



## Laboratoire Central de Surveillance de la Qualité de l'Air



### **Caractérisation du comportement spatio-temporel des COV à partir de l'exploitation des données canisters**

Stéphane SAUVAGE, Nadine LOCOGE

Novembre 2008

Version finale





## PREAMBULE

### **Le Laboratoire Central de Surveillance de la Qualité de l'Air**

**Le Laboratoire Central de Surveillance de la Qualité de l'Air est constitué de laboratoires de l'Ecole des Mines de Douai, de l'INERIS et du LNE. Il mène depuis 1991 des études et des recherches finalisées à la demande du Ministère chargé de l'environnement, sous la coordination technique de l'ADEME et en concertation avec les Associations Agréées de Surveillance de la Qualité de l'Air (AASQA). Ces travaux en matière de pollution atmosphérique supportés financièrement par la Direction des Préventions des Pollutions et des Risques du Ministère de l'Énergie, de l'Écologie, du Développement Durable et de l'Aménagement du Territoire sont réalisés avec le souci constant d'améliorer le dispositif de surveillance de la qualité de l'air en France en apportant un appui scientifique et technique aux AASQA.**

**L'objectif principal du LCSQA est de participer à l'amélioration de la qualité des mesures effectuées dans l'air ambiant, depuis le prélèvement des échantillons jusqu'au traitement des données issues des mesures. Cette action est menée dans le cadre des réglementations nationales et européennes mais aussi dans un cadre plus prospectif destiné à fournir aux AASQA de nouveaux outils permettant d'anticiper les évolutions futures.**



**ECOLE DES MINES DE DOUAI**

**DEPARTEMENT CHIMIE ET ENVIRONNEMENT**

**Caractérisation du comportement  
spatio-temporel des COV à partir  
de l'exploitation des données canisters**

**Convention n°0001190**

**Stéphane SAUVAGE, Nadine LOCOGE**

**Novembre 2008**



# SOMMAIRE

|  |           |
|--|-----------|
| <b>RESUME .....</b>  | <b>7</b>  |
| <b>1. INTRODUCTION.....</b>  | <b>10</b> |
| <b>2. PRESENTATION DE LA BASE CANIST'AIR .....</b>   | <b>11</b> |
| 2.1 Evolution du nombre d'échantillons entre les deux phases d'étude .....                               | 11        |
| 2.2 Représentativité temporelle des observations.....  | 13        |
| <b>3. COMPARAISON DES DES SITES DE MESURES .....</b>   | <b>15</b> |
| <b>3.1 Analyse univariée .....</b>   | <b>15</b> |
| 3.1.1 Extraction des observations atypiques .....  | 15        |
| 3.1.2 Profils saisonniers d'espèces caractéristiques .....   | 18        |
| 3.1.3 Comparaison des sites pour les espèces majeures .....  | 22        |
| 3.1.4 Focus sur les sites de Nancy et Charles III .....  | 25        |
| <b>3.2 Analyse bi-variée : vers l'identification de sources .....</b>                                    | <b>26</b> |
| 3.2.1 Méthode .....  | 26        |
| 3.2.2 Résultats .....  | 27        |
| <b>3.3 Etude de ratios d'alcane, âge des masses d'air .....</b>  | <b>29</b> |
| 3.3.1 Méthode .....  | 29        |
| 3.3.2 Résultats .....  | 31        |
| <b>4. IDENTIFICATION DES PROFILS DE SOURCES PAR APPLICATION DU<br/>MODELE SOURCES-RECEPTEUR PMF.....</b> | <b>35</b> |
| <b>4.1 Méthode .....</b>   | <b>36</b> |
| 4.1.1 Principe général du PMF .....  | 36        |
| 4.1.2 Principe d'itération du modèle .....   | 36        |
| 4.1.3 Données d'entrée du modèles.....   | 38        |
| 4.1.4 Choix du modèle d'incertitude .....  | 38        |
| 4.1.5 Traitement des valeurs extrêmes .....  | 39        |
| 4.1.6 Détermination du nombre de facteurs .....  | 40        |
| 4.1.7 Valeur ajoutée du modèle PMF par rapport à l'Analyse en Composante Principale<br>(ACP) 40          |           |
| <b>4.2 Résultats .....</b>   | <b>41</b> |
| 4.2.1 Site de Nancy Charles III .....  | 41        |
| 4.2.2 Site de Mairie .....   | 48        |
| <b>5 CONCLUSION .....</b>  | <b>55</b> |
| <b>6 BIBLIOGRAPHIE .....</b>   | <b>59</b> |
| <b>ANNEXES .....</b>   | <b>61</b> |





## RESUME

L'objectif de cette étude est d'étudier l'intérêt des mesures des Composés Hydrocarbonés Non Méthaniques (HCNM) par canisters à partir de l'exploitation scientifique des données de concentration en HCNM mesurées par les AASQA depuis 2005. L'étude a été réalisée en deux temps :

Une première phase pour laquelle le rapport a été émis en novembre 2007. Il s'agissait d'une analyse descriptive de la base après consolidation. Suite à l'examen de la représentativité temporelle des mesures, l'analyse a permis de comparer les sites suivant leur typologie avec d'autres sites nationaux et étrangers. D'une manière générale il est apparu que les teneurs mesurées sur les sites français de type urbain sont en excellent accord avec celles mesurées dans plusieurs agglomérations européennes. Les sites français de type « trafic » présentent des concentrations supérieures mais qui restent néanmoins bien en deçà de ce qui est mesuré sur les agglomérations d'Asie ou d'Amérique. Quel que soit le type de site, l'éthane et l'éthylène sont les composés majoritaires. Le butane est parmi les 5 composés prépondérants pour la plupart des sites. La présence de propane et de toluène traduirait plutôt un site de type « urbain » alors que la présence d'acétylène et d'isopentane caractériserait un site de type « Trafic ». En classant les espèces suivant leur réactivité (indirectement suivant leur potentiel de formation d'ozone), les cinq premiers composés sont l'éthylène, le propène, l'isoprène, les m,p-xylènes et le 1,2,4-triméthylbenzène indépendamment de la typologie des sites étudiés.

La deuxième phase de l'étude a eu pour but de pousser plus loin l'analyse notamment pour dégager des informations concernant la typologie des sites et les profils de sources urbaines. Il s'agissait d'étudier différents ratios d'espèces avant de mettre en œuvre une approche multi-variée avec le modèle sources-récepteur Positive Matrix Factorization (PMF).

La base de données a été complétée entre les deux phases avec les mesures de l'année 2008. L'analyse descriptive préalable n'a été menée que sur 8 des 17 sites présentant au moins 30 échantillons. Elle a permis notamment d'étudier les profils saisonniers de variation des concentrations pour caractériser le comportement de quelques espèces. La comparaison des sites espèce par espèce ou à partir de ratios spécifiques a permis de mettre en évidence quelques particularités : Le site de Nancy Charles III (AIRLOR) présente des concentrations

élevées en isopentane, en pentane, en styrène et en xylènes. Le site de Mairie de Reims (ATMO Champagne Ardennes) est marqué par des teneurs très fortes en butane, en 1,4-dichlorobenzène ou en trichloroéthylène. Au regard d'espèces caractéristiques, le site de Clémenceau (ASPA) est le site le plus exposé aux émissions automobiles. Le site industriel de Sochaux montre des concentrations en xylènes indiquant un usage de solvants. Une approche estimant l'âge photochimique des observations a montré que les sites de la base canist'air sont influencés quasi exclusivement par des masses d'air jeunes. D'après les données disponibles, la configuration des sites et le type d'échantillonnage ne permet pas de distinguer l'influence de sources locales des apports de masses d'air plus lointaines.

L'analyse multi-variée demande un nombre d'échantillons suffisant pour obtenir des résultats stables et significatifs. Aussi, seuls les sites de Mairie et de Nancy Charles III ont pu être considérés pour une application du modèle sources-récepteur PMF. 2 facteurs communs ont été modélisés. Le premier a été associé à la source « chauffage résidentiel » et contribue majoritairement (37%) aux concentrations mesurées sur les deux sites. Le deuxième correspond au profil de la source automobile, les contributions sont de 13% et de 25% respectivement pour Nancy et Mairie. Les autres profils sont plus difficiles à identifier, ils semblent correspondre à des sources de types évaporation de carburants, fuite de gaz ou usage de solvants. L'analyse des contributions, en suivi chronologique et par secteur de vent, montre le caractère particulier, intense et localisé des sources associées à ces profils.

La démarche d'analyse de données plus avancée a mis en lumière plusieurs écueils concernant la base étudiée. L'approche statistique pour ce type d'espèces nécessite d'une part, un nombre de données important pour une bonne significativité des résultats et, d'autre part, une bonne représentativité temporelle des données pour prendre en compte l'évolution temporelle des processus de production et de déplétion des espèces. Seuls les sites de Nancy et de Mairie (tout juste suffisant) ont permis ce type d'approche. Par ailleurs la fréquence d'échantillonnage de 24 heures intègre les variations des concentrations au cours d'une journée liées aux différents processus intervenant dans l'atmosphère, ce qui limite la détection de variabilité inter espèces pouvant conduire à l'identification plus fine de sources.

Toutefois, les résultats obtenus montrent l'intérêt de ce type d'observations. Une réflexion doit être menée pour identifier les objectifs des

mesures de ces espèces par canisters, objectifs qui peuvent être l'étude d'impact, la surveillance à long terme, l'étude d'épisodes ou encore la connaissance des relations sources-récepteurs. Suivant les objectifs envisagés, la stratégie d'échantillonnage pourra alors être revue et optimisée suivant le tableau ci-après.

### Stratégies d'échantillonnage suivant les objectifs de la mesure

| Objectifs   | Echantillonnage           |                                 | Typologie site  | Espèces                              |
|---|---------------------------|---------------------------------|---|--------------------------------------|
|   | Durée                     | Fréquence                       |   |                                      |
| Exposition de la population                                     | plusieurs heures          | multi hebdo, régulier           | Urbain/urbain de fond                                 | Réglementées + Recommandées OMS      |
| Représentativité/typologie de site                              |                           |                                 | Tout type   | Réglementées + précurseurs           |
| Etude d'impact  | Courte (suivant activité) | avant, après                    | Zone d'impact   | Traceurs de l'activité ciblée        |
| Etude d'épisode   | 1 heure                   | plusieurs/jours                 | Urbain, péri-urbain                                   | Précurseurs, secondaires (oxygénées) |
| Surveillance à long terme                                       | ≥4h (fixe)                | hebdo, régulier                 | péri-urbain, urbain de fond, rural                    | Majoritaires, traceurs               |
| Relation sources Récepteur, Confrontation inventaires émissions | 1 heure                   | plusieurs/jours (nb données >>) | péri-urbain, urbain de fond (suivant récepteur ciblé) | Majoritaires, traceurs               |

## 1. INTRODUCTION

Les Composés Hydrocarbonés Non Méthaniques (HCNM) sont des espèces de premières importances puisqu'ils ont un impact direct ou indirect sur l'homme et l'écosystème. Ils jouent notamment des rôles clés dans les processus de formation de polluants secondaires comme l'ozone ou les aérosols organiques secondaires. En conséquence la surveillance de ces espèces est cruciale d'une part, pour évaluer le niveau d'exposition de la population et, d'autre part, pour établir les liens entre les émissions et les événements de pollution. Il s'agit notamment de pouvoir confronter les inventaires d'émission et leurs évolutions avec les niveaux mesurés dans l'air ambiant, de participer à la compréhension des processus de formation de polluants secondaires pour améliorer les modèles de prévision d'épisode ou pour envisager des stratégies efficaces de réduction de la pollution.

Depuis 2005, des mesures des 31 espèces précurseurs de l'ozone sont réalisées par canister dans 17 sites de 8 Associations Agréées de Surveillance de la Qualité de l'Air (AASQA). La méthode de mesure consiste en un prélèvement de 24 heures par canister et une analyse différée au Laboratoire Interrégionale (LIC). Les échantillonnages sont réalisés par campagne ou de manière régulière. Ces mesures sont compilées dans la base de données Canist'air.

Les objectifs de cette étude LCSQA sont l'exploitation scientifique des données de la base canist'air et l'étude de l'apport potentiel au niveau national de ce type de mesure. L'étude a été menée en deux temps. Une première phase plus descriptive a été présentée dans le rapport LCSQA 2007 (étude n°29). La deuxième phase a mis en œuvre des outils d'analyse de données plus avancés pour étudier la représentativité des sites suivant leur typologie et pour identifier les principales sources à l'origine des HCNM mesurés.

Après une présentation de la base, ce rapport présente les résultats obtenus pour les méthodes d'exploitation déployées allant de l'analyse uni-variée à l'analyse multi-variée.

## 2. PRESENTATION DE LA BASE CANIST'AIR

### 2.1 Evolution du nombre d'échantillons entre les deux phases d'étude

La base de données Canist'air transmise fin 2008 pour réaliser cette étude comprend les résultats d'analyse des échantillons prélevés sur les sites de 8 AASQA : AIRLOR, ATMO-CA, ASQAB, AERFORM, ASPA, ARPAM, et ESPOL. Les analyses ont été réalisées par le Laboratoire Interrégional de Chimie (LIC). La liste des sites et leur typologie sont présentés dans le Tableau 1. Leur situation géographique est représentée Figure 1.

**Tableau 1 : Liste des sites de prélèvement de la base canist'air**

| AASQA                   | Site                          | Typologie du site | Type de voie |
|-------------------------|-------------------------------|-------------------|--------------|
| ARPAM                   | Audincourt centre             | trafic            | rue          |
| ASQAB                   | Besançon-Mégevand             | trafic            | canyon       |
| ASQAB                   | Besançon-Planoise             | urbaine           | voie large   |
| ESPOL                   | Carling                       | industrielle      | rue          |
| ESPOL                   | Cité Jeanne d'Arc             | industrielle      | rue          |
| ASPA                    | Clémenceau                    | trafic            | voie large   |
| ESPOL                   | Freyming-Merlebach            | urbaine           | rue          |
| ESPOL                   | L'Hôpital Bois-Richard        | industrielle      | rue          |
| ESPOL                   | L'Hôpital Puits II            | industrielle      | rue          |
| ESPOL                   | L'Hôpital Rond-point Sabatier | industrielle      | voie large   |
| ESPOL                   | L'Hôpital-Mairie              | industrielle      | rue          |
| ATMO-Champagne Ardennes | Mairie                        | urbaine           | rue          |
| AERFORM                 | Metz Borny                    | urbaine           | rue          |
| AIRLOR                  | Nancy-Charles III (30034)     | urbaine           | canyon       |
| ASPA                    | Robertsau                     | urbaine           | rue          |
| ARPAM                   | Sochoux Ateliers              | industrielle      | rue          |
| ASPA                    | STG Nord                      | urbaine           | voie large   |

D'autres mesures sont également réalisées par d'autres AASQA : Atmo Rhône Alpes, Air Languedoc Roussillon. Elles ne seront pas traitées dans cette seconde phase de l'étude.



**Figure 1 : Situation géographique des sites de la base canist'air. Les points rouges indiquent les sites pour lequel plus de 30 prélèvements ont été réalisés. Le site du Donon est un site rural géré par l'ASPA et intégrant le dispositif national de MERA**

La Figure 2 présente l'évolution du nombre d'échantillons par site de la base de données entre les 2 phases d'étude, à savoir de décembre 2006 à décembre 2008. Le nombre d'échantillons reste inchangé pour les sites de Metz Borny de AREFORM et pour les 5 sites d'ESPOL. Le site industriel de Sochaux est renseigné uniquement pour la deuxième phase de l'étude. Sur les autres sites, les échantillonnages ont continué avec, approximativement, la même fréquence pour doubler le nombre d'échantillons. Pour répondre aux objectifs de l'étude, seuls les sites ayant un nombre d'échantillons supérieur à 30 seront considérés pour une première analyse descriptive. Une analyse plus avancée des données demande une consistance suffisante pour assurer une bonne significativité statistique. En conséquence, l'application de techniques multi-variées ne pourra être envisagée que pour les sites de Nancy Charles III et de Mairie (de Reims).

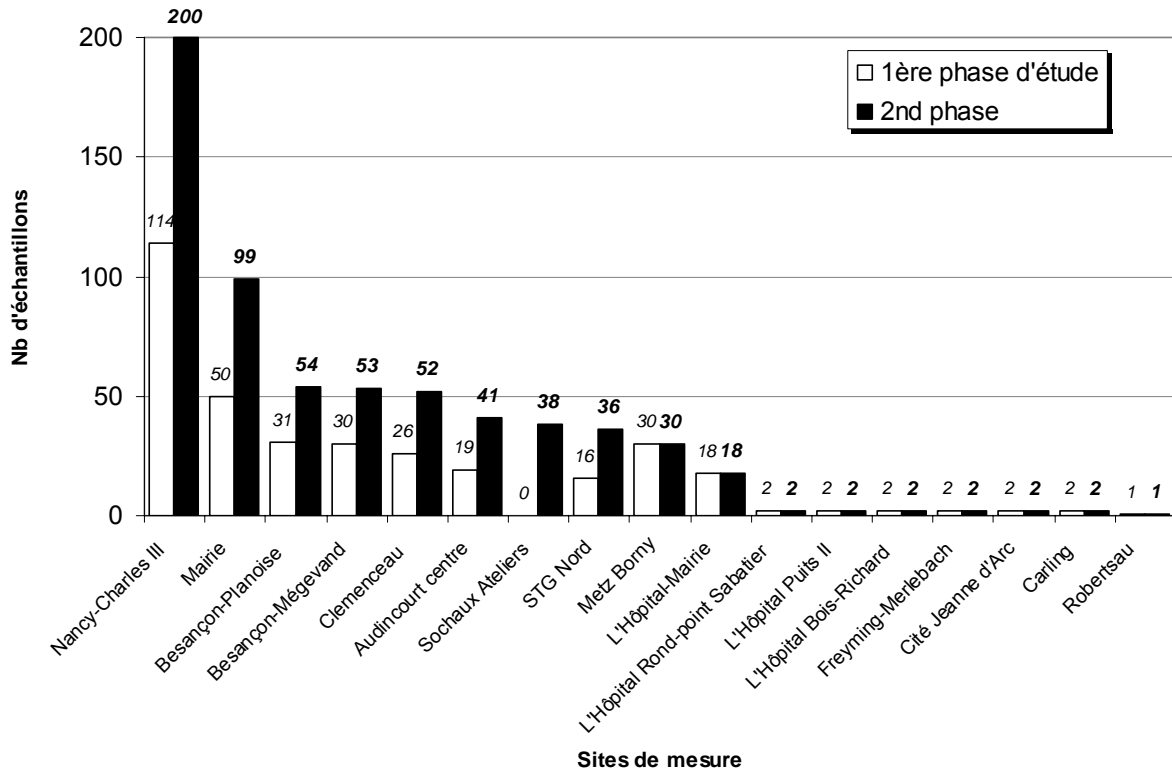
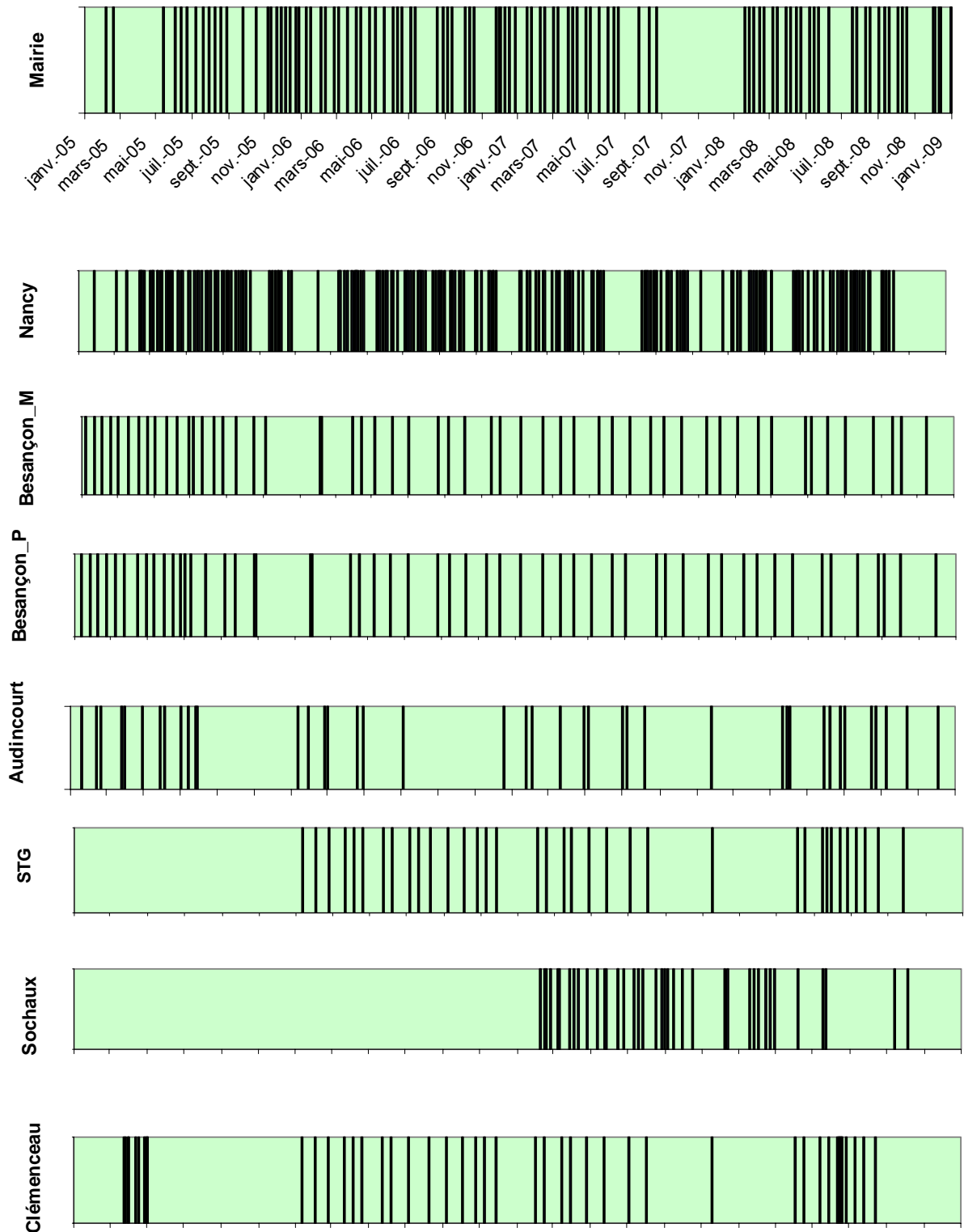


Figure 2 : Evolution du nombre d'échantillons par site de la base canist'air entre les deux phases d'étude

## 2.2 Représentativité temporelle des observations

Quels que soient l'objectif de surveillance et la typologie des sites, les séquences d'échantillonnage ont été de 24 heures. Les mesures intègrent donc toutes les variations journalière des niveaux de concentration des espèces qui pourraient être liées à la variabilité des sources urbaines associées à la densité du trafic automobile et à l'activité anthropique de manière générale, aux variations journalières de la couche de mélange ou encore à l'activité photochimique.

La Figure 3 permet de visualiser la répartition dans le temps des échantillons par site. Il apparaît une fréquence d'échantillonnage globalement régulière pour les sites de Maire, de Nancy Charles III, de Besançon-Megevan, de Besançon-planoise. En revanche, la repartition temporelle des échantillons est plus disparate pour les sites d'Audincourt, de STG, de Sochaux ou de Clémenceau avec notamment des périodes de plusieurs mois non renseignées ou une fréquence d'échantillonnage aléatoire. Il s'agira donc d'être prudent dans l'analyse comparative des sites compte tenu des variations saisonnières potentielles et des tendances des concentrations mesurées.



**Figure 3 : Représentativité temporelle de la base de données canist'air pour les différents sites**



### 3. COMPARAISON DES DES SITES DE MESURES

Au-delà des objectifs de surveillance propres à chacun des sites, l'étude de la base de données dans son ensemble devrait permettre de caractériser les sites suivant leur typologie. La démarche va consister, dans un premier temps, en une analyse comparative des niveaux et des variations de concentration des différentes espèces. Dans un second temps, l'analyse bi-variée d'espèces « traceurs » permettra de rechercher l'influence potentielle de certains types de sources.

#### 3.1 Analyse univariée

##### 3.1.1 Extraction des observations atypiques

Comme préalable à toute analyse statistique d'une base de données, les éléments atypiques doivent être mis en exergue. Ils pourront être attribués à des incidents techniques liés à la méthode de mesure ou à des événements de pollution isolés, associables à des sources ponctuelles. La distribution des données de HCNM étant log-normale, les valeurs sont considérées comme anormalement élevées si, en logarithme, elles sont supérieures à la moyenne plus trois fois l'écart-type sur les logarithmes des données de l'espèce pour un site considéré. Un échantillon est classé atypique si au moins une espèce présente une concentration anormalement élevée. Le nombre d'échantillons atypiques est reporté pour chaque site dans le Tableau 2. Leur nombre reste limité variant de 5 à 7 % du nombre d'échantillons total. Les concentrations observées pour chaque espèce sont visualisées Figure 4 pour les échantillons atypiques de trois sites.

**Tableau 2 : Nombre d'échantillons aux concentrations anormalement élevées**

| Site_mesure       | nb ech. Total | Ech. écartés |
|-------------------|---------------|--------------|
| Nancy-Charles III | 200           | 14           |
| Mairie            | 99            | 6            |
| Besançon-Planoise | 54            | 1            |
| Besançon-Mégevand | 53            | 4            |
| Clemenceau        | 52            | 3            |
| Audincourt centre | 41            | 1            |
| Sochaux Ateliers  | 38            | 2            |
| STG Nord          | 36            | 0            |

Pour le site de Mairie, les échantillons sont caractérisés par des associations d'espèces de concentrations particulièrement élevées comme le butane, le propane, l'isobutane, l'éthylène ou le 1,4 dichlorobenzène et le tétrachloroéthylène. Les concentrations en butane, en 1,4 dichlorobenzène et en tétrachloroéthylène peuvent dépasser 100 ppb. D'après les informations fournies par AIRLOR, un stockage de produits d'entretien du service technique de la ville de Reims se trouvait à proximité du site de prélèvement. Une grande partie du stock aurait été évacué courant 2007. Cependant, un certain nombre d'échantillons atypiques est encore identifié en 2008 (19/02/08 et 30/11/08)

Avant 2006, le site de Nancy Charles III présente des échantillons aux concentrations élevées en xylènes (supérieures à 10 ppb en méta et para xylènes) et en pentane. A partir de 2006, c'est l'isopentane qui présente de valeurs comprises entre 12 et 60 ppb, espèce à laquelle est associé le styrène avec des concentrations pouvant dépasser 2ppb. Les informations disponibles ne permettent pas d'identifier une source locale potentielle.

Le site de Clémenceau est marqué par trois échantillons atypiques, avec des concentrations élevées en trans-2-pentène pour le 01/04/05, en propane et propylène pour le 29/03/07 et en 1,4-dichlorobenzène (supérieures à 20 ppb) pour le 14/05/08.

Compte tenu des objectifs de l'étude, dans le but d'obtenir des informations plus générales sur ces sites, les éléments atypiques n'ont pas été considérés dans la suite de l'étude.

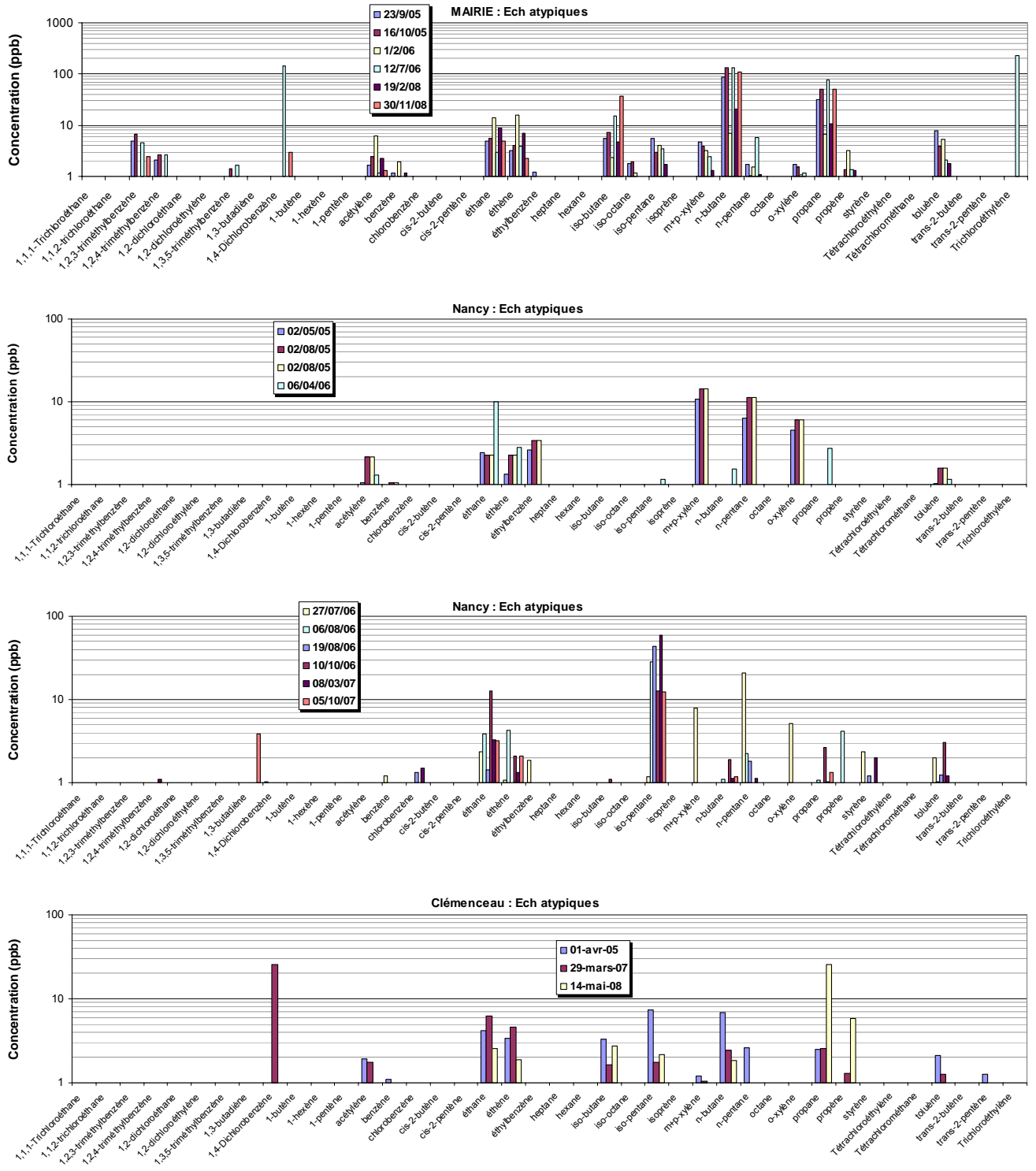


Figure 4 : Concentration mesurées par espèce pour les échantillons atypiques

### 3.1.2 Profils saisonniers d'espèces caractéristiques

La distribution des HCNM dans l'atmosphère résulte de la combinaison entre les processus chimiques de formation ou de destruction et les processus de mélange qui viennent disperser et transporter les composés. L'évolution des teneurs est donc généralement animée de variations saisonnières.

La période estivale, qui se caractérise par une durée d'ensoleillement plus importante et des températures élevées est propice à l'activité photochimique, notamment au travers de l'oxydation des composés par le radical OH. Par ailleurs, l'intensité des processus dynamiques au sein de la basse couche troposphérique est accrue en été au travers notamment de l'augmentation de la hauteur de mélange entraînant une dispersion plus importante des polluants. A cela vient s'ajouter la variabilité saisonnière de l'intensité des sources d'émission de ces composés. Cette variabilité se traduit par l'intensification, en hiver, du chauffage et de la distribution des combustibles associée, qui touchent surtout la fraction légère : éthane, propane, éthène, propène et acétylène. Dans le cadre des émissions de gaz d'échappement automobile, cette variabilité se traduit par l'augmentation due à l'effet du démarrage à froid en hiver où la stabilité thermique du moteur est plus longue à atteindre. Cet ensemble de constats est responsable de concentrations en HCNM supérieures, notamment pour les produits de combustion. De même, les sources d'HCNM liées au stockage et à la distribution des carburants ou à l'usage de solvants pourront être renforcées lorsque la température augmente en été en raison d'une plus forte évaporation des composés.

Les profils saisonniers des principaux HCNM pour le site le mieux renseignés de Nancy Charles III (Figure 5) sont visualisés par les distributions des données de concentration (représentées par des boîtes à moustaches) pour chaque mois. La signification d'une boîte à moustaches est présentée en Annexe 2.

Les profils de variation des concentrations en **éthylène** sont hivernaux assez marqués montrant d'une part l'impact d'un renforcement des émissions de combustions l'hiver, et, d'autre part l'impact de la photochimie en été pour cette espèce réactive.

Le profil de l'**acétylène** est assez marqué avec des niveaux moyens de concentration en été très inférieurs d'un facteur 3 par rapport aux niveaux mesurés l'hiver mettant en évidence les émissions plus intenses de la source automobile en

hiver. Le profil est du même type pour le benzène, montrant en outre l'impact des émissions liées au chauffage résidentiel et renforcées l'hiver.

Les profils saisonniers en **éthane** sont marqués par des valeurs minimales voisines de 2,5 ppb en été indiquant le niveau de fond élevé de cette espèce à plus longue durée de vie.

Certains composés se distinguent par des profils de variation s'opposant à ceux des hydrocarbures issus de processus de combustion. C'est le cas de l'**isoprène** dont le profil est marqué par des teneurs maximales en été (de mai à octobre). Un tel comportement traduit son origine biogénique (Jobson et al., 1994). Les émissions biogéniques croissantes d'isoprène en été sont dépendantes de la température, du niveau d'ensoleillement et de la densité foliaire (Simon et al., 2006). Au printemps et en hiver, les teneurs en isoprène tendent vers la valeur des limites de détection, indiquant le ralentissement de l'activité photosynthétique. Néanmoins cette espèce peut être émise en faible quantité à l'échappement automobile (Borbon et al., 2002).

Les concentrations en **pentane** sont également maximales en été. Pour cette espèce, à la source de type échappement automobile, vient s'ajouter la source évaporation de carburant renforcée l'été. Un profil de même type est également observé mais de manière moins marqué ou décalée vers l'automne pour l'**isopentane** ou **les mp-xylènes**. Une source du type évaporation ou usage de solvants peut être supposée assez proche du site.

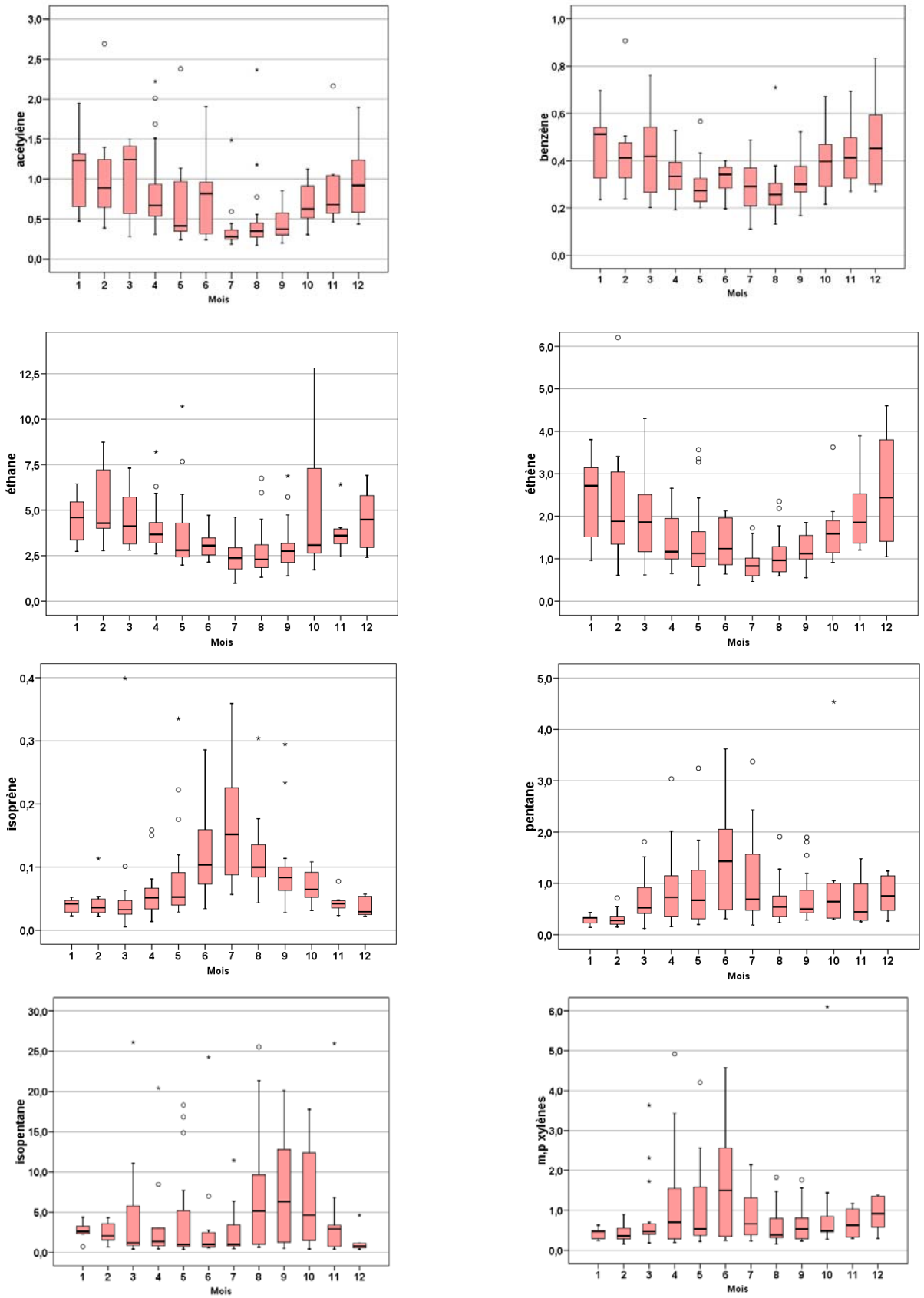
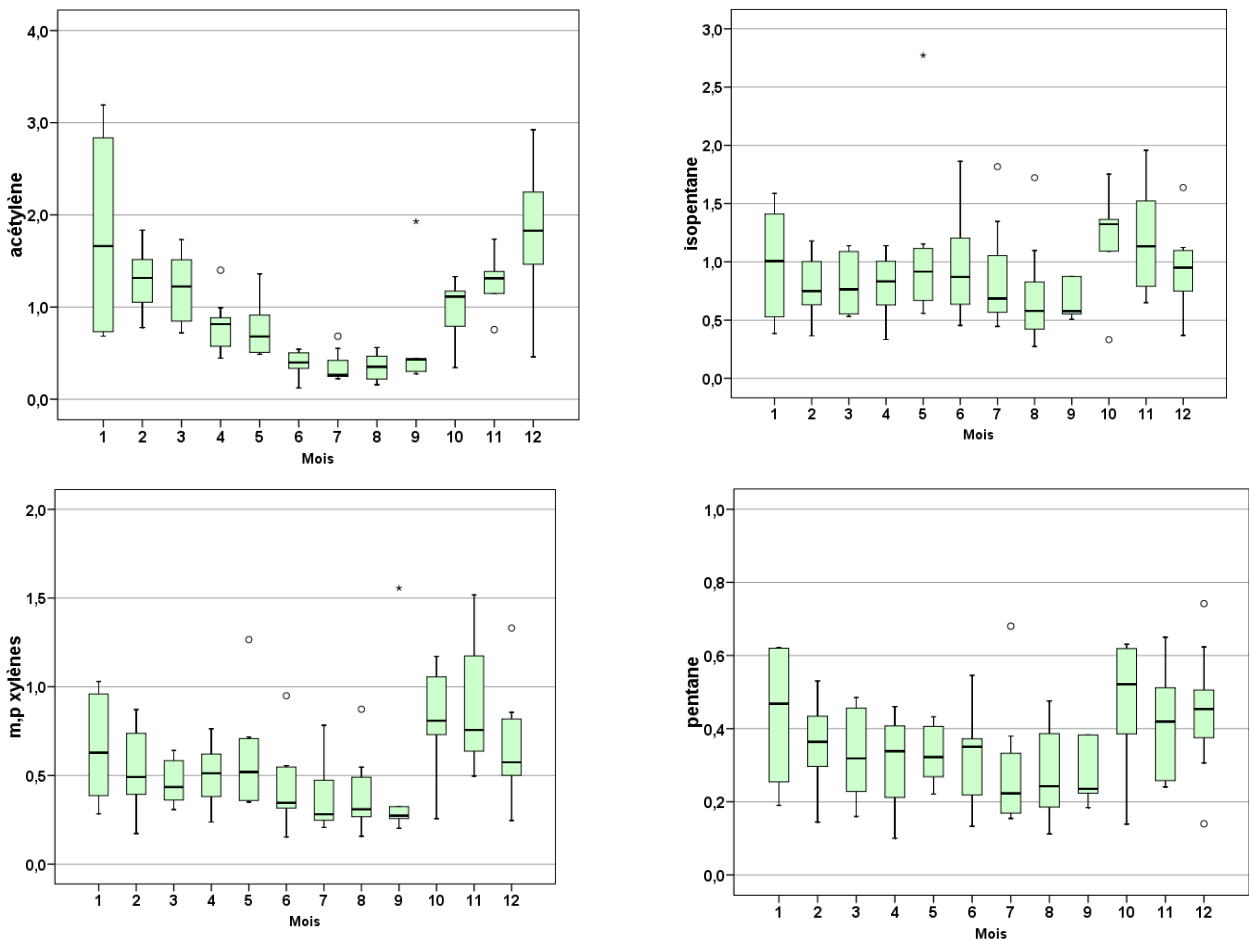


Figure 5 : Profils saisonniers des concentrations (ppb) de quelques espèces majeures pour le site de Nancy Charles III

L'examen des profils du site de Mairie montre un comportement similaire au site de Nancy Charles III pour les espèces issues de la combustion comme l'acétylène (cf Figure 6). En revanche les profils de l'isopentane, des mp-xylènes et du pentane sont particuliers. Des maxima sont observés en octobre novembre suivi d'une baisse lente pour des minimas en juillet août. L'examen des teneurs pour ces espèces montrent en effet des concentrations plus élevées en octobre et novembre pour les années 2005, 2006 et 2008 (Il n'y a pas eu de prélèvement en octobre et novembre 2007). Cette observation serait à relier à une activité particulière d'usage ou de stockage de solvants ou de carburants à cette période dans le local du service technique de la Mairie de Reims (Cf § 3.2.1).



**Figure 6 : Profils saisonniers des concentrations (ppb) de quelques espèces majeures pour le site de Mairie**

L'examen des profils de variation des concentrations mesurées permet de dégager des premières informations quant aux origines et aux comportements des espèces mesurées. Cependant, cette approche nécessite un nombre suffisant de

données sur plusieurs années ainsi qu'une bonne représentativité des saisons. L'analyse n'a été pertinente que pour les sites de Nancy Charles III et Mairie.

### 3.1.3 Comparaison des sites pour les espèces majeures

Les sites les plus renseignés peuvent être comparés pour les espèces majoritaires de manière à distinguer des caractéristiques liées à leur typologie urbain, trafic ou industrielle. Il faut toutefois tenir compte de la représentativité temporelle présentée précédemment. Pour comparaison, le site rural MERA du Donon a été ajouté aux sites de la base canist'air. Les mesures du Donon sont basées sur un échantillonnage de 4 heures (12-16hTU) par canister et analyse différée en laboratoire. Les données de l'année 2007 avec deux prélèvements hebdomadaires ont été considérées pour ce site.

La distribution obtenue pour une espèce et pour chaque site est représentée par une boîte à moustaches (Figure ). Nous distinguons la particularité du site de **Nancy** qui présente des teneurs significativement plus élevées en isopentane, chlorobenzène et styrène. De même le site de **Mairie** se distingue par des teneurs plus fortes en butane, trichloroéthylène et, dans une moindre mesure, en propane. La teneur moyenne en pentane pour ce site est en accord avec les autres sites, toutefois les concentrations fortes qualifiées de points extrêmes sont en nombre important. Les teneurs en éthane et dans une moindre mesure en propane du site de Starsbourg (STG) sont nettement supérieurs aux autres sites indiquant une source du type « fuite de gaz ». Le site industriel de Sochaux se distingue par des concentrations en méta et para xylènes élevées indiquant l'impact d'un usage de solvants à base de xylènes. Au regard des espèces typiques d'émissions automobiles comme l'acétylène, le benzène, le pentane, l'isopentane, le site de Clémenceau apparaît comme le plus influencé par le trafic automobile. Les concentrations des espèces anthropiques présentées sont nettement supérieures à celles mesurées sur le site rural du Donon montrant nettement l'impact des sources urbaines et industrielles. Le site rural se distingue par des teneurs en isoprène bien supérieures aux sites de la base canist'air. L'isoprène est une espèce très réactive de source biogénique principalement.



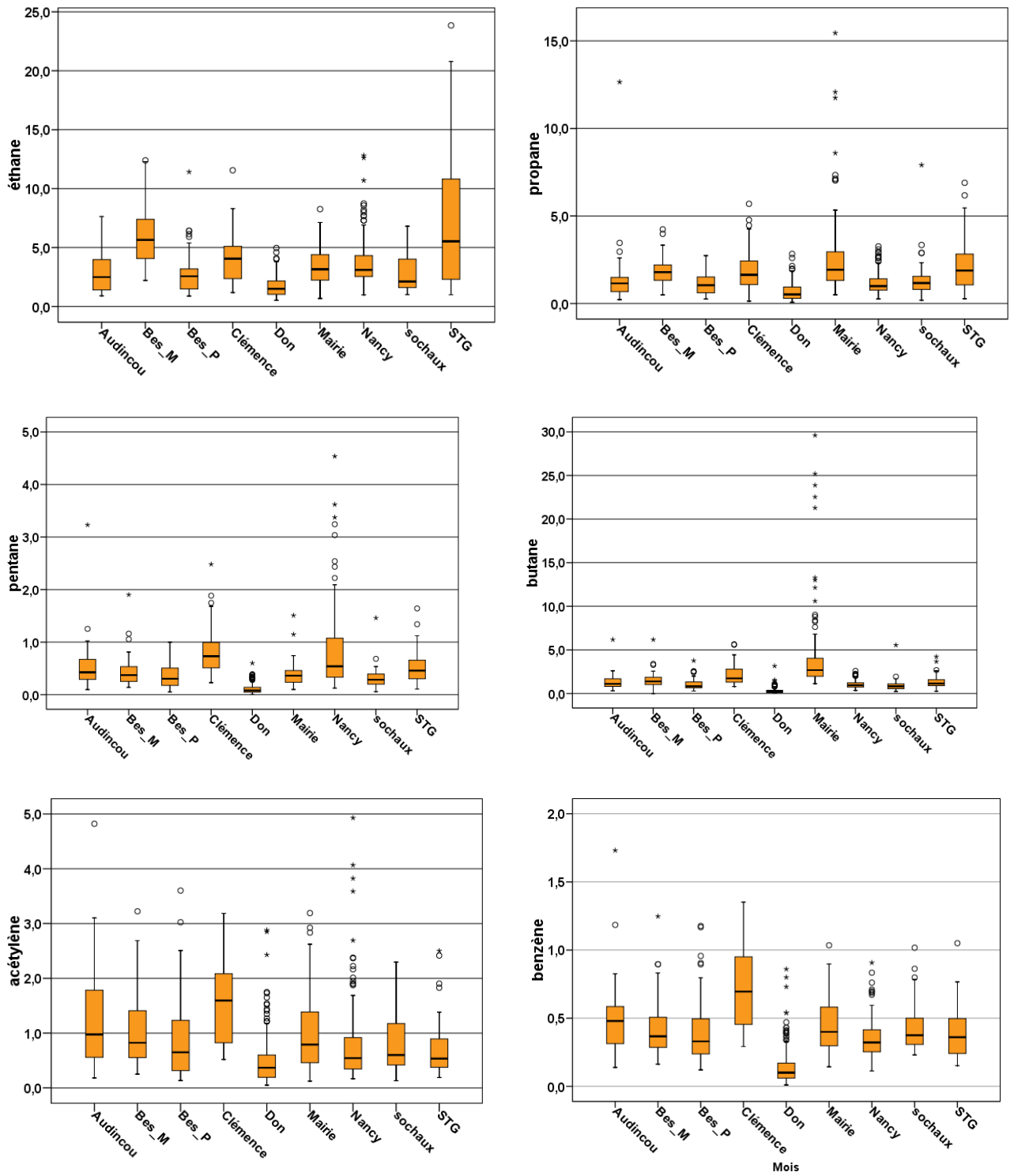


Figure 7 : Comparaison inter-sites des concentrations (ppb) de quelques espèces caractéristiques

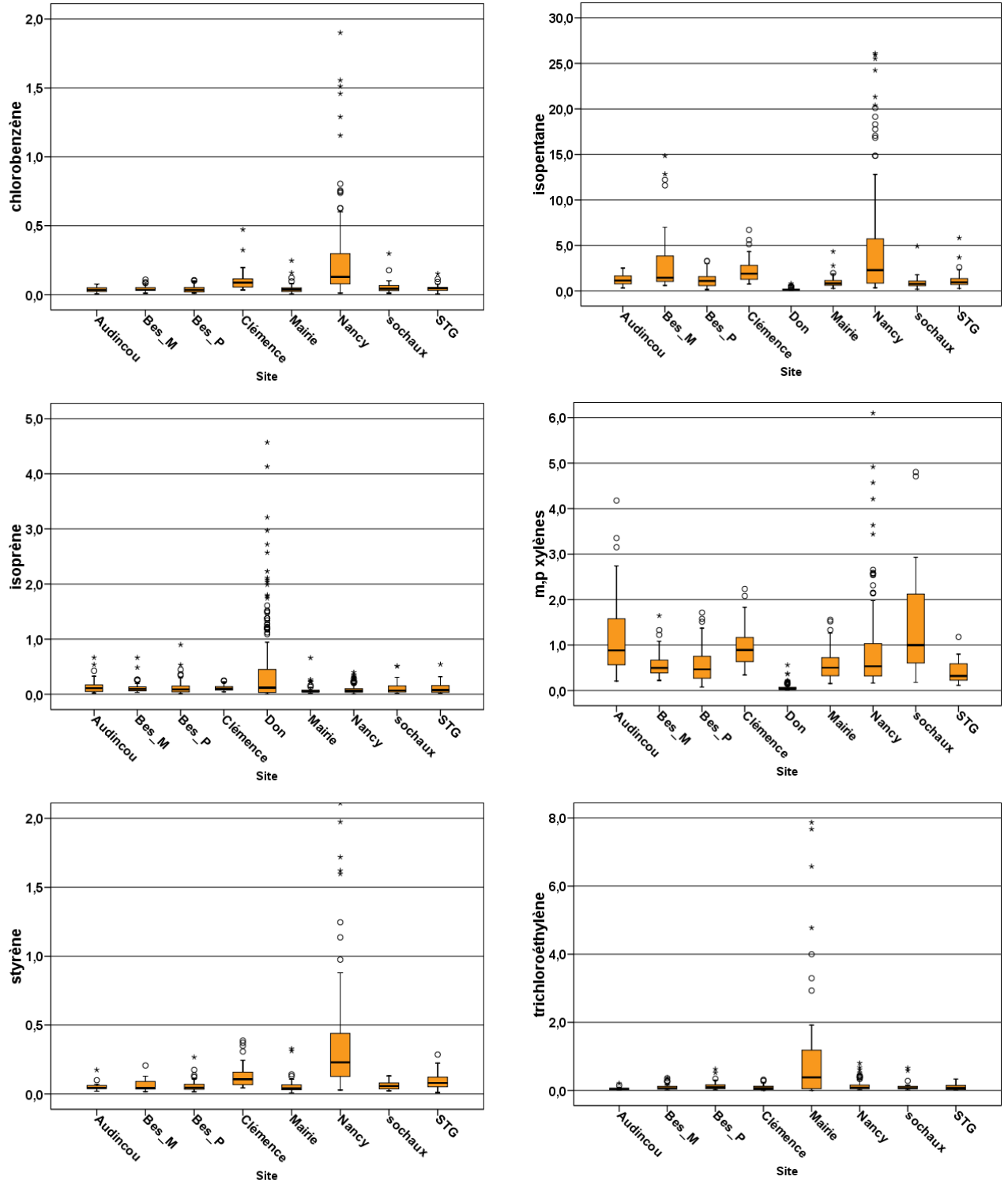


Figure 7 (suite) : Comparaison inter-sites des concentrations (ppb) de quelques espèces caractéristiques

### 3.1.4 Focus sur les sites de Nancy et Charles III

Les sites les plus renseignés de Nancy et de Mairie sont caractérisés par des distributions particulières pour quelques espèces. L'examen du suivi chronologique des espèces concernées permet de mettre en évidence des ruptures dans l'évolution des concentrations. Il faut noter que les composés chlorés n'ont été ajoutés à la liste des espèces mesurées qu'en juin 2006.

Le graphe présenté Figure 8 pour le site de Nancy Charles III montre que contrairement à la plupart des espèces, l'isopentane a augmenté fortement en juillet 2006, puis a diminué légèrement fin 2007 pour rester à un niveau assez élevé. Le styrène présente des concentrations remarquables dépassant le ppb entre septembre 2006 et avril 2007. La variation des teneurs de cette espèce semblent bien corrélée avec celle de l'isopentane. Après avril 2007, contrairement à l'isopentane, les teneurs en styrène restent faibles (inférieures à 0,3 ppb). L'hypothèse d'une contamination du système de prélèvement reste à exclure avant de mettre en avant un type de source potentielle.

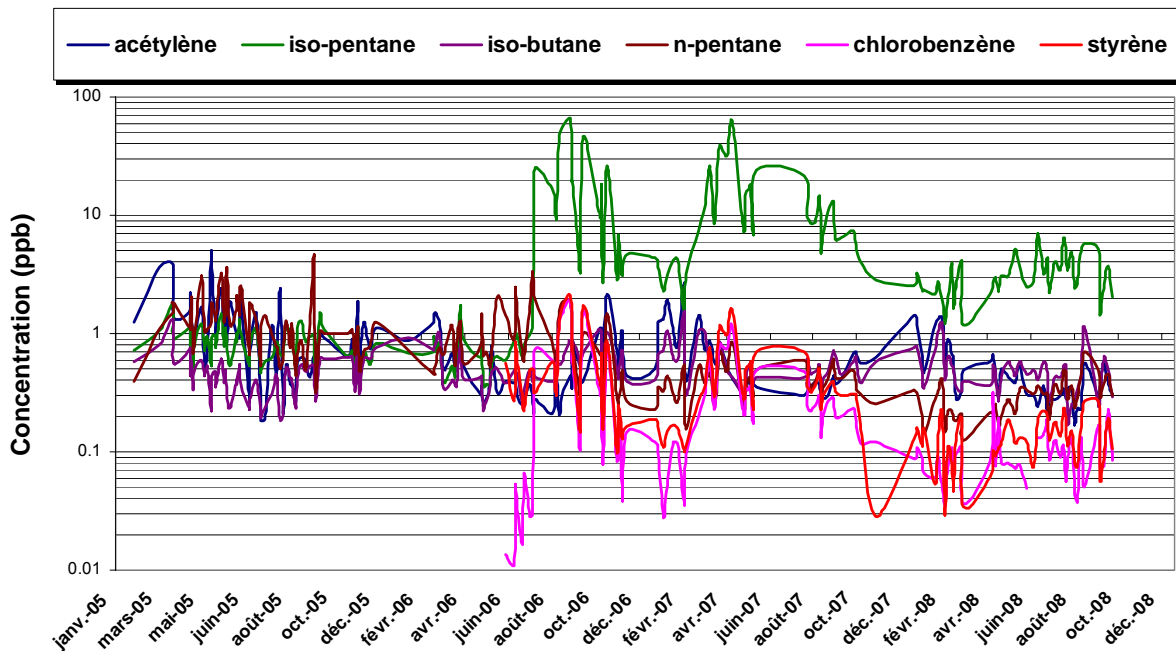


Figure 8 : Suivi chronologique de la concentration (ppb) pour quelques espèces caractéristiques pour le site de Nancy Charles III

Pour le site de Mairie (Figure 9), le trichloroéthylène a montré une nette baisse fin 2007 avec toutefois une nouvelle progression fin 2008. Les teneurs en

butane et en propane sont restées élevées avec des pointes importantes. Ces pointes pouvant atteindre 100 ppb ont conduit à écarter des échantillons de l'analyse. Des teneurs fortes en dichlorobenzène et trichloroéthylène ont également été observées avant septembre 2007. Une source liée à l'entreposage de produit d'entretien non loin du site a été identifiée. D'après les informations disponibles, une grande partie de ce stock a été évacuée courant 2007. Cependant, les teneurs en butane et propane restent élevées avec, de surcroît, une nouvelle tendance à la hausse.

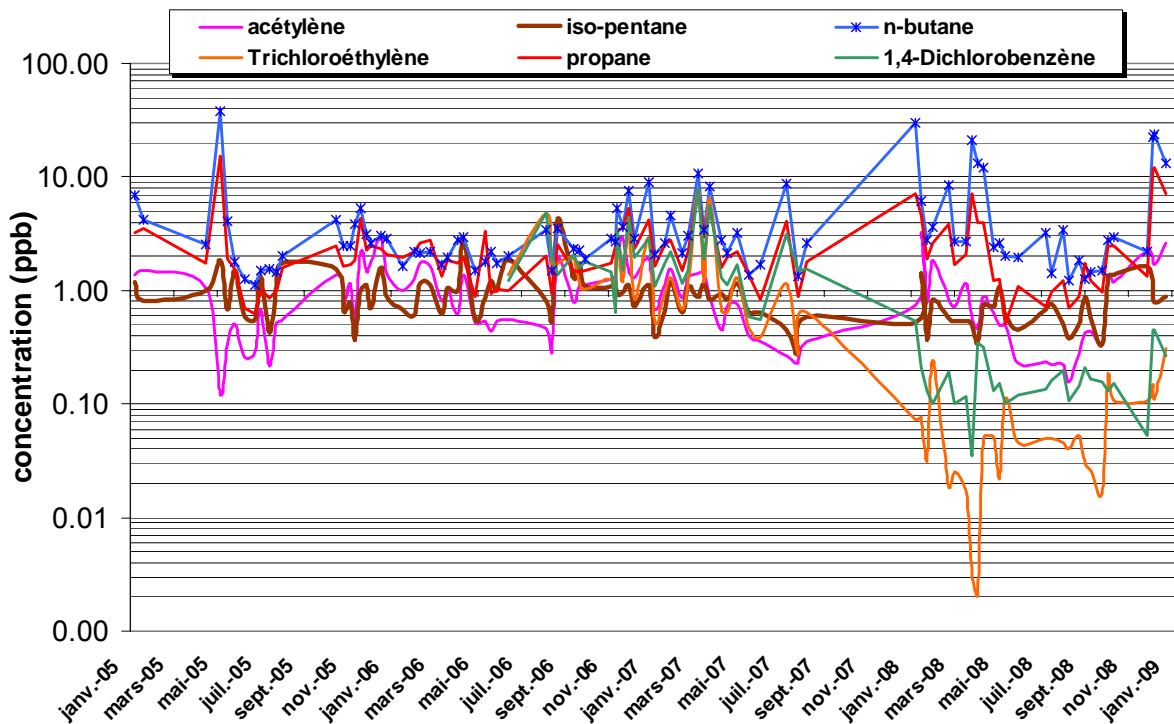


Figure 9 : Suivi chronologique de la concentration (ppb) pour quelques espèces caractéristiques pour le site de Mairie

### 3.2 Analyse bi-variée : vers l'identification de sources

#### 3.2.1 Méthode

Le ratio entre les teneurs en espèces chimiques mesurées simultanément permet de s'affranchir des phénomènes de dilution et fournit différentes informations. Il peut s'agir notamment d'étudier les ratios des teneurs ambiantes vis-à-vis d'un composé traceur d'une source. L'étude de l'évolution temporelle d'un ratio HCNM/traceur en fonction du temps permet alors de montrer l'apport d'une source

supplémentaire sur une période, une saison,... Cette approche permet de mettre en évidence des variabilités temporelles des sources.

Dans le but d'identifier des sources d'HCNM autres que l'échappement automobile, il est intéressant d'étudier le ratio HCNM/acétylène (Derwent et al., 2000; Borbon et al., 2004) en considérant l'acétylène comme composé de référence de ce type de composé. En effet l'acétylène est un composé stable au regard de sa chimie avec le radical OH (durée de vie de 13 jours pour  $[OH]^* = 10^6$  molécules.cm<sup>-3</sup>). C'est un produit de combustion et l'émission de gaz d'échappement automobile est généralement citée comme unique source significative d'acétylène en l'absence de source industrielle spécifique (Mayrsohn and Crabtree, 1976).

### 3.2.2 Résultats

La Figure 10 représente la comparaison entre les sites de la base canist'air de ratios présentant des particularités. Les ratios isopentane/acétylène sont assez similaires pour la plupart des sites hormis pour le site de Nancy où les valeurs sont particulièrement fortes. Dans une moindre mesure, les ratios isopentane/acétylène du site Besançon Megevan présentent une distribution également plus étendue. L'isopentane peut être issue de source d'évaporation de carburants. Mais ce type de source associe également des espèces comme l'isobutane, le butane ou encore le pentane. Pour le site de Nancy, seul l'isopentane se distingue dans cette analyse. En écartant une contamination du système de prélèvement, il s'agit probablement d'une source très spécifique (évaporation d'un solvant particulier) qui d'après le profil saisonnier présenté précédemment se renforce en période estivale. Les travaux d'étanchéité ou de carrelage sont des causes possibles à étudier.

L'examen du ratio butane/acétylène montre une distribution de valeurs particulières pour le site de Mairie. Comme déjà évoqué, ce site est sous influence d'une source de butane.

Le ratio mp-xylènes/acétylène permet également de confirmer des éléments d'information quant à l'identification de source. A la source automobile vient s'ajouter une source liée à l'usage de solvants pour les sites de Sochaux, de Nancy Charles III et d'Audincourt

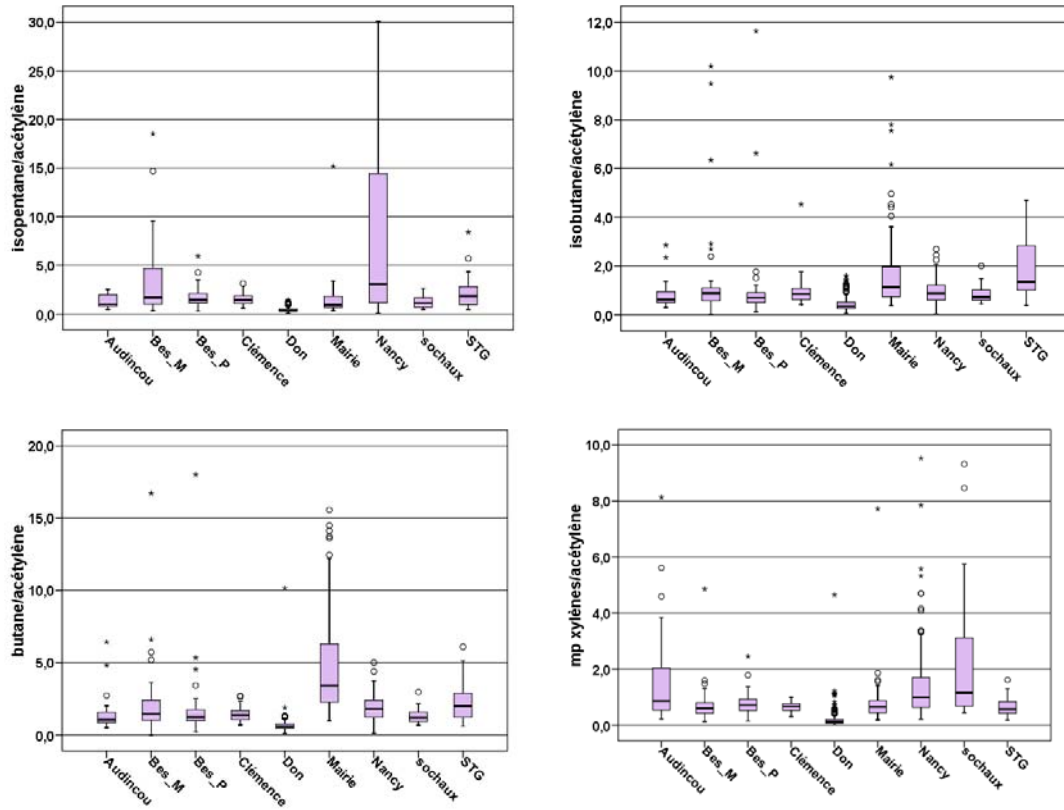


Figure 10 : Comparaison inter-sites de quelques ratios

Les ratios peuvent être considérés comme l’empreinte d’une source et permettre de l’identifier. Il est notamment intéressant de comparer le ratio de concentrations mesurées avec les ratios de mêmes espèces à l’émission. Le ratio toluène/benzène est souvent étudié en relation avec la source d’échappement automobile. Au regard du profil d’émission de l’échappement automobile d’un parc roulant français représentatif (Fontaine and Galloo, 2002), le ratio toluène/benzène est de 2,3. Il faut cependant considérer le chauffage résidentiel qui est une autre source significative pour ces espèces. Le ratio Toluène/benzène à l’émission peut être de 3,6 pour le chauffage au fioul contre 0,6 pour le chauffage au gaz (Passant, 2002). La Figure 11 présente les résultats obtenus pour les sites de la base canist’air. Les distributions sont assez homogènes et ne permettent pas de distinguer un site en particulier. Les valeurs moyennes sont proches de 2 ce qui correspond aux valeurs communément rencontrées dans d’autres zones urbaines (cf Tableau 3). Le site de STG présente un ratio moyen plus faible de 1,5, mais ce résultat ne peut être interprété compte tenu de la faible représentativité temporelle des mesures pour ce site.

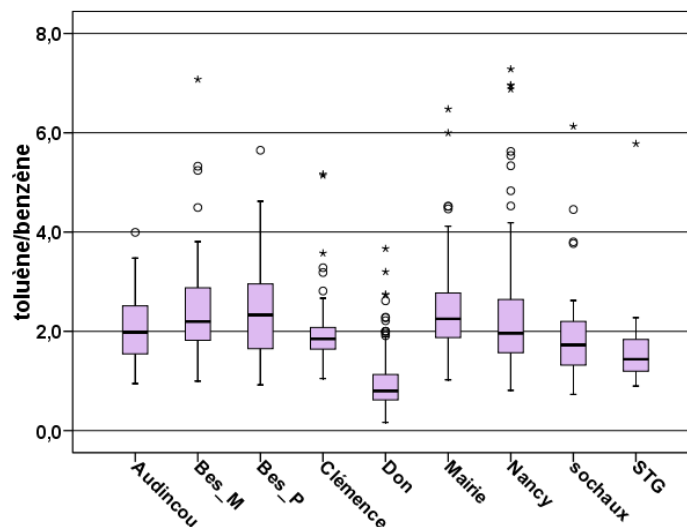


Figure 11 : Comparaison inter-sites du ratio toluène/benzène

Tableau 3 : Valeurs de ratios déterminées à partir de mesures en continu d'HCNM

| Site            | Marseille   | Strasbourg  | Lille Liberté | Dunkerque   |
|-----------------|-------------|-------------|---------------|-------------|
| Type            | Urbain      | Périurbain  | Trafic        | Urbain      |
| Période         | 06/01-12/03 | 03/02-12/03 | 05/97-04/99   | 06/01-12/03 |
| toluène/benzène | 2.87        | 2.76        | 2.49          | 2.36        |

### 3.3 Etude de ratios d'alcane, âge des masses d'air

#### 3.3.1 Méthode

L'étude des ratios pourra aussi permettre de qualifier l'âge d'une masse d'air. Si le ratio de deux espèces à l'émission permet de caractériser une source, celui-ci évolue au cours du temps lorsque les espèces n'ont pas la même durée de vie dans l'atmosphère.

Pour aider à l'identification des sources et apprécier leur éloignement, il est utile de connaître l'âge d'un hydrocarbure c'est-à-dire l'intervalle de temps écoulé entre son émission et son prélèvement. Cet âge photochimique des masses d'air peut être estimé par une approche cinétique et notamment sur des alcanes compte tenu de leur plus grande durée de vie dans l'atmosphère (Parrish et al., 2007).

En considérant que la principale voie de disparition des alcanes est la réaction de pseudo 1<sup>er</sup> ordre d'oxydation avec OH<sup>\*</sup> et en ne tenant pas compte de la

dilution des polluants dans l'atmosphère, la relation donnant la variation de la concentration d'un alcane A peut s'écrire :

$$\frac{d[A]}{dt} = -k_A [A] \cdot [OH^\bullet] \quad (\text{Eq. 1})$$

Avec comme solution :

$$[A] = [A]_0 e^{-k_A [OH^\bullet] \tau} \quad (\text{Eq. 2})$$

- $[A]_0$  est la concentration initiale en alcane A
- $k_A$  est la constante de vitesse
- $[OH^\bullet]$  est la concentration moyenne en radicaux  $OH^\bullet$  durant le temps  $t$
- $\tau$  représente l'intervalle de temps entre l'émission et le prélèvement

Il est possible de s'affranchir des phénomènes de dilution, en étudiant la variation simultanée de la concentration d'un autre alcane. Il faudra cependant considérer que :

- l'air de dilution n'apporte pas une quantité significative des deux alcanes étudiés.
- les deux espèces considérées sont émises dans la même parcelle d'air

Le rapport des concentrations de ces deux espèces pourra alors s'écrire :

$$\ln\left(\frac{[A]}{[B]}\right) = \ln\left(\frac{[A]_0}{[B]_0}\right) - (k_A - k_B) [OH^\bullet] \tau \quad (\text{Eq. 3})$$

$\tau$  est appelé l'âge photochimique. Il peut être déterminé suivant la relation précédente à partir de [A] et [B] sous réserve de connaître les concentrations initiales  $[A]_0$  et  $[B]_0$  des espèces A et B et  $[OH^\bullet]$ . L'introduction d'un troisième composé C permet de s'affranchir de  $\tau$  et  $[OH^\bullet]$  en écrivant la relation suivante :

$$\ln\left(\frac{[C]}{[A]}\right) = \alpha \cdot \ln\left(\frac{[B]}{[A]}\right) + \beta \quad (\text{Eq. 4})$$

avec  $\alpha = \frac{k_C - k_A}{k_B - k_A}$  et  $\beta = \ln\left(\frac{[C]_0}{[A]_0}\right) - \alpha \cdot \ln\left(\frac{[B]_0}{[A]_0}\right)$



Si le processus suit les hypothèses établies préalablement, le tracé de  $\ln([C]/[A])$  en fonction de  $\ln([B]/[A])$  doit être une droite de pente  $\alpha$  et d'ordonnée à l'origine  $\beta$ .

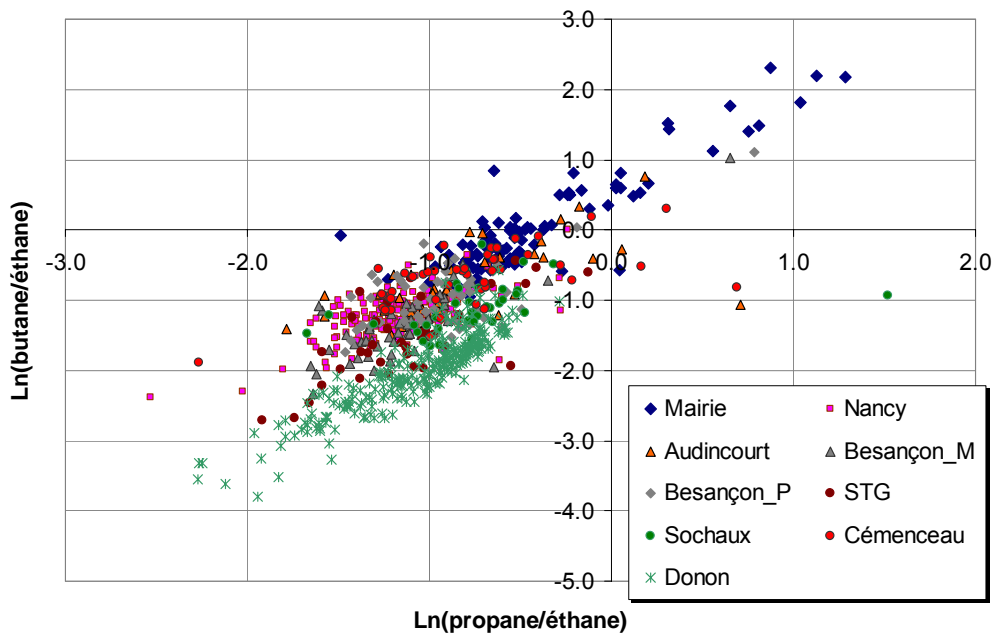
Les espèces choisies pour cette analyse sont l'éthane, le propane et le butane car :

- elles sont éliminées de l'atmosphère principalement par la réaction d'oxydation avec les radicaux OH
- elles ont des temps de vie suffisamment longs et différenciés (respectivement 46, 11 et 5 jours pour  $[^*OH]=10^6 \text{molécules.cm}^{-3}$ ) pour estimer l'âge photochimique d'une masse d'air ((Parrish, 1992).

### 3.3.2 Résultats

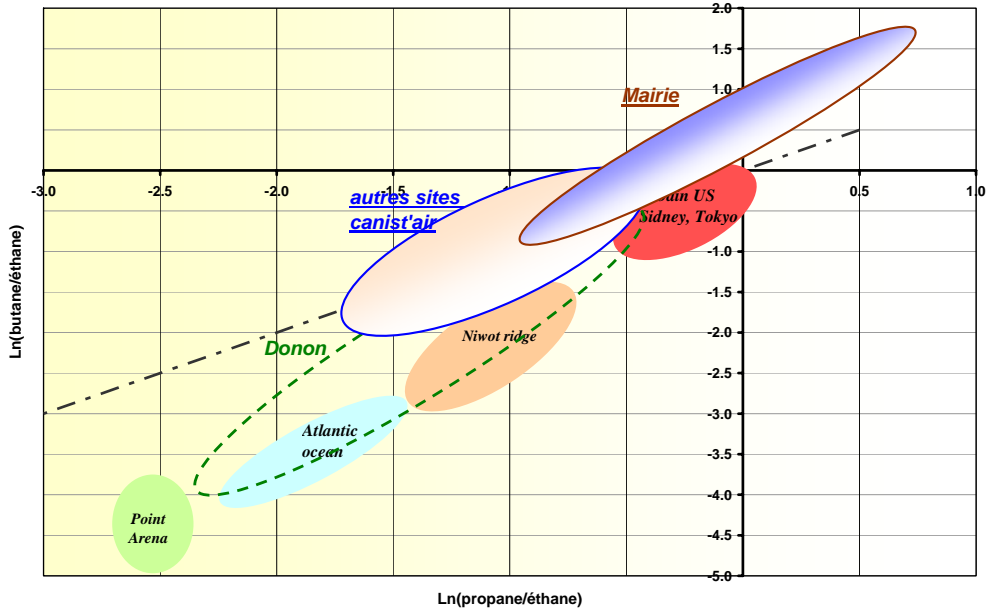
Le tracé de  $\ln([\text{butane}]/[\text{éthane}])$  en fonction de  $\ln([\text{propane}]/[\text{éthane}])$  pour les observations des sites de la base canist'air est présenté Figure 12. Le tracé a également été reporté pour les données du site rural du Donon. Pour ce site, le nuage de points est étiré vers le bas, le coefficient de corrélation de 0,89 indique un lien étroit dans la variation des deux ratios. Cela signifie que le temps de trajet depuis les principales sources est suffisamment long pour permettre une déplétion, liée à la photochimie, distincte entre les espèces considérées. La pente de la droite de régression obtenue pour ce site est de 1,34. Elle est éloignée de la pente théorique de 2,4 calculée sur la base des constantes cinétiques mais du même ordre que les pentes obtenues dans d'autres études : 1,47 pour des données de Point Arena (USA) d'avril à mai 1985 (Parrish, 1992); de 1,12 à 2,00 pour les données de Porspoder en 1992 (Boudries, 1994). L'écart avec la théorie montre que le modèle est insuffisant notamment s'agissant de l'hypothèse stipulant que la dilution des émissions s'effectue dans un air propre (McKeen, 1993; Boudries, 1994). Néanmoins, les valeurs de ratios obtenues permettent d'évaluer qualitativement la

typologie des masses d'air affectant les sites de mesures et notamment l'âge photochimique des masses d'air. Ce site de fond constitue un élément de comparaison avec les sites de la base canist'air. Le nuage de points du site de Mairie est étiré vers le haut, les ratios sont supérieurs à 1 indiquant plutôt l'apport d'une source plus ou moins intense associant le propane et le butane. Les autres sites présentent des nuages de points assez similaires et plutôt épars montrant que le modèle n'explique pas les variations combinées de ces espèces.



**Figure 12 : Comparaison des ratios d'alcane pour les sites étudiés de la base canist'air**

Les niveaux de ratios obtenus avec les mesures d'autres sites sont reportés sur la Figure 13 sous forme de zones graphiques. Cela permet de distinguer les sites urbains (Tokyo, Sydney), les sites de fond continentaux (Niwt Ridge), les sites océaniques (Atlantic Ocean) et les sites de fond très reculés (Point Arena). Les résultats des sites de la base canist'air, site de Maire mis à part, se situent assez proches des sites urbains ou du site de fond continental. Le site rural du Donon apparaît naturellement comme le site aux masses d'air les plus anciennes. Cette analyse montre donc que les sites de la base canist'air sont soumis principalement à des masses d'air jeunes.



**Figure 13 : Comparaison des zones graphiques couvertes par les nuages de points avec des données de la littérature**

Cette analyse peut toutefois être menée en fonction du secteur de provenance des masses d'air pour tenter de discerner des influences plus lointaines suivant la direction de vent. Cette analyse demande un nombre de données plus important, elle n'a donc pu être réalisée que pour les sites de Nancy Charles III et de Mairie. Le tracé du nuage de points  $\ln(\text{butane}/\text{éthane})$  en fonction de  $\ln(\text{propane}/\text{éthane})$  pour le site de Nancy est présenté Figure 14. Seul le secteur de vent Nord-Est se distingue des autres secteurs. Le nuage de points correspondant est plus étiré avec quelques valeurs de ratios faibles pouvant indiquer des masses d'air plus anciennes. Cette observation reste toutefois très hypothétique compte tenu du nombre limité de points par secteur.

La Figure 15 présente les résultats pour le site de Mairie. Elle ne met pas en évidence de secteur présentant des ratios faibles. En revanche, elle souligne l'impact d'une source de butane et de propane pour les secteurs Ouest et Sud-Ouest

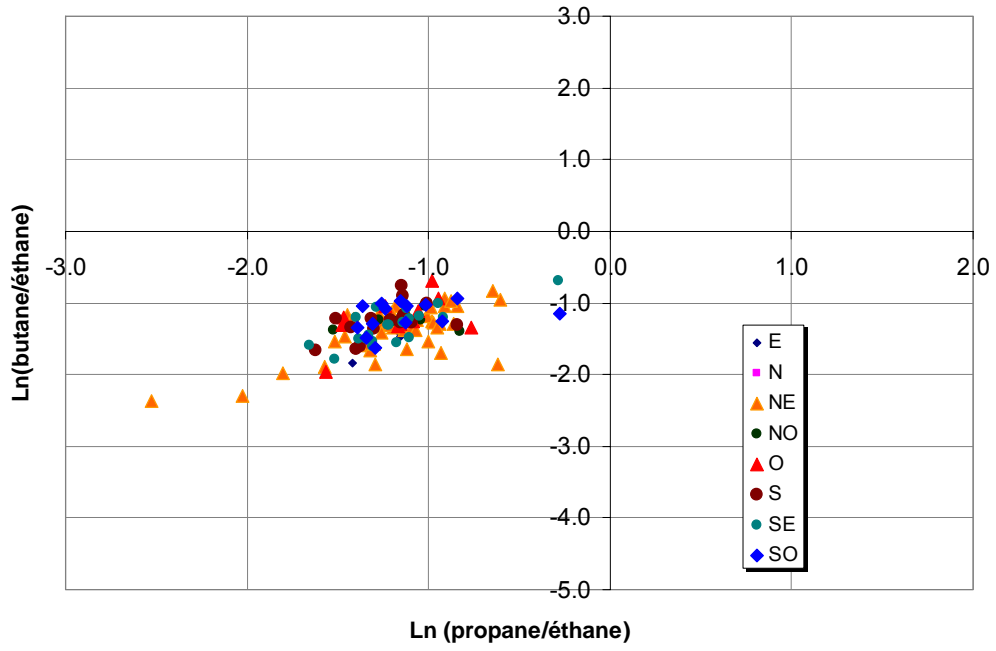


Figure 14 : Ratios d'alcane pour le site de Nancy Charles III

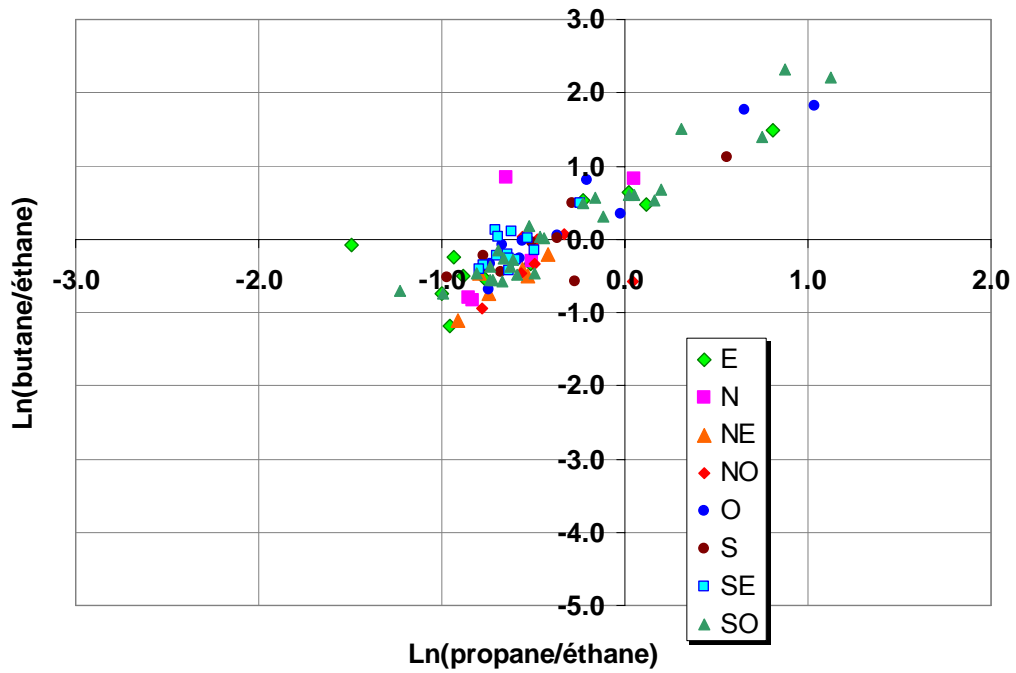


Figure 15 : Ratios d'alcane pour le site de Mairie

## 4. IDENTIFICATION DES PROFILS DE SOURCES PAR APPLICATION DU MODELE SOURCES-RECEPTEUR PMF

Les méthodes dites « modèles sources-récepteur » se focalisent sur le comportement des polluants aux sites récepteurs contrairement aux modèles de dispersion qui se concentrent sur le transport, la dilution et les transformations qui se produisent en commençant par la source jusqu'au site récepteur (Hopke 2003).

Le principe fondamental du modèle récepteur est que la conservation de masse peut être supposée et que l'analyse du bilan matière peut être utilisée pour identifier et répartir les sources de polluants dans l'atmosphère. Il est donc nécessaire de disposer d'une série de données importante comprenant plusieurs observations de plusieurs espèces. D'une manière générale le bilan matière peut être formalisé pour les m espèces des n observations suivant l'équation :

$$x_{ij} = \sum_{k=1}^p g_{ik} f_{kj} + e_{ij} \quad (\text{Eq. 5})$$

Où  $x_{ij}$  est la concentration mesurée de la  $j^{\text{ème}}$  espèce dans le  $i^{\text{ème}}$  échantillon,  $f_{kj}$  est la concentration de la  $j^{\text{ème}}$  espèce émise par la source k,  $g_{ik}$  est la contribution de la  $k^{\text{ème}}$  source pour l'observation i et enfin  $e_{ij}$  est la partie non expliquée par le modèle.

Deux types de modèles sources-récepteur se distinguent :

Les modèles sources-récepteur pour lesquels les profils des sources sont connus. La seule inconnue restante est la contribution des sources  $g_{kj}$ . Pour cela le modèle CMB (Chemical Mass Balance) utilise le principe de la variance effective par les moindres carrés

Les modèles pour lesquels les profils de sources ne sont pas connus ce qui correspond davantage à notre étude. L'outil le plus ancien et le plus accessible est l'Analyse en Composantes Principales. Les approches plus récentes sont les modèles UNMIX et PMF (Positive Matrix Factorization) qui permettent d'identifier plus finement les sources et leurs contributions.

## 4.1 Méthode

Le modèle PMF est apparu en 1994 (Paatero and Tapper, 1994) et a évolué en 1997 (Paatero, 1997). Un guide d'utilisation a été rédigé par Hopke (Hopke, 2000) et complété par des articles spécifiques pour certains paramètres du modèle (Paatero and Hopke, 2002; Paatero et al., 2002). Il a été appliqué dans différents travaux de recherche concernant les COVNM comme une alternative à l'analyse factorielle (Lanz et al., 2008; Sauvage et al., 2009).

### 4.1.1 Principe général du PMF

Toute matrice  $X$  ( $n \times m$ ), où  $n$  est le nombre d'observations et  $m$  le nombre d'espèces mesurées, peut être décomposée en un produit factoriel de 2 matrices  $G$  ( $n \times p$ ) et  $F$  ( $p \times m$ ),  $p$  étant le nombre de facteurs, et une part non expliquée  $E$  ( $n \times m$ ).  $X$  s'écrit alors :

$$X = F G + E$$

Le produit de  $F$  et de  $G$  peut expliquer les variations systématiques de  $X$ . Une solution est obtenue par minimisation de  $Q(E)$  :

$$Q(E) = \sum_{i=1}^n \sum_{j=1}^m \frac{e_{ij}^2}{s_{ij}^2}$$

avec :

$s_{ij}$  l'incertitude sur la mesure de l'espèce  $j$  dans l'échantillon  $i$

$$e_{ij} = x_{ij} - \sum_{k=1}^p g_{ik} f_{kj} \quad \text{avec } f_{ik} \geq 0 \text{ et } g_{kj} \geq 0$$

### 4.1.2 Principe d'itération du modèle

La solution du modèle ( $F$ ,  $G$ ) est obtenue itérativement en utilisant la méthode des moindres carrés « alternatifs ». Cela signifie qu'une itération s'effectue en deux temps :

- la matrice  $G$  étant fixée,  $Q(E)$  est alors minimisée en déterminant les éléments de  $F$  par les moindres carrés.

- La matrice F est fixée et les éléments de G sont déterminés par les moindres carrés pour minimiser Q(E).

Le processus continue ainsi jusqu'à convergence de Q(E)

Comme toute analyse factorielle, la solution n'est pas unique. En effet, la paire de matrice (G,F) peut être transformée en une autre paire ( $\tilde{G}, \tilde{F}$ ) telle que la même valeur de Q(E) correspond à (G,F) et à ( $\tilde{G}, \tilde{F}$ ).

La solution du modèle (F, G) est obtenue itérativement en utilisant la méthode des moindres carrés « alternatifs ». Cela signifie qu'une itération s'effectue en deux temps :

- la matrice G étant fixée, Q(E) est alors minimisée en déterminant les éléments de F par les moindres carrés.
- La matrice F est fixée et les éléments de G sont déterminés par les moindres carrés pour minimiser Q(E).

Le processus continue ainsi jusqu'à convergence de Q(E).

Pour améliorer les performances du programme, Paatero a élaboré en 1997 le PMF2. Le PMF2 permet d'imposer la contrainte non-négative et la réduction de la liberté de rotation en utilisant les fonctions de régularisation (ou de pénalités) P et R dans une fonction  $\bar{Q}$  plus élaborée :

$$\bar{Q}(E, G, F) = Q(E) + P(G) + P(F) + R(G) + R(F)$$

$$\bar{Q}(E, G, F) = \sum_{i=1}^n \sum_{j=1}^m \frac{e_{ij}^2}{s_{ij}^2} - \alpha \sum_{i=1}^m \sum_{k=1}^p \log g_{ik} - \beta \sum_{k=1}^p \sum_{j=1}^n \log f_{kj} + \gamma \sum_{i=1}^m \sum_{k=1}^p g_{ik}^2 + \delta \sum_{k=1}^p \sum_{j=1}^n f_{kj}^2$$

Les coefficients  $\alpha$ ,  $\beta$ ,  $\delta$  et  $\gamma$  diminuent à chaque itération jusqu'à devenir très petits. Un paramètre  $F_{PEAK}$  permet de changer la fonction Q(E) pour imposer une rotation R(G) et R(F). Une valeur positive entraîne une rotation avec un coefficient r positif. De même, une valeur de  $F_{PEAK}$  négative impose une rotation avec un coefficient r négatif. Lors de l'étude de données environnementales, l'ambiguïté de

rotation conduisant à plusieurs solutions possibles doit être estimée (Paatero et al., 2002). Le paramètre  $F_{PEAK}$  est étudié sur une gamme où  $Q(E)$  n'augmente pas trop.

#### 4.1.3 Données d'entrée du modèles

Pour le site de Nancy, la base de données d'entrée correspond aux mesures des 37 espèces pour les prélèvements de la période de 2005 à 2008 (N=199). Pour le site de Mairie, les données se limitent aux mesures de 30 espèces excluant les composés chlorés pour une période de prélèvement de 2005 à 2008 (N=90). En effet, les mesure de composés chlorés ont démarré en juin 2006, le nombre de données manquantes est donc trop important (>37%) relativement à la dimension de la base pour ce site.

Les mesures inférieures à la Limite de Détection (LD) de la méthode ont une signification et doivent être considérées comme non nulles. Par convention, ces mesures seront fixées à la valeur de la LD divisée par deux. Suivant les recommandations de Hopke (2000), les données manquantes sont remplacées par la moyenne géométrique de l'espèce considérée pour l'ensemble de la période prise en compte pour le site concerné.

#### 4.1.4 Choix du modèle d'incertitude

Le choix du modèle d'incertitude  $s_{ij}$  est crucial et peut influencer fortement les résultats. Le PMF2 propose différents modèles suivant le type de données et la connaissance ou non des incertitudes de mesure. Pour cette étude, la matrice des incertitudes a été élaborée manuellement, elle tient compte du niveau de concentration mesuré. En effet, l'incertitude est à considérer différemment lorsque les concentrations mesurées sont très faibles en particulier au voisinage de la limite de détection. Une contribution constante à l'incertitude globale doit être envisagée comme valeur seuil de l'incertitude (EURACHEM, 2000). L'incertitude peut alors être exprimée sous forme d'une combinaison entre une contribution proportionnelle à la concentration et une contribution constante basée sur la limite de détection de la méthode :

$$U(X) = s_0 + s_1 \cdot X$$



- $s_0 = \frac{LQ}{10}$  avec LQ la limite de quantification de la méthode de mesure pour une espèce donnée. Les limites de quantification de l'ensemble des espèces sont précisées en Annexe 3.
- $s_1$  représente l'incertitude relative de la méthode de mesure. L'estimation n'ayant pas été faite, nous la fixons à 15% pour toutes les espèces par analogie à celle déterminée pour la mesure des HCNM par canister dans le cadre du dispositif MERA.

Dans le cas des valeurs manquantes remplacées par la moyenne géométrique, l'incertitude sera fixée à quatre fois la moyenne géométrique pour l'espèce considérée. Cela permet de pondérer fortement ces données particulières et ainsi de limiter leur incidence sur les résultats du modèle. Pour les valeurs inférieures à la LD ( $3LQ/10$ ), l'incertitude associée sera égale aux  $5/6^{\text{èmes}}$  de la LD (Hopke, 2000).

#### 4.1.5 Traitement des valeurs extrêmes

Le PMF2 dispose d'un mode robuste pour prendre en compte les valeurs extrêmes. Pour cela, d'une part il prend en compte l'erreur sur la mesure et, d'autre part il applique une pondération itérative des données. Cette pondération réduit le poids de certains individus lorsque la prédiction s'avère médiocre. Il s'agit d'introduire une fonction filtre  $h_{ij}$  dans la minimisation de  $Q(E)$  :

$$Q = \sum_{i=1}^n \sum_{j=1}^m \left( \frac{e_{ij}}{h_{ij} s_{ij}} \right)^2$$

avec  $h_{ij}^2 = \begin{cases} 1 & \text{si } |e_{ij} / s_{ij}| \leq \alpha \\ |e_{ij} / s_{ij}| / \alpha & \text{sinon} \end{cases}$

$\alpha$  est paramétrable dans le modèle. D'après les préconisations du guide d'Hopke et les articles relatifs au PMF traitant de données environnementales, le mode robuste est en général activé avec  $\alpha=4,0$ .

Néanmoins, par cohérence avec le début de l'étude, les échantillons atypiques relevés au paragraphe 3.2.1 n'ont pas été considérés dans l'analyse.

#### 4.1.6 Détermination du nombre de facteurs

Le PMF2 est un modèle utilisable sans connaissance à priori du nombre de facteurs. Ce nombre doit être déterminé par l'utilisateur en examinant les résultats à l'aide de différents indicateurs. Le choix du nombre optimal de facteurs se fait dans un premier temps, en observant la variation de la valeur de Q en fonction du nombre de facteurs. De même, l'examen de la matrice des résidus permet de juger de la qualité du résultat. Une distribution symétrique des valeurs doit être observée. De plus, la plupart des résidus doit être compris entre  $\pm 2$  et supérieur à 0,3.

La recherche du nombre optimal de facteurs se fait également en calculant les indicateurs IM et IS relatifs aux résidus suivant les relations :

$$r_{ij} = \frac{e_{ij}}{s_{ij}} \quad (\text{Eq. 6})$$

$$IM = \max \left( \frac{1}{n} \sum_{i=1}^n r_{ij} \right) \text{ pour } j=1 \dots m \quad (\text{Eq. 7})$$

$$IS = \max \left( \sqrt{\frac{1}{n-1} \sum_{i=1}^n (r_{ij} - \bar{r}_j)^2} \right) \text{ pour } j=1 \dots m \quad (\text{Eq. 8})$$

Q, IM et IS sont alors tracés en fonction du nombre de facteurs (cf. Figure 16). Le nombre de facteurs pouvant être retenu correspond à une baisse accentuée de Q, IM et IS (Lee et al., 1999).

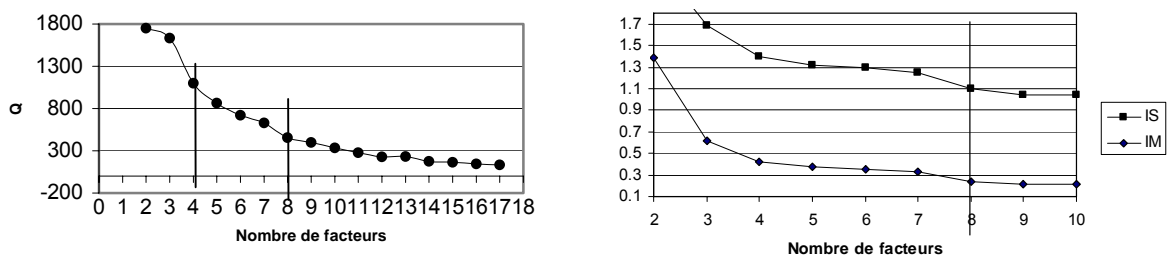


Figure 16 : Représentation graphique de Q, IM et IS en fonction du nombre de facteurs – Détermination du nombre de facteur optimal

#### 4.1.7 Valeur ajoutée du modèle PMF par rapport à l'Analyse en Composante Principale (ACP)

Une différence importante entre l'Analyse en Composante Principale (ACP) et la PMF est que la non-négativité des facteurs (poids et scores) est une

contrainte intégrée aux calculs du modèle PMF. Les facteurs issus du PMF sont donc interprétables physiquement en termes de profils de sources (matrice des poids F) et de contributions de ces sources (matrice des scores G).

Le PMF ne repose pas l'information uniquement sur la matrice des corrélations mais utilise un schéma de minimisation, par les moindres carrés, point par point. En conséquence, les résultats du PMF sont quantitatifs. En revanche, les résultats d'une ACP restent qualitatifs (Qin et al., 2002), ils indiquent les groupements de composés qui tendent à être statistiquement associés mais il ne s'agit pas de la composition des sources.

La standardisation des colonnes variables préalables à l'ACP introduit une certaine pondération non optimale (Paatero and Tapper, 1994). Les échantillons de fortes concentrations tendent à dominer l'analyse alors que les données de faibles concentrations sont virtuellement ignorées. En revanche le PMF valorise l'information de tous les échantillons en pondérant le carré des résidus avec le carré de l'écart-type (ou incertitude-type) sur les mesures.

Le PMF utilise l'erreur estimée sur les mesures pour tenir compte de la qualité des données d'entrée. Cela permet de traiter au mieux les valeurs proches ou égales à la limite de détection et les valeurs extrêmes.

## 4.2 Résultats

Un protocole d'application du PMF a été défini de manière à préciser le paramétrage du modèle ainsi que les démarches d'optimisation et de consolidation des résultats.

### 4.2.1 Site de Nancy Charles III

Les résultats du modèle **expliquent 95% de la concentration totale en HCNM** mesurée. **83% de la variabilité** des concentrations des espèces sont expliqués par **4 facteurs** modélisés. Les concentrations totales modélisées sont bien corrélées avec les concentrations totales mesurées puisque le coefficient de corrélation calculé sur l'ensemble des échantillons est de 0,95.

Les facteurs modélisés ont été comparés aux profils de sources connus issus de la littérature de manière à identifier des signatures de sources. L'examen du profil saisonnier des contributions modélisées permet de consolider l'identification le

cas échéant. Les facteurs modélisés et, dans le cas d'un type de source identifié, le profil de source d'émission issus de la littérature sont représentés sur un même graphe.

La Figure 17 représente la composition normalisée des facteurs modélisés pour le site de Nancy Charles III.

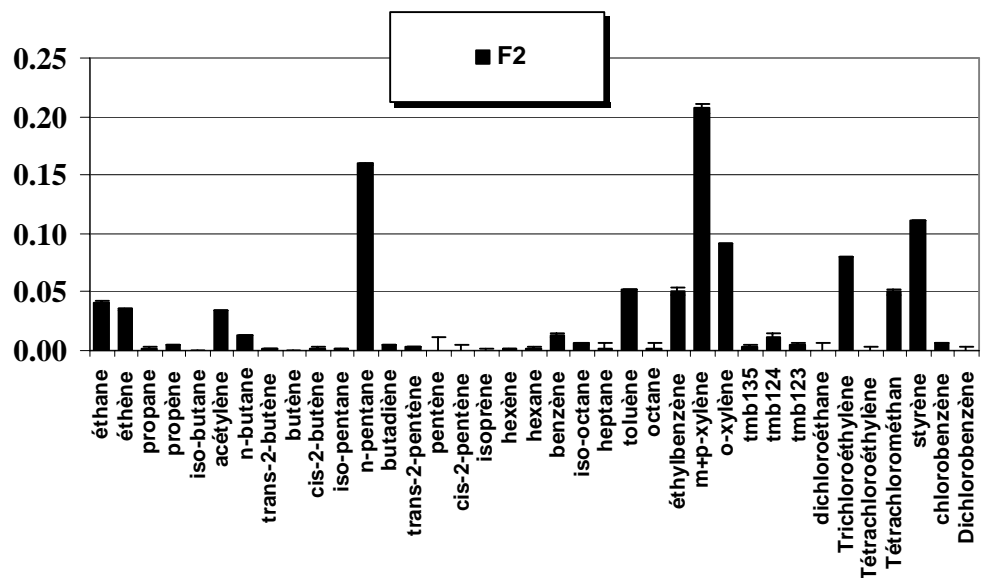
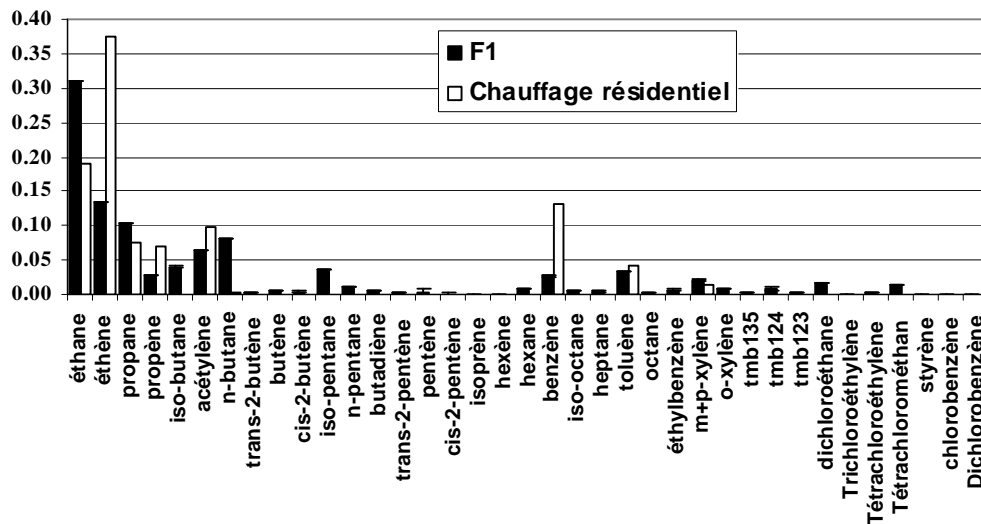
L'éthane, l'éthylène, le propane, l'acétylène, le butane, le benzène, et le toluène présentent des poids significatifs pour le facteur **F1** modélisé. Le profil ainsi modélisé est assez similaire au profil d'émission «**chauffage résidentiel**» calculé à partir de différents profils d'émission pour diverses sources d'énergie (GENEMIS). L'éthène et le benzène présentent des contributions assez faibles comparativement au profil d'émission. Cependant, la contribution de l'éthène et du benzène pour le profil d'émission « chauffage mixte » sont élevées (37,5% et 13,3% respectivement) à cause du mode de chauffage au bois (38,7% et 44,6%) moins représenté en zone urbaine. Celles du chauffage au gaz (<1% et 15,3%) ou au fioul (7,9% et 6,2%) sont plus modérées. En outre, les variations saisonnières des contributions modélisées (Figure 18) sont très marquées avec des valeurs très faibles en été et élevées en hiver confirmant l'identification de ce profil de source.

Le facteur **F2** est dominé par les mp-xylènes, le pentane, l'o-xylène, le styrène, le trichloroéthylène. Ce facteur n'a pu être rapproché d'un profil connu. Toutefois, les espèces contribuant à ce profil peuvent être issues de l'évaporation de solvants. Les contributions modélisées sont d'ailleurs plus élevées en été période favorable à l'évaporation. Il faut noter que ce facteur montre des contributions significatives avant août 2006 et beaucoup plus faibles sur le reste de la période. Ce comportement nécessite des éléments d'information complémentaires pour que la source puisse être déterminée.

Le facteur **F3** est exclusivement marqué par l'isopentane. Les contributions modélisées de ce facteur apparaissent brutalement après août 2006, de manière surprenante, en remplacement du facteur **F2**. Une investigation particulière est nécessaire pour permettre l'interprétation de ces résultats. Il s'agira d'abord d'écarter un problème de contamination du système de prélèvement ce qui est tout à fait possible avec le système utilisé (système de régulation de débit Veriflo en entrée de canister). Puis, une recherche d'éléments sera nécessaire quant aux changements autour du site. Il pourra s'agir de travaux de réfection ou d'aménagement de

bâtiments, d'un changement d'activités dans une structure environnante ou encore d'un stockage particulier de solvants.

Le quatrième facteur est dominé par l'éthane, l'éthylène, l'isopentane, l'acétylène le toluène et le propène. Hormis pour l'éthane, ce facteur peut être identifié comme un profil d'échappement automobile par comparaison avec les profils d'émission obtenus pour un parc automobile représentatif du parc français (Fontaine and Galloo, 2002). L'éthane est présent dans le profil d'échappement automobile mais avec une faible contribution (1,9%). La contribution du facteur modélisé est importante (22%). Compte tenu de sa longue durée de vie, cette espèce tend à s'accumuler dans l'atmosphère. Elle est donc caractérisée par un niveau de fond assez élevé, que ce soit en milieu rural ou urbain, et sa contribution dans les profils modélisés s'en trouve souvent surestimée.



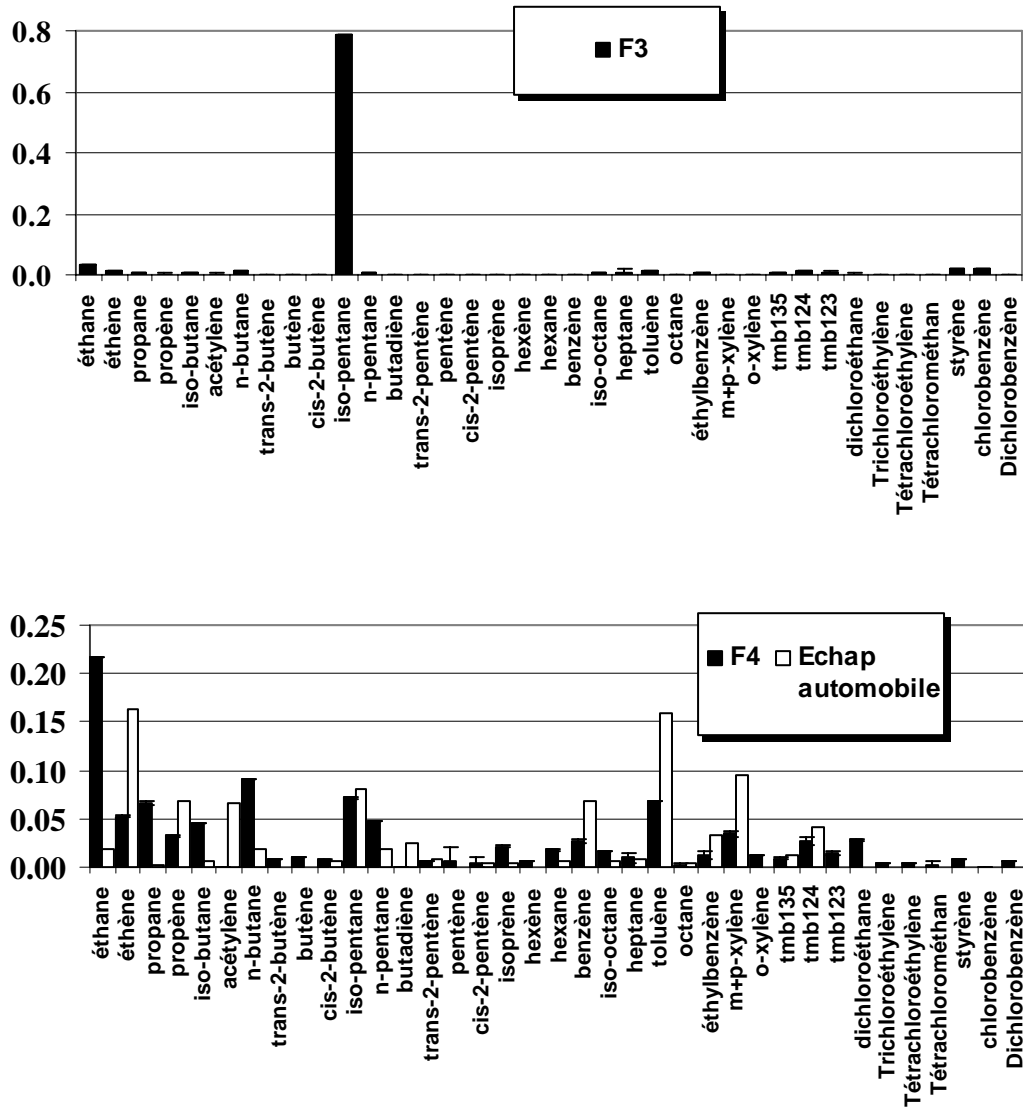
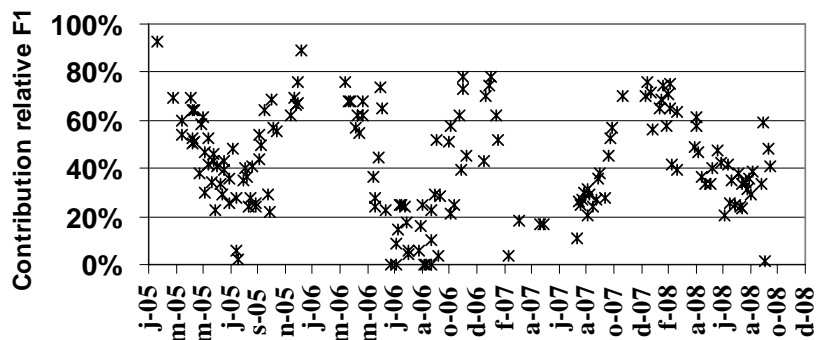


Figure 17 : Composition normalisée des profils de sources modélisés par le PMF pour le site de Nancy Charles III



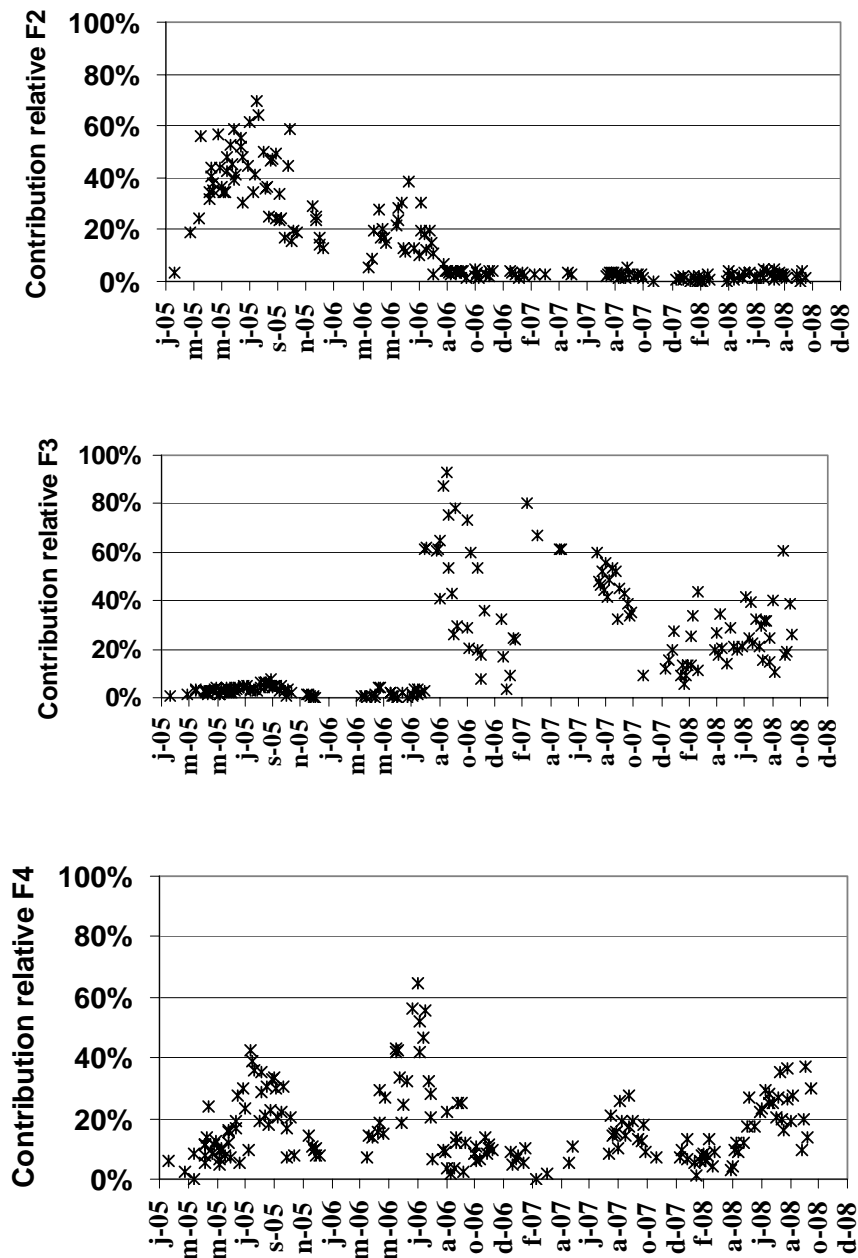


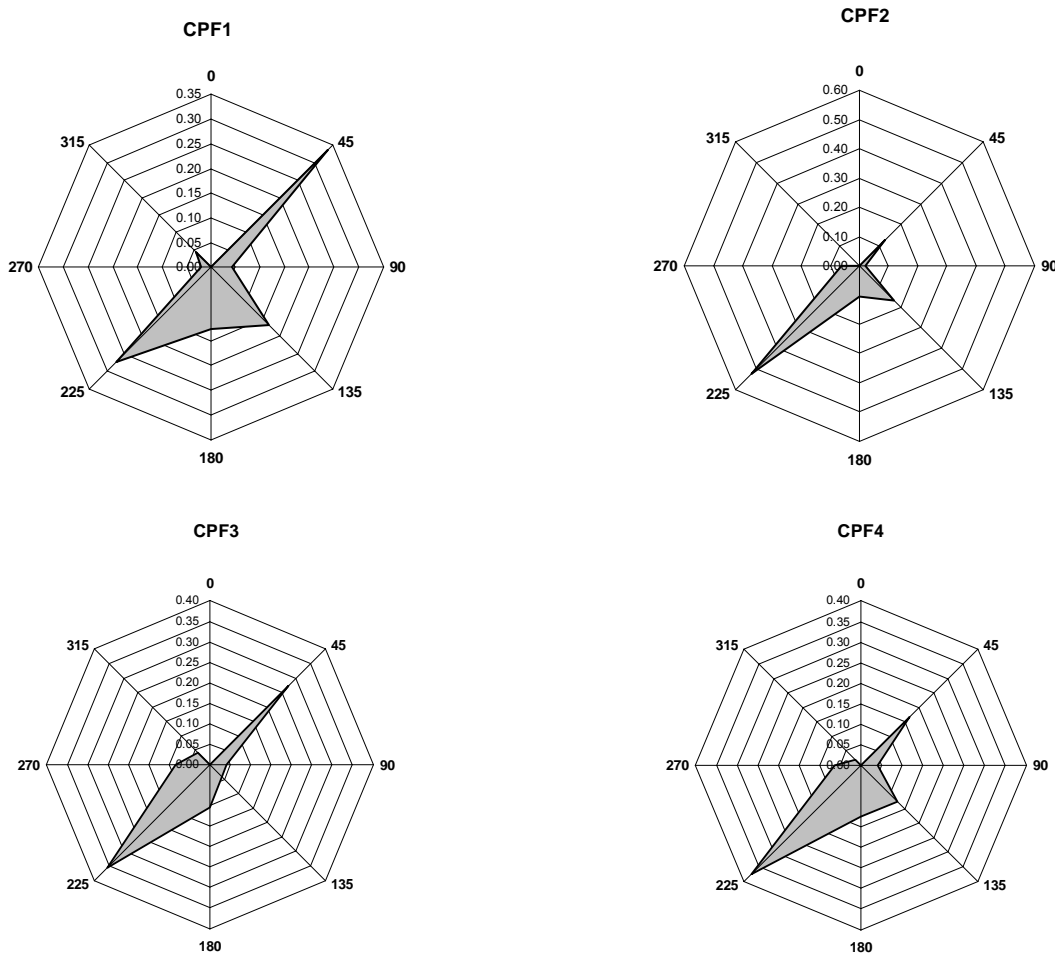
Figure 18 : Suivi chronologiques des contributions des facteurs modélisés pour le site de Nancy Charles III

Pour déterminer le secteur d'appartenance des profils de source modélisés, il est intéressant d'étudier la rose des contributions modélisés. Plus précisément, nous avons utilisé la **Fonction de Probabilité Conditionnelle (CPF)** qui est une rose de pollution plus élaborée qui permet d'identifier les secteurs de provenance de masses d'air associés à de fortes contributions d'un facteur (Ashbaugh et al., 1985). La CPF est définie suivant la relation suivante :

$$CPF = \frac{m_{\theta}}{n_{\theta}} \quad (\text{Eq. 9})$$

- $m_{\theta}$  est le nombre d'observations dans un secteur de vent donné ayant une contribution supérieure à un seuil défini. Pour cette étude, ce seuil est le percentile 75 des contributions modélisées pour un facteur
- $n_{\theta}$  est le nombre total d'observations dans le secteur.

Les résultats obtenus pour les quatre facteurs déterminés par le PMF sont reportés Figure 19. Comme la vue aérienne représentée Figure 20, le site est situé dans une rue canyon. Les roses CPF obtenues montrent donc deux secteurs privilégiés : Sud Ouest et Nord Est. La probabilité de fortes contributions sont plus importantes au Nord Est au Sud Ouest pour le premier facteur (Chauffage résidentiel). Les fortes contributions apparaissent plus fréquentes sur le secteur Sud-Ouest pour les 3 autres facteurs.



**Figure 19 : Rose des CPF des Contributions des facteurs modélisés pour le site de Nancy Charles III**





**Figure 20 : Vue aérienne du site de Nancy Charles III (point rouge) et de son environnement**

La contribution relative moyenne  $C_k$  de chaque facteur est calculée à partir de la contribution  $g_{ik}$  du facteur  $k$  pour l'échantillon  $i$  suivant la relation :

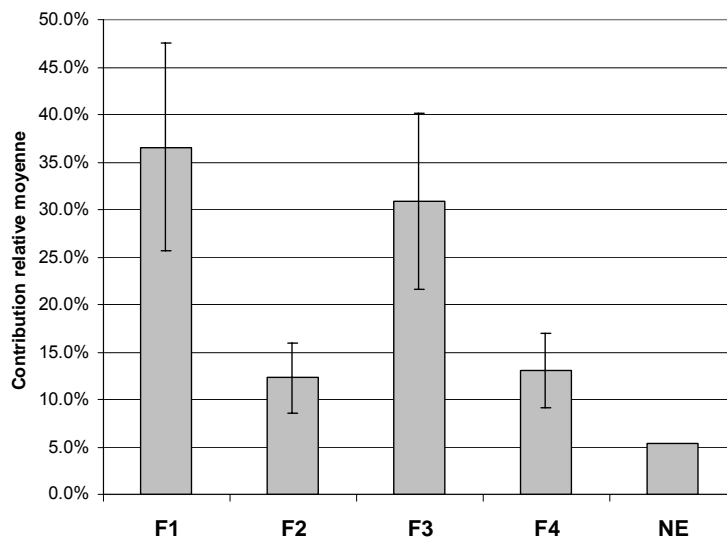
$$C_k = \frac{\sum_{i=1}^n g_{ik}}{\sum_{k=1}^p \sum_{i=1}^n g_{ik}} \times 100$$

Les résultats comparatifs pour les facteurs modélisés sont représentés par des histogrammes, avec des barres représentant les incertitudes des contributions (Figure 21). En effet, le modèle fournit une estimation des incertitudes sur les poids des espèces par facteur ( $F$ ) et sur les contributions modélisées ( $G$ ). Ces incertitudes sont calculées à partir des incertitudes sur les concentrations mesurées qui sont intégrées au calcul par régression.

C'est le facteur  $F1$  identifié « Chauffage résidentiel » qui contribue majoritairement (37%) aux concentrations en HCNM mesurées sur le site de Nancy Charles III. Le facteur  $F3$  principalement associé à l'isopentane contribue pour 31%. Les facteurs  $F2$  (Usage de solvants) et  $F4$  (Echappement automobile) ont des

contributions respectives de 12% et 13%. La contribution de la part non expliquée est de 5%.

Ces résultats peuvent difficilement être rapprochés des inventaires d'émission puisque des sources particulières très localisées semblent contribuer fortement aux observations et donner ainsi une répartition localisée très particulière des contributions de sources. Par ailleurs, si elle s'avérait pertinente, cette comparaison nécessiterait de convertir les données d'entrée du modèle de ppb en  $\mu\text{g}/\text{m}^3$  pour exprimer des contributions en % massique comparables aux données d'inventaire.



**Figure 21 : Contributions relatives moyennes et incertitudes associées des facteurs modélisés et de la part Non Expliquée (NE) pour le site de Nancy Charles III**

#### 4.2.2 Site de Mairie

Les résultats du modèle **expliquent 90% de la concentration totale en HCNM** mesurée. **84% de la variabilité** des concentrations des espèces sont expliqués par **3 facteurs** modélisés. Les concentrations totales modélisées sont bien corrélées avec les concentrations totales mesurées puisque le coefficient de corrélation calculé sur l'ensemble des échantillons est de 0,88.

La Figure 22 représente la composition normalisée des 3 facteurs modélisés pour le site de Mairie.

Le facteur F1 est marqué par une contribution prépondérante du butane suivi dans une moindre mesure par le propane et l'isobutane. Pour comparaison, le

profil type « évaporation de carburants » a été reporté, le butane et l'isobutane sont en général associés avec le pentane et l'isopentane. Ce n'est pas le cas pour le facteur modélisé, les analyses précédentes ont, en revanche, montré une association avec le 1,4-diclorobenzène et le trichloroéthylène, espèces qui n'ont pas été intégrées dans l'analyse PMF pour ce site. Ce profil peut provenir d'évaporation de produits particuliers ou encore de fuites de gaz. Un entreposage de produits d'entretien a été signalé dans un local technique de l'Hôtel de Ville non loin du site de prélèvement. L'examen chronologique des contributions modélisées pour ce profil (Figure 23) montre un profil estival confirmant le renforcement de l'évaporation ou des fuites avec la température. Les contributions sont caractérisées par une tendance à l'augmentation. Elles sont devenues très faibles à partir de mi 2007 pour s'accroître à nouveau en 2008. D'après les informations disponibles, les produits ont été évacués du local de stockage courant 2007. Des teneurs plus faibles en composés chlorés ont été mesurées ensuite avant de s'élever à nouveau courant 2008.

L'éthane, l'éthylène, le propane, l'acétylène et le butane contribuent significativement au facteur **F2** modélisé. Ce profil est en accord avec le profil d'émission «**chauffage résidentiel**» (GENEMIS). Tout comme pour le site Nancy Charles III traité précédemment, le benzène présente une contribution assez faible comparativement au profil d'émission. Les modes de chauffage urbains sont principalement au gaz ou au fioul moins émetteurs en benzène. Le profil de variation nettement hivernal corrobore l'identification de ce profil. Il est intéressant de noter la similitude des profils F1 de Nancy Charles III et F2 de Mairie reportés sur le même graphe (Figure 22). Le coefficient de corrélation calculé à partir du poids des 30 espèces communes pour ces deux profils est de 0,96 montrant le lien étroit entre ces deux profils. Ces profils peuvent être considérés comme des profils types urbains associés aux émissions du chauffage résidentiel.

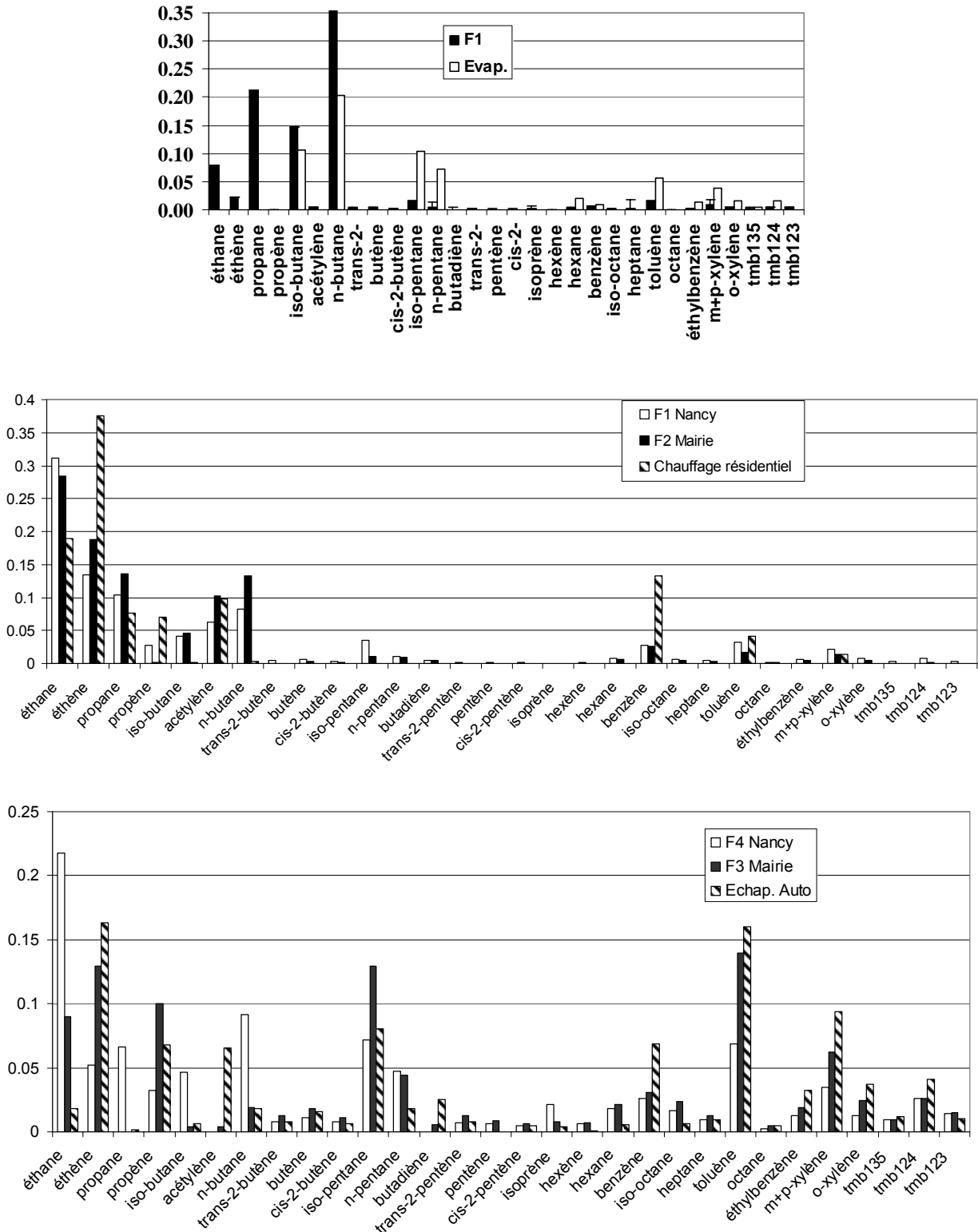


Figure 22 : Composition normalisée des profils de sources modélisés par le PMF pour le site de Mairie (1<sup>er</sup> facteur) et le site de Nancy pour comparaison (pour le 2<sup>nd</sup> et le 3<sup>ème</sup> facteur)

Le quatrième facteur est dominé par l'éthylène, l'isopentane, le toluène ; le propène et les mp-xylènes. Ce facteur peut être identifié comme un profil d'échappement automobile par comparaison avec le profil d'émission obtenu pour un parc automobile représentatif du parc français (Fontaine and Galloo, 2002). Tout comme pour le site de Nancy, l'éthane présente une contribution significative pour ce profil. A contrario l'acétylène souvent considéré comme marqueur des émissions d'échappement automobile est peu représenté avec une contribution de 0,4% pour le profil modélisé contre 6,4% pour le profil d'émission type. Les profils modélisés labelisés « échappement automobile » de Nancy et de Mairie concordent moyennement ( $R=0,56$ ), la correspondance avec le profil d'émission est moins nette que pour le profil « chauffage résidentiel ». Il faut noter que les émissions automobiles sont très variables sur 24 heures avec des périodes de pointe, contrairement aux émissions liées au chauffage plus lissées. Comme les mesures correspondent à des prélèvements 24 heures, les variations inter-espèces concernant le profil automobile sont plus difficiles à discerner. Le suivi chronologique de la contribution de ce facteur montre un profil plutôt estival. Alors que le profil de combustion « Chauffage résidentiel » domine en hiver, la contribution relative du profil échappement automobile est plus importante en été.

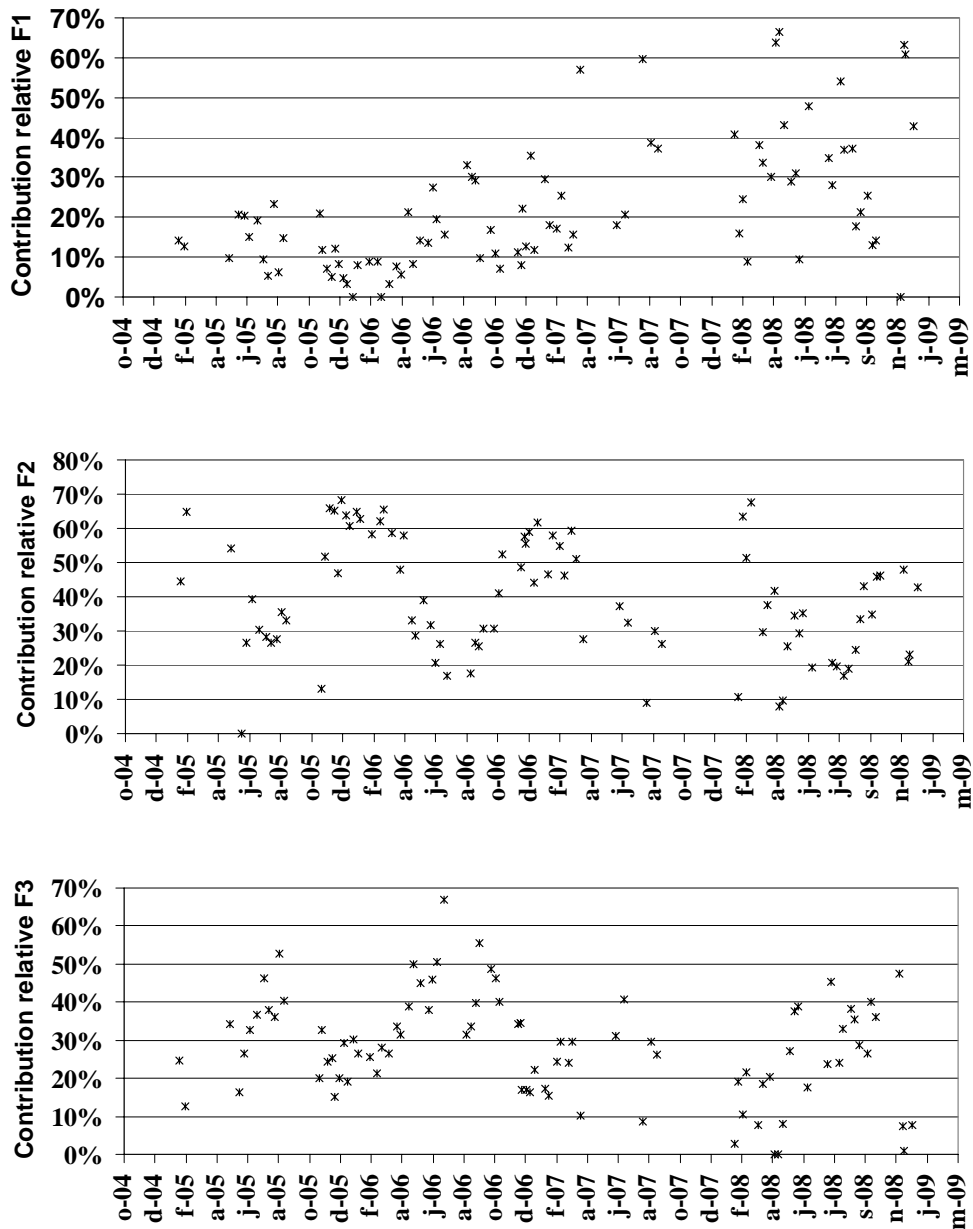
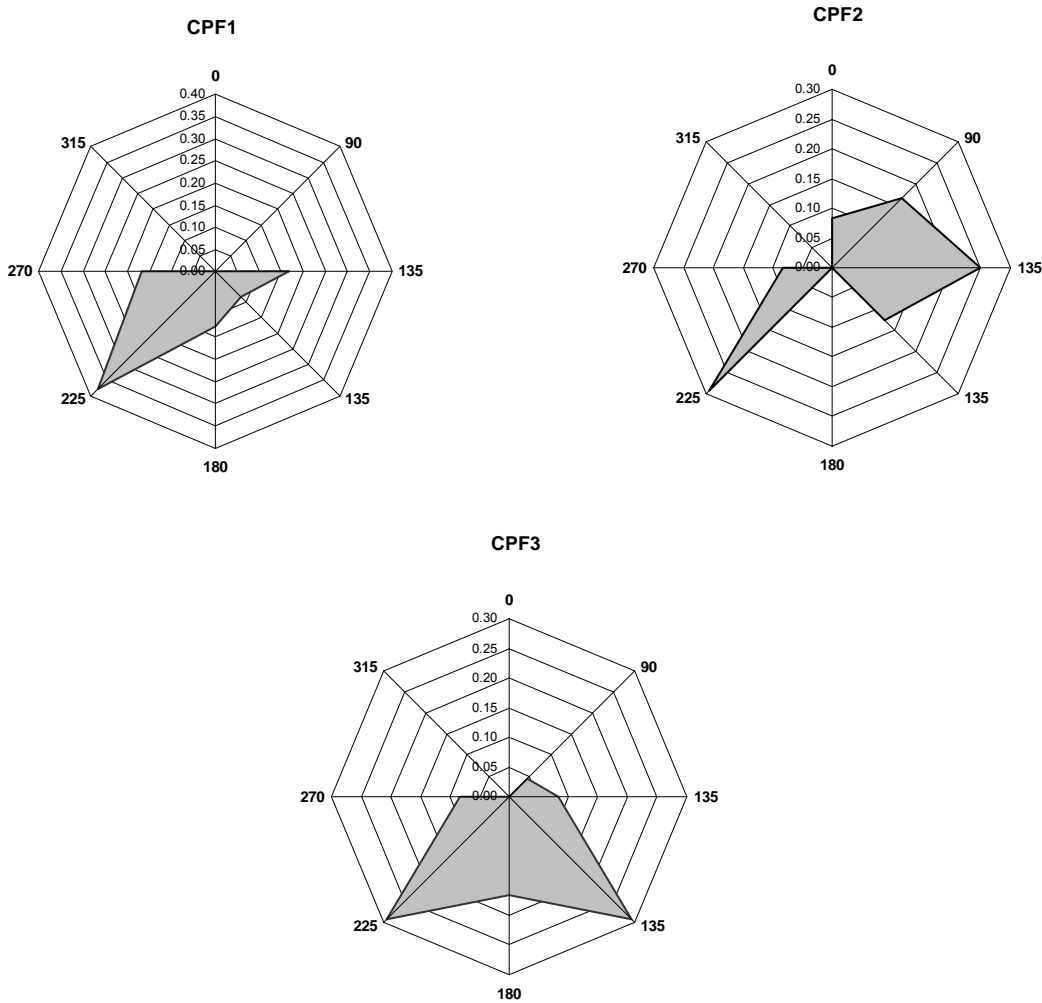


Figure 23 : Suivi chronologiques des contributions des facteurs modélisés pour le site de Mairie

L'examen des roses de CPF obtenues pour les trois facteurs déterminés par le PMF pour le site de Mairie sont reportés Figure 24. Les résultats sont à considérer avec précaution compte tenu du nombre limitées de données par secteur. Les roses CPF obtenues montrent deux secteurs privilégiés : Sud Ouest et Est-Sud-Est en accord avec la configuration du site de prélèvement (Figure 25). Le facteur F1 (Evaporations ou fuites) montre des contributions relatives fortes sur le secteur Sud Ouest permettant de situer le local technique où sont entreposés des produits

d'entretien. Pour le facteur labellisé « échappement automobile » la probabilité d'avoir des contributions fortes est identique pour les secteurs Sud-Est et Sud-Ouest. Concernant le facteur « chauffage résidentiel », les fortes contributions proviennent du secteur Sud-Ouest, et dans une moindre mesure de tout le secteur Nord à Sud-Est.



**Figure 24 : Rose des CPF des Contributions des facteurs modélisés pour le site de Mairie**

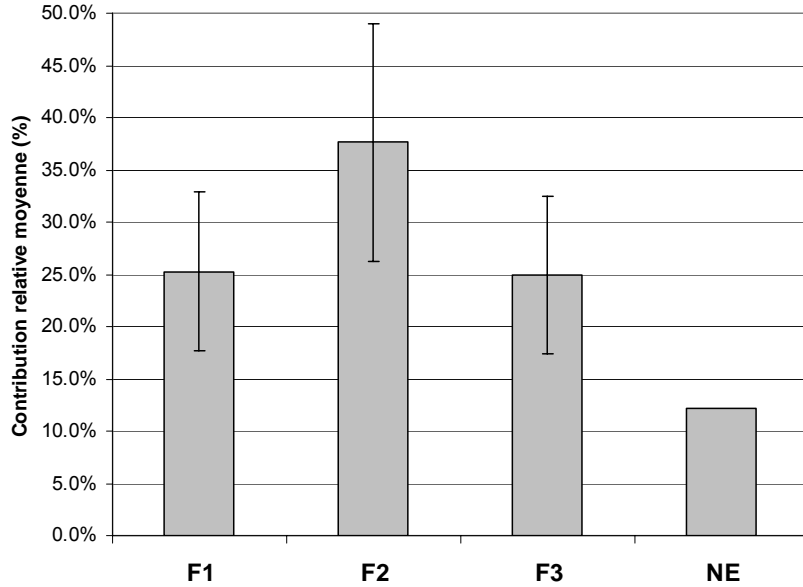




**Figure 25 : Vue aérienne du site de Mairie (point rouge) et de son environnement**

Les contributions relatives moyennes sont représentées avec leur incertitude Figure 26. C'est le facteur F2 identifié « Chauffage résidentiel » qui contribue majoritairement (38%) aux concentrations en HCNM mesurées sur le site Maire. Ce résultat est cohérent avec les résultats du site Nancy Charles III (facteur prépondérant contribuant à 37%). Le facteur F1 correspondant à l'évaporation de composés ou à des fuites de gaz contribue pour 25%. Le facteur F3 (échappement automobile) a une contribution relative de 25%. La contribution de la part non expliquée est assez élevée (12%) notamment en raison de la faible dimension de la base pour ce site (N=90).





**Figure 26 : Contributions relatives moyennes et incertitudes associées des facteurs modélisés et de la part Non Expliquée (NE) pour le site de Mairie**

## 5 CONCLUSION

Cette étude a consisté en l'analyse de la base de données canist'air comprenant les mesures de 37 espèces hydrocarbonées réalisées sur 17 sites en zone urbaine ou industrielle. La méthode de mesure est basée sur un échantillonnage d'une durée de 24 heures par canister et une analyse différée en laboratoire par chromatographie en phase gaz.

L'objectif de cette analyse est d'étudier l'intérêt de ce type de données dans les stratégies de surveillance déployées par les AASQA. L'étude a été réalisée en deux temps :

Une première phase pour laquelle le rapport a été émis en novembre 2007. Il s'agissait d'une analyse descriptive de la base après consolidation. Suite à l'examen de la représentativité temporelle des mesures, l'analyse a permis de comparer les sites suivant leur typologie avec d'autres sites nationaux et étrangers. D'une manière générale il est apparu que les teneurs mesurées sur les sites français de type urbain sont en excellent accord avec celles mesurées dans plusieurs agglomérations européennes. Les sites français de type « trafic » présentent des concentrations supérieures aux sites français du type « urbain » mais qui restent

néanmoins bien en deçà de ce qui est mesuré sur les agglomérations d'Asie ou d'Amérique. Quel que soit le type de site, l'éthane et l'éthylène sont les composés majoritaires. Le butane est parmi les 5 composés prépondérants pour la plupart des sites. La présence de propane et de toluène traduirait plutôt un site de type « urbain » alors que la présence d'acétylène et d'isopentane caractériserait un site de type « Traffic ». En classant les espèces suivant leur réactivité (indirectement suivant leur potentiel de formation d'ozone), les cinq premiers composés sont l'éthylène, le propène, l'isoprène, les m,p-xylènes et le 1,2,4-triméthylbenzène indépendamment de la typologie des sites étudiés.

La deuxième phase de l'étude a eu pour but de pousser plus loin l'analyse notamment pour dégager des informations concernant la typologie des sites et les profils de sources urbains. Il s'agissait d'étudier différents ratios d'espèces avant de mettre en œuvre une approche multi-variée avec un modèle sources-récepteur.

La base de données a été complétée entre les deux phases avec les mesures de l'année 2008. Une nouvelle analyse descriptive préalable a été menée sur 8 des 17 sites présentant au moins 30 échantillons. Elle a permis notamment d'étudier les profils saisonniers de variation des concentrations pour caractériser le comportement de quelques espèces.

Outre la confirmation des éléments déterminés avec la première phase d'étude, la comparaison des sites espèce par espèce ou à partir de ratios spécifiques a permis de mettre en évidence quelques particularités : Le site de Nancy présente des concentrations élevées en isopentane, en pentane, en styrène et en xylènes. Le site de Mairie est marqué par des teneurs très fortes en butane, en 1,4-dichlorobenzène ou en trichloroéthylène. Au regard d'espèces caractéristiques (acétylène, toluène), le site de Clémenceau est le site le plus exposé aux émissions automobiles. Le site industriel de Sochaux montre des concentrations en xylènes indiquant un usage de solvants. L'approche permettant d'estimer l'âge photochimique des observations a montré que les sites de la base canist'air sont influencés quasi exclusivement par des masses d'air jeunes. D'après les données disponibles, la configuration des sites et le type d'échantillonnage ne permet pas de distinguer l'influence de sources locales des apports de masses d'air plus lointaines.

L'analyse multi-variée demande un nombre d'échantillons suffisant pour obtenir des résultats stables et significatifs. Aussi, seuls les sites de Mairie et de

Nancy Charles III ont pu être considérés pour une application du modèle sources-récepteur PMF. 2 facteurs communs ont été modélisés. Le premier a été associé à la source « chauffage résidentiel » et contribue majoritairement (37%) aux concentrations mesurées sur les deux sites. Le deuxième correspond au profil de la source automobile, les contributions sont de 13% et de 25% respectivement pour Nancy et Mairie. Les autres profils sont plus difficiles à identifier, ils semblent correspondre à des sources de types évaporation de carburants, fuite de gaz ou usage de solvants. L'analyse des contributions, en suivi chronologique et par secteur de vent, montre le caractère particulier, intense et localisé des sources associées à ces profils.

La démarche d'analyse de données plus avancée met en lumière plusieurs écueils concernant la base étudiée. L'approche statistique pour ce type d'espèces nécessite d'une part, un nombre de données important pour une bonne significativité des résultats et, d'autre part, une bonne représentativité temporelle des données pour prendre en compte l'évolution temporelle des processus de production et de déplétion des espèces. Seuls les sites de Nancy et de Mairie (tout juste suffisant) ont permis ce type d'approche. Par ailleurs la fréquence d'échantillonnage de 24 heures intègre les variations potentielles des concentrations au cours d'une journée liées aux différents processus intervenant dans l'atmosphère, ce qui limite la détection de variabilité inter espèces pouvant conduire à l'identification plus fine de sources.

Toutefois, les résultats obtenus montrent l'intérêt de ce type d'observations s'agissant de la surveillance des niveaux d'exposition et de l'identification des sources et de leur contribution. Une réflexion doit être menée pour identifier clairement les objectifs de mesure de ces espèces par canisters, objectifs qui peuvent être l'étude d'impact, la surveillance à long terme, l'étude d'épisode ou encore la connaissance des relations sources-récepteurs. Suivant les objectifs envisagés, la stratégie d'échantillonnage pourra alors être revue et optimisée suivant le tableau support proposé ci-après (Tableau 4).

**Tableau 4 : Stratégies d'échantillonnage suivant les objectifs de la mesure**

| Objectifs   | Echantillonnage           |                                 | Typologie site  | Espèces                              |
|---|---------------------------|---------------------------------|---|--------------------------------------|
|   | Durée                     | Fréquence                       |   |                                      |
| Exposition de la population                                     | plusieurs heures          | multi hebdo,<br>régulier        | Urbain/urbain de fond                                 | Réglementées +<br>Recommandées OMS   |
| Représentativité/typologie de site                              |                           |                                 | Tout type   | Réglementées + précurseurs           |
| Etude d'impact  | Courte (suivant activité) | avant, après                    | Zone d'impact   | Traceurs de l'activité ciblée        |
| Etude d'épisode   | 1 heure                   | plusieurs/jours                 | Urbain, péri-urbain                                   | Précurseurs, secondaires (oxygénées) |
| Surveillance à long terme                                       | ≥4h (fixe)                | hebdo, régulier                 | péri-urbain, urbain de fond, rural                    | Majoritaires, traceurs               |
| Relation sources Récepteur, Confrontation inventaires émissions | 1 heure                   | plusieurs/jours (nb données >>) | péri-urbain, urbain de fond (suivant récepteur ciblé) | Majoritaires, traceurs               |

## 6 BIBLIOGRAPHIE

- Borbon, A., P. Coddeville, N. Locoge and J.-C. Galloo** (2004). Characterising sources and sinks of rural VOC in eastern France. *Chemosphere* 57(8): 931-942.
- Borbon, A., N. Locoge, M. Veillerot, J. C. Galloo and R. Guillermo** (2002). Characterisation of NMHCs in a French urban atmosphere: overview of the main sources. *The Science of The Total Environment* 292(3): 177-191.
- Boudries, H.** (1994). Mesure et comportement des hydrocarbures non méthaniques en atmosphère naturelle - expérimentation à la pointe de Bretagne. PARIS, PARIS 7. Doctorat: 164.
- Derwent, R. G., T. J. Davies, M. Delaney, G. J. Dollard, R. A. Field, P. Dumitrean, P. D. Nason, B. M. R. Jones and S. A. Pepler** (2000). Analysis and interpretation of the continuous hourly monitoring data for 26 C2-C8 hydrocarbons at 12 United Kingdom sites during 1996. *Atmospheric Environment* 34(2): 297.
- Fontaine, H. and J.-C. Galloo** (2002) Volatile organic compound emission factors in the exhaust gases of light duty vehicles: uncertainties, influence of fuel formulation and new data, Technical report, Department of Chemistry and Environment, Ecole des Mines de Douai, Douai, France, Final report PRIMEQUAL-PREDIT 1999, March 2002
- GENEMIS** Emission dataset, Institute for Energy Economics and rational use of energy, Stuttgart University.
- Hopke** (2000). A guide to positive Matrix Factorization. EPA Workshop proceedings materials from the workshop on UNMIX and PMF as applied to PM2.5.
- Jobson, B. T., Z. Wu and H. Niki** (1994). Seasonal trends of isoprene, C2-C5 alkanes and acetylene at a remote boreal site in Canada. *Journal of Geophysical Research* 99: 1589-1699.
- Lanz, V. A., C. Hueglin, B. Buchmann, M. Hill, R. Locher, J. Staehelin and S. Reimann** (2008). Receptor modeling of C2-C7 hydrocarbon sources at an urban background site in Zurich, Switzerland: changes between 1993-1994 and 2005-2006. *Atmos. Chem. Phys.* 8(9): 2313-2332.
- Lee, E., C. K. Chan and P. Paatero** (1999). Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. *Atmospheric Environment* 33(19): 3201.
- Mayrsohn, H. and J. H. Crabtree** (1976). Source reconciliation of atmospheric hydrocarbons. *Atmospheric Environment* 10: 137-143.
- McKeen, S. A.** (1993). Hydrocarbon ratios and photochemical history of air masses. *Geophysical research letters* 20(21): 2363-2366.
- Paatero, P.** (1997). A weighted non-negative least squares algorithm for three-way 'PARAFAC' factor analysis. *Chemometrics and Intelligent Laboratory Systems* 38(2): 223-242.
- Paatero, P. and P. K. Hopke** (2002). Utilizing wind direction and wind speed as independent variables in multilinear receptor modeling studies. *Chemometrics and Intelligent Laboratory Systems* 60(1-2): 25.
- Paatero, P., P. K. Hopke, X.-H. Song and Z. Ramadan** (2002). Understanding and controlling rotations in factor analytic models. *Chemometrics and Intelligent Laboratory Systems* 60(1-2): 253.

- Paatero, P. and U. Tapper** (1994). Positive Matrix Factorization : a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* 5: 111-126.
- Parrish** (1992). Indication of photochemical histories of Pacific air masses from measurements of atmospheric trace species at Point Arena, California. *Journal Of Geophysical Research-Atmospheres*.
- Parrish, D. D., A. Stohl, C. Forster, E. L. Atlas, D. R. Blake, P. D. Goldan, W. C. Kuster and J. A. de Gouw** (2007). Effects of mixing on evolution of hydrocarbon ratios in the troposphere. *Journal of Geophysical Research-Atmospheres* 112(D10).
- Passant, N. R.** (2002) Speciation of UK emissions of non-methane volatile organic compounds, AEA Technology, Report No. AEAT/ENV/R/0545 Issue 1,
- Qin, Y., K. Oduyemi and L. Y. Chan** (2002). Comparative testing of PMF and CFA models. *Chemometrics and Intelligent Laboratory Systems* 61(1-2): 75.
- Sauvage, S., H. Plaisance, N. Locoge, A. Wroblewski, P. Coddeville and J. C. Galloo** (2009). Long term measurement and source apportionment of non-methane hydrocarbons in three French rural areas. *Atmospheric Environment* In Press, Corrected Proof.
- Simon, V., L. Dumergues, J.-L. Ponche and L. Torres** (2006). The biogenic volatile organic compounds emission inventory in France: Application to plant ecosystems in the Berre-Marseilles area (France). *Science of The Total Environment* 372(1): 164-182.

## **ANNEXES**





**ANNEXE 1 : FICHE D'ETUDE LCSQA****THEME 8 : MODELISATION - TRAITEMENTS NUMERIQUES****Caractérisation du comportement spatio-temporel des COV à partir de l'exploitation des données canisters**

*Responsable de l'étude : EMD*

**Objectif**

L'objectif de cette étude est d'étudier l'intérêt des campagnes de mesure des 31 COV précurseurs à l'aide de canisters à partir de l'exploitation scientifique des données de concentrations en COV mesurés par les AASQA depuis 2005 dans différentes régions. Elle devra permettre un transfert à l'ensemble des AASQA d'éléments scientifiques relatifs à l'apport des canisters à la stratégie de surveillance des COV.

**Contexte et travaux antérieurs**

Depuis plusieurs années des travaux ont été menés dans la caractérisation du comportement spatio-temporel des COV à partir des données acquises à partir d'analyseurs automatiques mis en place dans plusieurs AASQA. En effet, pour trois d'entre elles, la mesure automatique et continue des COV depuis mai 2001.

Ces travaux ont porté essentiellement sur :

- l'exploitation des résultats en vue d'identifier au travers de l'analyse statistique les différentes sources influençant le site récepteur avec l'analyse d'une part temporelle des résultats de mesures actuellement en cours et ce sur plusieurs pas de temps (horaire, saisonnier) et d'autre part spatiale compte tenu de la dispersion géographique des différents sites concernés (Marseille, Grenoble, Strasbourg, ...) notamment en terme de répartition des COV dans les différentes familles.
- le développement des différentes approches (approches monovariée, bivariée, multivariée au travers de l'analyse en composantes principales) permettant de conduire l'analyse fine des évolutions temporelles des COV. Ce travail a permis de mieux comprendre le comportement temporel des COV sur plusieurs sites soumis à des sources différentes compte tenu des typologies des sites très différentes. Cette analyse a également permis d'identifier les principales sources de COV influençant chacun des sites.
- la réalisation de différentes modélisations visant à évaluer la contribution des différentes sources influençant les teneurs ambiantes. Plusieurs modèles (modèle de régression linéaire simple, CMB, PMF) ont été testés afin de vérifier la cohérence des résultats et de choisir le modèle le mieux adapté à l'exploitation des données de COV. Afin de comparer les deux approches que sont les approches sources-récepteur et les inventaires d'émission, il paraît particulièrement intéressant de comparer les données du cadastre aux résultats obtenus à l'aide de méthodes permettant la quantification des contributions de ces sources et ce sur au moins un site de mesure. En effet, les résultats

des applications CMB ont déjà été utilisées à des fins de validation des inventaires d'émissions.

Depuis quelques années, l'ASPA réalise l'analyse des 31 COV précurseurs de l'ozone présents dans l'air prélevé à l'aide de canisters dans différentes régions (Alsace, Lorraine, Franche Comté, Champagne Ardenne, Rhône Alpes et Languedoc Rousillon).

Une base de données appelée CANISTAIR a été construite par l'ASPA pour compiler l'ensemble des résultats. D'autres AASQA (ATMOPACA, ...) ont également réalisé des campagnes de mesures des COV à l'aide de canisters dans différents cadres.

En 2007, à partir de la base de données CANISTAIR fournie par l'ASPA et complétée par des données compatibles d'autres AASQA volontaires, nous avons réalisé une exploitation relative à :

- **l'étude de la répartition spatiale des COV sur des sites de typologies différentes et de l'évolution temporelle des concentrations**
- **la cohérence entre la classification des sites selon leur typologie et les niveaux et ratios des différents COV spécifiques.**

### Travaux proposés en 2008

Nous compléterons les travaux 2007 par :

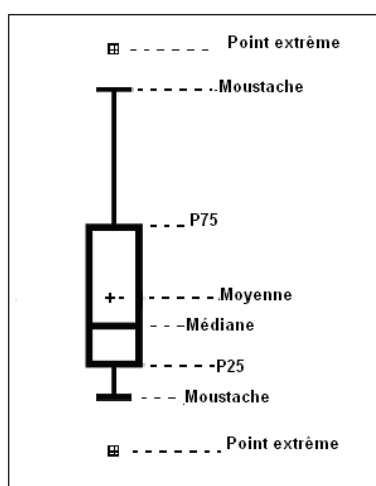
- **l'étude de la représentativité des sites** à partir de l'établissement des différentes typologies de sites
- **l'identification des sources.**

### Renseignements synthétiques

|                                   |   |  |       |
|-----------------------------------|---|--|-------|
| Titre de l'étude                  |   | Caractérisation du comportement spatio-temporel des COV à partir de l'exploitation des données canisters |       |
| Personne responsable de l'étude   |   | Nadine LOCOGE  |       |
| Travaux                           | pluri-annuels   |  |       |
| Durée des travaux pluriannuels    | 2 ans (2007-2008)                                     |  |       |
| Collaboration AASQA               | ASPA + AASQA réalisant la mesure de COV par canisters |  |       |
| Heures d'ingénieur                | EMD : 300   | INERIS :   | LNE : |
| Heures de technicien              | EMD :   | INERIS :   | LNE : |
| Document de sortie attendu        | Rapport   |  |       |
| Lien avec le tableau de suivi CPT | 3) Traitements numériques des données                 |  |       |
| Lien avec un groupe de travail    | non   |  |       |
| Matériel acquis pour l'étude      | -   |  |       |

## ANNEXE 2 : SIGNIFICATION DE LA BOITE A MOUSTACHES

Les boîtes à moustaches qui sont des représentations graphiques permettant de réunir sur un même tracé plusieurs caractéristiques pour un jeu de variables numériques. La boîte à moustache se présente sous la forme d'un rectangle représentant les valeurs comprises entre le 25<sup>ème</sup> percentile (P25) et le 75<sup>ème</sup> percentile (P75) de l'ensemble des données. Les moustaches, segments issus de la boîte, s'étirent jusqu'à 1,5 fois l'espace interquartile. Une croix représente la moyenne arithmétique des données et le segment qui partage la boîte symbolise la médiane. Enfin les points hors des moustaches sont appelés points extrêmes.



*Signification de la représentation d'une boîte à moustaches*

### ANNEXE 3 : LIMITE DE QUANTIFICATION (LQ) DE LA METHODE DE MESURE PAR CANISTER

| Composés               | LQ (ppb) |
|------------------------|----------|
| Ethane                 | 0,33     |
| Ethylène               | 0,33     |
| Propane                | 0,23     |
| Propène                | 0,17     |
| i-Butane               | 0,10     |
| n-Butane               | 0,10     |
| Acétylène              | 0,13     |
| Trans-2-butène         | 0,10     |
| 1-Butène               | 0,10     |
| Cis-2-butène           | 0,10     |
| i-Pentane              | 0,07     |
| n-Pentane              | 0,07     |
| 1,3-Butadiène          | 0,10     |
| Trans-2-pentène        | 0,07     |
| 1-Pentène              | 0,07     |
| Cis-2-pentène          | 0,07     |
| Isoprène               | 0,07     |
| 1,1-dichloroéthane     | 0,21     |
| 1-Hexène               | 0,10     |
| 1,2-dichloroéthylène   | 0,19     |
| n-Hexane               | 0,10     |
| 1,2-dichloroéthane     | 0,24     |
| 1,1,1-trichloroéthane  | 0,23     |
| Benzène                | 0,10     |
| tétrachlorométhane     | 0,35     |
| Trichloroéthylène      | 0,21     |
| i-octane               | 0,07     |
| n-Heptane              | 0,07     |
| 1,1,2-trichloroéthane  | 0,18     |
| Toluène                | 0,07     |
| n-Octane               | 0,07     |
| Tétrachloroéthylène    | 0,14     |
| Chlorobenzène          | 0,07     |
| Ethylbenzène           | 0,07     |
| m+p-Xylène             | 0,07     |
| styrène                | 0,03     |
| o-Xylène               | 0,03     |
| 1,3,5-Triméthylbenzène | 0,07     |
| 1,2,4-Triméthylbenzène | 0,07     |
| 1,4-dichlorobenzène    | 0,09     |
| 1,2,3-Triméthylbenzène | 0,07     |