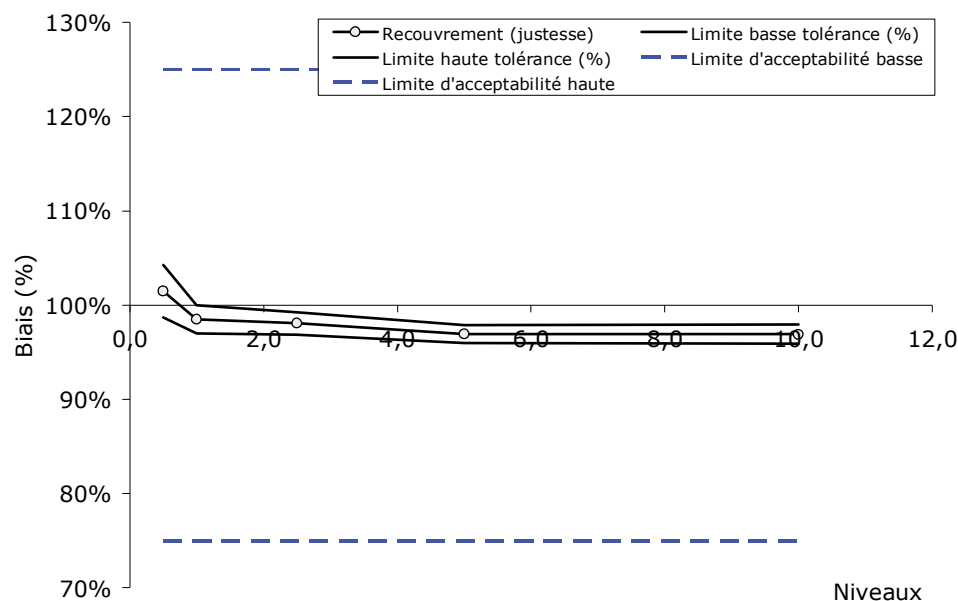


Figure 1 : profil d'exactitude de la méthode de titration de l'alcalimétrie complète d'une eau

Le profil d'exactitude de cette méthode indique que la méthode est validée par rapport aux exigences pour l'ensemble du domaine de validation, de 0,5 à 10,0 méq/l. Le fichier du profil d'exactitude calcule l'incertitude de la méthode pour chaque niveau en fonction de la fidélité mesurée. L'incertitude relative de cette méthode varie entre 4 et 2 %.

Conclusion

La méthode de détermination du TAC a été validée dans le domaine de 0,5 à 10,0 méq/l avec une limite d'acceptabilité de 25 % et un intervalle de tolérance de 80 %.

Bibliographie

- AFNOR (1996) Norme NF T90-036 / NF EN ISO 9963-1 : Détermination de l'alcalinité. Qualité de l'eau, T2 :115-125
- Feinberg Max (2010^b) Mise en œuvre du profil d'exactitude *in* Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 27-44
- Feinberg M. (2010^c) Interprétation du profil d'exactitude *in* Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial : 45-59
- Lazzarotto J., Rapin F. (2008) Évolution physico-chimique des eaux du Léman. Rapports sur les études et recherches entreprises dans le bassin lémanique. Programme quinquennal 2006-2010. Campagne 2007, 31-55
- SEQ-Eau (2003) Système d'évaluation de la qualité des cours d'eau, rapport de présentation SEQ-Eau. Agences de l'eau, les études des agences de l'eau n° 64, p. 59

Méthode d'approche globale de l'estimation des incertitudes de mesure

Annie Guerin et Henri Ciesielski¹

Résumé : Afin de répondre aux besoins des chercheurs ainsi qu'aux exigences de la norme ISO 17025, à l'Inra, le LAS, unité de service accréditée par le COFRAC (comité français d'accréditation), a développé une méthode originale d'estimation des incertitudes de mesures. Cette méthode, basée sur le traitement des données issues de cartes de contrôle et de résultats d'essais sur des sols d'un circuit d'essais d'aptitude, permet d'établir une relation simple entre un résultat (concentration en un élément) et l'incertitude associée, pour l'ensemble des analyses proposées.

Mots clefs : incertitude de mesure, analyse de sols

Introduction

Le laboratoire d'analyses des sols d'Arras (LAS) est une unité de services de l'Inra, accréditée par le COFRAC. Au LAS, les critères de validation d'une méthode d'analyse sont choisis en fonction des objectifs de la méthode, des besoins exprimés par nos partenaires et des exigences réglementaires auxquelles nous devons nous conformer. Parmi ces critères, l'incertitude des mesures occupe une place particulière non seulement pour son intérêt scientifique mais aussi pour la difficulté à l'estimer de façon rapide, économique et compatible avec le nombre d'échantillons et de données à traiter.

L'incertitude des mesures est le reflet de leur variabilité. Les sources de variabilité sont diverses (Feinberg et Laurentie, 2006) et notamment : la préparation des échantillons (conditions ambiantes, réactifs), la méthode de mesure (étalonnage, instrument) et l'expérience de l'opérateur ainsi que la nature même de l'échantillon et de l'analyte considéré. L'incertitude totale (I_t) est donc propre à un laboratoire, à une matrice et à un niveau de concentration. Son calcul repose sur la répétition des analyses au sein d'une même série mais aussi sur la répétition des séries.

Le LAS traite plus de 25 000 échantillons de sols par an, correspondant à environ 200 000 analyses par an. Dans ces conditions, on comprendra aisément qu'il n'est pas envisageable d'estimer expérimentalement les incertitudes de mesure par des répétitions systématiques nécessitant du temps, une quantité d'échantillons importante et pour un coût non négligeable. L'idéal serait de disposer d'un modèle mathématique simple reliant l'incertitude à la mesure. Mais ce modèle n'existe pas dans la littérature. C'est pourquoi nous avons développé une stratégie d'évaluation des incertitudes de mesure avec pour objectif de définir une relation entre la concentration en analyte et l'incertitude associée. Afin d'illustrer cette démarche, nous exposons les résultats obtenus pour le dosage du Cadmium total dans les sols, dosé après minéralisation totale par de l'acide fluorhydrique (suivant la norme NF X 31-147).

¹ US0010 LAS - Laboratoire d'analyses des sols, INRA -F- 62000 ARRAS – www.lille.inra.fr/las

☎ 03 21 21 86 38

✉ Annie.Guerin@arras.inra.fr

1. Principe et méthode

1.1 Démarche générale

Au LAS, le suivi de la qualité des résultats est réalisé d'une part à l'aide de cartes de contrôle élaborées pour chaque analyse avec un sol témoin (échantillon de contrôle interne) et d'autre part en participant régulièrement à des circuits inter-laboratoires. Nous avons émis l'hypothèse que les données ainsi obtenues peuvent être exploitées pour apporter des informations sur les incertitudes de mesure et établir un modèle mathématique de calcul. On peut également supposer qu'en étudiant une population représentative des sols, composée de sols issus d'un circuit d'essais d'aptitude, il est envisageable d'estimer une incertitude s'adaptant à l'ensemble des sols.

L'incertitude totale (I_t) est définie mathématiquement par la relation : $I_t = k \times s_{\text{total}}$ (1)

avec k , facteur d'élargissement, = 2 ou 3 selon de niveau de confiance choisi

(2 pour $P = 0.05$ et 3 pour $P = 0.01$)

et s_{total} est l'écart type global.

Cet écart type global est tel que $s_{\text{total}} = (s_{\text{intra}}^2 + s_{\text{inter}}^2)^{1/2}$ (2)

où s_{intra} est l'écart type intra série (répétitions au sein d'une même série)

et s_{inter} est l'écart type inter séries (lié aux répétitions des séries).

Ainsi, d'après les relations (1) et (2), pour estimer l'incertitude totale il nous faut évaluer les écarts types intra et inter séries correspondants.

L'approche originale développée au LAS consiste à estimer d'une part s_{intra} en fonction de la concentration de l'analyte et d'autre part s_{inter} en fonction de s_{intra} . Elle est basée sur les résultats d'essais de répétitions sur une population de sols d'un circuit d'essais d'aptitude judicieusement choisis ainsi que sur l'étude de nos cartes de contrôle.

Trois hypothèses sont avancées pour justifier cette démarche. D'une part, on considère que la meilleure estimation disponible des influences respectives des facteurs de répétabilité et de fidélité intermédiaire exprimée par $r = s_{\text{inter}} / s_{\text{intra}}$ est issue des cartes de contrôle du LAS.

D'autre part, ayant observé sur nos cartes de contrôle que le rapport r est stable pour des analyses de même nature (ensemble des teneurs totales en éléments en traces dans les sols par exemple), on considère qu'il peut être généralisé à l'ensemble des échantillons et à divers niveaux de concentration.

Enfin, les éventuels écarts en justesse ne sont pas quantifiés et sont considérés comme non significatifs, dans la mesure où les résultats fournis pour des essais d'aptitude s'inscrivent dans les fourchettes de tolérance des circuits correspondants.

1.2 Evaluation de s_{intra}

Afin de mettre en évidence une relation entre l'écart type intra série, s_{intra} , et la concentration de l'analyte, C , nous avons étudié la répétabilité des analyses de 20 sols du circuit d'essais d'aptitude du Bipea (Bureau InterProfessionnel d'Etudes Analytiques) de caractéristiques physico-chimiques variées (**tableau 1**). Il s'agit de sols des années 2005, 2006 et 2007. L'étude de la répétabilité des mesures sur un nombre restreint de sols de caractéristiques contrastées permet d'évaluer s_{intra} à différents niveaux de concentration. Les déterminations étudiées sont de type agronomique (pH, oligo-éléments, cations échangeables,...) et environnementale (éléments en traces métalliques). Pour chaque type d'analyse, 3 répétitions ont été réalisées. A partir des données obtenues, les s_{intra} de chaque échantillon ont été calculés

puis des relations linéaires de type $s_{\text{intra}} = a \times C + b$ (3) ont été recherchées pour chaque paramètre déterminé (a et b étant respectivement la pente et l'ordonnée à l'origine de la droite de régression linéaire).

Tableau 1 : Quelques caractéristiques physico-chimiques globales des sols étudiés.

	pH eau	CEC (cmol+/kg)	Cd Total (mg/kg)	Argiles (g/kg)	Limons (g/kg)	Sables (g/kg)	C (g/kg)	N (g/kg)
Mini	5.48	3.27	0.06	82	96	34	4.49	0.42
Maxi	8.38	16.2	1.04	374	825	819	112	2.44
Médiane	7.42	13.1	0.25	180	554	217	16.4	1.35
Moyenne	7.28	11.3	0.31	203	527	270	25.0	1.39

1.3 Evaluation de s_{inter}

L'intérêt de définir une méthode d'évaluation de s_{inter} en fonction de s_{intra} est de pouvoir s'affranchir de « répéter les répétitions » des analyses de sols du Bipea réalisées pour évaluer s_{intra} .

Le LAS dispose de cartes de contrôle pour chaque type de paramètre proposé, construites à partir des données de répétitions intra-série et inter-séries d'un échantillon de contrôle interne. A partir de ces données, nous avons recherché des relations de type $s_{\text{inter}} = K \times s_{\text{intra}}$ (4) entre les écarts types intra-série et les écarts types inter-séries obtenus pour un groupe d'éléments cohérent (différents éléments dosés par une même méthode analytique par exemple) sur une période d'un semestre. Dans le cas du dosage du cadmium total, nous avons recherché une relation entre les s_{inter} et les s_{intra} obtenus pour Cd, Cr, Cu, Co, Mo, Ni, Pb, Tl et Zn, ces éléments étant tous dosés par des techniques spectrométriques (ICP-AES et ICP-MS) dans le même minéralisat de sol.

1.4 Evaluation de l'incertitude de mesure

Si l'on reprend les équations (1) et (2), nous obtenons :

$$I_t = k \times (s_{\text{intra}}^2 + s_{\text{inter}}^2)^{1/2}$$

En intégrant l'équation (4), cette relation devient :

$$I_t = k \times [s_{\text{intra}}^2 + (K \times s_{\text{intra}})^2]^{1/2}$$

$$\text{d'où } I_t = k \times [s_{\text{intra}}^2 \times (K^2 + 1)]^{1/2}$$

$$\text{et } I_t = k \times s_{\text{intra}} \times (K^2 + 1)^{1/2}$$

En intégrant la relation (3), nous obtenons :

$$I_t = k \times (a \times C + b) \times (K^2 + 1)^{1/2}$$

$$\text{d'où } I_t = k \times a \times (K^2 + 1)^{1/2} \times C + k \times b \times (K^2 + 1)^{1/2}$$

que l'on peut reformuler en :

$$I_t = a' \times C + b'$$

$$\text{avec } a' = k \times a \times (K^2 + 1)^{1/2} \text{ (5)}$$

$$\text{et } b' = k \times b \times (K^2 + 1)^{1/2} \text{ (6)}$$

Ainsi, en évaluant expérimentalement « a' » et « b' » à partir de « a », « b » et « K » nous pouvons estimer l'incertitude totale de mesure en fonction de la concentration mesurée.

2. Résultats

Afin d'alléger la présentation des résultats, seuls quelques exemples représentatifs sont exposés ci-après et plus particulièrement celui du dosage de la teneur en cadmium total dans les sols. La **figure 1** montre la bonne relation obtenue entre les écarts types intra-série s_{intra} et les teneurs en Cd des 20 sols de circuit d'essais d'aptitude étudiés. Ainsi, pour Cd, $s_{intra}=0.026 \times C - 0.001$. Le coefficient a (pente de la droite) est étroitement lié au coefficient de variation des mesures (CV) : une relation satisfaisante entre C et s_{intra} est le reflet d'un CV relativement constant sur l'ensemble de la gamme de concentration étudiée. De fait, lorsque la relation obtenue n'est pas satisfaisante, a est calculé à partir du CV moyen obtenu pour l'ensemble des 17 sols ($a = CV_m/100$) et b (ordonnée à l'origine) est fixé comme étant égal au plus faible s_{intra} de la série. Il est également possible d'avoir $a = 0$; dans ce cas l'incertitude est indépendante de la concentration, comme constaté pour le pHeau (**figure 2**). La valeur de b est alors fixée comme étant égale au s_{intra} de l'échantillon de contrôle interne du LAS.

La **figure 3** représente la relation entre les s_{inter} et les s_{intra} de l'échantillon de contrôle interne du LAS issus des cartes de contrôle de Cd, Cr, Cu, Co, Mo, Ni, Pb, Tl et Zn (dosés après minéralisation totale suivant la norme Afnor NF X 31-147). Pour l'ensemble de ces éléments, nous considérons donc que $s_{inter} = 0.95 \times s_{intra}$, d'où $K \approx 1$.

Au final, si nous reprenons comme exemple le cas de Cd total, qu'en est-il de l'estimation de l'incertitude des mesures ?

Nous avons défini que :

$$s_{intra} = 0.026 \times C - 0.001 \text{ (figure 1)}$$

$$s_{inter} = 1 \times s_{intra} \text{ (figure 3)}$$

$$\text{avec } a = 0.026, b = -0.001 \text{ et } K = 1$$

d'où, si l'on reprend les relations (5) et (6) :

$$a' = k \times 0.026 \times 2^{1/2} = k \times 0.037$$

$$b' = k \times (-0.001) \times 2^{1/2} = k \times (-0.001)$$

Avec $k = 3$, nous obtenons :

$$I_t = 0.11 \times C + 0.004$$

Ainsi, à partir de cette formule, il nous est possible d'estimer l'incertitude de mesure de la concentration C en cadmium total dans les sols.

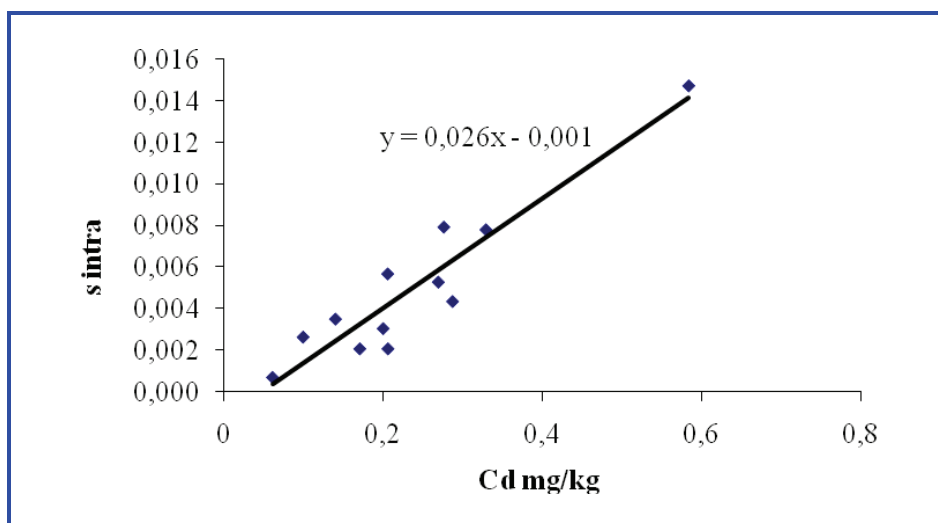


Figure 1 : relation entre la concentration en Cd total dans les sols et l'écart type intra série

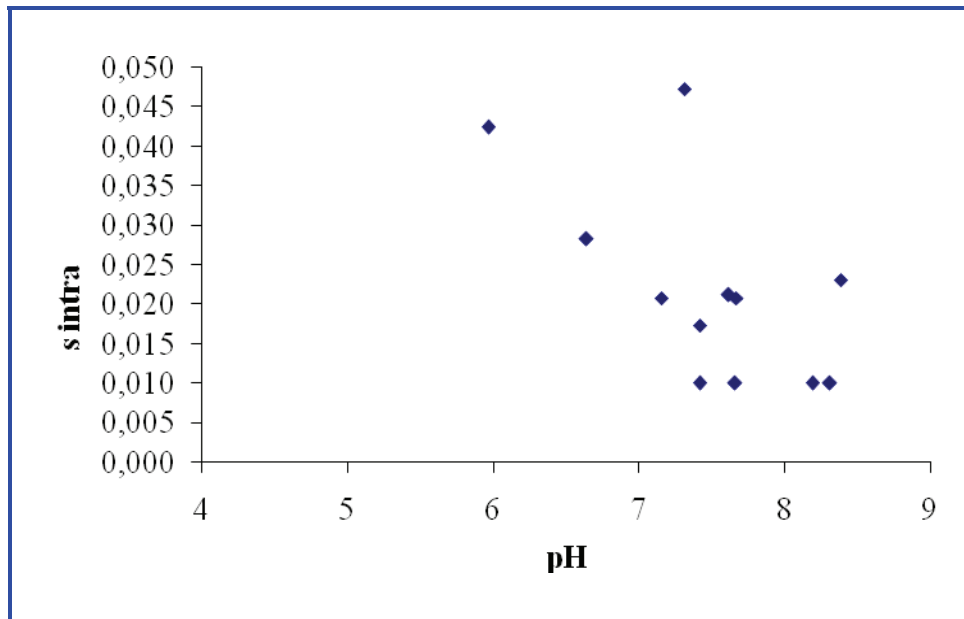


Figure 2 : relation entre le pHeau des sols et l'écart type intra série.

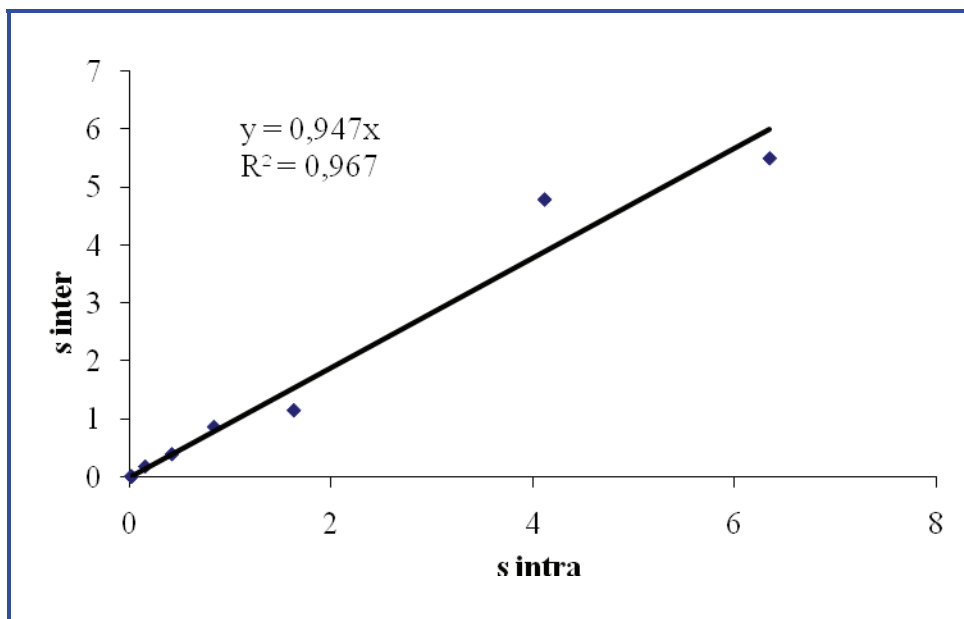


Figure 3 : relation entre s_{intra} et s_{inter} de éléments en traces (Cd, Cr, Co, Cu, Mo, Ni, Pb, Tl et Zn) dans les minéralisats HF de l'échantillon de contrôle interne du LAS.

3. Conclusion

A partir de l'étude de 20 sols de nature variée, issus d'un circuit d'essais d'aptitude et en exploitant les données de ses cartes de contrôle internes, le LAS a défini une méthode propre d'estimation des incertitudes de mesure d'une partie des déterminations qu'il propose. Afin de couvrir l'ensemble des déterminations du catalogue du LAS, des études similaires ont été

réalisées avec d'autres sols de circuits inter-laboratoires ainsi qu'avec des boues. De fait, les coefficients a' et b' utilisés pour estimer l'incertitude des mesures sont disponibles et fournis sur demande auprès du LAS.

Bibliographie

Bipea, rapports 2005, 2006, 2007

Feinberg M., Laurentie M. (2006) A global approach to method validation and measurement uncertainty. *Accred Qual Assur*, 11: 3-9.

Norme NF X 31-147 : (Qualité des sols) : Sols, sédiments – Mise en solution totale par attaque acide

Validation d'une méthode de chimie analytique

Application au dosage des anions fluorure, chlorure, nitrite, bromure, nitrate, phosphate et sulfate par chromatographie ionique

Monique Mayor, Guilhem Bourrié¹

La validation d'une méthode de chimie analytique passe par l'estimation de l'erreur de justesse, ou biais systématique, et de l'erreur aléatoire, qu'il faut combiner pour vérifier si la méthode remplit les objectifs qui lui ont été assignés. La construction d'un modèle statistique permet d'évaluer quantitativement ces notions, qui peuvent être visualisées graphiquement, suivant la méthode du profil d'exactitude (Feinberg, 2010a, b et c), et de se donner les moyens de contrôler le maintien dans le temps des performances de la chaîne de mesure, par la construction des cartes de contrôle.

Résumé : *La validation d'une méthode d'analyse passe par différentes étapes qui visent : (i) à tester la normalité de la distribution des mesures ; (ii) à estimer les composantes de l'incertitude de mesure, fidélité, justesse, spécificité ; (iii) à définir des tests de contrôle de la non dégradation des performances de la méthode. Après un rappel des définitions et des concepts métrologiques, cet article présente les différentes étapes qui permettent de valider une méthode d'analyse. La méthode consiste en la construction d'un modèle statistique, d'un plan d'expérience et l'utilisation d'un matériau de référence certifié (MRC), utilisé (i) pour évaluer le biais (ou « erreur systématique »), (ii) pour calculer « l'incertitude composée », somme de l'incertitude sur la moyenne (« erreur aléatoire ») et de l'incertitude sur la valeur exacte du MRC et (iii) pour construire les cartes de contrôle. Cette méthode est appliquée au dosage des anions contenus en solution aqueuse par chromatographie ionique, et le modèle choisi est linéaire. Les résultats montrent : (i) que les distributions ne sont pas significativement différentes de la loi normale ; (ii) qu'il existe généralement un biais significativement différent de zéro et qu'il faut corriger ; (iii) qu'il n'existe aucune liaison entre le biais et la concentration de l'analyte ; (iv) que l'intervalle de confiance peut être calculé par application de la loi de Student avec l'incertitude composée. La représentation graphique suivant la méthode du « profil d'exactitude » aide à visualiser les performances de la méthode pour chacun des analytes. La méthode des ajouts dosés permet de vérifier la spécificité de la méthode et d'établir le pourcentage de recouvrement. Les droites de régression obtenues permettent de corriger les concentrations brutes de l'erreur systématique. Les cartes de contrôle permettent de déceler la dégradation des performances dans le temps et la stabilité des étalonnages.*

Mots clés : Anions, bromure, chimie analytique, chimiométrie, chromatographie ionique, chlorure, étalonnage, fluorure, incertitude, nitrate, nitrite, phosphate, statistiques, sulfate, validation.

¹ UR1119, Géochimie des sols et des eaux - INRA – F-13545 Aix-en-Provence

☎ 04 42 90 85 42

✉ monique.mayor@aix.inra.fr

Introduction

Toute analyse chimique nécessite des procédures de mise au point de méthodes de dosage, d'étalonnage et de validation des résultats. La démarche Qualité a pour objectif d'assurer la fiabilité et la traçabilité des résultats. Nous présentons ici le protocole expérimental et les méthodes statistiques utilisées pour valider une méthode d'analyse. La démarche est appliquée ici au dosage d'anions en solution par chromatographie ionique et peut être étendue sans difficultés à d'autres méthodes. Plus précisément, l'objectif de cet article est d'évaluer statistiquement la validité de cette méthode de dosage dans le domaine de concentration de 10^{-5} à 10^{-3} M, et de décrire les procédures d'étalonnage multiparamétrique, de correction de biais, de calcul d'incertitude par l'utilisation d'étalons certifiés et la méthode d'ajouts dosés.

1. Principe de la méthode et description de la technique

Le choix d'une méthode analytique doit être fait en fonction du problème posé. Au centre Inra d'Avignon, les travaux de l'unité « Géochimie des sols et des eaux » nécessitent l'analyse des eaux dans les sols hydromorphes et les sols salés. Dans ces milieux, la salinité totale varie très largement. La chromatographie ionique a l'avantage d'être multiparamétrique, de nécessiter peu d'échantillon, environ 2 ml, et d'être rapide. Le détecteur utilisé ici est conductimétrique et peu coûteux.

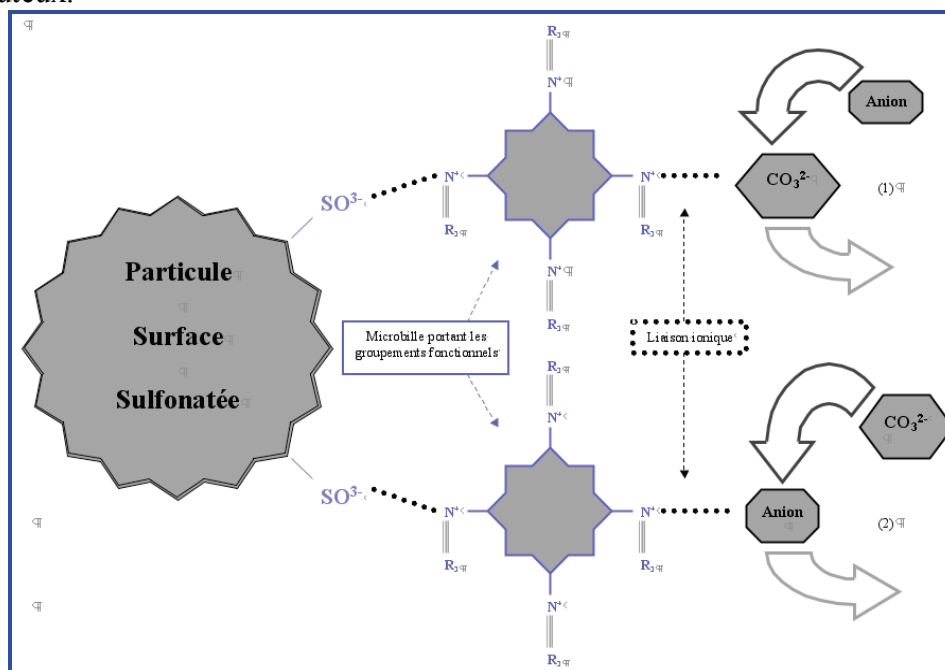


Figure 1 : l'échange ionique : (1) l'anion déplace l'éluant ; (2) l'éluant déplace l'anion

1.1 Principe de la chromatographie ionique

La chromatographie est une technique de séparation d'espèces dans un mélange par partage entre une phase mobile (gaz ou liquide) et une phase stationnaire (liquide ou solide). Chaque soluté injecté dans la colonne est soumis à deux effets antagonistes : l'entraînement par la phase mobile dans laquelle il est soluble et la rétention par la phase stationnaire avec laquelle il interagit. La séparation est possible grâce aux différences d'affinités de ces composés pour

les phases mobile et stationnaire. La phase stationnaire, la colonne, est un échangeur d'ions constitué de particules de résine (**figure 1**).

L'appareil utilisé est un chromatographe ionique Dionex hautes performances monocanal modèle DX 100 muni d'un passeur diluteur Dionex modèle AS50, le tout piloté par le logiciel de gestion « *Chroméléon™ CM-CHM-1-IC version 6.5* ».

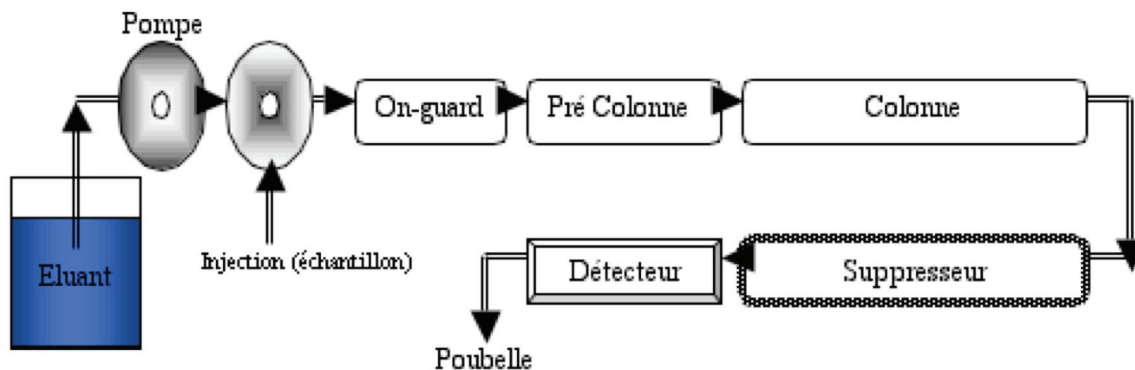


Figure 2 : schéma de principe d'un équipement de chromatographie ionique

1.2 Mise au point de la méthode

Il est nécessaire avant de valider une méthode d'optimiser les conditions d'analyse, en faisant varier le débit, la température, en modifiant la nature ainsi que la concentration de l'éluant, en optimisant le choix de la colonne. On peut ainsi minimiser le temps de l'analyse, fixé par le temps du dernier composé élué, tout en préservant une bonne résolution pour que la méthode reste spécifique, c'est-à-dire qu'elle produise une réponse uniquement pour l'analyte d'intérêt.

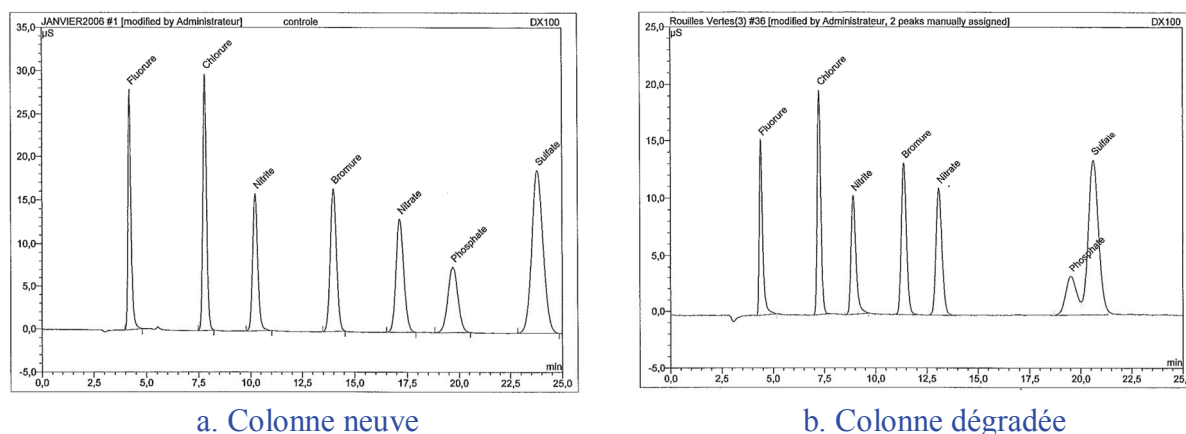


Figure 3 : dégradation des performances d'une colonne de chromatographie ionique.

La résolution de la méthode est le pouvoir de séparation des composés, elle est quantifiable par exemple par le rapport $R = d/m$, où d est la distance séparant les sommets des 2 pics, et m la moyenne des largeurs à la base des pics. Si on considère des pics rapprochés, par exemple les pics du phosphate et du sulfate, il apparaît qu'une résolution $R = 1,12$ n'est pas suffisante, mais qu'avec une résolution supérieure à 2, les pics sont bien résolus. **La figure 3** montre la différence entre une bonne résolution et une résolution incorrecte due à la dégradation des performances de la colonne avec le temps.

2. Théorie, modèle statistique et méthodes de traitement des données

2.1 Modèle statistique

La construction d'un modèle statistique permet d'évaluer quantitativement les grandeurs métrologiques définies ci-dessus à partir de séries de résultats expérimentaux. Le modèle choisi ici est additif : Résultat = valeur vraie + erreur systématique + erreur aléatoire.

La démarche choisie consiste d'abord à tester la normalité de la distribution des résultats, puis à évaluer l'erreur systématique, l'erreur aléatoire, l'incertitude composée, la spécificité de la méthode, les limites de détection et de quantification, l'intervalle de confiance du résultat de mesure, enfin à construire la procédure de carte de contrôle.

2.2 Méthodes

2.2.a Normalité

La normalité est testée en construisant la droite de Henry : si les points sont alignés sur la droite, on ne peut pas infirmer l'hypothèse de normalité de la distribution.

2.2.b Évaluation du domaine d'analyse

La détermination du domaine d'analyse permet de vérifier si la technique est adaptée aux valeurs ou aux concentrations recherchées. La courbe d'étalonnage peut être linéaire mais cela ne fait pas partie des exigences. L'exigence de linéarité s'applique à la relation entre la concentration calculée et la concentration introduite et non pas à la fonction de réponse qui est la relation entre le signal et la concentration introduite. Il est en revanche nécessaire que la fonction de réponse soit strictement croissante ou décroissante. Le calcul de la courbe d'étalonnage par la méthode des moindres carrés suppose la constance de la variance de la réponse à toutes les concentrations.

2.2.c Calcul des limites de détection et de quantification

À partir d'une mesure faite sur un blanc analytique, on calcule l'écart type du signal. La limite de détection LOD correspond à la concentration qui conduit à un signal dont l'intensité est égale à 3 fois celle de l'écart type du blanc $y_d = y_b + 3 s_b$. La limite de quantification LOQ est calculée en prenant comme valeur du signal 10 fois la valeur de l'écart type du blanc $y_q = y_b + 10 s_b$. (ACS, 1980).

2.2.d Calcul des incertitudes

L'incertitude de la mesure, qui caractérise la dispersion, est exprimée sous la forme d'un écart-type s qui est la racine carrée de la somme de l'ensemble des variances associées aux sources d'erreur :

- *le biais du résultat ou erreur systématique* est la différence entre l'espérance des résultats d'essais et la valeur de référence acceptée, et dépend de la méthode, de l'opérateur et de l'instrument. L'estimation de l'erreur systématique, c'est-à-dire l'incertitude systématique, est évaluée en utilisant un étalon certifié. On obtient alors un facteur de correction, qu'on applique au résultat de mesure. Mais comme l'étalon certifié (MRC) n'est connu qu'avec

une certaine incertitude, soit u_{mrc} , celle-ci doit être comptée dans l'incertitude de mesure. Cette incertitude est donnée par l'organisme certificateur, sous la forme de l'incertitude élargie, U_{mrc} , telle que : $U_{mrc} = k_{mrc}u_{mrc}$. Sauf indication contraire, dûment documentée, $k_{mrc} = 2$ pour un niveau de confiance de 95 % (Taylor et Kuyatt, 1994) ;

- *la fidélité ou erreur aléatoire*, est définie comme la dispersion de la moyenne. Son estimation, ou incertitude aléatoire est :

$$u_a = s/\sqrt{n}, \quad (1)$$

Où n est le nombre de mesures. Le matériau de référence certifié doit avoir une matrice identique à celle des échantillons. Ceci implique l'égalité des variances et peut également être testé par la méthode des ajouts dosés. Dans ces conditions, s est mesuré sur le MRC. On obtient donc l'incertitude composée :

$$u_c = \sqrt{u_{mrc}^2 + u_a^2}, \text{ ou encore } u_c = \sqrt{\left(\frac{U_{mrc}}{k_{mrc}}\right)^2 + \left(\frac{s}{\sqrt{n}}\right)^2}. \quad (2)$$

2.2.e Calcul de l'intervalle de confiance du résultat

L'intervalle de confiance au seuil α (niveau de confiance $1-\alpha$) délimite la zone dans laquelle se trouve probablement la moyenne de la population d'où est extrait l'échantillon. On fixe le risque acceptable pour cette affirmation. L'intervalle de confiance à 95 % signifie que l'on accepte un risque de 5 % de se tromper en affirmant que la moyenne de la population est comprise entre [moyenne - IC et moyenne + IC].

Si x est la grandeur recherchée (exemple la concentration) et y la grandeur mesurée (exemple la hauteur du pic) et si l'étalonnage linéaire est donné par la droite de régression de x par rapport à y , alors $x = C_1 y + C_2$, et l'intervalle de confiance de la moyenne de n mesures est :

$$\bar{x} \pm k(\alpha, \nu)u_c, \quad (3)$$

Où \bar{x} est la moyenne corrigée de toutes les erreurs systématiques identifiées, k suit une **loi de Student** à ν degrés de liberté, $\nu = n - 1$, α est le niveau de confiance, généralement $\alpha = 0,05$ pour $P = 0,95$, n est la taille de l'échantillon, et u_c est l'incertitude composée définie ci-dessus. Le produit $U_c = k u_c$ est l'incertitude élargie.

2.2.f Spécificité

La spécificité est la capacité à établir l'existence de la substance à analyser en présence d'autres composants présents. Il s'agit de démontrer que la substance analysée au sein de la matrice est bien l'analyte recherché. La spécificité se fonde sur une absence d'interférences. Elle sera estimée par la méthode des ajouts dosés en calculant le pourcentage de recouvrement.

3. Protocole expérimental

Le protocole expérimental est présenté dans le **tableau 1**, et l'ensemble des étapes dans le **tableau 2**.

Tableau 1 : *protocole expérimental*

Jour 1	Jour 2	Jour 3	Jour 4	Jour 5
Mettre les sels à déshydrater dans l'étuve à 105 °C pendant 24 h en vue de la préparation de la solution multiéléments	Préparation des solutions multiéléments à 0, 1 ; 0,25 ; 0,5 et 1 mmol/l en Fluorure, Chlorure, Nitrite, Bromure, Nitrate, Phosphate, Sulfate	Plan d'étalonnage Fidélité 3 répétitions des solutions multiéléments, niveaux 1, 2, 3, 4	Plan d'étalonnage Fidélité 3 répétitions des solutions multiéléments, niveaux 1, 2, 3, 4	Plan d'étalonnage Fidélité 3 répétitions des solutions multiéléments, niveaux 1, 2, 3, 4
	Dilution de la solution certifiée multiélément à 1 mmol/l au 1/10 ; 1/4 ; 1/2 en Fluorure, Chlorure, Nitrite, Bromure, Nitrate, Phosphate, Sulfate	Plan de validation Justesse 5 répétitions des solutions certifiées, niveaux 1, 2, 3, 4	Plan de validation Justesse 5 répétitions des solutions certifiées, niveaux 1, 2, 3, 4	Plan de validation Justesse 5 répétitions des solutions certifiées, niveaux 1, 2, 3, 4
	Normalité de la distribution : 30 répétitions de la solution certifiée à 0,5 mmol/L	Spécificité Doser la matrice seule et la matrice + 3 concentrations d'ajouts dosés		

- Pour vérifier **la normalité** de la distribution, nous mesurons 30 fois le matériau de référence certifié sous des conditions de répétabilité.

- Pour apprécier **la fidélité**, nous utilisons le plan d'étalonnage en faisant 9 répétitions des solutions multiéléments préparées sans la matrice, à 4 niveaux de concentrations, (0,1 mmol/l; 0,25 mmol/l ; 0,5 mmol/l et 1 mmol/l) ; les dosages sont réalisés sur 3 jours avec le même opérateur. L'équation de la droite d'étalonnage journalière est utilisée pour calculer les concentrations qui serviront pour le calcul de justesse. Le coefficient de variation *CV* des aires est calculé pour les 7 analytes, ceci pour les différents niveaux de concentration.

- Pour apprécier **la justesse**, nous mesurons 15 répétitions du matériau de référence certifié multiélément (MRC) de valeurs connues, réalisées sur 3 jours ceci aux 4 niveaux de concentrations ci-dessus (**tableau 1**), ce qui permet d'apprécier la correction à apporter au résultat obtenu, ainsi que d'effectuer les calculs d'incertitudes.

- Pour apprécier **la spécificité** et les effets de matrice, nous dosons un échantillon de solution du sol représentatif du contexte géochimique étudié, d'abord seul, puis après addition de l'analyte à trois concentrations différentes.

Tableau 2 : résumé des différentes étapes de la validation

Principales étapes	Protocole	Tests et calculs
Vérification de la normalité de la distribution	30 répétitions du matériau de référence certifié : MRC multiélément à une concentration de 0,5 mmol/l en Fluorure, Chlorure, Nitrite, Bromure, Nitrate, Phosphate, Sulfate sous des conditions de répétabilité.	Vérification visuelle et droite de Henry
Plan d'étalonnage : Fidélité	9 répétitions des Solutions multiéléments préparées sans la matrice, à 4 niveaux de concentrations, dosages réalisés 3 jours différents avec le même opérateur.	Calcul de la moyenne, variance, écart-type des 9 mesures et du coefficient de variation $CV = \frac{s}{\bar{x}} \times 100$.
Plan de validation : Justesse Spécificité Vérification de l'absence d'un effet de matrice	15 répétitions du matériau de référence certifié multiélément de valeurs connues réalisés sur 3 jours différentes, ceci à 4 niveaux de concentrations, ce qui permet d'apprécier la correction à apporter au résultat obtenu ainsi que les calculs des incertitudes. Par la méthode des ajouts dosés, en dosant un échantillon de solution du sol sans ajout et avec 3 ajouts successifs.	Biais ou erreur systématique : Différence entre résultat obtenu et valeur fournie par le laboratoire certifié. Erreur aléatoire : $u_a = s / \sqrt{n}$ tenir compte de l'incertitude de l'étalon $U_{crm} = k_{crm} u_{crm}$ Incertitude de la mesure : $u_c = \sqrt{u_{crm}^2 + u_a^2}$ Intervalle de confiance du résultat (t de Student) $\bar{x} \pm t(\alpha, \nu) u_c$ Calculer le taux de recouvrement , à partir des pentes des droites de régression et démontrer la linéarité entre les concentrations calculées et les concentrations introduites.
Évaluation du domaine d'analyse , Limite de Détection et de Quantification	A l'aide de mesures faites sur un blanc analytique, on déduit la valeur du bruit de fond y_b qui est la réponse au blanc « terrain » Et on calcule l'écart-type de la réponse au blanc « terrain » s_b	$LOD : y_d = y_b + 3 s_b$, $LOQ : y_d = y_b + 10 s_b$, Déduire les concentrations de LOD et LOQ à partir de la valeur de leur signal grâce à l'équation : $x = C_1 y + C_2$ de la droite d'étalonnage de chaque élément. Courbe d'étalonnage : s'assurer que la fonction de réponse est strictement croissante ou décroissante.

Pour valider une méthode, il faut :

- que les CV dans des conditions de fidélité intermédiaire soient inférieurs à une valeur de tolérance, par exemple 10 % ;
- que la régression permettant de relier les concentrations calculées aux concentrations introduites soit parfaitement linéaire dans la gamme de concentrations étudiées ;
- que la différence entre le résultat obtenu et la valeur vraie inconnue de l'échantillon soit petite ou du moins inférieure à une limite d'acceptation : pour cela, les estimateurs de biais et de variance obtenus grâce à des expériences effectuées sur des échantillons connus nous renseignent si la procédure va pouvoir quantifier avec une exactitude suffisante.

4. Résultats et discussion

Les résultats (histogramme et droite de Henry) montrent que l'hypothèse de **normalité** de la distribution ne peut pas être rejetée.

4.1 La fidélité

Les résultats complets pour la **fidélité** sont reportés dans le **tableau 3**. Toutes les valeurs des coefficients de variation sont inférieures à 4 %. La valeur admise dans notre protocole est de 10 %.

Tableau 3 : résultats du test de fidélité, aux quatre niveaux de concentration des solutions multiéléments.
Valeurs en aires de pics ($\mu S \text{ min}$)

	F ⁻	Cl ⁻	NO ₂ ⁻	Br ⁻	NO ₃ ⁻	P-PO ₄ ³⁻	SO ₄ ²⁻
Solution multiélément 0,10 mmol/l							
Moyenne	0,547	0,688	0,526	0,655	0,658	0,502	1,324
Écart type	0,015	0,017	0,018	0,013	0,021	0,019	0,034
Variance	0,000	0,000	0,000	0,000	0,000	0,000	0,001
CV	2,74 %	2,44 %	3,40 %	1,98 %	3,25 %	3,79 %	2,59 %
Solution multiélément 0,25 mmol/l							
Moyenne	1,439	1,883	1,416	1,730	1,661	1,324	3,463
Écart type	0,032	0,038	0,027	0,032	0,028	0,026	0,066
Variance	0,001	0,001	0,001	0,001	0,001	0,001	0,004
CV	2,24%	2,02%	1,88%	1,83%	1,66%	1,99%	1,90%
Solution multiélément 0,5 mmol/l							
Moyenne	2,974	3,750	2,900	3,514	3,409	2,747	7,104
Écart type	0,094	0,122	0,089	0,104	0,103	0,065	0,210
Variance	0,009	0,015	0,008	0,011	0,011	0,004	0,044
CV	3,17%	3,25%	3,09%	2,97%	3,03%	2,38%	2,95%
Solution multiélément 1 mmol/l							
Moyenne	6,167	8,184	6,168	7,682	7,369	5,815	15,358
Écart type	0,155	0,165	0,105	0,147	0,140	0,126	0,338
Variance	0,024	0,027	0,011	0,022	0,020	0,016	0,114
CV	2,51%	2,02%	1,70%	1,91%	1,90%	2,17%	2,20%

4.2 La justesse

Les résultats du test de justesse montrent (non présentés ici, données disponibles sur demande) que **le biais** (estimation de l'erreur systématique) est significativement différent de zéro, en général positif, ce qui correspond à une surestimation ; il est négatif pour le chlorure et pour le fluorure et pour ce dernier, le biais maximum est observé à la concentration 0,1 et atteint - 6,2 %, ce qui est considérable. On n'observe pas de variation ordonnée du biais avec la concentration. Il faut donc corriger les valeurs mesurées de ce biais estimé.

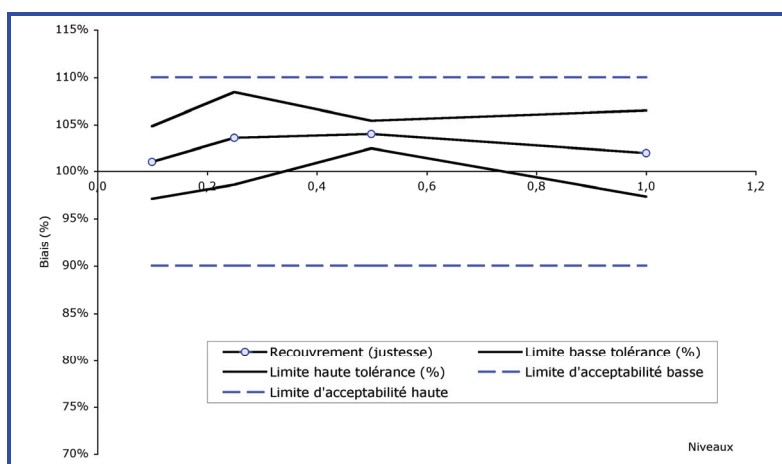
Une fois corrigées de ce biais, les valeurs sont accompagnées d'une **incertitude composée** obtenue selon l'équation ci-dessus (éq. 2), qui permet de calculer l'intervalle de confiance du résultat. Ici, le nombre de mesures effectuées pour estimer la moyenne de la concentration d'un échantillon inconnu est de 5. L'intervalle de confiance est donc obtenu par application de l'équation (3) avec $n = 5$, $k(0,95; 4) = 2,78$ et u_c l'incertitude composée. Les résultats obtenus montrent que les moyennes journalières sont comprises dans les intervalles de confiance de la moyenne ainsi calculés. Elles peuvent donc être considérées comme des estimations indépendantes de la même grandeur. Ceci veut dire que les performances instrumentales et que l'étalonnage sont restés stables pendant les 3 jours consécutifs des mesures. On peut en déduire également le nombre de mesures nécessaires pour obtenir un intervalle de confiance sur la moyenne donnée, puisqu'on est dans les conditions de validité de la loi normale.

4.3 Les profils d'exactitude

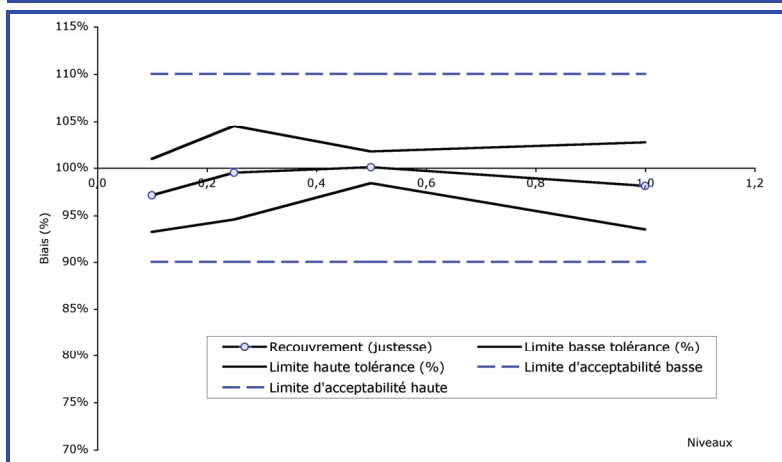
Une autre représentation possible est celle du « **profil d'exactitude** » (Feinberg, 2000a et b, 2006, 2010a, b, c.). On se fixe une limite d'acceptabilité (ici 10 %) et une probabilité de tolérance (95 %). Par exemple pour le sulfate (**tableau 4, figure 4**) au niveau 3, concentration moyenne du MRC, le biais a une valeur de 3,9 % avant correction et le taux de recouvrement est de 103,9 %.

Tableau 4 : profil d'exactitude du sulfate

	Avant correction du biais				Après correction du biais			
	95 %				95 %			
Limite d'acceptabilité	10 %				10 %			
Niveaux	1	2	3	4	1	2	3	4
Valeur cible	0,100	0,250	0,500	1,000	0,100	0,250	0,500	1,000
Moyenne niveau	0,101	0,259	0,519	1,020	0,097	0,249	0,500	0,981
Écart-type de fidélité (s_{FI})	0,002	0,006	0,003	0,021	0,002	0,006	0,004	0,021
Valeur basse tolérance	0,097	0,246	0,512	0,974	0,093	0,236	0,492	0,935
Valeur haute tolérance	0,105	0,271	0,527	1,066	0,101	0,261	0,509	1,027
Biais (%)	1,07	3,49	3,89	2,00	-2,90	-0,51	0,05	-1,90
Recouvrement (justesse) (%)	101,07	103,49	103,9	102,00	97,10	99,49	100,05	98,10
Limite basse tolérance (%)	97,18	98,57	102,38	97,42	93,27	94,57	98,37	93,52
Limite haute tolérance (%)	104,95	108,42	105,40	106,58	100,93	104,42	101,74	102,68
Limite d'acceptabilité basse	90,00	90,00	90,00	90,00	90,00	90,00	90,00	90,00
Limite d'acceptabilité haute	110,00	110,00	110,00	110,00	110,00	110,00	110,00	110,00



a)



b)

Figure 4 : profil d'exactitude du sulfate
a) avant correction du biais ; b) après correction du biais

La correction du biais (figure 4b) permet de satisfaire plus largement au critère d'acceptabilité, et ramène le taux de recouvrement à 100 %. Les profils d'exactitude pour les autres analytes sont représentés sur les figures 5 à 10 et démontrent que la méthode pour ces éléments est parfaitement adaptée aux objectifs fixés : elle est donc valide sans avoir à affecter un facteur de correction aux futures mesures.

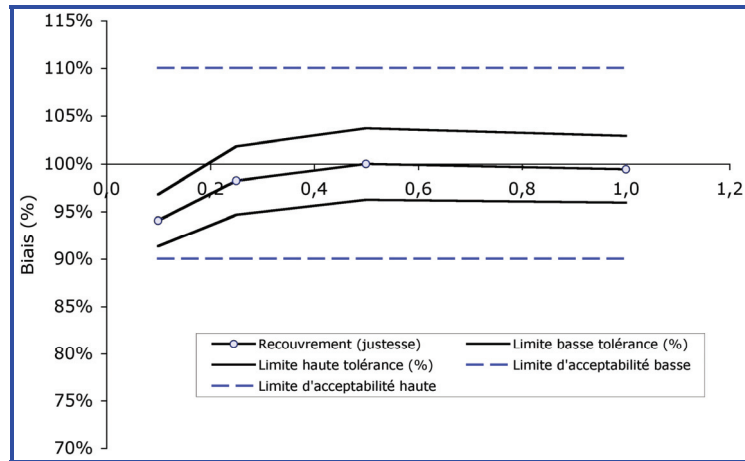


Figure 5 : profil d'exactitude du fluorure

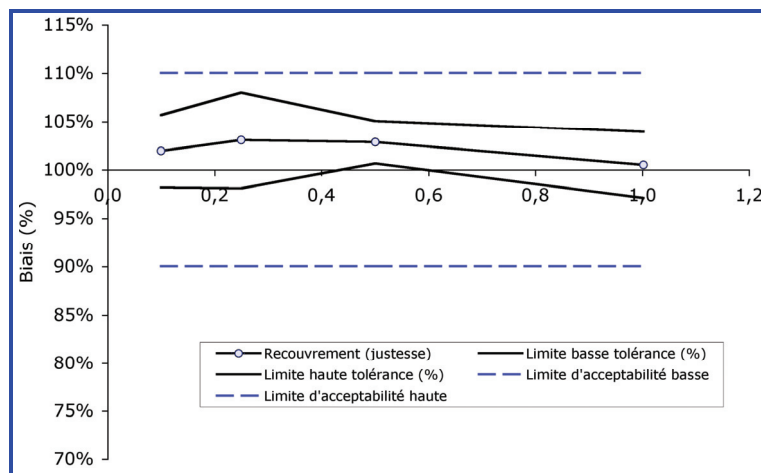


Figure 6 : profil d'exactitude du bromure

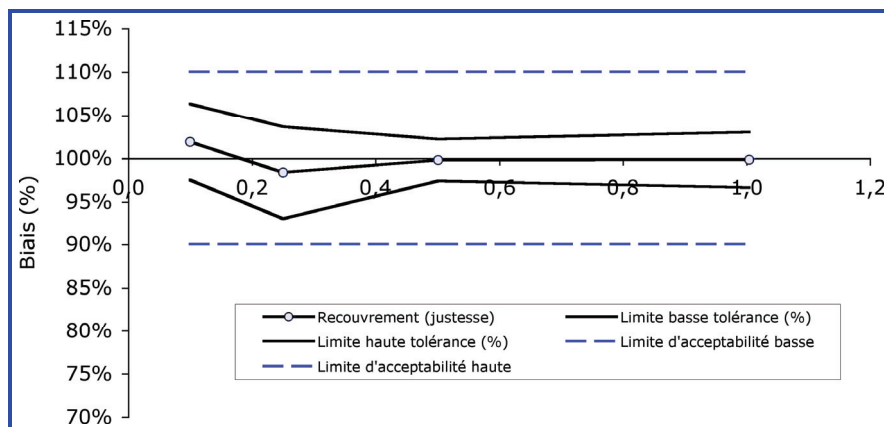


Figure 7 : profil d'exactitude du chlorure

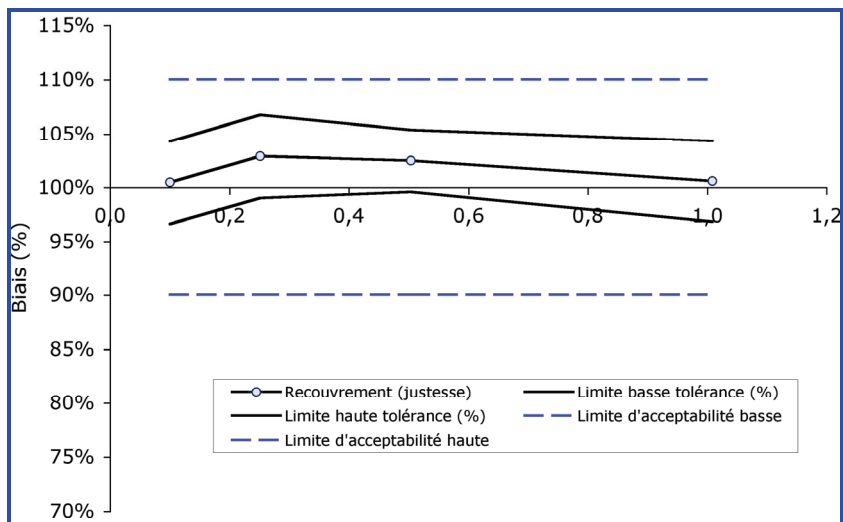


Figure 8 : profil d'exactitude du nitrate

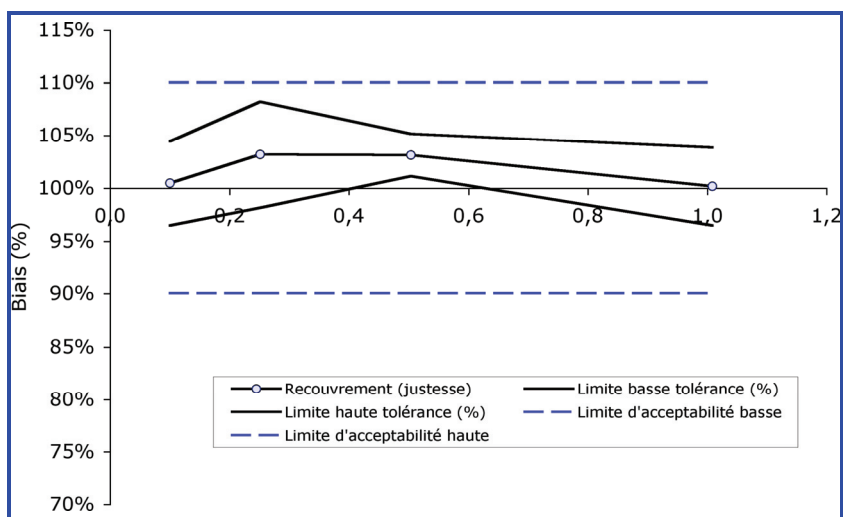


Figure 9 : profil d'exactitude du nitrite

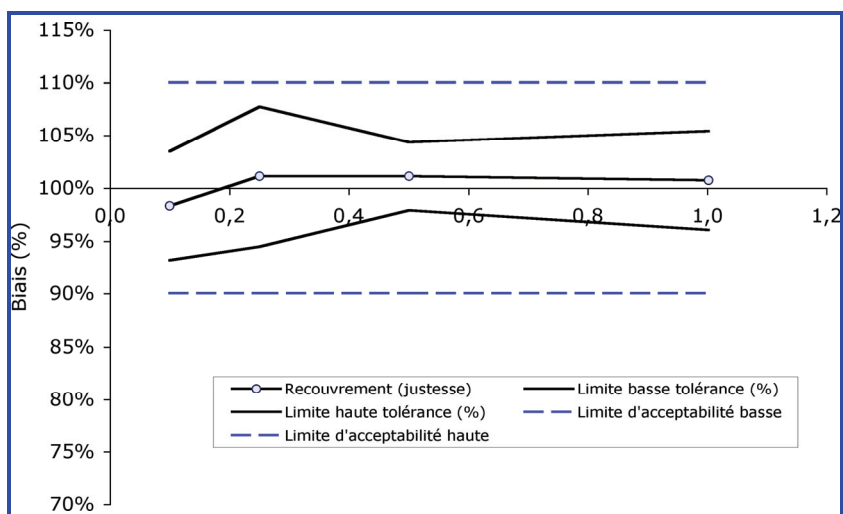


Figure 10 : profil d'exactitude du phosphate

4.4 La spécificité et l'absence d'effets de matrice

Elles sont estimées par la méthode des ajouts dosés (**figure 11** et **tableau 5**). Dans tous les cas, on observe une très bonne linéarité entre la concentration retrouvée et la concentration ajoutée. Toutes les valeurs du test de Fisher sont largement supérieures à la valeur limite. Les pourcentages de recouvrement obtenus par la méthode des ajouts dosés sont déduits des pentes des droites de régression. Des tests sur les ordonnées à l'origine ont montré que celles-ci n'étaient pas significativement différentes de zéro. Les pentes des droites de régression ont donc été recalculées en forçant le passage par l'origine. Les pentes des droites de régression donnent alors directement le pourcentage de recouvrement.

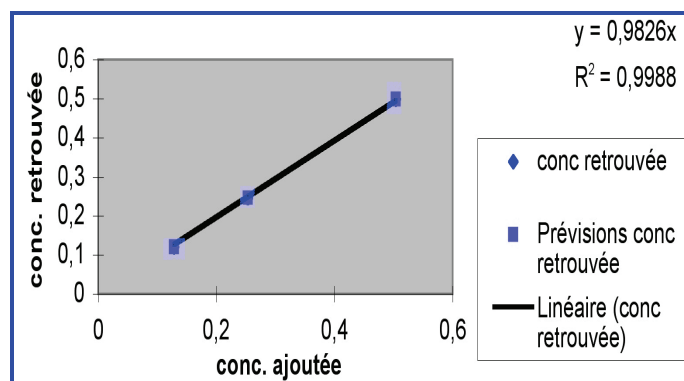


Figure 11 : relation linéaire entre concentration retrouvée et concentration ajoutée par la méthode des ajouts dosés. Cas du fluorure

Tableau 5 : pourcentages de recouvrement obtenus par la méthode du profil d'exactitude et par la méthode des ajouts dosés.

a. moyenne des 4 valeurs de gamme ; b. pente de la droite de régression forcée à passer par l'origine

Analyte	Profil d'exactitude ^a	Ajouts dosés ^b
Fluorure	98,0	98,3
Chlorure	100,0	102,2
Nitrite	102,0	96,5
Bromure	102,0	95,0
Nitrate	101,6	95,4
Phosphate	100,5	97,9
Sulfate	102,6	100,2

Les valeurs obtenues par les pentes des droites de régression sont différentes des valeurs obtenues par la méthode du profil d'exactitude, et plus éloignées de 100 %. Ceci peut être attribué au fait que la méthode du profil d'exactitude a été appliquée à des solutions multiéléments en solution pure, tandis que la méthode des ajouts dosés intègre l'effet de matrice sur un échantillon naturel représentatif du milieu géochimique étudié. On peut considérer que la méthode du profil d'exactitude (**figures 4 à 10**) permet de visualiser l'ordre de grandeur des biais des méthodes, mais qu'il vaut mieux utiliser les droites de régression obtenues (**figure 11**) pour corriger les concentrations brutes, car ces droites tiennent compte de l'effet de matrice.

Dans ces conditions, le résultat d'analyse est simultanément corrigé des effets de matrice et du biais suivant la relation :

$$\bar{x} = (C_1 y + C_2) / C_3 \pm k(\alpha, \nu) u_c, \quad (4)$$

Où C_1 et C_2 sont les coefficients de la courbe de réponse, ici supposée linéaire, et C_3 le pourcentage de recouvrement divisé par 100.

4.5 Les limites de détection (LOD) et de quantification (LOQ)

Elles sont reportées dans le **tableau 6**. Par convention :

□ $y < LOD$: l'analyte n'est pas détecté ;

$LOD < y < LOQ$: l'analyte est détecté, mais non quantifiable ;

$LOQ < y$: l'analyte est quantifiable.

Tableau 6 : limites de détection et de quantification des méthodes (mmol/l)

Analyte	Fluorure	Chlorure	Nitrite	Bromure	Nitrate	Phosphate	Sulfate
LOD	0,018	0,013	0,018	0,014	0,015	0,019	0,007
LOQ	0,060	0,045	0,060	0,047	0,050	0,063	0,024

Les LOD sont toutes de l'ordre de 15 à 20 $\mu\text{mol/l}$, sauf celle du sulfate, qui est deux fois plus faible. Ceci est dû à la pente plus grande de la droite d'étalonnage, qui traduit une sensibilité plus grande, du fait que le sulfate est un ion bivalent. La chromatographie ionique semble donc plus performante pour le dosage des sulfates, qui pose problème avec les méthodes traditionnelles comme la néphélométrie après précipitation du sulfate de baryum.

5. Tests de conformité

Les tests de conformité sont des tests permettant de vérifier que, pendant l'analyse, les performances du système restent compatibles avec le cahier des charges de la méthode.

La vérification permet de s'assurer que les écarts entre les valeurs indiquées par un appareil de mesure et les valeurs connues correspondantes d'une grandeur mesurée sont tous inférieurs aux erreurs maximales tolérées.

5.1 Construction de la carte de contrôle et calcul des erreurs maximales tolérées.

La carte de contrôle est un enregistrement utilisé pour contrôler la stabilité d'une méthode dans le temps. On utilise comme « étalon de surveillance » le Matériau de Référence Certifié (MRC) qui a servi au contrôle de la justesse. On réalise un nombre assez important de mesures groupées en séries, les résultats sont reportés sur un graphique. Il est préférable de faire une moyenne sur une trentaine de répétitions au moins et dans ces conditions, les valeurs du t de Student peuvent être prises égales à 2 et 3 respectivement.

Les limites de surveillance LS et de contrôle LC sont calculées suivant les formules :

$$LS = \text{moyenne} \pm 2 u_c, \quad (5)$$

$$LC = \text{moyenne} \pm 3 u_c, \quad (6)$$

et reportées sur la carte.

5.2 Surveillance

La *surveillance* complète les opérations de vérification et d'étalonnage afin de *détecter les dérives éventuelles* de l'équipement entre ces mesures.

Il s'agit d'enregistrer les écarts entre les valeurs d'indication de l'équipement de mesure et les valeurs vraies du contrôle certifié. Ces opérations de surveillance sont consignées sur des enregistrements tels que la carte de contrôle. A chaque nouvelle analyse, le résultat du contrôle est ajouté, ce qui permet de vérifier s'il est bien à l'intérieur des limites admises, et ainsi d'entreprendre des actions correctives si nécessaire. Cela permet également de contrôler au cours du temps que les valeurs expérimentales sont réparties d'une façon aléatoire par rapport à la moyenne et de nous alerter sur une éventuelle dérive de l'appareil (**Tableau 7**).

Tableau 7 : résumé des étapes de la surveillance

Principales étapes	Protocole	Tests
1-Vérification d'un taux de résolution correct $R > 1,2$ d'une bonne répétabilité $CV < 5\%$		$R = \frac{d}{m}$, d est la distance séparant les sommets des 2 pics, et m la moyenne des largeurs à la base des pics. coefficient de variation $CV = \frac{s}{x} \times 100$,
2-Construction de la carte de contrôle	30 répétitions du matériau de référence certifié MRC multiélément , à une concentration de 0,5 mmol/L en Fluorure, Chlorure, Nitrite, Bromure, Nitrate, Phosphate, Sulfate, sous des conditions de reproductibilité	Calculer, Moyenne, Variance s^2 - Erreur aléatoire : $u_a = s / \sqrt{n}$ - Incertitude de la mesure : $u_c = \sqrt{u_{mrc}^2 + u_a^2}$ tenir compte de l'incertitude de l'étalon $U_{mrc} = k_{mrc} u_{mrc}$ - Intervalle de confiance du résultat (t de Student) $\bar{x} \pm t(\alpha, \nu) u_c$ Évaluer les limites de surveillance et de contrôle $LS = \text{moyenne} \pm 2 u_c$, $LC = \text{moyenne} \pm 3 u_c$ Construire la carte de contrôle avec les limites inférieures et supérieures des LS et LC en fonction de la « valeur moyenne de la solution certifiée ».
3-Surveillance quotidienne	Insérer dans la séquence analytique le MRC à peu près tous les 5 échantillons, il sert de contrôle	S'assurer que les valeurs du MRC sont comprises à l'intérieur des limites de surveillance Pour les échantillons à analyser, s'assurer que la concentration obtenue est supérieure à la Limite de Quantification et se trouve bien à l'intérieur du domaine d'analyse

On fait correspondre sur cet enregistrement un point à chacune des valeurs d'une statistique calculée sur des échantillons successifs ; chacun de ces points a pour abscisse le numéro de l'échantillon et pour ordonnée la valeur de la statistique calculée sur cet échantillon.

Le but de ces tests de surveillance est de nous alerter en cas d'anomalie. Reste à définir quels seront les critères utilisés pour considérer ce qui sera anormal. Nous décidons dans notre cas d'être vigilants lorsqu'une mesure sort des limites de contrôle, ce qui déclenchera la répétition du contrôle pour confirmer ou infirmer la dérive et nous alertera lorsque la répartition des mesures fait suspecter un comportement non aléatoire. Un exemple de test de surveillance donné par le logiciel *Chroméléon*TM est reporté sur la **figure 12**. On peut constater que le contrôle ne sort pas des limites autorisées.

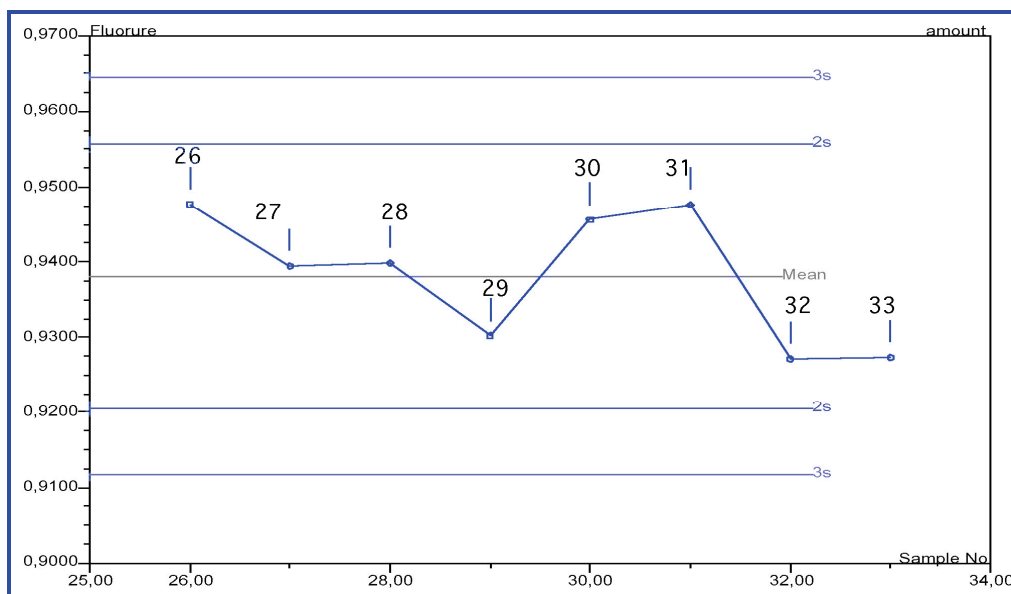


Figure 12 : test de surveillance quotidienne

Conclusion

La démonstration de la normalité des distributions ne nécessite pas d'être répétée. Lors de la mise au point de la méthode d'analyse, ce test est nécessaire et permet d'assurer la validité de l'emploi des statistiques classiques et de leur puissance. Nous avons utilisé ici un modèle statistique linéaire, qui permet d'estimer le biais ou incertitude systématique (« erreur systématique») et l'incertitude aléatoire. Les résultats montrent que le biais n'est pas négligeable et doit être corrigé. L'emploi d'un Matériau de Référence Certifié (MRC) s'impose à cette fin. Ce MRC étant lui-même connu avec une incertitude, donnée par le fabricant, cette incertitude doit être ajoutée à l'incertitude aléatoire, pour donner l'incertitude composée. C'est cette dernière qui est utilisée pour calculer l'intervalle de confiance de la mesure. Ici, comme aucune relation n'a été observée entre l'incertitude composée et la concentration, la valeur de l'incertitude composée obtenue pour une valeur de concentration moyenne (milieu de gamme) a été retenue par souci de simplification. La méthode du profil d'exactitude visualise l'intérêt de corriger du biais et les limites d'acceptabilité. La méthode des ajouts dosés permet d'estimer le pourcentage de recouvrement et de vérifier la linéarité entre concentration retrouvée et concentration ajoutée dans la matrice même des échantillons analysés, c'est-à-dire pour des échantillons représentatifs des milieux étudiés. Les droites de régression obtenues sont utilisées pour corriger des biais systématiques les concentrations

brutes données par la courbe de réponse. L'incertitude composée est également utilisée pour calculer les limites de l'intervalle de confiance acceptables, par une mesure régulière du MRC, qui sert ainsi à effectuer les tests de surveillance et à alerter en cas de dégradation des performances de la chaîne de mesure. L'échantillon doit bien sûr d'abord se situer au dessus de la limite de quantification *LOQ*.

Matériel supplémentaire :

L'ensemble des données et des tableaux peuvent être fournis sur demande.

Bibliographie

- ACS (1980) Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry, *Analytical chemistry*, **52**, 14, 2242-2249
- AFNOR (1994a) Norme française NF X 07-001 -Vocabulaire International des termes fondamentaux et généraux de métrologie
- AFNOR (1994b) Norme NF ISO 5725-1 : Exactitude (justesse et fidélité) des résultats et méthodes de mesure. Partie 1 : Principes généraux et définitions.
- AFNOR (1994c) Norme NF ISO 5725-2 : Exactitude (justesse et fidélité) des résultats et méthodes de mesure. Partie 2 : Méthodes de base pour la détermination de la répétabilité et de la reproductibilité d'une méthode de mesure normalisée.
- AFNOR (1994d) Norme NF ISO 5725-4 : Exactitude (justesse et fidélité) des résultats et méthodes de mesure. Partie 4 : Méthodes de base pour la détermination de la justesse d'une méthode de mesure normalisée.
- AFNOR (1996) Métrologie dans l'entreprise : outil de la qualité.
- AFNOR (1998) Norme NF ISO 5725-5 : Exactitude (justesse et fidélité) des résultats et méthodes de mesure. Partie 5 : Méthodes alternatives pour la détermination de la fidélité d'une méthode de mesure normalisée.
- AFNOR (1999) Norme XP T90-210 : Protocole d'évaluation d'une méthode alternative d'analyse physico-chimique quantitative par rapport à une méthode de référence.
- Feinberg M. (2000) La validation des méthodes d'analyse. *Le cahier des Techniques de l'Inra*, 44, 19-49.
- Feinberg M. (2003) L'incertitude des mesures analytiques. *Le cahier des Techniques de l'Inra*, 50, 3-10
- Feinberg, M. (2006) Approche globale et harmonisée de la validation. *Spectra Analyse*, **249**, avril-mai 2006, 16-23.
- Feinberg M. (2010a) Principes et vocabulaire in Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial, 13-25
- Feinberg M. (2010b) Mise en œuvre du profil d'exactitude in Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial, 27-44

Feinberg M. (2010c) Interprétation du profil d'exactitude *in* Validation des méthodes d'analyse quantitative par le profil d'exactitude. *Le Cahier des Techniques de l'Inra*, numéro spécial, 45-59

ISO (1997) Guide 32 : Étalonnage en chimie analytique et utilisation de matériaux de référence certifiés.

ISO (2000) Guide 33 : Utilisation des matériaux de référence certifiés.

Taylor B.N. et Kuyatt C.E. (1994) Guidelines for Evaluating and expressing the Uncertainty of NIST Measurement Results, *NIST Technical Note 1297, 1-20*



Conclusion

François Jacquet¹

La validation des méthodes d'analyse quantitative physico-chimique utilisées pour produire des mesures justes et fiables est une préoccupation constante des analystes et des destinataires des résultats des analyses, car elle permet de renforcer la confiance dans la qualité de ces résultats mesurables en assurant leur fiabilité.

La validation des méthodes est un deuxième pilier, complémentaire à la maîtrise des instruments et des équipements, pour atteindre l'objectif de validité des résultats mesurables qui est l'un des objectifs de la politique qualité de l'Inra. La réflexion est en cours pour l'intégration de la validation des méthodes d'analyse quantitative dans une future version du référentiel *qualité Inra*.

Afin de décrire et d'expliquer cette démarche de validation des méthodes dans les laboratoires, deux écoles des techniques « Validation des méthodes d'analyse quantitative physico-chimique », en 2007 et 2008, ont rassemblé plus de cent participants.

Je remercie tous les collègues qui ont contribué à ces écoles, en les organisant et / ou en y intervenant : Sandra Arrault, Olivier Berdeaux, Laurent Debrauwer, Marjolaine Deschamps, Martine Dupupet, Denys Durand, Max Feinberg, Christophe Garin, Annie Guerin, Guy Jamet (DGCCRF), Anne Jaulin, Michel Laurentie (AFSSA), Jérôme Lazzarotto, Patricia Manzanno, Marie-Andrée Piedallu, Thierry Prunet et Agnès Thomas. Ces écoles ont été soutenues par les Départements ALIMH, CEPIA, EA et PHASE, la Formation permanente nationale et la Mission qualité / Délégation à la qualité.

¹ Délégué à la qualité de l'INRA - F- 75338 Paris cedex 03 ✉ francois.jacquet@paris.inra.fr

Plusieurs approches, normes, méthodologies sont décrites pour la validation des méthodes. Le profil d'exactitude est l'approche, tout récemment intégrée dans une norme (V03-110:2010), que nos experts et praticiens ont jugée la mieux adaptée au contexte des laboratoires de recherche.

Ce numéro du *Cahier des Techniques de l'Inra* sera une aide précieuse pour les équipes dans la mise en œuvre de la validation de leurs méthodes d'analyse par le profil d'exactitude.

Je remercie les auteurs des articles de ce numéro du *Cahier des Techniques*, et plus particulièrement Max Feinberg et Michel Laurentie (AFSSA) qui ont mis leur expertise au service de la relecture des articles, ainsi que Marie Huyez-Levrat, cheville ouvrière du *Cahier des Techniques de l'Inra*, et la Délégation au partenariat avec les entreprises de la Direction de la valorisation de l'Inra.

La Délégation à la qualité soutient les actions des départements de recherche qui vont favoriser et accompagner cette mise en œuvre, en particulier par des formations, et contribuera à toute autre action utile pour la généralisation de la validation des méthodes d'analyse :

- mise en application de la validation des méthodes d'analyse qualitative physico-chimique telle que présentée dans ce cahier ;
- adaptation de ces méthodes à d'autres natures d'analyse quantitative ;
- travaux en vue de la construction d'outils et méthodes de validation de méthodes d'analyse « semi-quantitatives » ou qualitatives.

Pour la mise en œuvre, vous pourrez aussi faire appel, via la liste de diffusion validation-analytique-et-maitrise :

<http://listes.inra.fr/www/info/validation-analytique-et-maitrise>²

et, aux experts et praticiens de la validation des méthodes, que je remercie de poursuivre ainsi leur implication dans la diffusion de ces méthodes, et de contribuer à l'amélioration de la garantie de la fiabilité des résultats sur lesquels nombre de publications se fondent.

Avec tous mes encouragements et vœux de réussite pour vos mises en application dans vos laboratoires.

² Pour s'y abonner, contacter Christophe.Garin@paris.inra.fr

Index des auteurs

Nom	Prénom	page
Bourdat-Deschamps	Marjolaine	85
Bourrié	Guilhem	117,
Ciesielki	Henri	111
Durand	Denys	73
Feinberg	Max	5, 13, 27, 45
Guerin	Annie	111
Laurentie	Michel	61
Lazzarotto	Jérôme	99, 107
Maherou	Jennifer	99
Mayor	Monique	117
Ortigue-Marty	Isabelle	73

Index des articles

page

3

Introduction

5 *Témoignages de directeurs d'unité sur les écoles techniques*

11

Chapitre 1 – Validation des méthodes

13 Principe et vocabulaire, *Max Feinberg*

27 Mise en œuvre du profil d'exactitude, *Max Feinberg*

45 Interprétation du profil d'exactitude, *Max Feinberg*

61 Complément sur la quantification : pondération et transformation, *Michel Laurentie*

71

Chapitre 2 – Application dans les laboratoires à l'Inra

73 Enjeux de la validation pour l'expérimentation en nutrition animale, *Agnès Thomas, Denys Durand et Isabelle Ortigues-Marty*

85 Validation de l'analyse des HAP dans les eaux de pluie par la méthode du profil d'exactitude, *Marjolaine Bourdat-Deschamps*

99 Validation de la méthode de dosage du pigment chlorophylle a dans l'eau douce par extraction et par HPLC-DAD, *Jérôme Lazzarotto et Jennifer Maherou*

107 Validation d'une méthode automatisée de titration de l'alcalimétrie complète dans l'eau douce, *Jérôme Lazzarotto*

111 Méthode d'approche globale de l'estimation des incertitudes de mesure, *Annie Guerin et Henri Ciesielki*

117 Validation d'une méthode de dosage des anions fluorure, chlorure, nitrite, bromure, nitrate, phosphate et sulfate par chromatographie ionique, *Monique Mayor et Guilhem Bourrié*

Conclusion



Numéro spécial

Année 2010

Validation des méthodes d'analyse quantitative par le profil d'exactitude

Sommaire

Introduction	p. 03
Témoignages sur les écoles techniques	p. 05
Chapitre I : Validation des méthodes	p. 11
Chapitre II : Application dans les laboratoires à l'Inra	p. 71
Conclusion	p. 135
Index des auteurs	p. 137
Index des articles	p. 139