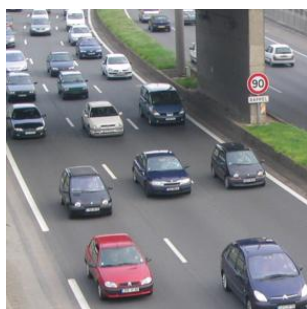


Recherche de l'échantillonnage temporel optimal pour estimer une concentration annuelle de HAP

Prélèvements continus durant 1 an des HAP dans l'air ambiant sur la station urbaine Lyon Centre

2009-2010



L'Association Agréée de Surveillance de la Qualité de l'Air de la Région Rhône-Alpes (AASQA), COPARLY fait partie du dispositif français de surveillance et d'information de la qualité de l'air. Sa mission s'exerce dans le cadre de la loi sur l'air du 30 décembre 1996 et de ses décrets d'application notamment le décret 98-361 du 6 mai 1998 relatif à l'agrément des organismes de surveillance de la qualité de l'air.

A ce titre et compte tenu du statut d'organisme non lucratif, l'association COPARLY est garante de la transparence de l'information sur le résultat de ses travaux.

Conditions de diffusion :

- Les données recueillies tombent dès leur élaboration dans le domaine public. Le rapport d'étude est mis à disposition sur www.atmo-rhonealpes.org, un mois après validation interne.
- Les données contenues dans ce document restent la propriété de l'association. Données non rediffusées en cas de modification ultérieure des données.
- Toute utilisation partielle ou totale de ce document doit faire référence à l'association en termes de «COPARLY (2010) : Recherche de l'échantillonnage temporel optimal pour estimer une concentration annuelle de HAP ».
- COPARLY n'est en aucune façon responsables des interprétations et travaux intellectuels, publications diverses résultant de ses travaux et pour lesquels aucun accord préalable n'aurait été donné.

Cette étude a été réalisée grâce au concours financier de la DRASS Rhône-Alpes (Agence Régionale de Santé depuis 2010)

Sommaire

1. CONTEXTE ET OBJECTIFS	6
2. PRÉSENTATION DES DONNÉES	7
2.1. PRÉLÈVEMENTS.....	7
2.1.1. <i>Appareillage spécifique</i>	7
2.1.2. <i>Date des prélèvements</i>	7
2.1.3. <i>Composés analysés</i>	7
2.1.4. <i>Données manquantes</i>	7
2.2. VARIABILITÉ TEMPORELLE.....	8
2.2.1. <i>Variabilité quotidienne</i>	8
2.2.2. <i>Profils hebdomadaires</i>	9
2.2.3. <i>Profils mensuels</i>	10
2.3. RÉPARTITION DES DONNÉES, VARIABILITÉ.....	11
2.4. INFLUENCE DES MAXIMA SUR LA CONCENTRATION MOYENNE ANNUELLE	12
3. COMPARAISON DES DIFFÉRENTS PLANS D'ÉCHANTILLONNAGE	13
3.1. QUELQUES DÉFINITIONS D'ÉCHANTILLONNAGE	13
3.1.1. <i>Échantillonnage systématique ou aléatoire</i>	13
3.1.2. <i>Échantillonnage par strates</i>	13
3.1.3. <i>Échantillonnage par grappe</i>	13
3.1.4. <i>Plans d'échantillonnage testés</i>	13
3.2. ÉCHANTILLONNAGES SYSTÉMATIQUES.....	14
3.2.1. <i>Échantillonnage simple : 1 jour sur 3</i>	14
3.2.2. <i>Échantillonnage n°2 avec 4 strates</i>	15
3.2.3. <i>Échantillonnage n°3 avec 4 strates</i>	16
3.2.4. <i>Échantillonnage n°4 avec 2 strates</i>	17
3.2.5. <i>Conclusions sur les échantillonnages systématiques</i>	18
3.3. ÉCHANTILLONNAGES ALÉATOIRES.....	18
3.3.1. <i>Échantillonnage aléatoire stratifiés : 30 jours aléatoires par trimestre</i>	18
3.3.2. <i>Échantillonnage stratifié par grappe : 2 fois 15 jours consécutifs sélectionnés aléatoirement dans chaque strate</i>	20
3.3.3. <i>Échantillonnage aléatoire stratifié : 2 saisons</i>	20
4. DISCUSSION	22
4.1. CONCLUSIONS DES COMPARAISONS DES DIFFÉRENTES MÉTHODES D'ÉCHANTILLONNAGES.....	22
4.2. COMPARAISON À D'AUTRES ÉTUDES	23
4.3. CONTRAINTES TECHNIQUES	23
5. CONCLUSION	24
BIBLIOGRAPHIE	26
ANNEXE 1 : PROFILS MENSUELS DES HAP CITÉS DANS LA DIRECTIVE	27
ANNEXE 2 : REPRÉSENTATION PAR BOÎTE À MOUSTACHE (OU BOX PLOT)	29
ANNEXE 3 : INFLUENCE DES MAXIMA DANS L'ESTIMATION DE LA CONCENTRATION MOYENNE	30

Tables des figures

Figure 1 : Préleveurs haut débit DigitelDA80	7
Figure 2 : Pourcentage de données valides par mois	8
Figure 3 : Concentration quotidienne du benzo[a]pyrène en 2009	8
Figure 4 : Concentration quotidienne de la somme totale des HAP	8
Figure 5 : Corrélation linéaire entre les différents HAP étudiés.....	9
Figure 6 : Profils hebdomadaires des HAP	10
Figure 7 : Profils mensuels du Benzo[a]pyrène et de la somme des 19 HAP	10
Figure 8 : Histogramme et boîte à moustache pour le Benzo[a]pyrène	11
Figure 9 : Histogramme et boîte à moustache pour la somme des 19 HAP	11
Figure 10 : Influence des maxima sur la concentration moyenne du BaP	12
Figure 11 : Estimation sur 3 simulations de la moyenne annuelle et du maximum à partir des prélèvements journaliers réalisés 1 jour sur 3	15
Figure 12 : Estimation sur 3 simulations de la moyenne annuelle et du maximum à partir des prélèvements journaliers réalisés 2 jours sur 3 sur la première moitié des 4 strates.....	16
Figure 13 : Estimation sur 3 simulations de la moyenne annuelle et du maximum à partir des prélèvements journaliers réalisés 2 jours sur 3 sur la deuxième moitié des 4 strates.....	17
Figure 14 : Estimation sur 3 simulations de la moyenne annuelle et du maximum échantillonnés sur 2 strates. La moyenne estimée correspond au barycentre des moyennes des 2 strates pondérées par la couverture annuelle des 2 strates.	18
Figure 15 : Répartition par boxplot de l'estimation de la concentration moyenne annuelle pour 5 HAP à partir de 30 prélèvements sélectionnés aléatoirement par trimestre...19	19
Figure 16 : Statistiques sur les erreurs relatives calculées à partir de 30 prélèvements sélectionnés aléatoirement par trimestre.....	19
Figure 17 : Répartition par boxplot de l'erreur relative estimée de 5 HAP à partir de 30 prélèvements sélectionnés aléatoirement par trimestre.....	20
Figure 18 : Statistiques sur les erreurs relatives calculées à partir de 30 prélèvements sélectionnés aléatoirement par trimestre.....	20
Figure 19 : Statistiques sur les erreurs relatives calculées à partir de 102 prélèvements sélectionnés aléatoirement en période froide et 21 en période chaude	21
Figure 20 : Répartition par boxplot de l'erreur relative estimée à partir de 30 prélèvements sélectionnés aléatoirement par trimestre.....	21

Résumé

Tout au long de l'année 2009, COPARLY a réalisé un suivi continu quotidien des Hydrocarbures Aromatiques Polycycliques (HAP particuliers et gazeux) sur le site urbain de Lyon Centre. La mise en œuvre de ces mesures avait pour principal but de constituer une série annuelle de prélèvements à destination d'un travail d'estimation de la méthode d'échantillonnage optimale permettant d'accéder à une estimation de la concentration moyenne annuelle.

Cette valeur moyenne annuelle constitue le résultat à comparer à la valeur européenne cible concernant le benzo[a]pyrène (BaP). La réglementation requiert pour cette surveillance des mesures représentant 33% des jours de l'année, également répartis sur les saisons. La présente étude a visé ainsi à comparer la performance de différentes méthodes d'échantillonnage des HAP dans l'air ambiant en configuration urbaine pour une couverture annuelle de 33%.

A l'issue de l'année 2009, 356 résultats valides sur 365 ont été récupérés, soit plus de 97% de l'année. Ce bon résultat a ainsi permis d'offrir une vision quasiment complète de l'année et une série solide de données pour l'analyse statistique qui a suivi.

En particulier, il apparaît que la variabilité des données augmente de manière significative entre mi-octobre et mi-mars.

Différentes méthodes d'échantillonnage ont ensuite été évaluées : systématique, aléatoire, stratifiée (avec 2 et 4 strates), par grappe. Au total 7 plans d'échantillonnage ont été testés et comparés.

La méthode actuellement utilisée en Rhône-Alpes consistant à mesurer un jour sur trois (échantillonnage 1) présente des résultats convenables au vu des erreurs relatives des autres méthodes (erreur relative inférieure à 22% pour l'ensemble des HAP et inférieure à 17 % pour le BaP). Il convient de préciser que la méthode actuelle présente un risque relativement élevé de manquer des pointes importantes en particulier en période hivernale. En témoignent les différences non négligeables entre les 3 simulations (erreurs relatives de -16% à +10% pour le BaP).

Les méthodes d'échantillonnage testées réduisant les erreurs relatives sont les plans d'échantillonnage divisés en 2 périodes ou « strates » (5 mois de période « froide » et 7 mois de période « chaude »). Elles permettent d'observer des erreurs relatives en moyenne deux fois plus faibles qu'avec la méthode actuellement mise en place (environ 6% au lieu de 11%). Elles consistent à augmenter le nombre de prélèvements (102 prélèvements contre 51 prélèvements) lorsque les variations de concentrations sont plus importantes (entre mi-octobre et mi-mars) et à le réduire (21 prélèvements contre 71) lorsque les concentrations sont stables (entre mi-mars et mi-octobre). Ces prélèvements peuvent être choisis suivant un échantillonnage systématique ou un échantillonnage aléatoire.

Néanmoins, l'estimation de la concentration moyenne annuelle avec ces méthodes ne doit pas être calculée en prenant la moyenne arithmétique de l'ensemble des prélèvements car la durée des deux périodes n'étant pas identique, le résultat serait alors surestimé. L'estimation est réalisée en prenant le barycentre des moyennes des 2 strates pondérées par leur couverture annuelle.

Ces méthodes, qui présentent des erreurs relatives faibles, apparaissent donc à privilégier dans le cas d'environnement urbain (tel que Lyon centre) si les contraintes techniques et les textes réglementaires le permettent.

1. Contexte et objectifs

COPARLY, a débuté en novembre 2001 un suivi des Hydrocarbures Aromatiques Polycycliques.

La directive 2004/107/CE établit une valeur cible pour le benzo(a)pyrène de 1ng.m^{-3} , dans l'air ambiant en moyenne annuelle applicable en 2012. Elle impose par ailleurs la surveillance d'autres HAP : benzo(a)anthracène, benzo(b)fluoranthène, benzo(j)fluoranthène, benzo(k)fluoranthène, indéno(1,2,3-cd)pyrène, dibenzo(a,h)anthracène.

D'un point de vue réglementaire, le suivi consiste en des prélèvements échantillonnés de manière régulière dans le temps afin d'estimer une moyenne annuelle.

La durée d'échantillonnage est de 24h et la directive précise une période minimale de prise en compte des données :

- 33 % pour des mesures fixes
- 14 % pour les mesures indicatives

La directive ne précise pas comment doit être mené cet échantillonnage. Néanmoins, l'incertitude sur la moyenne annuelle ne doit pas dépasser 50%.

En 2009, COPARLY a réalisé un suivi continu quotidien sur le site de Lyon centre. En exploitant cette série annuelle de prélèvements HAP, l'objectif est de déterminer quel plan d'échantillonnage conviendrait le mieux à l'estimation de la moyenne annuelle. Cette étude a donc pour principal but de comparer différents plans d'échantillonnage afin de conclure quant à la meilleure méthode d'estimation de la moyenne annuelle.

Le Groupe de Travail (GT) national « Plans d'échantillonnage et reconstitutions des données » a déjà pu étudier les différences d'échantillonnage dans le cas de mesure indicative (échantillonnage de 14% de l'année) à l'aide d'une série temporelle de HAP d'AIRPARIF de 1998. Cette étude montre que « pour une même couverture annuelle la réalisation de prélèvements isolés permet d'estimer plus précisément la moyenne annuelle que des prélèvements groupés par semaine ».

Les prélèvements de HAP effectués en 2009 permettront d'une part de valider cette conclusion sur le site de Lyon centre et d'autre part d'évaluer la meilleure méthode d'échantillonnage pour une mesure fixe (échantillonnage de 33% de l'année).

2. Présentation des données

2.1. Prélèvements

2.1.1. Appareillage spécifique

Les prélèvements de HAP se font sur des filtres (phase particulaire) et des mousses (phase gazeuse) à l'aide d'un préleveur haut débit durant 24 heures. Après analyse par un laboratoire agréé, les résultats permettent d'accéder à une concentration moyenne journalière de 19 HAP.



Figure 1 : Préleveurs haut débit
DigitelDA80

2.1.2. Date des prélèvements

Les prélèvements ont eu lieu pendant toute l'année 2009 sur le site de référence de Lyon centre en typologie urbaine. Il s'agit de prélèvements quotidiens de 24h, soit au total 365 prélèvements visés dans le cadre de cette étude.

A l'issue d'une telle série, la moyenne annuelle peut être calculée de manière non biaisée et servir de référence aux différents tests de plan d'échantillonnage.

2.1.3. Composés analysés

Sur les 19 HAP analysés, l'étude se concentrera sur le benzo(a)pyrène, seul HAP réglementé et à la somme de l'ensemble des HAP afin d'avoir une vision globale de tous les HAP.

A titre de comparaison avec l'étude du GT, les HAP cités dans la directive 2004/107/CE et pour lesquels les données sont disponibles seront aussi étudiés :

- benzo(a)anthracène, BaA
- benzo(b)fluoranthène, BbFa
- benzo(k)fluoranthène, BkFa
- benzo(a)pyrène, BaP
- indénopyrène, IP

2.1.4. Données manquantes

La série contient peu de données manquantes. Au total, 356 résultats sur 365 sont valides constituant un taux de représentativité annuelle de plus de 97%.

La série de données comprend peu de plages manquantes. Les mesures non disponibles ne concernent que des prélèvements isolés et n'ont par conséquent que peu d'influence sur le calcul de la moyenne ou sur la mise en place des plans d'échantillonnage.

Ce bon résultat permet d'offrir une vision quasiment complète de l'année et une série solide de données pour l'analyse statistique.

Mois	Janvier	Février	Mars	Avril	Mai	Juin
Couverture mensuelle	100%	100%	93%	97%	93%	100%

Mois	Juillet	Aout	Septembre	Octobre	Novembre	Décembre
Couverture mensuelle	100%	93%	93%	100%	100%	100%

Figure 2 : Pourcentage de données valides par mois

2.2. Variabilité temporelle

2.2.1. Variabilité quotidienne

BaP

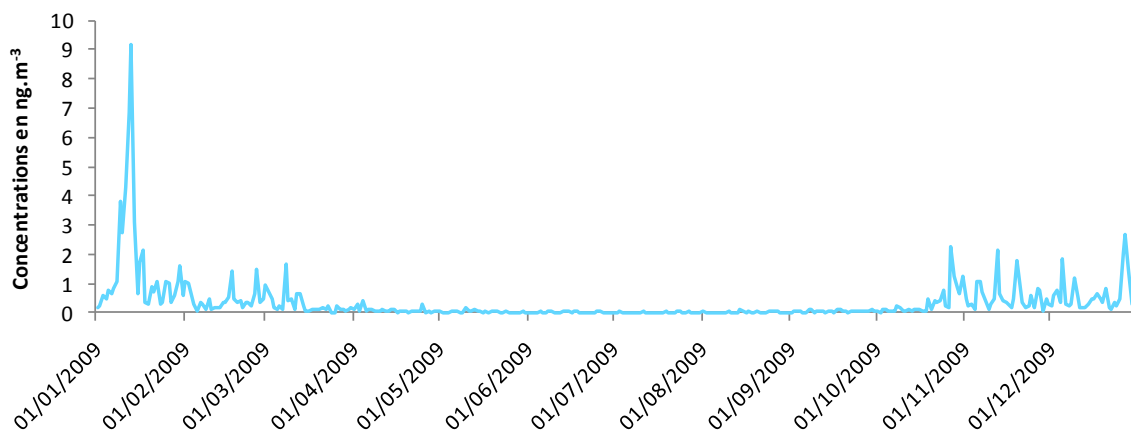


Figure 3 : Concentration quotidienne du benzo[a]pyrène en 2009

Somme des HAP

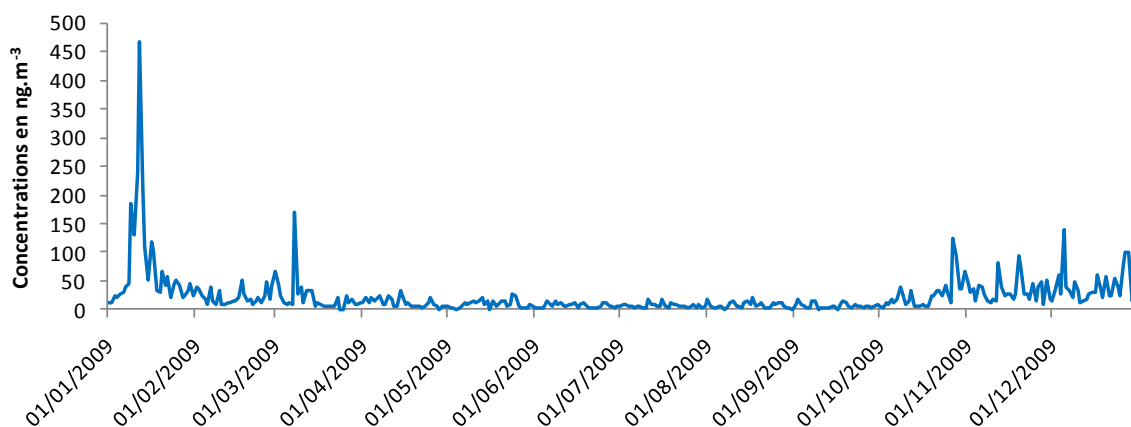


Figure 4 : Concentration quotidienne de la somme totale des HAP

La série des mesures journalières pour le BaP et la somme totale des 19 HAP mesurés sont représentées sur les graphiques précédents. Il apparaît clairement une variabilité temporelle marquée. Les concentrations sont en effet plus élevées en fin d'automne et en hiver (d'octobre à mars) que pour le reste de l'année. Comparé aux autres années, cette

hausse hivernale des concentrations a été particulièrement marquée au cours de l'année 2009.

Par ailleurs, les deux séries temporelles montrent une très forte corrélation. Cette corrélation avec la somme totale des 19 HAP est notable pour l'ensemble des HAP cités dans la directive.

Le tableau suivant résume les coefficients de corrélation linéaire entre la somme des HAP et les différents HAP étudiés. Il existe globalement une très forte corrélation linéaire entre les composés étudiés :

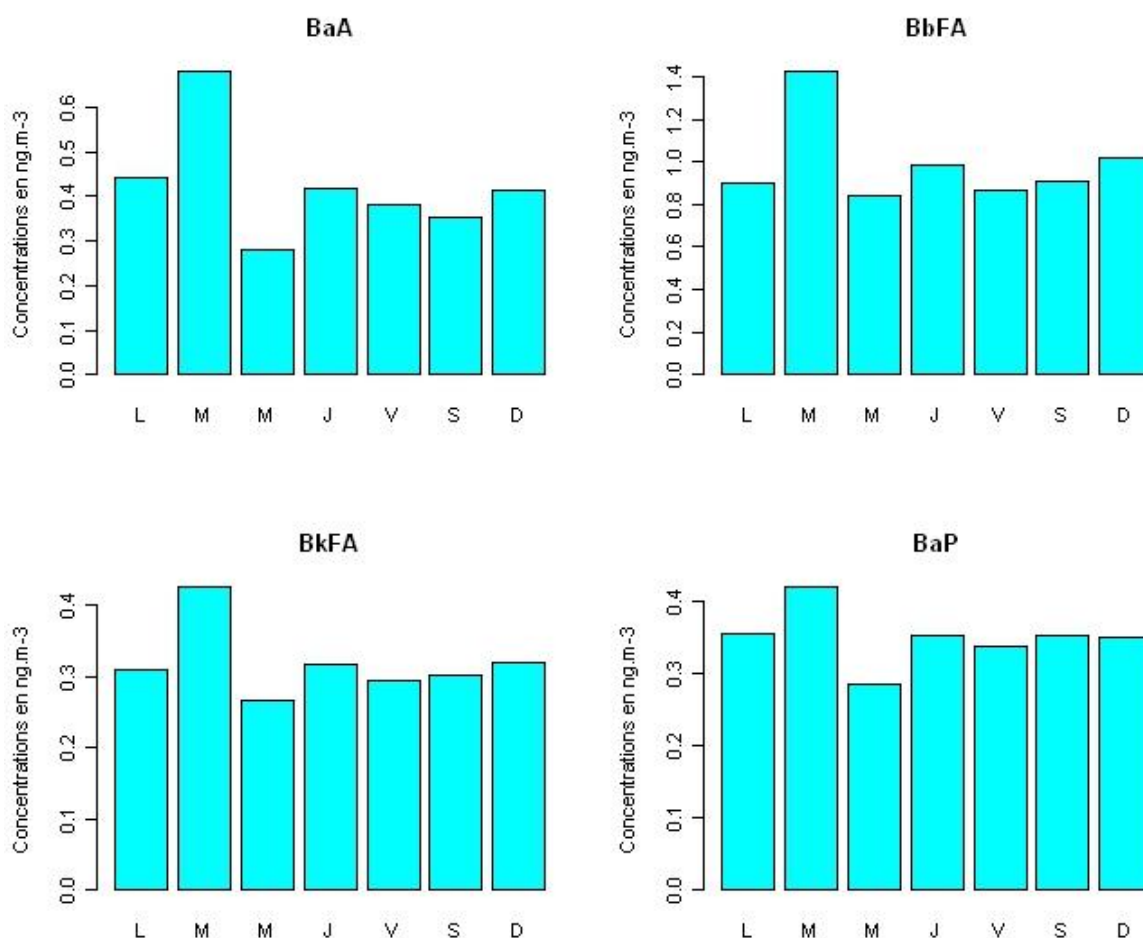
Corrélation linéaire entre :	Somme	BaP	BaA	BbFA	BkFA	IP
Somme	1,00	0,89	0,76	0,78	0,85	0,90
BaP	0,89	1,00	0,93	0,91	0,97	0,98
BaA	0,76	0,93	1,00	0,93	0,95	0,76
BbFA	0,78	0,91	0,93	1,00	0,99	0,89
BkFA	0,85	0,97	0,95	0,99	1,00	0,94
IP	0,90	0,98	0,76	0,89	0,94	1,00

Figure 5 : Corrélation linéaire entre les différents HAP étudiés

2.2.2. Profils hebdomadaires

Les 5 HAP de la directive européenne montrent des profils hebdomadaires similaires avec une concentration minimale le mercredi et maximale le mardi. Du jeudi au lundi, les profils sont relativement stables, et on observe peu de variations.

En ce qui concerne le profil hebdomadaire de la somme des 19 HAP, la moyenne minimale est relevée le mercredi et la moyenne maximale le lundi.



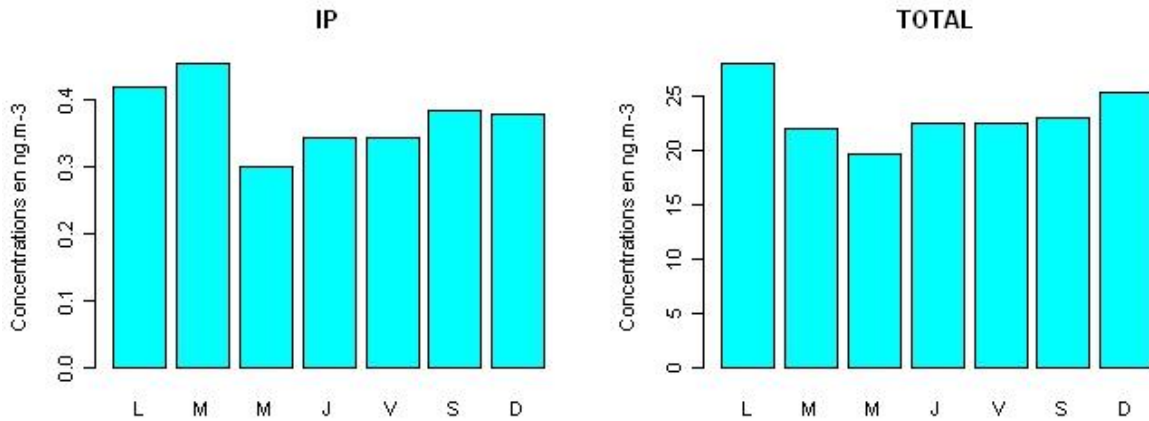


Figure 6 : Profils hebdomadaires des HAP

2.2.3. Profils mensuels

Pour le benzo[a]pyrène, il apparaît clairement une hausse des concentrations mensuelles mesurées en hiver entre le mois d'octobre et le mois de mars. Entre avril et septembre, les concentrations sont très basses.

Les maxima sont aussi atteints en hiver et plus précisément pendant le mois de janvier où les concentrations journalières peuvent dépasser les 9 ng.m⁻³.

Pour la somme totale des HAP, les résultats sont comparables, la hausse des concentrations en hiver est toujours visible et le maximum journalier (467 ng.m⁻³) est aussi atteint en janvier.

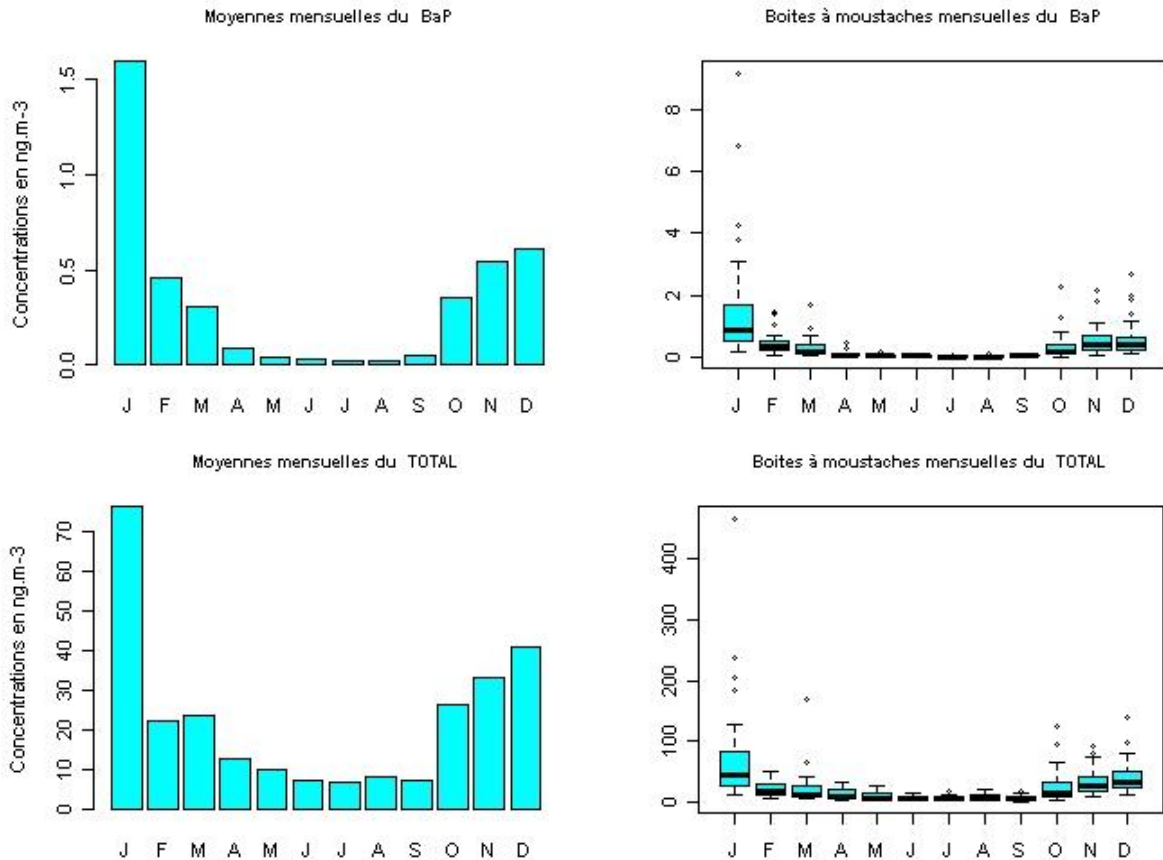


Figure 7 : Profils mensuels du Benzo[a]pyrène et de la somme des 19 HAP

Le profil mensuel des 4 autres HAP cités dans la directive sont comparables à ceux présentés plus haut et sont disponibles en annexe 1. Les représentations par boîte à moustache sont explicitées en annexe 2.

2.3. Répartition des données, Variabilité

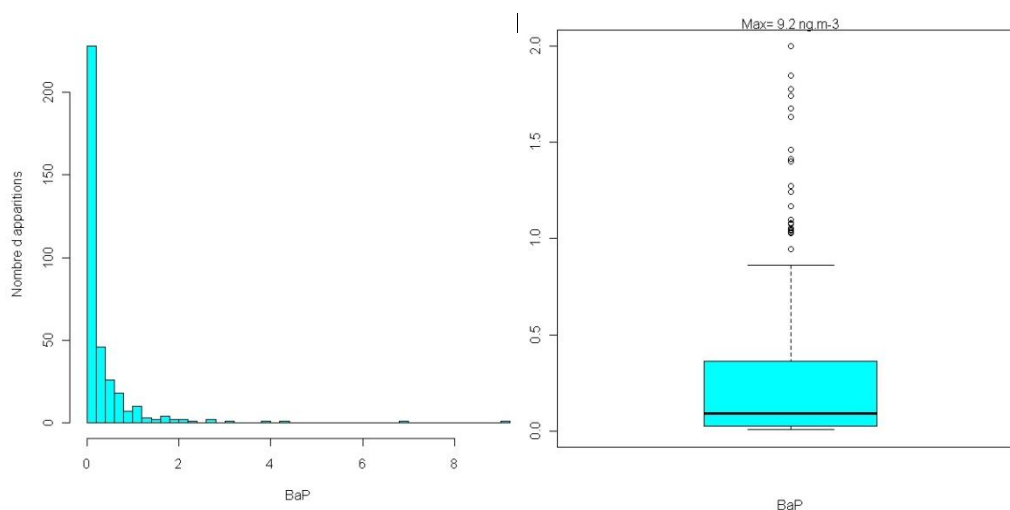


Figure 8 : Histogramme et boîte à moustache pour le Benzo[a]pyrène

La répartition des concentrations de benzo[a]pyrène montre que :

- 64 % des données sont inférieures à 0.2 ng.m^{-3}
- 50 % des données sont comprises entre 0.028 ng.m^{-3} et 0.36 ng.m^{-3}
- 8.7% des données sont supérieures à 1 ng.m^{-3}

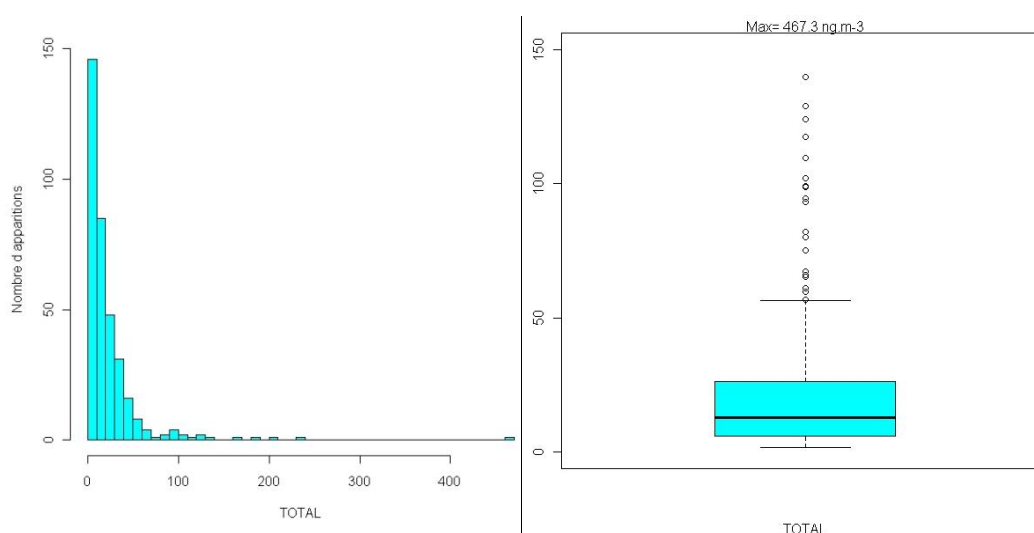


Figure 9 : Histogramme et boîte à moustache pour la somme des 19 HAP

En ce qui concerne la totalité des HAP, la répartition des données est assez comparable à celle du benzo[a]pyrène. Les précédentes représentations de la répartition des concentrations pour la somme des 19 HAP montre que :

- 41 % des données sont inférieures à 10 ng.m^{-3}

- 50 % des données sont comprises entre 6.1 ng.m^{-3} et 26.2 ng.m^{-3}
- 3.1% des données sont supérieures à 100 ng.m^{-3}
- Le maximum journalier relevé est de 467 ng.m^{-3}

2.4. Influence des maxima sur la concentration moyenne annuelle

Les maxima ont un impact certain sur l'estimation de la concentration moyenne annuelle dans le cas des HAP.

Le cas du BaP a été étudié plus précisément :

En enlevant la concentration journalière la plus élevée, l'estimation de la concentration moyenne passe de 0.35 ng.m^{-3} à 0.33 ng.m^{-3} et chute de 7%. En retirant les 2 plus hautes concentrations, la concentration moyenne chute de 13% en passant de 0.35 ng.m^{-3} à 0.31 ng.m^{-3} . Ces calculs ont ainsi été effectués pour l'ensemble de la série annuelle de données, et permettent de tracer le graphe suivant. Ce dernier illustre ainsi les variations de la concentration moyenne annuelle au fur et à mesure que l'on enlève les maxima. Il apparaît clairement une décroissance exponentielle de la concentration moyenne annuelle.

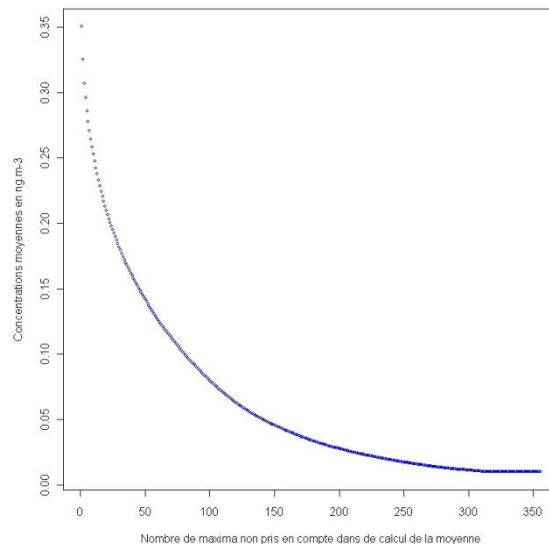


Figure 10 : Influence des maxima sur la concentration moyenne du BaP

Plus précisément, en supprimant le tiers des valeurs les plus élevées, la concentration moyenne annuelle chute de 82%. Elle passe en effet de 0.35 ng.m^{-3} à 0.06 ng.m^{-3} . Ces observations montrent donc qu'il est primordial d'estimer au mieux les maxima dans les méthodes d'échantillonnage et confirment l'intérêt d'étudier une série annuelle de concentrations de HAP, avec notamment une période hivernale la plus complète possible.

Ces conclusions sont semblables pour l'ensemble des HAP étudiés (Annexe 3).

3. Comparaison des différents plans d'échantillonnage

Le but est d'étudier et de comparer différents plans d'échantillonnage qui couvrent **33% de l'année**, condition correspondant à une mesure dite « fixe ou continue » au sens des directives européennes et réalisée de manière générale en station fixe. Il faudra ensuite déterminer quelle stratégie d'échantillonnage conduit à l'estimation la plus précise de la moyenne annuelle réelle pour les HAP.

Certaines méthodes d'échantillonnage sont identiques à celles testées par le groupe de travail : *Plans d'échantillonnage et reconstitutions de données*. Néanmoins, les conditions ne sont pas complètement comparables (33% de couverture annuelle au lieu de 14%), ce qui pourrait éventuellement amener à des conclusions légèrement différentes.

3.1. Quelques définitions d'échantillonnage

3.1.1. Échantillonnage systématique ou aléatoire

On appelle **échantillon systématique** un échantillon obtenu en sélectionnant des individus (prélèvements) également espacés dans le temps dans la série complète annuelle.

A l'inverse, dans un **sondage aléatoire** les prélèvements sont tirés au sort individuellement et de manière aléatoire.

3.1.2. Échantillonnage par strates

Dans le **sondage stratifié**, l'année est découpée en grandes périodes, appelées strates présumées devoir être, pour la plupart, plus homogènes que l'année totale par rapport aux caractères étudiés (par exemple les saisons). Dans chacune des strates est ensuite tiré un échantillon préalablement dimensionné.

3.1.3. Échantillonnage par grappe

Dans le **sondage par grappes**, les prélèvements constitutifs de l'échantillon ne sont pas tirés individuellement, mais par groupes de mesures consécutives qui reçoivent la dénomination de grappes (par exemple 5 prélèvements consécutifs).

3.1.4. Plans d'échantillonnage testés

7 plans d'échantillonnage ont été testés : les 4 premiers correspondent à un échantillonnage systématique et les 3 suivants sont des échantillonnages aléatoires. Tous les plans d'échantillonnage sont détaillés dans les paragraphes suivants.

Échantillonnage 1 : systématique simple

Mesures 1 jour sur 3. C'est la méthode qui est actuellement utilisée pour estimer la concentration moyenne annuelle en HAP dans le cas de mesures dites « fixes ».

Échantillonnage 2 : systématique n°1, 4 strates

Mesures 2 jours sur 3 dans chaque première moitié de trimestre.

Échantillonnage 3 : systématique n°2, 4 strates

Mesures 2 jours sur 3 dans chaque seconde moitié de trimestre.

Échantillonnage 4 : systématique 2 strates

Mesures 2 jours sur 3 pendant 5 mois de période « froide » (grande variabilité des concentrations). Puis mesures 1 jour sur 10 pendant 7 mois en période « chaude » (faible variabilité des concentrations). Cet échantillonnage se base sur l'observation d'augmentation des concentrations entre le mois d'octobre et le mois mars.

Échantillonnage 5 : aléatoire 4 strates

30 prélèvements sont sélectionnés aléatoirement dans chaque trimestre. Cette stratégie a été évaluée en réalisant 200 simulations d'échantillonnage.

Échantillonnage 6 : aléatoire par grappe

2 grappes de 15 prélèvements consécutifs sélectionnés dans chaque trimestre. Cette stratégie a été évaluée en réalisant 200 simulations d'échantillonnage.

Échantillonnage 7 : aléatoire 2 strates

102 prélèvements sont sélectionnés aléatoirement en période « froide ». 21 prélèvements sont sélectionnés aléatoirement en période « chaude ». Cette stratégie a été évaluée en réalisant 200 simulations d'échantillonnage.

Remarque : d'une manière générale, il est courant d'utiliser la moyenne arithmétique des prélèvements échantillonnés comme estimation de la concentration moyenne annuelle. Or les fréquences d'échantillonnage des méthodes 4 et 7 n'étant pas également réparties sur l'année, l'utilisation de la moyenne arithmétique dans ce cas ne serait pas adaptée et risquerait de biaiser les résultats. Ces deux stratégies d'échantillonnage (méthodes 4 et 7) sont testées ici en estimant la concentration moyenne annuelle comme barycentre des moyennes des 2 strates pondérées par la couverture annuelle des 2 strates.

3.2. Échantillonnages systématiques

Pour chacune des stratégies systématiques testées (échantillonnage 1,2,3,4), on effectue 3 simulations d'échantillons décalées chacune d'une journée. A partir de ces échantillons, la moyenne estimée avec chaque stratégie et pour chaque simulation peut être comparée à la moyenne réelle.

3.2.1. Échantillonnage simple : 1 jour sur 3

Simulation 1	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	118	118	118	118	118	118
Couv. annuelle (%)	32.3	32.3	32.3	32.3	32.3	32.3
Moyenne estimée (ng.m ⁻³)	0.39	0.52	1.18	0.36	0.41	23.96
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	205.09
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Écart type (ng.m ⁻³)	0.96	2.18	3.25	0.94	0.95	31.31
Biais (ng.m ⁻³)	0.04	0.09	0.19	0.05	0.04	0.64
Erreur relative (%)	10.30	21.85	19.03	14.38	9.82	2.76

Simulation 2	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	121	121	121	121	121	121
Couv. annuelle (%)	33.1	33.1	33.1	33.1	33.1	33.1
Moyenne estimée (ng.m ⁻³)	0.29	0.32	0.71	0.25	0.32	20.07
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	4.26	7.45	8.26	3.14	4.68	238.39
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Écart type (ng.m ⁻³)	0.57	0.81	1.19	0.44	0.56	27.90
Biais (ng.m ⁻³)	-0.06	-0.11	-0.28	-0.07	-0.06	-3.25
Erreur relative (%)	-16.30	-25.45	-28.44	-22.21	-15.62	-13.95

Simulation 3	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	117	117	117	117	117	117
Couv. annuelle (%)	32.0	32.0	32.0	32.0	32.0	32.0
Moyenne estimée (ng.m ⁻³)	0.37	0.44	1.09	0.35	0.40	26.03
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	6.82	11.12	13.07	4.97	8.10	467.35
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Écart type (ng.m ⁻³)	0.81	1.23	2.02	0.67	0.88	49.24
Biais (ng.m ⁻³)	0.02	0.02	0.10	0.03	0.02	2.71
Erreur relative (%)	6.47	4.29	10.22	8.48	6.25	11.64

Figure 11 : Estimation sur 3 simulations de la moyenne annuelle et du maximum à partir des prélèvements journaliers réalisés 1 jour sur 3

Les estimations de la moyenne annuelle sont relativement bonnes avec une erreur relative maximale de 16% pour le BaP et de 28% sur l'ensemble des HAP étudiés.

Avec les simulations 1 et 3, les erreurs relatives sont toutes positives, ce qui traduit une tendance à la surestimation de la concentration moyenne annuelle. A contrario les erreurs relatives avec la 2^{ème} simulation sous estiment plus largement les moyennes annuelles (erreurs relatives toutes négatives) car des pointes importantes ne sont pas pris en compte lors de cet échantillonnage.

Ces résultats montrent qu'un décalage d'un jour sur le premier prélèvement suffit à conduire à des différences significatives d'estimation de la moyenne annuelle. Pour le BbFA par exemple, en fonction du choix du premier jour de prélèvement, l'estimation peut être inférieure de 28% ou supérieure de 19% à la moyenne réelle.

Le choix initial de la simulation, c'est-à-dire le choix du premier jour de prélèvement a donc une forte influence sur les erreurs observées. Il s'agit pourtant d'un facteur non contrôlé lors de l'estimation de la moyenne annuelle étant donné qu'il est déterminé, dès le début des mesures, par le choix du 1^{er} jour de prélèvement.

3.2.2. Échantillonnage n°2 avec 4 strates

Les erreurs d'estimations de la moyenne annuelle sont plus élevées avec cette stratégie qu'avec le premier échantillonnage (1 jour sur 3). L'estimation de la moyenne annuelle a tendance à surestimer la moyenne réelle. Ces surestimations de 24%, 23% et 30% pour le BaP par exemple, sont dues à l'influence des valeurs extrêmes de janvier.

Simulation 1	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	119	119	119	119	119	119
Couv. annuelle (%)	32.6	32.6	32.6	32.6	32.6	32.6
Moyenne estimée (ng.m ⁻³)	0.44	0.57	1.19	0.39	0.43	24.92
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	238.39
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Écart type (ng.m ⁻³)	1.03	2.26	3.36	1.00	1.03	35.34
Biais (ng.m ⁻³)	0.09	0.14	0.20	0.07	0.06	1.60
Erreur relative (%)	24.34	33.19	20.65	22.31	16.18	6.88

Simulation 2	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	121	121	121	121	121	121
Couv. annuelle (%)	33.1	33.1	33.1	33.1	33.1	33.1
Moyenne estimée (ng.m ⁻³)	0.43	0.49	1.12	0.37	0.43	27.68
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	6.82	11.12	13.07	4.97	8.10	467.35
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Écart type (ng.m ⁻³)	0.92	1.36	2.09	0.72	0.98	52.13
Biais (ng.m ⁻³)	0.08	0.07	0.13	0.05	0.05	4.36
Erreur relative (%)	23.07	15.87	12.64	15.50	14.69	18.69

Simulation 3	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	117	117	117	117	117	117
Couv. annuelle (%)	32.0	32.0	32.0	32.0	32.0	32.0
Moyenne estimée (ng.m ⁻³)	0.46	0.63	1.24	0.40	0.48	27.44
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Écart type (ng.m ⁻³)	1.15	2.43	3.39	1.03	1.22	51.52
Biais (ng.m ⁻³)	0.10	0.20	0.25	0.08	0.10	4.12
Erreur relative (%)	29.94	48.06	25.03	25.34	27.43	17.66

Figure 12 : Estimation sur 3 simulations de la moyenne annuelle et du maximum à partir des prélèvements journaliers réalisés 2 jours sur 3 sur la première moitié des 4 strates

3.2.3. Échantillonnage n°3 avec 4 strates

Simulation 1	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	118	118	118	118	118	118
Couv. annuelle (%)	32.3	32.3	32.3	32.3	32.3	32.3
Moyenne estimée (ng.m ⁻³)	0.24	0.25	0.73	0.23	0.29	18.98
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	2.69	2.99	7.79	2.09	2.42	168.42
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Écart type (ng.m ⁻³)	0.39	0.48	1.17	0.36	0.38	22.50
Biais (ng.m ⁻³)	-0.11	-0.17	-0.26	-0.09	-0.08	-4.34
Erreur relative (%)	-31.17	-40.33	-25.85	-26.73	-22.32	-18.63

Simulation 2	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	118	118	118	118	118	118
Couv. annuelle (%)	32.3	32.3	32.3	32.3	32.3	32.3
Moyenne estimée (ng.m ⁻³)	0.26	0.29	0.82	0.26	0.30	19.84
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	2.00	2.96	7.74	2.16	2.02	139.77
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Écart type (ng.m ⁻³)	0.39	0.55	1.22	0.37	0.38	21.87
Biais (ng.m ⁻³)	-0.09	-0.13	-0.17	-0.06	-0.07	-3.48
Erreur relative (%)	-25.58	-30.49	-17.17	-19.32	-19.28	-14.93

Simulation 3	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	117	117	117	117	117	117
Couv. annuelle (%)	32.0	32.0	32.0	32.0	32.0	32.0
Moyenne estimée (ng.m ⁻³)	0.28	0.32	0.84	0.27	0.31	21.12
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	2.69	2.99	7.79	2.16	2.42	168.42
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Écart type (ng.m ⁻³)	0.45	0.58	1.39	0.42	0.43	25.84
Biais (ng.m ⁻³)	-0.07	-0.11	-0.15	-0.05	-0.06	-2.20
Erreur relative (%)	-20.88	-25.53	-14.68	-16.70	-17.20	-9.44

Figure 13 : Estimation sur 3 simulations de la moyenne annuelle et du maximum à partir des prélèvements journaliers réalisés 2 jours sur 3 sur la deuxième moitié des 4 strates

Les erreurs d'estimation sont comparables à la deuxième stratégie d'échantillonnage et l'estimation est donc moins performante que la 1^{ère} stratégie d'échantillonnage. L'ensemble des simulations pour cette stratégie sous estime la moyenne annuelle (31%, 26% et 21% pour le BaP) étant donné que les valeurs extrêmes de janvier ne sont pas compris dans l'échantillonnage.

3.2.4. Échantillonnage n°4 avec 2 strates

Les erreurs d'estimation de la moyenne n'excèdent pas 16 % pour l'ensemble des HAP étudiés avec cette méthode. En effet, les erreurs pour le BaP sont de 5%, 8% et 5% suivant les simulations ce qui est faible au vu des autres simulations. Aucune tendance à la surestimation ou à la sous estimation des moyennes n'est observée. Les maxima ont aussi plus de chance d'être convenablement estimés étant donné que la fréquence d'échantillonnage est plus élevée en période froide.

Simulation 1	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	124	124	124	124	124	124
Couv. annuelle (%)	34.0	34.0	34.0	34.0	34.0	34.0
Moyenne estimée (ng.m ⁻³)	0.33	0.41	0.91	0.30	0.36	22.02
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	238.39
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Biais (ng.m ⁻³)	-0.02	-0.02	-0.08	-0.02	-0.01	-1.30
Erreur relative (%)	-5.41	-4.18	-7.68	-6.51	-3.80	-5.57

Simulation 2	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	122	122	122	122	122	122
Couv. annuelle (%)	33.4	33.4	33.4	33.4	33.4	33.4
Moyenne estimée (ng.m ⁻³)	0.33	0.41	0.91	0.30	0.36	22.02
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	6.82	11.12	13.07	4.97	8.10	467.35
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Biais (ng.m ⁻³)	-0.03	-0.06	-0.16	-0.04	-0.03	-1.80
Erreur relative (%)	-7.60	-13.47	-15.65	-11.88	-8.61	-7.72

Simulation 3	BaP	BaA	BbFA	BkFA	IP	TOTAL
Effectif	120	120	120	120	120	120
Couv. annuelle (%)	32.9	32.9	32.9	32.9	32.9	32.9
Moyenne estimée (ng.m ⁻³)	0.33	0.41	0.91	0.30	0.36	22.02
Moyenne réelle (ng.m ⁻³)	0.35	0.42	0.99	0.32	0.37	23.32
Max échantillon (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Max réel (ng.m ⁻³)	9.17	23.17	32.92	9.39	9.22	467.35
Biais (ng.m ⁻³)	0.02	0.04	0.09	0.02	0.01	0.57
Erreur relative (%)	4.60	9.57	9.48	6.55	3.45	2.44

Figure 14 : Estimation sur 3 simulations de la moyenne annuelle et du maximum échantillonnés sur 2 strates. La moyenne estimée correspond au barycentre des moyennes des 2 strates pondérées par la couverture annuelle des 2 strates.

3.2.5. Conclusions sur les échantillonnages systématiques

L'échantillonnage 4 avec 2 strates « période chaude » et « période froide » semble être la méthode la plus fiable en termes d'estimation de la moyenne annuelle pour un échantillonnage systématique. Globalement, l'échantillonnage 4 présente des erreurs relatives en moyenne deux fois plus faibles qu'avec la méthode actuellement mise en place. Pour le BaP, par exemple, la moyenne des erreurs relatives des 3 simulations passe de 11% (avec l'échantillonnage 1) à 5.8% (avec l'échantillonnage 4).

3.3. Échantillonnages aléatoires

Trois plans d'échantillonnage aléatoires ont été mis en place. Afin d'étudier la répartition de l'estimation de la moyenne, 200 tirages d'échantillons ont été simulés pour chacune des stratégies d'échantillonnage. Pour chaque simulation, la moyenne estimée ainsi que l'erreur relative sont calculées pour l'ensemble des HAP étudiés. Il est ensuite possible d'étudier la répartition de ces 200 moyennes estimées et 200 erreurs relatives.

3.3.1. Échantillonnage aléatoire stratifiés : 30 jours aléatoires par trimestre

Pour cet échantillonnage, la moyenne des 200 moyennes estimées par simulation est proche de la moyenne réelle. Par ailleurs plus le nombre de simulations augmente, plus l'estimation de la moyenne tend vers la moyenne réelle. En ce qui concerne le BaP, pour

50% des simulations, la moyenne estimée est comprise entre 0.305 ng.m⁻³ et 0.387 ng.m⁻³ sachant que la moyenne réelle est de 0.350 ng.m⁻³.

La dispersion des moyennes estimées est plus élevée pour le BbFA que pour les autres HAP étudiés. La moyenne du BbFA correspond aussi à la moyenne la plus élevée des 5 HAP avec presque 1 ng.m⁻³.

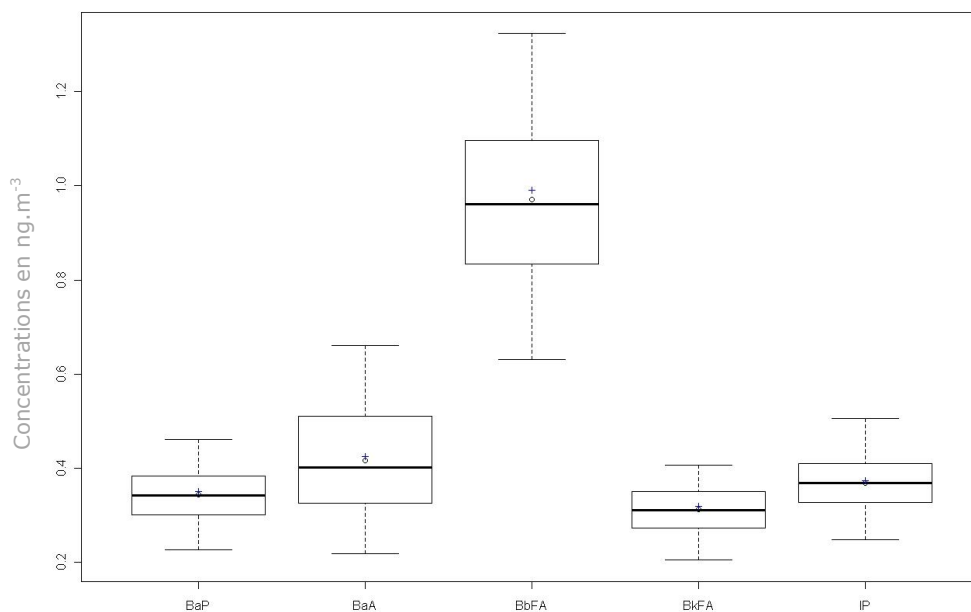


Figure 15 : Répartition par boxplot de l'estimation de la concentration moyenne annuelle pour 5 HAP à partir de 30 prélèvements sélectionnés aléatoirement par trimestre

Les répartitions suivantes concernent les erreurs relatives d'estimation de la moyenne pour les 200 simulations. Il apparaît que l'erreur moyenne sur les 200 simulations de la moyenne estimée est comprise entre 15% et 22% pour l'ensemble des HAP, avec une dispersion des erreurs plus élevée pour le BaA.

Pour 75% des simulations, l'erreur relative se situe en dessous de 20% pour le BaP et le BbFA, 33% pour le BaA, 19% pour le BkFA et le IP, et 13% pour la somme totale des HAP.

Répartition des erreurs relatives	BaP	BaA	BbFA	BkFA	IP	TOTAL
Moyenne des erreurs relatives	12.85	22.59	13.39	12.62	12.79	9.38
Écart type des erreurs relatives	9.09	13.99	9.29	9.03	8.79	6.02
Minimum des erreurs relatives	0.00	0.61	0.13	0.07	0.01	0.08
P25 des erreurs relatives	4.53	10.40	5.25	4.30	4.75	4.56
P50 des erreurs relatives	11.75	22.50	12.96	11.92	12.13	8.91
P75 des erreurs relatives	19.70	33.02	19.66	18.92	18.98	13.04
Maximum des erreurs relatives	38.63	59.95	42.37	38.67	38.97	27.00

Figure 16 : Statistiques sur les erreurs relatives calculées à partir de 30 prélèvements sélectionnés aléatoirement par trimestre

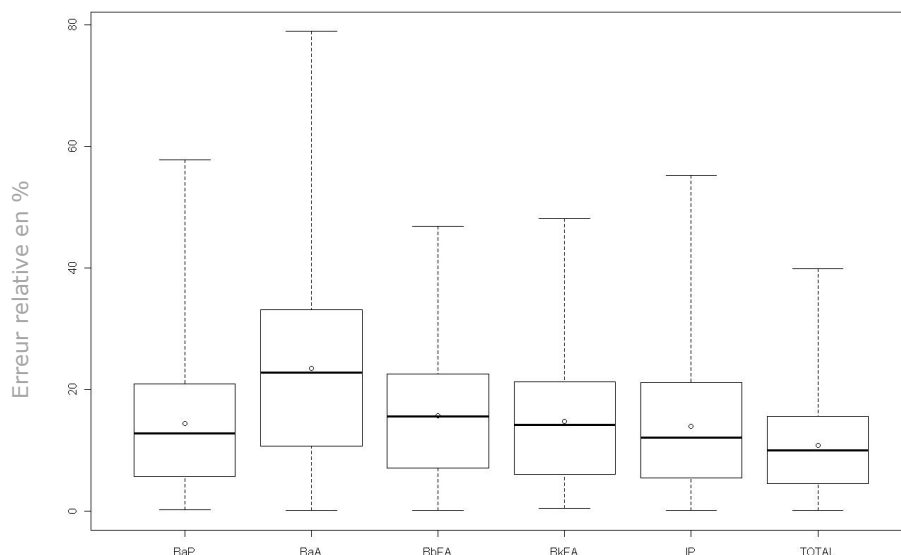


Figure 17 : Répartition par boxplot de l'erreur relative estimée de 5 HAP à partir de 30 prélèvements sélectionnés aléatoirement par trimestre

3.3.2. Échantillonnage stratifié par grappe : 2 fois 15 jours consécutifs sélectionnés aléatoirement dans chaque strate

Pour cet échantillonnage, 2 grappes de 15 prélèvements consécutifs ont été sélectionnées dans chaque trimestre.

Répartition des erreurs relatives	BaP	BaA	BbFA	BkFA	IP	TOTAL
Moyenne des erreurs relatives	31.43	49.04	27.05	29.01	32.43	21.26
Écart type des erreurs relatives	17.12	22.34	17.03	16.98	16.38	12.78
Minimum des erreurs relatives	1.66	7.70	0.27	0.44	0.53	0.05
P25 des erreurs relatives	16.85	31.60	14.07	15.11	19.62	11.26
P50 des erreurs relatives	29.50	42.16	24.42	25.94	29.25	18.31
P75 des erreurs relatives	44.90	69.58	39.52	42.69	43.72	31.05
Maximum des erreurs relatives	69.52	99.99	68.78	66.87	77.22	50.19

Figure 18 : Statistiques sur les erreurs relatives calculées à partir de 30 prélèvements sélectionnés aléatoirement par trimestre

Avec cette méthode d'échantillonnage, les erreurs relatives d'estimation de la moyenne sont élevées et varient entre 20% et 50 %. Les erreurs maximales peuvent atteindre des niveaux aussi très élevés (100% pour le BaA). Par rapport à l'échantillonnage aléatoire précédent, les centiles, l'écart type et le biais moyen sont plus élevés avec cette méthode.

3.3.3. Échantillonnage aléatoire stratifié : 2 saisons

Cette méthode d'échantillonnage correspond à 102 prélèvements en période froide (de mi-octobre à mi-mars) et 21 prélèvements en période chaude (de mi-mars à mi-octobre) sélectionnés aléatoirement. Les prélèvements sélectionnés de manière aléatoire couvrent 33% de l'année.

Répartition des erreurs relatives	BaP	BaA	BbFA	BkFA	IP	TOTAL
Moyenne des erreurs relatives	6.00	9.62	7.26	6.36	6.07	5.14
Écart type des erreurs relatives	4.97	7.36	5.54	5.12	4.81	3.60
Minimum des erreurs relatives	0.03	0.01	0.02	0.01	0.00	0.01
P25 des erreurs relatives	1.72	3.49	2.80	2.03	2.33	2.31
P50 des erreurs relatives	5.11	8.19	5.67	5.06	4.93	4.69
P75 des erreurs relatives	9.15	14.06	11.28	9.78	8.91	7.18
Maximum des erreurs relatives	22.18	36.51	21.97	21.33	24.55	17.73

Figure 19 : Statistiques sur les erreurs relatives calculées à partir de 102 prélèvements sélectionnés aléatoirement en période froide et 21 en période chaude

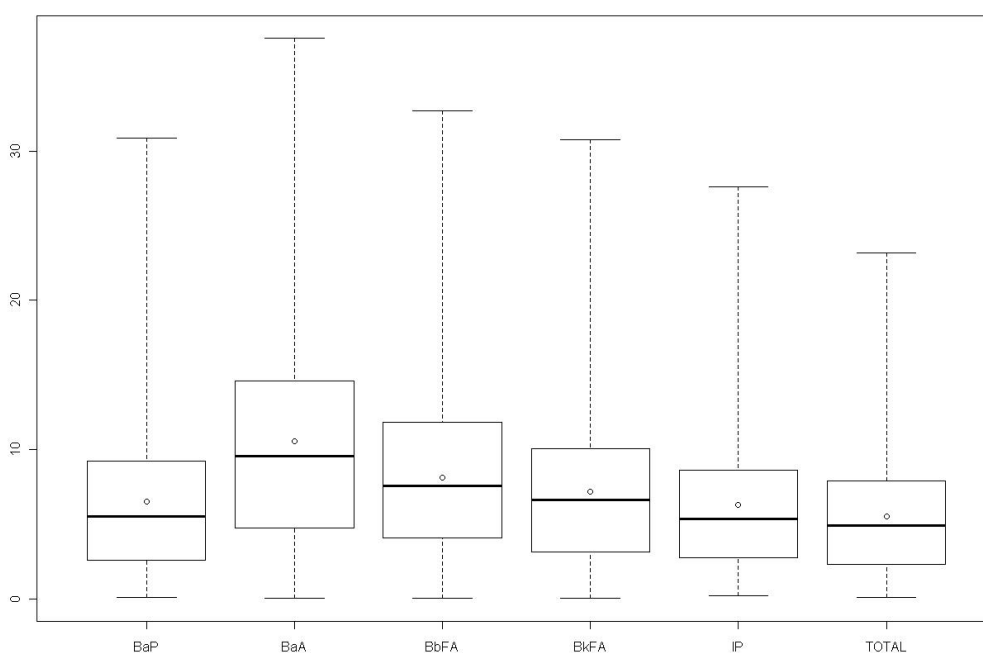


Figure 20 : Répartition par boxplot de l'erreur relative estimée à partir de 30 prélèvements sélectionnés aléatoirement par trimestre

Les erreurs moyennes d'estimation n'excèdent pas 10% pour l'ensemble des 5 HAP étudiés (9.6% pour le BaA). Les erreurs maximales sont aussi relativement faibles avec un maximum de 36% pour le BaA.

D'après les statistiques d'erreurs, cette dernière méthode présente les meilleures estimations en termes d'erreurs relatives parmi les échantillonnages aléatoires testés. En ce qui concerne le BaP, pour 50% des simulations, les erreurs d'estimation de la moyenne sont comprises entre 2% et 9%.

4. Discussion

La base de données des HAP constituée au cours de l'année 2009 sur le site urbain de référence de Lyon centre permet d'étudier de manière précise la variabilité temporelle des HAP avec 97% de représentativité.

4.1. Conclusions des comparaisons des différentes méthodes d'échantillonnages

Il convient avant tout de préciser qu'il est difficile de comparer les méthodes d'échantillonnage systématiques et les méthodes d'échantillonnage aléatoire. En effet, les méthodes aléatoires peuvent être étudiées à l'aide de simulations qui permettent une vue d'ensemble des différentes possibilités d'estimation de la moyenne annuelle. A contrario, les méthodes systématiques n'offrent seulement que quelques cas d'études bien précis.

Néanmoins certains points ressortent de la comparaison des différentes méthodes d'échantillonnage testées :

- L'année 2009 présente des pics hivernaux élevés de concentrations de HAP et de particules par rapport aux années précédentes. Les prélèvements de 2009 permettent donc d'évaluer les différentes méthodes d'échantillonnage pendant une année présentant des spécificités propres qui peuvent possiblement dégrader le score de l'ensemble de ces méthodes.
- La méthode actuellement mise en place d'un jour sur 3 (échantillonnage 1) présente des résultats convenables au vu des erreurs relatives comparées aux autres méthodes (erreur relative inférieure à 28% pour l'ensemble des HAP et inférieure à 17 % pour le BaP). Il convient néanmoins de préciser que des différences non négligeables sont observées entre les 3 simulations, ces dernières ne différant que par le choix du premier jour de prélèvement de l'année.
- L'importante variabilité des données et plus particulièrement les fortes concentrations relevées au moins de janvier impliquent que l'échantillonnage par grappe n'est pas toujours approprié dans le cadre des HAP. Pour une meilleure estimation de la moyenne annuelle, les prélèvements sélectionnés de manière discontinue (échantillonnage 5) sont à favoriser par rapport à l'échantillonnage par grappe (échantillonnage 6).
- L'échantillonnage de 2 jours sur 3 par moitié de trimestre (échantillonnages 2 et 3) semble moins performant que l'échantillonnage d'un 1 sur 3 tout au long de l'année. De grosses différences de variabilité peuvent, en effet, être observées entre le début et à la fin du trimestre (ex : 1^{er} trimestre entre janvier 2009 et mars 2009).
- Les tests des différents échantillonnages montrent que les erreurs relatives sont plus faibles avec les méthodes d'échantillonnage à 2 strates (période chaude de mi-mars à mi-octobre et période froide de mi-octobre à mi-mars). Les erreurs relatives (quasiment toutes inférieures à 10%) sont du même ordre de grandeur entre la méthode systématique et la méthode aléatoire. Ces méthodes permettent ainsi d'améliorer l'estimation de la concentration moyenne annuelle même en cas de hautes concentrations hivernales. Rappelons que l'estimation de la moyenne annuelle est calculée à partir du barycentre des moyennes de chaque strate pondérées par leur couverture annuelle respective ; et que, par conséquent, ces méthodes d'estimation nécessitent de ne plus utiliser la moyenne arithmétique pour l'estimation de la moyenne annuelle.

4.2. Comparaison à d'autres études

Le groupe de travail national *Plans d'échantillonnage et reconstitutions de données* s'est intéressé à une série temporelle de concentrations de HAP provenant d'Airparif, dans le but de déterminer le plan d'échantillonnage adapté à une couverture annuelle de **14%** (et non 33% comme cette étude). Les données HAP datent de 1998 et concernent un site urbain de Paris sous l'influence du trafic routier. La série de mesures couvrait 75% de l'année 1998 : elles débutaient le 10 février 1998 et se terminaient le 31 décembre 1998 avec une plage de données manquantes entre le 1^{er} juillet et le 27 juillet.

Comme ce qui a été montré plus haut, les données d'AirParif mettent aussi en évidence que l'erreur d'estimation commise avec un échantillonnage régulier (1 jour sur 6) uniformément repart, est généralement plus faible qu'avec un échantillonnage plus fréquent mais restreint à une moitié de trimestre (1 jour sur 3 pendant une moitié de trimestre).

En ce qui concerne les méthodes d'échantillonnage aléatoires, les deux séries montrent que l'échantillonnage par grappes de 7 et 15 jours est moins performant qu'un échantillonnage avec des mesures journalières ponctuelles.

Pour ce qui est du BaP, la série de mesures d'AirParif montre que l'échantillonnage d'un jour sur 6 sur l'année entière conduit à une erreur d'estimation de la moyenne annuelle inférieure à 12% dans 5 cas sur 6 et de 33% pour l'un des échantillons.

Ces résultats sont relativement comparables aux résultats obtenus sur Lyon avec une fréquence des prélèvements 2 fois plus élevée (couverture annuelle de 33%) : les erreurs relatives sont en effet de 10%, 16%, et 6%. Il en va de même pour les méthodes d'échantillonnage de prélèvements ponctuels. En effet, la moyenne des erreurs relatives estimée sur les 200 simulations sont comparables sur les 2 séries de données pour la plupart des HAP (15%, 9%, 13%, 11%, 12% pour les données d'AirParif contre 13%, 23%, 13%, 13%, 13% pour les données de Lyon centre) alors que la fréquence des prélèvements est 2 fois plus élevée sur le site de Lyon centre (15 prélèvements aléatoires contre 30 prélèvements aléatoires dans chaque trimestre).

Cette différence s'explique par un écart type plus élevée des données de Lyon centre (0.793 ng.m^{-3} contre 0.290 ng.m^{-3}). Ce paramètre plus élevé sur Lyon illustre une variabilité plus importante des données et implique donc la nécessité d'une fréquence d'échantillonnage plus élevée pour arriver aux mêmes précisions. En effet, plusieurs années de surveillance en Rhône-Alpes ont montré que les concentrations moyennes de HAP y sont plus élevées, notamment en raison du recours au mode de chauffage au bois en hiver plus répandu dans cette région.

L'absence de données pour le mois de janvier (mois d'hiver le plus impacté en 2009 à Lyon par des concentrations élevées) sur le site d'AirParif peut aussi contribuer à la sous estimation de l'écart type annuel.

4.3. Contraintes techniques

Les comparaisons des différentes méthodes d'échantillonnage ont été réalisées indépendamment des contraintes techniques de terrain. Les conclusions exposées jusqu' alors ne prennent pas en compte l'aspect technique qui doit être considéré dans le choix final du plan d'échantillonnage. Il convient de préciser que les plans d'échantillonnages systématiques sont, d'un point de vue technique, plus facile à mettre en place qu'un plan d'échantillonnage aléatoire. En effet, le plan d'échantillonnage systématique est déterminé de façon régulière et permet de prévoir longtemps à l'avance les déplacements sur le terrain.

5. Conclusion

Durant l'année 2009, COPARLY a mis en œuvre une série de mesures des Hydrocarbures Aromatiques Polycycliques visant à prélever un échantillon d'air (gaz et particules) de 24h tous les jours de l'année. La mise en œuvre de ces mesures avait pour principal but de constituer une série annuelle complète de prélèvements de 19 HAP à destination d'un travail statistique d'étude de la méthode d'échantillonnage optimale permettant d'accéder à une estimation de la concentration moyenne annuelle.

Plusieurs plans d'échantillonnage ont été testés et comparés sur le site urbain de Lyon centre dans le courant de l'année 2009 pour les 5 HAP ciblés par la directive 2004/107/CE dont le benzo[a]pyrène. Le profil mensuel de cette station est typique d'un site urbain avec des pointes en hiver et des concentrations faibles pendant une grande partie de l'été. Celui d'un site de proximité industrielle comme Vénissieux Village étant différent, car les pointes peuvent y être enregistrées à tout moment de l'année, les conclusions de la présente étude ne pourront pas être considérées comme transposables à une surveillance industrielle.

Cette étude a montré que la méthode d'échantillonnage actuellement utilisée pour l'estimation de la concentration annuelle des HAP dans l'air ambiant (échantillonnage systématique régulier d'un jour sur 3) est un bon compromis entre les erreurs dues à la méthode d'échantillonnage (entre 4% et 29% sur l'ensemble des HAP étudiés, et entre 6% et 16% pour le BaP), les contraintes techniques et logistiques, ainsi que le respect des contraintes réglementaires. Cette méthode n'est malgré tout pas complètement optimale compte tenu des fortes variations des concentrations pendant l'hiver (novembre à mars, et particulièrement durant janvier 2009) et pouvant engendrer un risque de sous-estimation en ne mesurant pas les jours pollués.

Les méthodes d'échantillonnage testées réduisant les erreurs relatives sont les plans d'échantillonnage divisés en 2 périodes ou « strates » (5 mois de période « froide » et 7 mois de période « chaude »). A nombre d'échantillons et coût constants, elles permettent d'observer des erreurs relatives en moyenne deux fois plus faibles qu'avec celle actuellement mise en place. Elles consistent à augmenter le nombre de prélèvements lorsque les variations de concentrations sont plus importantes (période froide) et à le réduire lorsque les concentrations sont stables (période chaude). 102 prélèvements sont ainsi échantillonnés entre mi-octobre et mi-mars et 21 sont effectués entre mi-mars et mi-octobre.

Ces prélèvements peuvent être choisis suivant un échantillonnage systématique ou un échantillonnage aléatoire, mais peu de différences sont notables entre les 2 types d'échantillonnage pour ce type de stratification. En effet, les erreurs relatives varient entre 3% et 16% pour l'échantillonnage systématique alors qu'elles varient entre 5% et 10% pour l'aléatoire. Même s'il n'apparaît pas de grande différence entre ces deux méthodes, il est d'un point de vue technique, plus simple et moins coûteux en temps de personnel, de mettre en place un échantillonnage systématique plutôt qu'un échantillonnage aléatoire.

Néanmoins, il est important de préciser que l'estimation de la concentration moyenne annuelle avec ces méthodes ne peut pas être calculée comme moyenne arithmétique de l'ensemble des prélèvements, étant donné que la durée des deux périodes n'est pas égale. L'estimation de la concentration moyenne annuelle doit être calculée à partir du barycentre des moyennes des 2 strates pondérées par leur couverture annuelle.

Ces méthodes, qui présentent des erreurs relatives faibles et meilleures que les méthodes en vigueur pour l'estimation des concentrations moyennes, peuvent donc être privilégiées pour la surveillance de l'exposition moyenne des populations des agglomérations en Rhône-Alpes si les contraintes techniques et les textes réglementaires le permettent. En revanche, concernant la surveillance en proximité industrielle, il

conviendrait de réaliser une étude équivalente avant d'envisager une modification de la méthode d'échantillonnage actuelle (1 prélèvement tous les 3 jours).

Enfin, il serait intéressant de poursuivre ces travaux pour tester des méthodes d'échantillonnage plus poussées. En effet, compte-tenu de la bonne qualité de la série de données recueillie en 2009 (97% de représentativité), il semble possible par exemple de fixer une précision sur l'estimation de la moyenne annuelle par la théorie des sondages. Cette dernière permettrait de déterminer le nombre optimal de grappes par strates et d'améliorer éventuellement les résultats d'erreur sur les méthodes à 2 strates (période chaude et période froide).

Bibliographie

ADEME, Guide d'élaboration de plan d'échantillonnage temporel et de reconstitution de données.

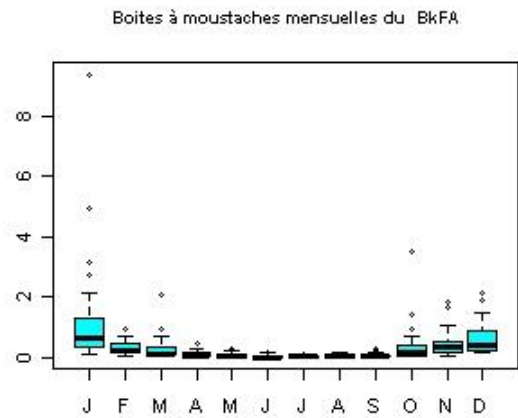
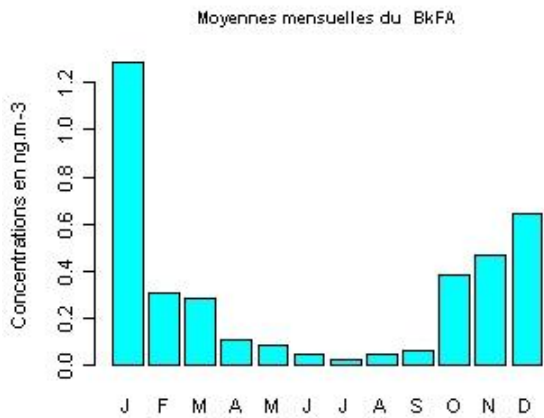
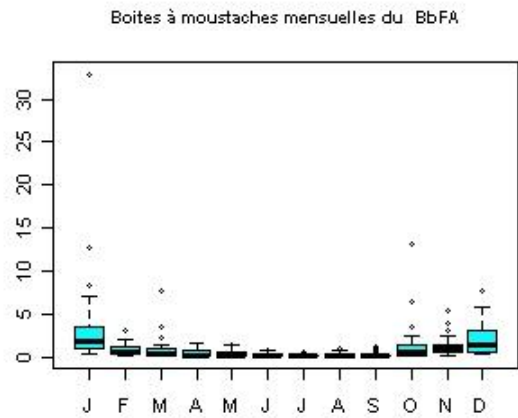
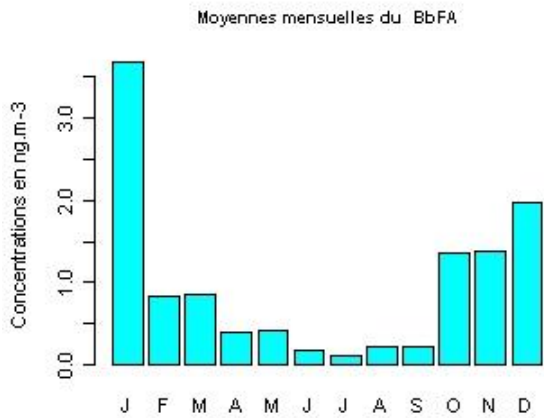
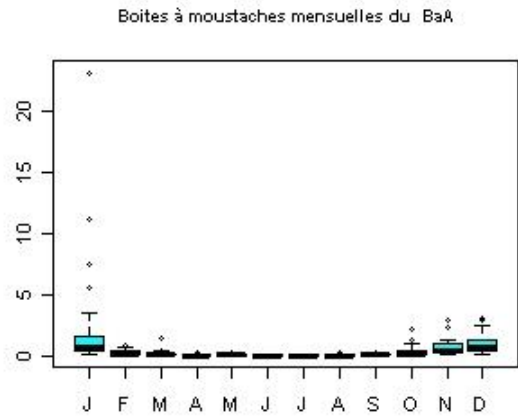
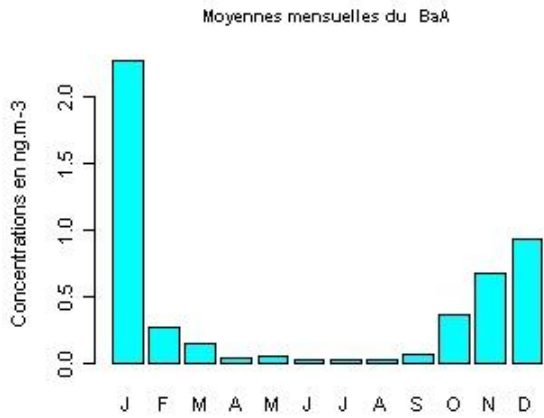
COPARLY, Surveillances des HAP dans l'environnement au voisinage d'un émetteur industriel – 2006/2007

COPARLY, ASCOPARG, Bilan de la surveillance des HAP en air ambiant dans le Rhône et l'Isère - 2002/2006.

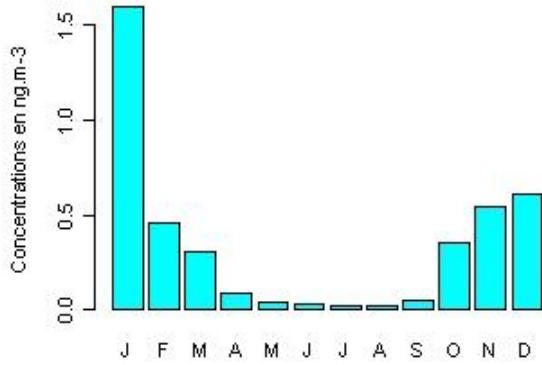
COPARLY, Lexique des abréviations pour les HAP.

SUPAIR, ASCOPARG, Rapport activité 2007.

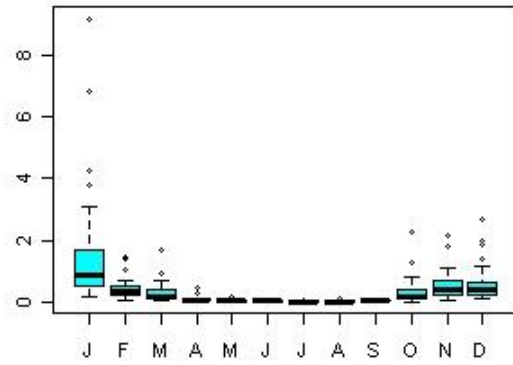
Annexe 1 : Profils mensuels des HAP cités dans la directive



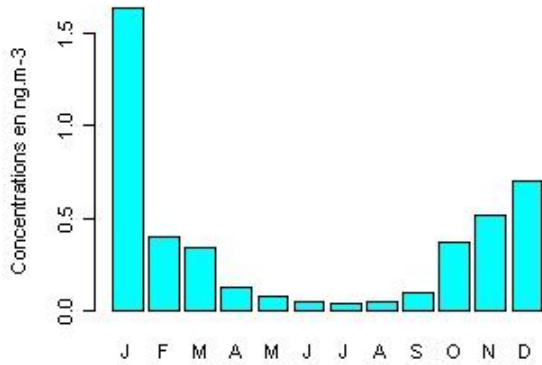
Moyennes mensuelles du BaP



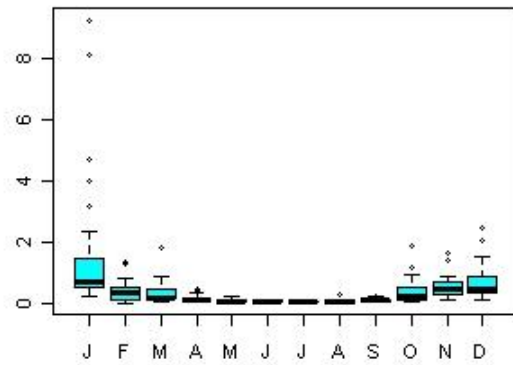
Boites à moustaches mensuelles du BaP



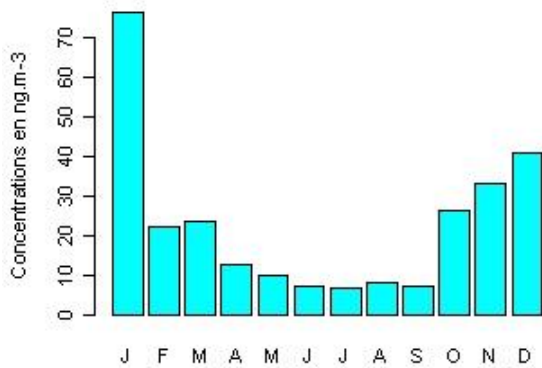
Moyennes mensuelles du IP



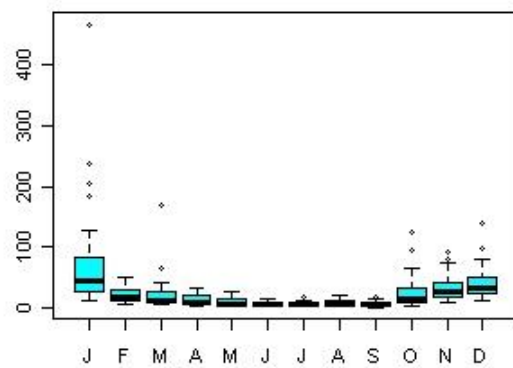
Boites à moustaches mensuelles du IP



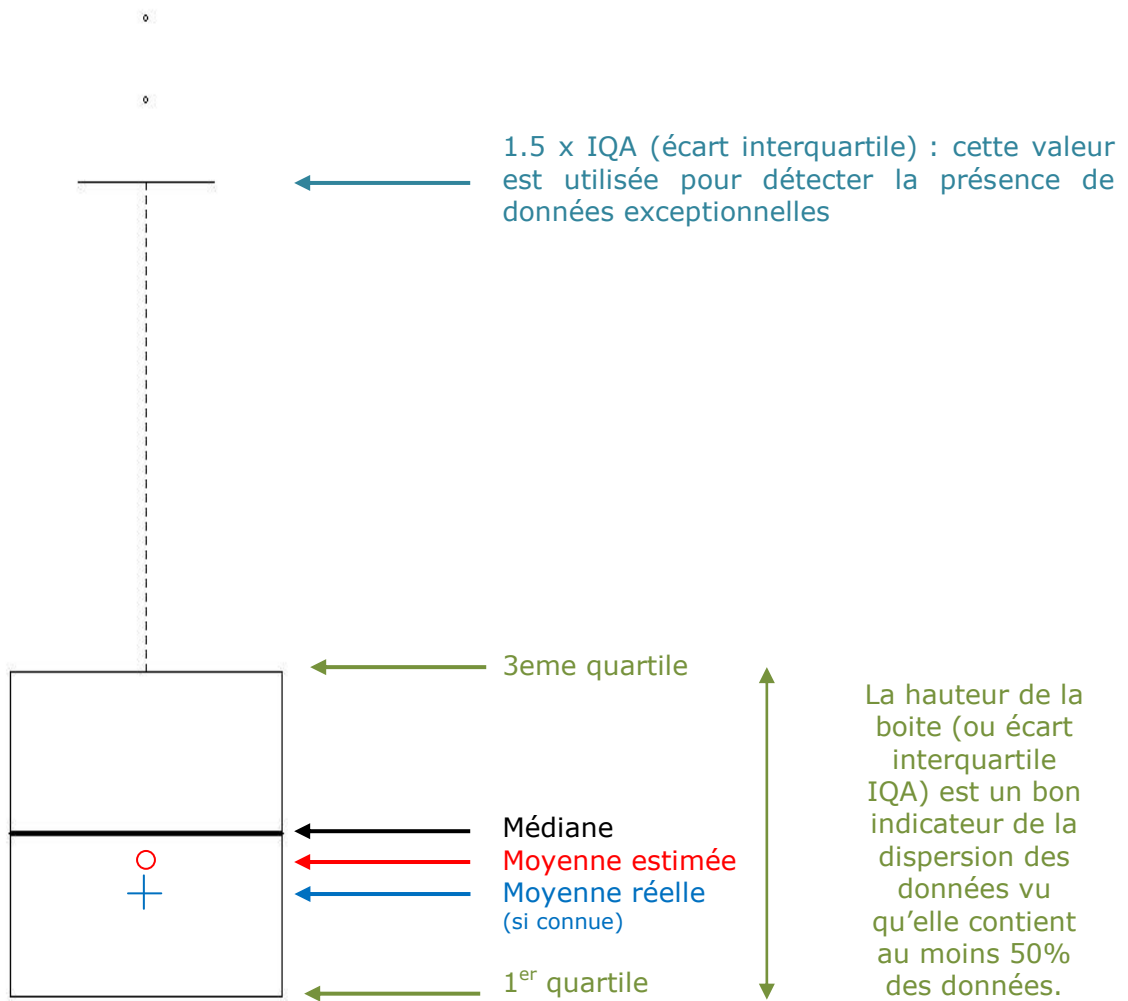
Moyennes mensuelles du TOTAL



Boites à moustaches mensuelles du TOTAL



Annexe 2 : Représentation par Boite à moustache (ou Box Plot)



Annexe 3 : Influence des maxima dans l'estimation de la concentration moyenne

