ARTICLES

Exercices d'intercomparaison pour la mesure du dioxyde d'azote dans l'air ambiant par tubes à diffusion

Intercomparison exercises for the monitoring of ambient nitrogen dioxide using diffusive samplers

Michel GERBOLES*, Daniela BUZICA*, Lucio AMANTINI*,
Hervé PLAISANCE**, Hélène MARFAING***

Résumé

Cette étude présente un essai d'aptitude pour les Associations agréées de surveillance de la qualité de l'air effectuant la mesure du NO_2 par échantillonneur à diffusion en France. L'évaluation des données montre que les participants obtiennent de bons résultats pour les expositions sur sites. Pour les essais sur tubes dopés et matériaux de référence, environ la moitié des participants obtiennent de bons résultats alors que l'objectif de qualité était aussi plus ambitieux. L'intercomparaison a également été planifiée pour déterminer la répétabilité et la reproductibilité de la méthode de mesure. Pour les essais sur site, la répétabilité diminue avec le niveau de concentration mesurée, passant d'environ 20 à 10 % quand les concentrations évoluent de 15 à 40 µg/m³. En revanche, la reproductibilité semble stable à 45 %. La répétabilité (10 %) et la reproductibilité (20 %) obtenues pour les essais sur les tubes dopés et les matériaux de référence sont inférieures aux répétabilité et reproductibilité des essais sur site.

La reproductibilité est toujours supérieure à la répétabilité, ce qui suggère la présence d'erreurs systématiques dans la préparation et/ou l'analyse des échantillonneurs. De plus, le contrôle qualité efffectué sur la méthode d'analyse montre que la valeur du blanc, l'écart de linéarité et la variation de la pente d'étalonnage peuvent être diminués.

Globalement, les résultats de cette étude apparaissent plus précis que ceux qui sont obtenus dans le cadre du programme de contrôle qualité WASP et du réseau national NO_2 du Royaume-Uni. Cette étude montre également une amélioration des résultats par rapport à l'intercomparaison 1999 réalisée par l'École des mines de Douai.

Le résultat le plus notable reste que la médiane des valeurs moyennes de tous les laboratoires pour les essais sur site est en accord parfait avec les mesures des analyseurs par chimiluminescence. De plus, les nouveaux matériaux de référence préparés par l'ERLAP en exposition sous atmosphère contrôlée, ont démontré leur validité pour l'utilisation dans ce type d'intercomparaison

** Centre Commun de Recherche, Institut de l'Environnement Durable,Via E. Fermi, 21020 Ispra, Italie. michel.gerboles@jrc.it*

*** École des Mines de Douai, Département Chimie et Environnement, 941, rue Charles Bourseul – BP 838 – 59508 Douai Cedex.*

**** AIRPARIF, 7 rue Crillon, 75004 Paris.*

Abstract

This study presents the results of a proficiency testing by intercomparison exercise for NO_2 measured using diffusive samplers by the French air quality monitoring networks. The participants obtained good results for the field exposition. In the case of spiked tubes and reference materials, half of the participants obtained good results while the objective was more ambitious.

The aim of the intercomparison exercise was also to determine the repeatability and reproducibility of the method. For the field experiments, the repeatability decreased from 20 to 10% when the NO_2 concentration increased from 15 to 40 µg/m^3, while the reproducibility was stable. The reproducibility was around 45%. For the experiments on spiked tubes and reference materials, the repeatability (10%) and reproducibility (20%) were smaller than those obtained for the field expositions.

The reproducibility was always higher than the repeatability, indicating the presence of systematic errors during the preparation and/or the analysis. Moreover, the quality control of the analytical procedure showed that the blank, linearity error and slope of the calibration line could be improved.

The results of this intercomparison are more precise than those obtained with the WASP quality control and NO_2 national network of United Kingdom. This study also shows an improvement of the results compared with those obtained by the intercomparison study organised by the École des mines de Douai, France, in 1999.

The most important result is that the median of all laboratories for the field expositions is in good agreement with the measurements of the chemiluminiscence analyser. Moreover, the new reference materials, which were prepared by ERLAP by exposition in a chamber under controlled atmosphere, showed their validity for this type of intercomparison.

Introduction

En 2000, environ 2 000 sites d'échantillonnage ont été étudiés par les Associations agréées de surveillance de la qualité de l'air (AASQA) pour effectuer des mesures du dioxyde d'azote (NO_2) au moyen d'échantillonneurs à diffusion. L'échantillonneur le plus utilisé était le tube de Palmes fabriqué par la compagnie Gradko International LTD (UK) et l'échantillonneur fabriqué par la compagnie Passam ag. (Suisse). En 2001, environ 1 500 sites d'échantillonnage ont encore été équipés pour la surveillance du NO_2. Les AASQA utilisant cette technique de mesure étaient au nombre de 34. Dans l'Union européenne, le Royaume-Uni est un autre État membre où les tubes de Palmes sont utilisés à grande échelle. La surveillance du NO_2 y est réalisée dans le cadre d'un réseau national qui est coordonné par AEA Technology au nom du *Department of the Environment, Transport and the Regions* (DETR), et implique environ 300 autorités locales, et plus de 1 200 sites de contrôle [1].

En 1999, l'École des mines de Douai (EMD) a réalisé un exercice d'intercomparaison [2] sur site pour évaluer la qualité des mesures du NO_2 obtenues par tubes à diffusion. Les conclusions principales de cette étude furent que la préparation des échantillonneurs était la principale source d'erreur de la mesure du NO_2 par cette technique et qu'il était nécessaire d'améliorer la technique de préparation des échantillonneurs, en particulier le nettoyage et l'injection dans le tube de la solution d'absorption. En 2001, toujours pour améliorer les mesures réalisées à l'aide des échantillonneurs diffusifs, le Groupe de travail national sur les échantillonneurs à diffusion (GT) a rédigé un protocole harmonisé [3] d'analyse, de préparation et de traitement des données.

Deux années après la publication du protocole harmonisé et trois années après la réalisation de la première intercomparaison, les réseaux de mesure qui utilisent les tubes de Palmes ont émis le désir d'évaluer de nouveau la qualité des mesures du NO_2 réalisées au moyen de ces tubes. La coordinatrice du GT, Mme Fromage-Mariette, a proposé au Laboratoire de référence européen de la pollution de l'air (ERLAP) du Centre commun de recherche d'Ispra, en Italie, d'organiser une série d'exercices d'intercomparaison de mesures du NO_2 par tubes à diffusion pour les AASQA.

L'objectif de cette étude était d'organiser un essai d'aptitude par essais d'intercomparaison [4, 5]. La participation à de tels exercices d'intercomparaison est une exigence du référentiel d'accréditation ISO 17025 pour les laboratoires d'essai [6]. En même temps, la majorité des laboratoires participants utilisant un protocole analytique harmonisé (celui du GT) ou un protocole sensiblement équivalent, l'intercomparaison a aussi été planifiée pour déterminer la répétabilité et la reproductibilité [7] de la méthode. L'évolution de la qualité des mesures depuis l'intercomparaison réalisée en 1999 par l'EMD a été étudiée. Des comparaisons avec les résultats du système de contrôle qualité du Royaume-Uni WASP [8] et du réseau national de surveillance du NO_2 par tube de Palmes au Royaume-Uni [9] sont aussi présentées.

De façon à mettre en évidence différentes sources d'erreur dans la mesure du NO_2, l'étude a été divisée en plusieurs séries d'échantillonneurs à analyser par les laboratoires participants :

- une série d'échantillonneurs préparés par chaque participant, dopés en nitrite par l'ERLAP ;
- une série de matériaux de référence préparés par l'ERLAP en chambre d'exposition sous atmosphère contrôlée en NO_2 ;
- enfin deux séries d'échantillonneurs préparés par chaque participant et exposées sur le terrain en deux sites différents.

Finalement, l'influence du transport par avion des tubes de Palmes et l'effet de la boîte de protection sur la réponse des échantillonneurs ont été évalués.

Abréviations

- AASQA : Association agréée de surveillance de la qualité de l'air.
- AIRPARIF : Association interdépartementale pour la gestion du réseau automatique de surveillance de la pollution atmosphérique et d'alerte en région Ile-de-France.
- CERMATAIR : matériaux de référence certifiés pour la mesure de polluants gazeux dans l'air ambiant, *Certified reference materials for the measurement of gaseous pollutants in ambient air.*
- CV : coefficient de variation.
- DETR : *Department of the Environment, Transport and the Regions.*
- EMD : École des mines de Douai.
- ERLAP : Laboratoire de référence européen de la pollution de l'air. *European Reference Laboratory of Air Pollution.*
- GT : groupe de travail national sur les échantillonneurs à diffusion.
- GUM : guide pour l'expression de l'incertitude de mesure. *Guide to the expression of Uncertainty in Measurement.*
- HSL : *Health and Safety Laboratory.*
- ISO : Organisation internationale de normalisation.
- NIST : *National Institute of Standards and Technologies.*
- NO_2 : dioxyde d'azote.
- r : répétabilité.
- R : reproductibilité.
- RM : matériaux de référence.
- σt : écart type cible.
- TEA : triéthanolamine.
- WASP : *Workplace analysis scheme for proficiency.*

Laboratoires participants, type d'échantillonneur et organisation des essais d'intercomparaison

Sept des laboratoires membres du GT, l'EMD et l'ERLAP ont participé aux exercices d'intercomparaison (Tableau 1). À part AIRPARIF qui a utilisé le tube fabriqué par la société Passam ag., les autres laboratoires ont préparé des tubes à diffusion de type Palmes, fabriqués par la société Gradko International Ltd. mais en injectant eux-mêmes la solution absorbante. Le tube Passam, en revanche, est commercialisé déjà prêt à l'emploi. Il est à noter que le matériel distribué par Gradko a changé à trois reprises au cours des 10 dernières années. Aujourd'hui, le tube original, fabriqué jusqu'au milieu des années 1990 a totalement disparu du parc d'utilisation. Par contre, les tubes de type 2 (distribués depuis le milieu des années 1990 et jusqu'en 2000 environ), et les tubes de type 3 (distribués en 2002) et 4 (distribués depuis 2002) coexistent. Le tube de type 3 a été retiré de la distribution en raison d'un défaut sur de nombreuses capsules colorées (fentes), il doit être évité. La capsule et le tube de type 4 (identifiés comme XDIF…) assurent les meilleures performances mécaniques lors de l'utilisation.

Tableau 1.
Laboratoires participant à l'intercomparaison, échantillonneurs et protocole analytique utilisés.
Participating laboratories, implemented diffusive sampler and analytical protocol.

Laboratoires	Échantillonneur / Protocole analytique
AIRLOR, Association pour la surveillance et l'étude de la pollution atmosphérique en Lorraine, Vandœuvre-lès-Nancy (Nancy)	Tube Gradko, protocole harmonisé
AIR LR, Air Languedoc-Roussillon, Montpellier	Tube Gradko, protocole harmonisé
AIRPARIF, Surveillance de la qualité de l'air en Ile-de-France, Paris	Tube Passam, protocole harmonisé
ATMO CHAMPAGNE-ARDENNE, Association régionale pour la surveillance de la qualité de l'air en Champagne-Ardenne, Reims	Tube Gradko, protocole harmonisé
ATMO PICARDIE, Association pour la surveillance de la qualité de l'air en Picardie, Amiens	Tube Gradko, protocole harmonisé
ATMO POITOU-CHARENTES, Association régionale pour la mesure de la qualité de l'air en Poitou-Charentes, Périgny	Tube Gradko, protocole harmonisé
École des mines de Douai, Département Chimie et Environnement, Douai	Tube Gradko, chromatographie ionique
European Reference Laboratory for Air Pollution, Health and Emission Unit – Institute for Environment and Sustainability, Ispra, Italie	Tube Gradko, colorimétrie et chromatographie ionique
ORA, Observatoire Réunionnais de l'Air, Sainte-Clotilde, La Réunion	Tube Gradko, protocole harmonisé

Pour la série d'échantillonneurs dopés en nitrite par l'ERLAP, chaque participant a préparé neuf tubes à diffusion, incluant trois blancs et six échantillonneurs à doper. Les tubes dopés ont été préparés par gravimétrie. Après injection d'une solution aqueuse de 20 µl de nitrite correspondant à une masse de 0,640 µg, l'ERLAP a renvoyé par courrier express ces tubes aux participants pour analyse. L'évaluation suivant la méthode du « Guide pour l'expression de l'incertitude de mesure [10] (GUM) » de la masse de nitrite injectée a donné une estimation de l'incertitude élargie égale à ± 0,0034 µg, avec un facteur de couverture k de 2. Cette évaluation tient compte de l'incertitude sur l'étalon de nitrite utilisé pour l'étalonnage, celui-ci étant traçable à l'étalon du *National Institute of Standards and Technologies* (NIST), et des incertitudes sur le volume des solutions évalué par gravimétrie, ainsi que de l'influence de la dérive en température sur la densité des solutions et de l'incertitude sur le volume de la verrerie de laboratoire. Le laboratoire ATMO-Picardie a préparé 18 échantillonneurs, incluant une série de neuf échantillonneurs devant faire un aller-retour entre Ispra et Amiens avec analyse finale réalisée par l'ERLAP de façon à évaluer l'effet du transport par avion sur la réponse des échantillonneurs.

La préparation des matériaux de référence (RM) est décrite dans le cadre du projet européen CERMATAIR [11]. L'exposition de ces RM a été réalisée en chambre d'exposition par l'ERLAP. Les matériaux de référence sont des tubes de Palmes sur lesquels une membrane est placée à l'extrémité normalement ouverte pour stabiliser le parcours de diffusion. La vitesse de vent était limitée pour éviter les turbulences. En même temps, la température et l'humidité relative étaient régulées avec précision. La masse de nitrite contenue dans les RM était de 0,83 µg avec une incertitude élargie de 3 %. Après exposition, l'ERLAP a envoyé par courrier rapide ces tubes aux participants pour analyse.

Deux stations de surveillance du réseau d'AIRPARIF disposant d'un analyseur de NO_2 par chimiluminescence (un site rural « Fontainebleau » et un site urbain « Gennevilliers ») ont été utilisées en période hivernale pour les expositions sur le terrain. Chaque laboratoire a préparé deux séries de six échantillonneurs qui ont été exposés dans leur boîte de protection (de type Passam) par AIRPARIF dans les deux stations de surveillance. Une série d'échantillonneurs de l'ERLAP a été placée dans une boîte de protection et une seconde série a été placée hors de la boîte pour évaluer la différence entre les deux types d'échantillonnage. Pendant la durée de l'exposition, un échantillonneur « blanc » était conservé au réfrigérateur pour soustraire le blanc analytique. AIRPARIF a ensuite envoyé l'ensemble des échantillonneurs à l'ERLAP qui les a redistribués à chaque laboratoire pour analyse. Tous les transports ont été effectués par courrier rapide en avion. Trois vérifications de l'étalonnage effectuées sur les analyseurs durant la période d'exposition ont montré que leur dérive était au maximum de 1,5 %. Durant la période d'exposition la température moyenne était de 10 °C et l'humidité relative était de 75 % environ. Les pertes dans les lignes d'échantillonnage des stations ont été évaluées à moins de 2 %. Les éventuelles erreurs de mesure des analyseurs, dues à une sous-estimation du rendement de conversion interne du NO_2 en NO ont été corrigées par calcul dans les analyseurs. AIRPARIF est accrédité suivant la norme ISO 17025 pour la mesure sur site du NO_2 par chimiluminescence.

Résultats et traitement statistique

L'ensemble des résultats des essais est présenté dans les figures 1 à 4. La valeur moyenne des tubes « blancs » a systématiquement été soustraite des mesures présentées. Les figures sont divisées en deux parties : d'une part les moyennes des six tubes analysés par laboratoire avec leur écart type représenté par une barre d'erreur et d'autre part les valeurs du z-score par participant. Sur les graphes des moyennes, la ligne continue indique la médiane des moyennes.

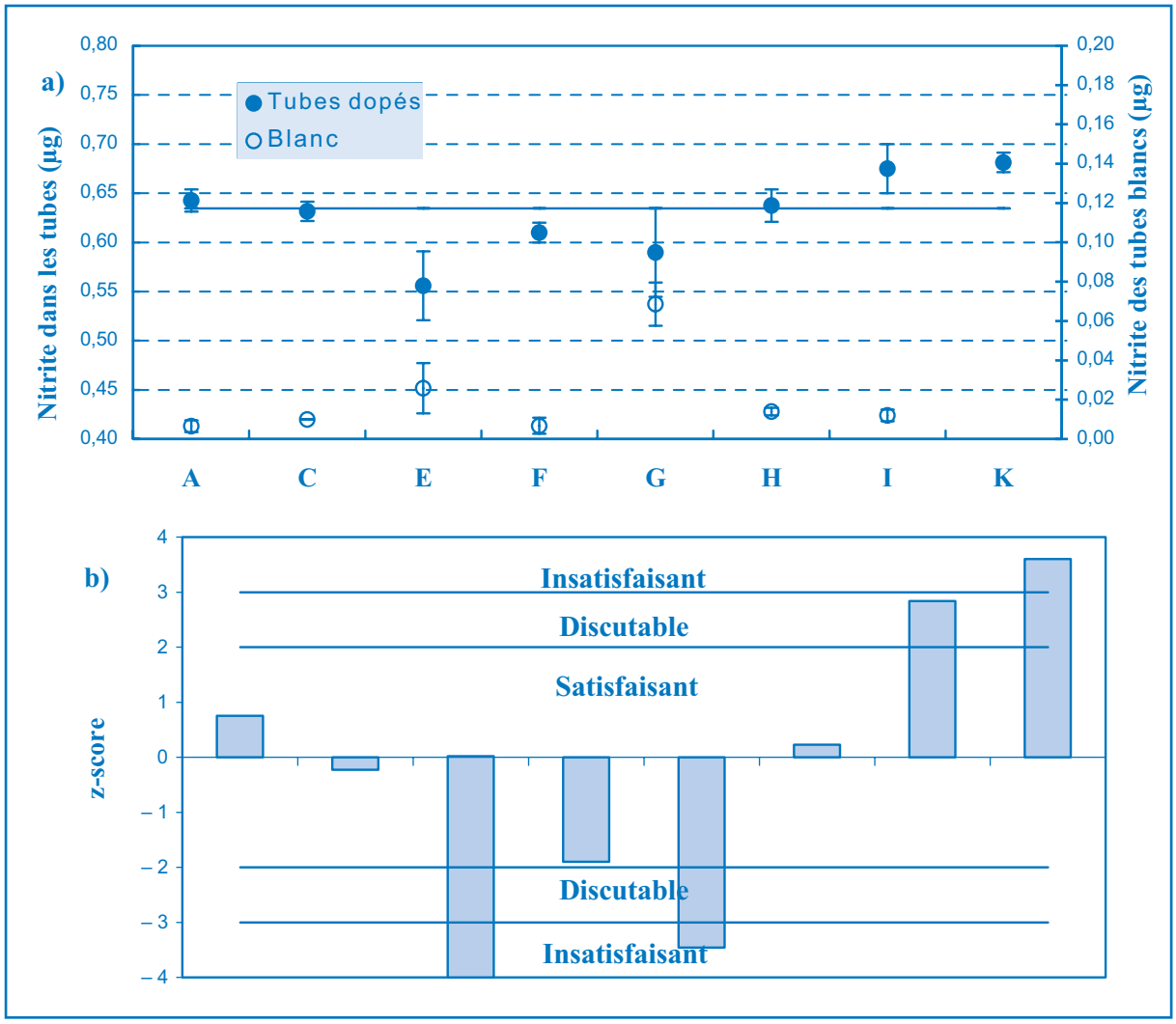

Figure 1.

Essai sur les tubes dopés.
a) les ronds bleus représentent les moyennes des six tubes analysés par chaque laboratoire, les ronds blancs représentant les valeurs du blanc avec leur écart type indiqué par les barres d'erreur. La ligne continue indique la médiane des moyennes ;
b) valeurs du z-score par participant calculées en prenant la médiane des moyennes comme valeur vraie consensuelle et un écart type cible de 5 %.

Test on spike tubes.
a) the blue circles show the averages of six replicates samplers *versus* participating laboratories, the white circles being the value of blank. The errors bars indicate standard deviations. The solid line shows the median of averages;
b) z-score values *versus* participants computed using the median of averages as true value and a target standard deviation of 5%.

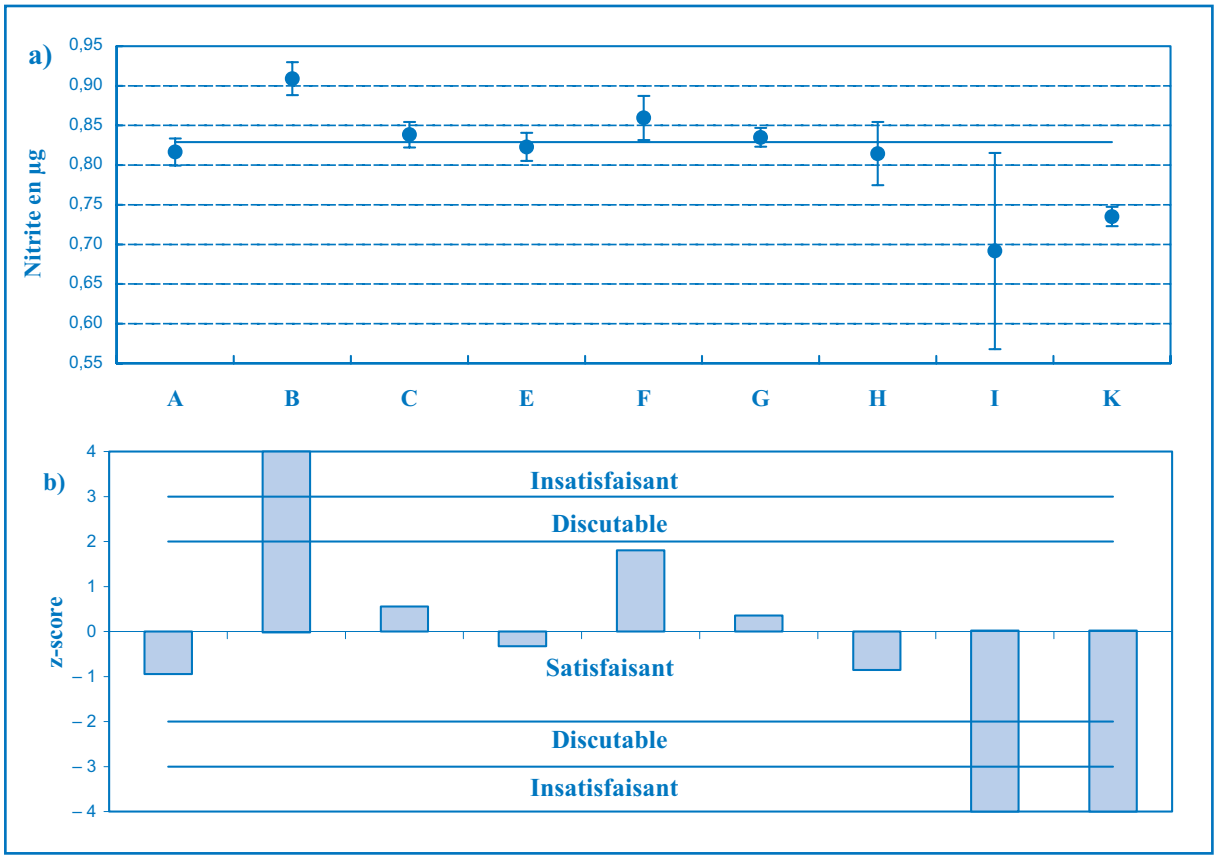

Figure 2.

Essai sur les matériaux de référence.
a) les ronds bleus représentent les moyennes des six tubes analysés par chaque laboratoire avec leur écart type indiqué par les barres d'erreur. La ligne continue indique la médiane des moyennes ;
b) valeurs du z-score par participant calculées en prenant la médiane des moyennes comme valeur vraie consensuelle et un écart type cible de 5 %.

Test on reference materials.
a) the blue circles show the averages of six replicates samplers *versus* participating laboratories. The errors bars indicate standard deviations. The solid line shows the median of averages;
b) z-score values *versus* participants computed using the median of averages as true value and a target standard deviation of 5%.

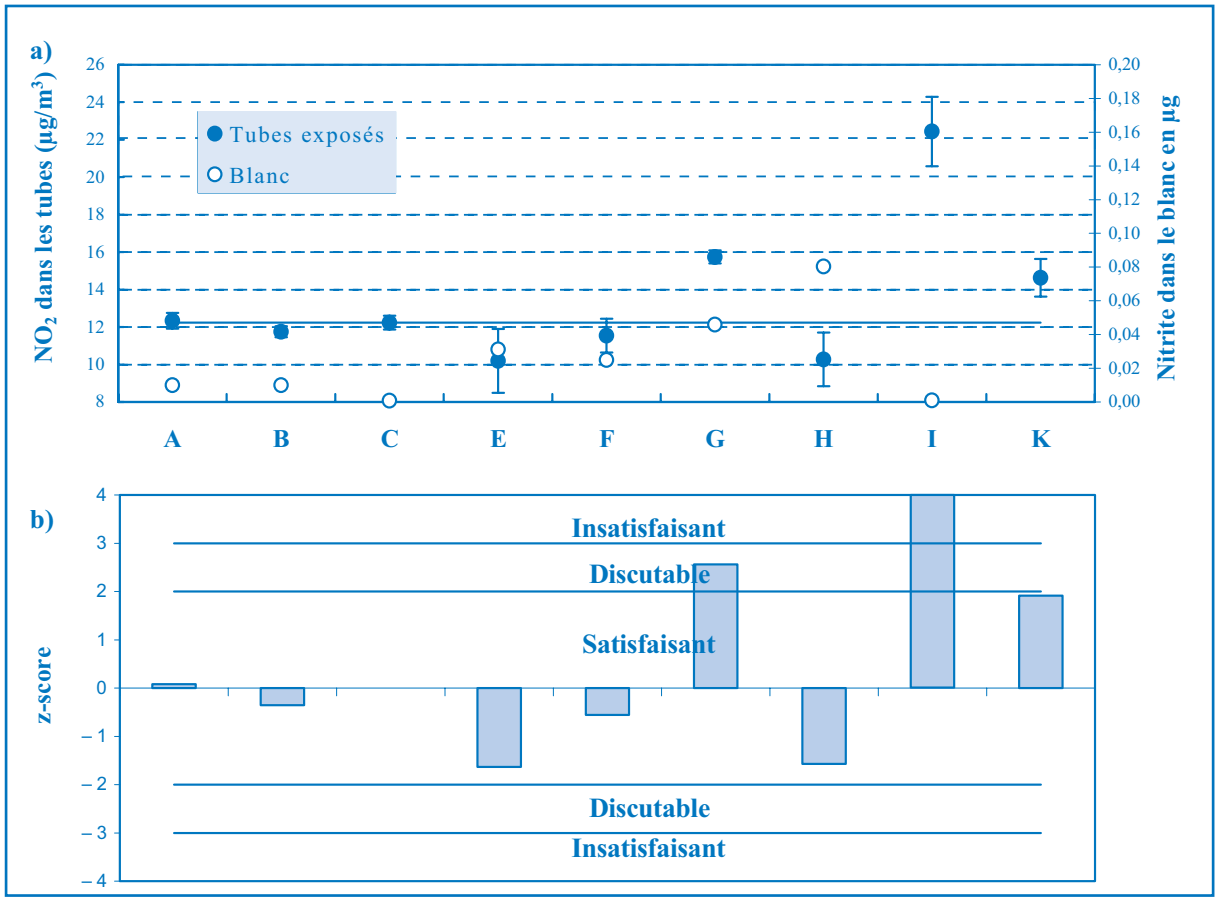

Figure 3.

Essai sur site rural à Fontainebleau.
a) les ronds bleus représentent les moyennes des six tubes analysés par chaque laboratoire, les ronds blancs représentant les valeurs du blanc avec leur écart type indiqué par les barres d'erreur. La ligne continue indique la médiane des moyennes ;
b) valeurs du z-score par participant calculées en prenant la médiane des moyennes comme valeur vraie consensuelle et un écart type cible de 25 %.

Field test at Fontainebleau (rural site).
a) the blue circles show the averages of six replicates samplers *versus* participating laboratories, the white circles being the value of blank. The errors bars indicate standard deviations. The solid line shows the median of averages;
b) z-score values *versus* participants computed using the median of averages as true value and a target standard deviation of 25%.

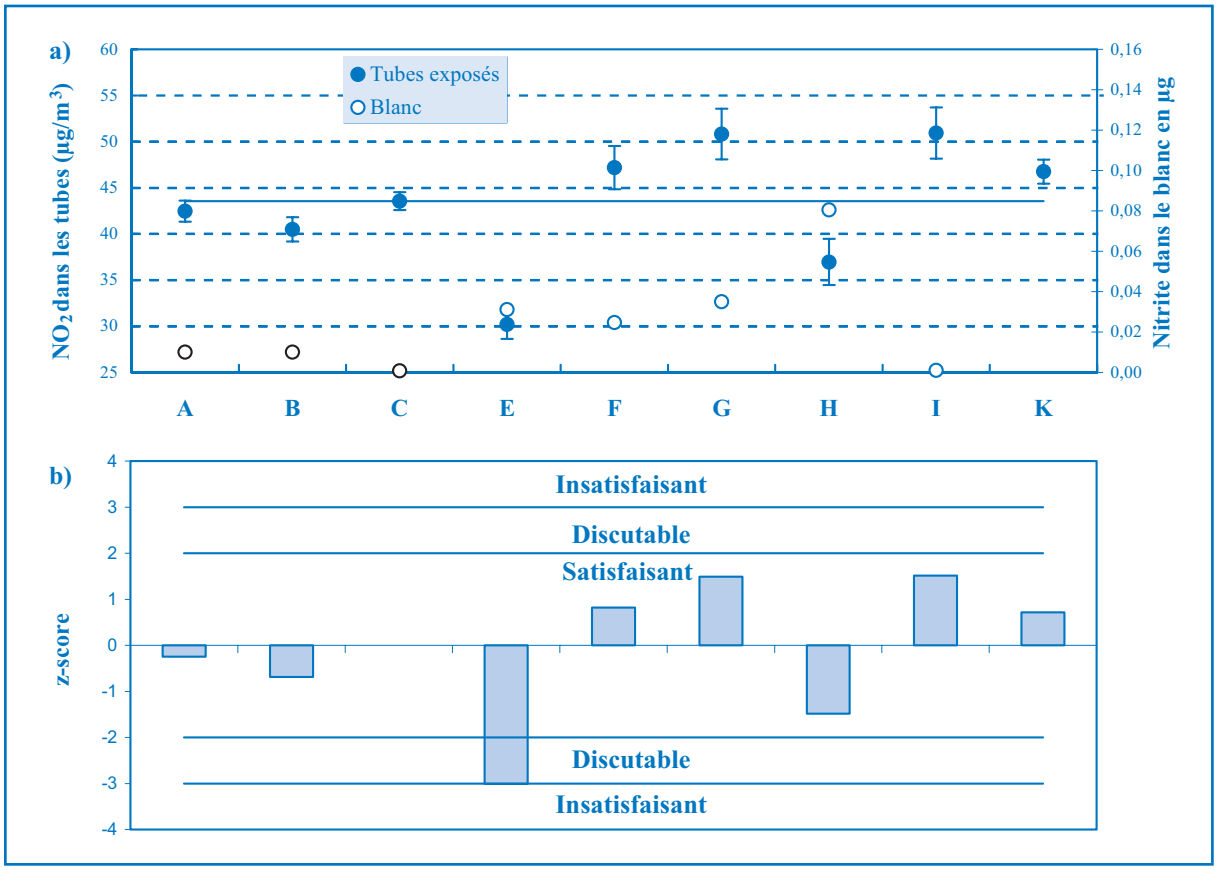

Figure 4.

Essai sur site urbain à Gennevilliers.
a) les ronds bleus représentent les moyennes des six tubes analysés par chaque laboratoire, les ronds blancs représentant les valeurs du blanc avec leur écart type indiqué par les barres d'erreur. La ligne continue indique la médiane des moyennes ;
b) valeurs du z-score par participant calculées en prenant la médiane des moyennes comme valeur vraie consensuelle et un écart type cible de 25 %.

Field test at Gennevilliers (urban site).
a) the blue circles show the averages of six replicates samplers *versus* participating laboratories, the white circles being the value of blank. The errors bars indicate standard deviations. The solid line shows the median of averages;
b) z-score values *versus* participants computed using the median of averages as true value and a target standard deviation of 25%.

Essai d'aptitude

Les essais d'aptitude sont évalués en utilisant la méthode du z-score. Le score z de chaque laboratoire est calculé en appliquant la formule suivante :

$$z = \frac{x - X}{\sigma_t / \sqrt{n}}$$

où : x est la moyenne de chaque laboratoire pour un essai donné,

X est la valeur vraie consensuelle de l'inter-comparaison,

σ_t est la valeur cible pour l'écart type,

n est le nombre de mesures par essai.

Le but de chaque participant doit être d'obtenir un score compris entre – 2 et + 2 et en tout cas jamais supérieur à 3 ou inférieur à – 3. Pour la méthode du z-score, il a été consensuellement décidé de choisir la médiane des moyennes de tous les laboratoires comme valeur vraie de référence. Pour les inter-comparaisons sur les échantillonneurs dopés et avec matériaux de référence, l'écart relatif cible a été fixé à 5 % de la valeur de référence. Pour les tubes exposés sur site, l'écart relatif cible a été fixé à 25 % de la valeur de référence conformément à l'objectif qualité pour les mesures indicatives de NO_2 de la première Directive fille européenne [12].

Étude de la répétabilité et de la reproductibilité

L'ensemble de la procédure de traitement des données de la norme ISO 5725 [7] a été mise en œuvre pour le calcul de la répétabilité et de la reproductibilité de la méthode de mesure. En particulier, le traitement des aberrants nous a amenés à éliminer certaines mesures ou laboratoires dans l'application directe des tests pour aberrants. Les tests pour la

recherche d'aberrants sont appliqués de nouveau sur les données restantes après avoir éliminé une série de données et ainsi de suite.

L'observation des graphiques de Mandel k (Figure 5) suggère la présence de données aberrantes concernant la dispersion des mesures obtenues par les laboratoires G, H et I. Ces observations sont complétées par l'application du test de Cochran [7] (Tableau 2) qui conduit à éliminer les valeurs aberrantes suivantes :

- les mesures du laboratoire I sont retirées pour l'essai sur les matériaux de référence ; une valeur de la série de tubes dopés est exclue ;
- une valeur du laboratoire H pour l'essai sur les tubes dopés et une pour l'exposition sur site de fond ont été éliminées. Après avoir éliminé les données du laboratoire I, la série de mesures sur les materiaux de référence est aberrante au seuil de probabilité de 95 % mais pas au seuil de 99 %. N'ayant pas de raison de l'éliminer, cette série est conservée ;
- les mesures du laboratoire G de la série de tubes dopés sont éliminées ;
- pour le laboratoire E, les mesures de la série de tubes dopés sont éliminées suite à l'élimination des valeurs des laboratoires G et H de cette intercomparaison.

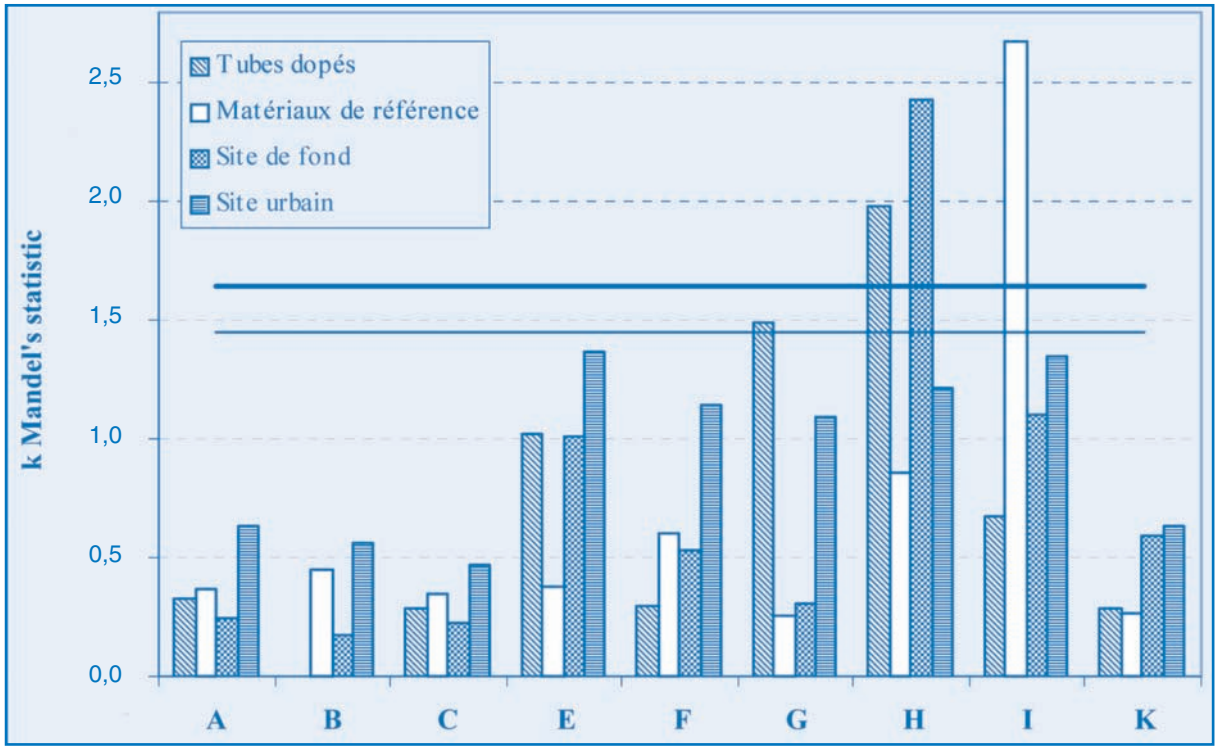

Figure 5.

Statistique k de Mandel par laboratoire et par essai d'intercomparaison avec la valeur critique pour la statistique k au seuil de 99 % (ligne en gras) et la valeur critique au seuil de 95 % (ligne continue).

Mandel's k statistic *versus* laboratories and intercomparison exercices. The bold line shows the critic value for k with 99% probability while the solid line shows the one with 95% probability.

Tableau 2.

Test de Cochran pour la recherche de valeurs aberrantes (notées en gras) dans la dispersion des données.

Cochran test for outliers showing too large data scattering. Values in bold font indicate outliers.

Laboratoires	Tubes dopés	Matériaux de référence	Site de fond	Site urbain
A	0,013	0,015	0,007	0,044
B		0,023	0,003	0,035
C	0,010	0,013	0,005	0,025
E	0,130	0,016	0,113	0,207
F	0,011	0,040	0,032	0,145
G	0,279	0,007	0,011	0,132
H	**0,490**	0,083	**0,655**	0,164
I	0,057	**0,795**	0,134	0,203
K	0,010	0,008	0,039	0,045
Valeur critique à 99 %	0,423	0,387	0,387	0,387
Valeur critique à 95 %	0,360	0,329	0,329	0,329

Le graphique de Mandel h (Figure 6) indique, lui, la présence d'aberrants en ce qui concerne les moyennes des laboratoires B, E, I, et K. Ces observations sont vérifiées par l'application d'une part du test de Grubb sur les moyennes par laboratoire et d'autre part du même test sur les mesures maximales et minimales de chaque laboratoire (Tableaux 3 à 5). Ces tests de Grubb conduisent à éliminer les données suivantes :

- une valeur du laboratoire A pour la série de tubes dopés et une valeur de l'exposition à Paris en site urbain ;
- la valeur maximale pour l'essai sur les tubes dopés du laboratoire H a déjà été éliminée suite au test de Cochran ;
- à la suite à l'élimination des valeurs des deux points précédents, l'ensemble des données du laboratoire K pour la série de tubes dopés est alors éliminé (car le test de Grubb sur les moyennes donne alors une valeur de 2,2).

La série de mesures du laboratoire I pour l'exposition en site de fond n'est pas tout à fait une valeur

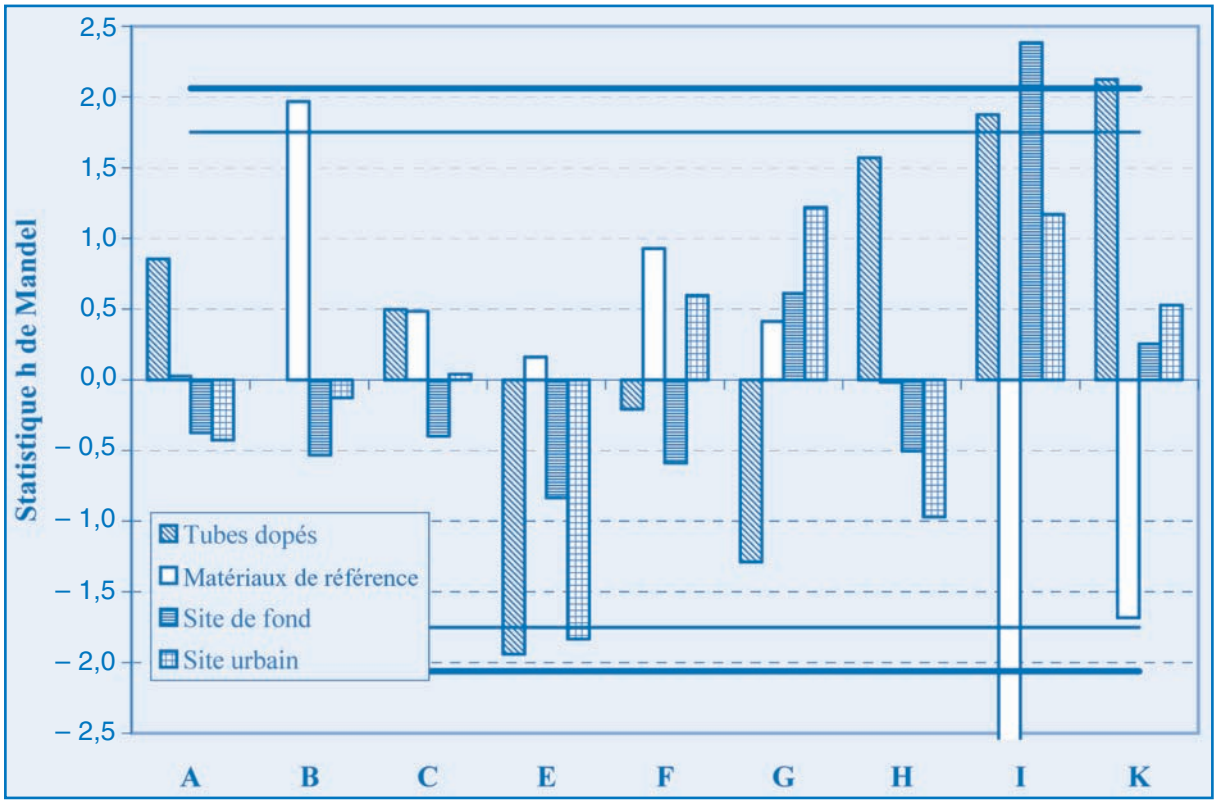

Figure 6.

Statistique h de Mandel par laboratoire et par essai d'intercomparaison avec la valeur critique pour la statistique h au seuil de 99 % (ligne en gras) et la valeur critique au seuil de 95 % (ligne continue).

Mandel's h statistic versus laboratories and intercomparison exercices. The bold line shows the critic value for h with 99% probability while the solid line shows the one with 95% probability.

Tableau 3.

Test de Grubb pour la recherche de valeurs aberrantes (notées en gras) dans les moyennes par laboratoire.

Grubbs test for outliers within laboratory averages. Values in bold font indicate outliers.

Laboratoires	Tubes dopés	Matériaux de référence	Site de fond	Site urbain
A	0,575	0,019	– 0,373	– 0,428
B		1,450	– 0,534	– 0,128
C	0,336	0,356	– 0,400	0,041
E	– 1,307	0,117	– 0,835	– 1,835
F	– 0,139	0,684	– 0,589	0,597
G	– 0,869	0,305	0,610	1,221
H	1,058	– 0,012	– 0,502	– 0,970
I	1,264	– 1,914	**2,384**	1,170
K	1,432	– 1,240	0,253	0,528
Valeur critique à 99 %	1,887	2,215	2,215	2,215
Valeur critique à 95 %	1,973	2,387	2,387	2,387

aberrante au seuil de probabilité de 99 %. Toutefois son influence sur la reproductibilité (qui passe de 85 à 45 % sans ces valeurs) est prépondérante. En même temps, la différence entre les mesures du laboratoire I et la valeur de référence de cet essai est trop importante et ne semble pas réaliste. Il a donc été décidé d'éliminer les valeurs du laboratoire I pour l'essai en site de fond.

Les estimations de la répétabilité et de la reproductibilité figurent dans le tableau 6.

Tableau 4.

Test de Grubb pour la recherche de valeurs aberrantes (notées en gras) dans les valeurs maximales de chaque série de mesures par laboratoire.

Grubbs test for outliers in the maximum of each laboratory series. Values in bold font indicate outliers.

Laboratoires	Tubes dopés	Matériaux de référence	Site de fond	Site urbain
A	0,75	2,23	1,72	0,62
B		1,58	1,04	1,80
C	0,85	1,35	1,63	1,25
E	1,88	1,52	1,82	1,84
F	1,24	1,10	1,60	1,45
G	0,70	1,45	1,46	1,61
H	**1,99**	1,24	**1,94**	1,70
I	1,55	1,04	1,37	1,10
K	0,848	1,713	1,64	1,31
Valeur critique à 99 %	1,887	1,887	1,887	1,887
Valeur critique à 95 %	1,973	1,973	1,973	1,973

Tableau 5.

Test de Grubb pour la recherche de valeurs aberrantes (notées en gras) dans les valeurs minimales de chaque série de mesures par laboratoire.

Grubbs test for outliers in the minimum of each laboratory series. Values in bold font indicate outliers.

Laboratoires	Tubes dopés	Matériaux de référence	Site de fond	Site urbain
A	**– 2,546**	– 0,774	– 1,190	**– 2,004**
B		– 1,242	– 1,193	– 0,999
C	– 1,187	– 1,144	– 0,816	– 0,975
E	– 1,618	– 1,122	– 0,772	– 0,929
F	– 1,831	– 1,168	– 0,809	– 1,341
G	– 1,301	– 1,134	– 0,902	– 1,004
H	– 0,652	– 1,165	– 0,929	– 0,896
I	– 1,057	– 1,628	– 1,263	– 1,111
K	– 1,187	– 0,967	– 1,376	– 1,079
Valeur critique à 99 %	– 1,887	– 1,887	– 1,887	– 1,887
Valeur critique à 95 %	– 1,973	– 1,973	– 1,973	– 1,973

Tableau 6.

Niveau de chaque essai avec la valeur absolue et relative de la répétabilité r, de la reproductibilité R et nombre de laboratoires par niveau de z-score.

Results of each test including the absolute and relative value of the repeatability, reproducibility and number of laboratories according to the z-score.

	Valeur	r	R	r %	R %	z < 2	2 < z < 3	z > 3
Tubes dopés (µg)	0,637	0,03	0,06 µg	4,7 %	9,5 %	4	1	3
Matériaux de référence	0,830	0,06	0,15 µg	7,5 %	17,8 %	6	1	1
Essai en site de fond (µg/m^3)	12,4	2,7	5,8 µg/m^3	22,2 %	47,1 %	7	1	1
Essai en site urbain (µg/m^3)	43,4	5,7	19,0 µg/m^3	13,1 %	43,9 %	8	1	0

Discussion

Tubes dopés

L'évaluation par la méthode du z-score laisse penser que les laboratoires peuvent obtenir un bon z-score (z < 2) avec un écart type cible fixé à 5 %. De plus, la moyenne du laboratoire G est affectée par un blanc trop élevé. Le blanc du laboratoire K est nul. Pour ces deux laboratoires, le score supérieur à 2 pourrait être expliqué par une valeur incorrecte du blanc. Après traitement des valeurs aberrantes, la répétabilité est de 5 % et la reproductibilité est de 10 % ; les résultats sont donc très satisfaisants.

Lors de la préparation, la solution de nitrite ne doit absolument pas être injectée sur les parois du tube mais sur les grilles imprégnées de triéthanolamine (TEA). L'EMD a observé une destruction partielle et une dispersion majeure des résultats d'analyse du nitrite lorsque celui-ci est injecté sur les parois du tube et après une conservation des échantillonneurs pendant une semaine au réfrigérateur. Dans ce cas, les taux de recouvrement du nitrite analysé passent de 99 ± 1 % à 97 ± 7 %. L'EMD et l'ERLAP ont vérifié qu'il ne s'agit pas d'une oxydation de nitrite en nitrate qui serait indétectable par colorimétrie. La façon dont le nitrite est injecté (sur les parois ou sur la TEA) peut être à l'origine d'une extrême dispersion couplée à un biais important avec la masse de nitrite déterminée par gravimétrie.

De plus, des tubes supplémentaires ayant fait un aller-retour ERLAP-ATMO Picardie-ERLAP en avion dans la soute, ne présentent pas de différence significative avant et après le transport (0,813 ± 0,016 µg puis 0,798 ± 0,008).

Essai sur les matériaux de référence

L'évaluation des données suivant la méthode du z-score montre qu'il est possible pour les laboratoires d'obtenir un bon score en se fixant un écart type cible de 5 %, comme pour les tubes dopés. Ici aussi la répétabilité et la reproductibilité sont satisfaisantes, 7,5 et 17,8 % respectivement ; d'autant plus que le processus de fabrication introduit une source de variabilité dont l'écart type relatif est évalué à environ 3 % [11].

Essai sur site

L'évaluation suivant la méthode du z-score montre que les participants peuvent obtenir un score satisfaisant en se fixant un écart type cible de 25 %. Pour les deux sites de mesure, l'indice de reproductibilité est trop élevé (R d'environ 50 %). En revanche, dans les deux cas la médiane des valeurs moyennes de tous les laboratoires est cohérente avec les mesures des analyseurs par chimiluminescence d'AIRPARIF. En effet, dans le site de fond, l'analyseur a donné 14,4 contre 12,2 µg/m³ pour les tubes de Palmes et dans le site urbain 44 contre 44 µg/m³.

Les résultats des mesures effectuées dans les boîtes de protection par l'ERLAP sont légèrement inférieurs aux mesures sans boîtes, sans toutefois faire apparaître de différence significative : 42 ± 1,1 µg/m³ contre 41 ± 1,3 µg/m³ en site urbain et 12 ± 0,4 µg/m³ contre 12 ± 0,3 µg/m³ en site de fond. Toutefois, lors d'un essai en chambre d'exposition dans des conditions de vent important (de 3 à 7 m/s), six tubes non protégés ont présenté des résultats de 35 % supérieurs à ceux de six tubes placés dans une boîte de protection.

Relation entre la répétabilité et la masse de nitrite dans les échantillonneurs

Le tableau 6 résume pour chaque essai les valeurs obtenues pour la répétabilité et la reproductibilité. Il est nécessaire de rechercher une éventuelle relation entre la répétabilité et le niveau de concentration de tous les essais pour pouvoir mieux exprimer le résultat des intercomparaisons. La figure 7, indiquant la valeur de l'écart type de répétabilité en fonction du niveau de concentration en µg de nitrite par essai, ne montre pas de relation claire entre la répétabilité et la masse de nitrite en considérant globalement les quatre essais. Toutefois, si l'on prend à part les deux essais sur site, alors il apparaît que l'écart type de répétabilité augmente avec la masse de nitrite absorbée dans l'échantillonneur.

L'écart type de répétabilité des matériaux de référence est supérieur à celui qui est obtenu pour les tubes dopés. Cette différence est vraisemblablement due à la phase d'extraction supplémentaire dans l'analyse des matériaux de référence, car les modes de préparation des deux types d'échantillonneurs incluent une variabilité similaire. Ces conclusions demanderaient à être confirmées par un plus grand nombre d'essais.

Contrôle qualité

Durant l'étude, le débit d'échantillonnage n'était pas unifié : par exemple le laboratoire G utilisait 72 ml/h alors que le laboratoire H utilisait 72,8 ml/min. Pour les tubes dopés, il n'a pas été tenu compte du volume de la TEA et du nitrite injecté, soit une dilution de 2 % à ajouter au volume de réactif dans les calculs des masses de nitrite.

Afin d'évaluer la maîtrise de chaque laboratoire pour la préparation et l'analyse des tubes à diffusion, la valeur du blanc et son écart type, la pente et l'ordonnée à l'origine de la droite d'étalonnage et l'erreur de linéarité ont été calculés. Ils sont présentés dans le tableau 7. Un processus maîtrisé sera caractérisé par un blanc, un écart type du blanc, un écart type de la pente de la droite d'étalonnage et une erreur de linéarité limités. L'erreur de linéarité est calculée suivant la formule :

$$S_{y/x} = \left(\frac{\sum (y_i - \hat{y}_i)^2}{n - 2} \right)$$

où y_i est la concentration de nitrite de chaque étalon de la droite d'étalonnage, $\hat{y}_i$ est la concentration estimée et n est le nombre d'étalons utilisés pour l'étalonnage.

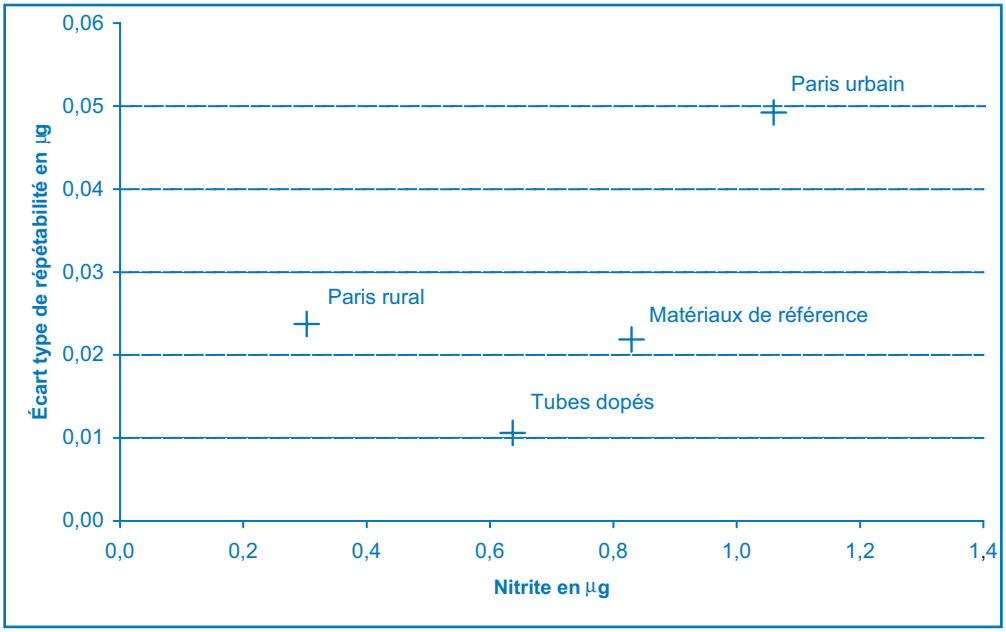

Figure 7.

Relation entre l'écart type de répétabilité et la masse de nitrite dans les échantillonneurs pour chaque essai d'intercomparaison.

Repeatability standard deviation *versus* the mass of nitrite in a sampler for each test of the intercomparison.

Tableau 7.

Contrôle qualité du blanc et de l'étalonnage : valeur du blanc et son écart type (s blanc), ordonnée de la droite d'étalonnage, pente et écart type (s pente) de la droite d'étalonnage, erreur de linéarité (Sy/x).

Quality control of the blank and the calibration: values of the blank and their standard deviation (s blanc), the intercept of the calibration curve, the slope and its standard deviation (s pente) of the calibration curve, linearity errors (Sy/x).

Lab.	A	C	E	F	G	H	I	J	K
Blanc (µg)	0,010	0,010	0,032	0,014	0,029	0,043	0,013	0,024	
s blanc (µg)	0,003	0,005	0,010	0,002	0,013	0,014	0,014	0,029	
Ordonnée (µg/ml)	– 0,001	0,000	0,000		– 0,011	0,002	– 0,003	– 0,003	– 0,019
Pente (µg/ml/Abs)	0,905	0,921	0,975		0,938	0,915	0,927	0,964	0,933
s pente (µg/ml/Abs)	0,2 %	0,6 %	0,4 %		5,1 %	1,1 %	1,9 %		3,2 %
Sy/x (µg/ml)	0,001	0,004	0,003	0,005	0,011	0,007	0,005	0,059	0,005

Abs : absorbance.

En observant le tableau 7, on peut noter que pour certains laboratoires, les valeurs du blanc (laboratoires E, G, H, I), de l'écart type de la pente de la droite d'étalonnage (laboratoires G et K) et de l'erreur de linéarité (laboratoires G, H et J) peuvent être améliorés. Il s'agit des laboratoires qui ont présenté des valeurs aberrantes dans l'étude de la répétabilité et de la reproductibilité. De plus, il a été noté que les meilleurs résultats ont été obtenus par les laboratoires ayant le matériel en meillleur état : tubes non rayés, capsules non fendues ou capsules placées sur la bonne extrémité du tube. En particulier, les nouveaux tubes et capsules (code XDIF…) ont obtenu les valeurs les plus proches des valeurs de référence à une exception près, due à un étalonnage incorrect.

Comparaison avec d'autres exercices d'intercomparaison

Au Royaume-Uni, les laboratoires impliqués dans les analyses du NO_2 par tube à diffusion pour le réseau NO_2 britannique, peuvent participer à l'essai d'aptitude dénommé WASP [8]. Celui ci est maintenant géré par le *Health and Safety Laboratory* (HSL).

Il a été introduit en avril 1999. Il comporte l'analyse, par chaque laboratoire participant, d'un tube contenant une masse connue de nitrite. Chaque mois des tubes à diffusion dopés sont distribués à chaque participant. Ils sont analysés et le résultat est renvoyé au HSL pour traitement suivant la méthode du z-score. WASP est conçu comme un exercice d'essai d'aptitude. De 1999 à 2001, dans le cadre de l'exercice WASP, pour le dopage entre 0,5 et 1 µg de nitrite l'écart type relatif publié est entre 10 et 18 %. Cette valeur est à comparer avec le coefficient de variation (CV) inférieur à 3,5 % de la présente étude, obtenu en divisant la reproductibilité de 9,5 % de l'exercice sur les tubes dopés par 2,83 ($2\sqrt{2}\ s_R$). La confrontation montre de meilleurs résultats pour l'étude présente même pour la reproductibilité des essais sur les matériaux de référence (CV = 6,3 %) et comparables pour l'essai sur site urbain (CV = 15 %).

Dans le cadre du réseau national britannique du NO_2 [9], des solutions étalons et des tubes dopés sont régulièrement analysés par les laboratoires analytiques. Pour l'exercice réalisé avec des tubes dopés, l'écart type relatif est au mieux de 21 %. La comparaison est, une fois encore, à l'avantage de l'étude présentée ici.

Les résultats de l'intercomparaison de l'EMD de 1999 [2] peuvent être directement comparés avec l'essai réalisé en site urbain. L'amélioration notable en 2003 est que la majorité des participants (8 laboratoires sur 9) ont une différence inférieure à 25 % par rapport à la valeur de référence, ce qui n'était pas le cas en 1999 (10 laboratoires sur 19 seulement).

Conclusion

Le traitement des données suivant la méthode du z-score montre que les participants obtiennent de bons résultats pour les expositions sur site en se fixant un écart type cible de 25 %. Toutefois, dans les exercices futurs, il faudra diminuer de moitié cet écart type (12,5 %) de façon à obtenir une incertitude élargie de 25 % en incluant un facteur d'élargissement de 2 suivant les prescriptions du GUM.

Pour les essais sur les tubes dopés et les matériaux de référence, l'objectif, en terme d'écart type cible fixé à 5 %, était plus ambitieux. Néanmoins, plus de la moitié des participants se trouve en dehors de la zone des mauvais résultats, démontrant que l'écart type de 5 % peut rester un objectif pour les essais à venir. De plus les nouveaux matériaux de référence préparés par l'ERLAP, en exposition sous atmosphère contrôlée, ont démontré leur validité pour l'utilisation dans ce type d'intercomparaison.

La répétabilité et la reproductibilité de la méthode de mesure par tube de Palmes ont été déterminées durant cette intercomparaison. Pour les essais sur site, la répétabilité relative diminue avec le niveau des concentrations mesurées, passant d'environ 20 à 10 % quand les concentrations évoluent de 15 à 40 µg/m³. En revanche, la reproductibilité semble stable à 45 %. Pour les essais sur les tubes dopés et les matériaux de référence, les valeurs de la répétabilité (< 10 %) et de la reproductibilité (< 20 %) sont nettement inférieures. Les tubes dopés présentent eux-mêmes des valeurs de répétabilité et reproductibilité deux fois plus petites que celles des matériaux de référence préparés par exposition. Ces valeurs de 10 et 20 % peuvent être considérées comme l'objectif de répétabilité et reproductibilité à atteindre dans les intercomparaisons futures. De plus, les exercices de contrôle qualité devraient toujours inclure des mesures sur site et pas seulement des tubes dopés comme c'est généralement le cas actuellement. Globalement, les résultats de cette étude apparaissent plus précis que ceux qui sont obtenus dans le contrôle qualité WASP et le réseau national NO_2 du Royaume-Uni. Cette étude montre également une amélioration des résultats par rapport à l'intercomparaison 1999 réalisée par l'École des mines de Douai.

Toutefois, la reproductibilité est toujours nettement supérieure à la répétabilité, ce qui suggère la présence d'erreurs systématiques dans la préparation et/ou l'analyse. En fait, de meilleurs résultats ont été obtenus avec les matériaux de référence que sur le terrain, ce qui met de nouveau en cause la préparation des échantillonneurs. De plus, le contrôle qualité montre que la valeur du blanc, la linéarité et la variation de la pente d'étalonnage peuvent être améliorées.

Le transport par avion des tubes de Palmes dopés n'a pas d'influence significative sur le résultat d'analyse. L'utilisation des boîtes de protection provoque une légère diminution des concentrations, significative en présence de turbulences. Le tube de Palmes avec extrémité ouverte est sensible aux turbulences dues au vent.

Le résultat le plus notable reste que la médiane des valeurs moyennes de tous les laboratoires pour les essais sur site est en accord parfait avec les mesures des analyseurs par chimiluminescence.

Rermerciements :

Les auteurs souhaitent remercier :

- Mme Anne Fromage-Mariette coordinatrice du Groupe de travail « Échantillonneurs à diffusion » à l'origine de cette étude pour la revue du rapport ;
- AIRPARIF et en particulier MM. Patrick Garnoussi, Laurent Gauvin et Christophe Ampe pour la réalisation des expositions sur site et pour la mise à disposition des mesures des analyseurs en station ;
- les AASQA pour leur disponibilité à effectuer les analyses des échantillonneurs à diffusion et plus particulièrement : Nathalie Marquis, d'AIRLOR ; Bruno Dadolle, d'AIR LR ; Esthel Le Bronnec, d'AIRPARIF ; Eve Chrétien, d'ATMO Champagne-Ardenne ; Benoît Rocq et Élodie Roullet, d'ATMO Picardie ; Sébastien Morenilla, ATMO Poitou-Charentes ; Chatrapatty Bhugwant, de l'Observatoire de l'air réunionnais.

Mots clés

Dioxyde d'azote. Tube de Palmes. Essai d'aptitude. Exercice d'intercomparaison. Répétabilité. Reproductibilité. Méthode indicative. Échantillonneur à diffusion.

Keywords

Nitrogen dioxide. Palmes tube. Proficiency testing. intercomparison exercice. Repeatability. Reproducibility. Indicative method. Diffusive sampler.

Références

1. Campbell GW, Stedman JR, Stevenson KJ. A Survey of Nitrogen Dioxide Concentrations in the United Kingdom Using Diffusion Tubes, July-December 1991. *Atmos Environ* 1994 ; 28 (3) : 477-86.
2. Garcia-Fouqué S, Houdret J-L, Plaisance H. Synthèse de l'expérience acquise par les réseaux sur l'échantillonnage du NO_2, École des Mines de Douai, Département Chimie et Environnement, Étude n° 7, mai 2000.
3. Échantillonneurs passifs pour le dioxyde d'azote. Guide rédigé par le Groupe de Travail français « Échantillonneurs passifs », octobre 2001.
4. Norme Internationale « Essais d'aptitude des laboratoires par intercomparaison – Partie 1 : Développement et mise en œuvre de systèmes d'essais d'aptitude », ISO/IEC Guide 43-1 : 1997, Genève.
5. Norme Internationale « Essais d'aptitude des laboratoires par intercomparaison – Partie 2 : Sélection et utilisation de systèmes d'essais d'aptitude par des organismes d'accréditation de laboratoires », ISO/IEC Guide 43-2 : 1997, Genève.
6. Norme Internationale « Prescriptions générales concernant la compétence des laboratoires d'étalonnages et d'essais », ISO/IEC 17025 : 1999, Genève.
7. Norme Internationale « 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 normalisée », ISO 5725-4 : 1994, Genève.
8. Stacey P. Proficiency testing for benzene, Toluene, m-xylene and nitrogen dioxide in the air: experience in the WASP quality assurance scheme. *In*: Proceedings of International Conference Measuring Air Pollutants by Diffusive Sampling, Montpellier, France 26-28 September 2001, European Commission, Joint Research Centre, EUR 20242 EN.
9. Stevenson K, Bush T, Mooney D. Five years of nitrogen dioxide measurements with diffusion tube samplers at over 1000 sites in the UK. *Atmos Environ* 2001 ; 35 : 281-7.
10. Norme Internationale « ISO Guide to the expression of uncertainty in measurement », 1995.
11. NMi Report S-CH-02-10, CERMATAIR, Certified Reference Materials for the measurement of gaseous pollutants in ambient air. Desk study into the technological possibilities of preparing and certifying reference materials, Compilation of reports of all partners. http://www.cermatair.com/
12. Conseil de l'Union Européenne. Directive 1999/30/CE du Conseil du 22 avril 1999 relative à la fixation de valeurs limites pour l'anhydride sulfureux, le dioxyde d'azote et les oxydes d'azote, les particules et le plomb dans l'air ambiant. Journal officiel des Communautés européennes 1999 ; L 163/41 : 1.