

# Identification et contribution des sources de HCNM en zone rurale

## Source apportionment of Non-Methane Hydrocarbons in French rural area

Stéphane SAUVAGE<sup>\*, \*\*, \*\*\*</sup>, Nadine LOCOGE<sup>\*, \*\*, \*\*</sup>,  
Hervé PLAISANCE<sup>\*, \*\*, \*\*</sup>, Patrice CODDEVILLE<sup>\*, \*\*, \*\*</sup>, Jean-Claude GALLOO<sup>\*, \*\*, \*\*</sup>

### Résumé

Les composés hydrocarbonés non méthaniques (HCNM) jouent un rôle prépondérant dans la chimie de l'atmosphère puisqu'ils interviennent en tant que précurseurs de la formation de polluants secondaires comme l'ozone ou les aérosols organiques secondaires. Disposant d'une base de données conséquente d'au moins cinq années de mesures systématiques de 46 HCNM sur trois sites de l'observatoire MERA (Mesure et Évaluation en zone Rurale de la pollution Atmosphérique à longue distance), le modèle sources-récepteur PMF (*Positive Matrix Factorization*) a été appliqué pour mettre en évidence des associations inter-espèces et identifier des profils de sources à l'origine des concentrations mesurées sur les sites récepteurs. L'application de ce type de modèle sur des sites ruraux, éloignés des sources et à des espèces réactives comme les HCNM a constitué une nouvelle approche. L'hypothèse de la conservation des masses sur lequel s'appuie le modèle est discutable compte tenu de la réactivité différente des espèces. C'est pourquoi la réactivité a été prise en compte par l'ajout d'une composante d'incertitude comme données d'entrée du modèle. Une étude de sensibilité du modèle a permis de consolider les résultats obtenus par rapport aux choix méthodologiques.

Cinq profils de sources communs aux trois sites ont été identifiés : les profils « chauffage résidentiel », « échappement automobile », « évaporation de carburants », « biogénique » et « sources lointaines ». Les contributions relatives de ces sources aux teneurs en HCNM sont cohérentes avec les données d'inventaire d'émissions et avec les résultats d'autres travaux. Elles font apparaître une hiérarchie différente et logique pour les trois sites. L'analyse de ces contributions par classe de rétrotrajectoires de masses d'air met en évidence les parts d'influence locale et continentale où les profils de sources urbaines contribuent davantage alors que la part d'influence océanique est dominée par le profil « sources lointaines ».

### Mots clés

Hydrocarbures non méthaniques (HCNM).  
Modèle sources-récepteur. Contributions de sources.  
Rétrotrajectoire de masses d'air. Classification  
« K-Means ».

### Abstract

It has been well established that non-methane hydrocarbons (NMHCs) play an important role as precursors of ozone and other secondary pollutants like Oxygenated Volatile Organics Compounds (OVOC), organic nitrates or particulate organics matter. Forty-six non-methane hydrocarbon (NMHC) concentrations have been measured between 1997 and 2006 at three MERA sites (Mesure des Retombées Atmosphériques, French acronym for background air pollution monitoring) according to the EMEP programme. VOCs monitoring has been carried out firstly to estimate their background concentration levels and trends, and secondly to provide observation data for model validation and quality check of VOC emission inventories. The objective of this work was to use this extensive dataset in order to identify NMHC source types and to apportion their contributions.

Data analysis was performed using Positive Matrix Factorization (PMF). The general concept of this receptor model being mass conservation, the computed factors may not be directly related to emission profiles when reactive species are measured far from the sources but can be interpreted as aged emission profiles. Beside, in this work the uncertainty taken into account as PMF input data combined uncertainty of the measurement method and photochemical reactivity of each considered species. The PMF parameters and the uncertainty model were established and discussed by examining the sensibility of the model.

Six factors were considered for the 3 sites. The model results explain on average 86.5% of the total concentrations of the 46 hydrocarbons measured for the 3 sites. About 80% of the variability of the VOC species concentrations could be explained by the model for all sites. The calculated factors have been identified as potential source categories by comparing them to known source profiles from the literature, taking into account species lifetimes and examining time series of source contributions. Five of the six PMF source profiles are common to all sites and all factors are identified as known and aged plume profiles: vehicle exhaust, evaporative sources, residential heating, biogenic and remote sources. Contribution of evaporative sources (29.5%) remains predominant for the southern site while residential heating (28.0% and 21.7%) is the main

\* Université de Lille Nord de France – F-59000 Lille – France.

\*\* École des Mines de Douai – Département Chimie et environnement – F-59508 Douai – France.

\*\*\* Correspondant.

factor for the two others. The resultant information allowed comparison with emission inventories usually given by activity sectors. Contributions of vehicle exhaust and evaporative sources ( $36.6 \pm 8.4\%$  on average for the 3 sites) are in agreement with emissions coming from road transport and extraction & distribution of fossil fuels ( $31.6\%$  on average over the period 2002-2006 period), and so is residential heating ( $19.4 \pm 6.4\%$ ) which is close to non-industrial combustion emissions ( $13.7\%$ ). Biogenic and background factors contribute respectively at  $10.5 \pm 3.5\%$  and  $12.7 \pm 3.9\%$  on average for the 3 sites.

This work was completed with an air mass trajectory analysis in order to apportion source contribution depending on air mass origins. Each sample was associated with its 3-day back-trajectory arriving at the sampling station at 12:00 GMT (P = 950 hPa) and calculated by the British Atmospheric Data Centre. The K-means clustering algorithm was used to group these trajectories depending on their direction, their speed and their altitude. Two main groups have been distinguished: (1) older and diluted air

masses from an oceanic origin where "remote sources" present the main contribution; and (2) anthropogenic and closer sources indicating continental influence. In this group, some clusters represent trajectories coming from the north and distant urban areas with high wind speed, and are marked by high contributions of residential heating and evaporative sources. Trajectories with short pathways corresponding to low wind speed allow to discern more local influences from urban areas with vehicle exhaust and residential heating profiles as main contributions.

Further works concerning the potential source locations of NMHC observed in these French rural sites will be published.

### Keywords

Ambient non-methane hydrocarbons. Rural atmosphere. Source apportionment. Air mass trajectories. K-means clustering.

## 1. Introduction

Les composés hydrocarbonés non-méthaniques (HCNM) sont parmi les espèces ayant des impacts directs ou indirects sur l'homme et la nature. En présence d'oxydes d'azote et sous irradiation solaire, ces composés sont à l'origine de la formation de polluants secondaires oxydants comme l'ozone ( $O_3$ ), le peroxyde d'hydrogène ( $H_2O_2$ ) les particules organiques, ou les PéroxyAcyl Nitrates (PAN). Ils sont également précurseurs de la formation de monoxyde et de dioxyde de carbone. En conséquence, ils sont impliqués dans plusieurs phénomènes de pollution comme la pollution photo-oxydante, la formation d'aérosols organiques ou encore le changement global de l'atmosphère. Une part importante des HCNM provient des différentes activités humaines, en particulier l'usage de combustibles fossiles. Les transports (échappement automobile, évaporation des carburants), le chauffage résidentiel, les activités industrielles (production et stockage de carburants, usage de solvants, procédés de combustion) constituent les sources anthropiques prépondérantes. Une part importante des émissions est également attribuable aux sources biogéniques. Compte tenu de ces divers effets directs et indirects, la connaissance des relations entre les émissions et les teneurs en HCNM mesurées sur des sites d'observation reste cruciale pour la compréhension du devenir de ces espèces.

Le programme EMEP (*co-operative Programme for Monitoring and Evaluation of the long range transmission of air Pollutants in Europe*) a été créé suite à la ratification, par 34 gouvernements et la Commission Européenne, de la Convention de Genève (1979) sur la pollution transfrontière. Ce programme porte sur trois points principaux : (1) l'établissement d'un inventaire des émissions européennes, (2) la surveillance de la qualité de l'air et des dépôts dans les zones éloignées de sources de pollution et (3) la modélisation du transport atmosphérique et du dépôt

des polluants. Le principal objectif de ce programme est de fournir régulièrement aux états membres des informations scientifiques pour permettre le développement et l'évaluation des protocoles internationaux visant à la réduction des émissions. Sur le plan français, le dispositif MERA (Mesure et Évaluation en zone Rurale de la pollution Atmosphérique à longue distance) a été mis en place en 1984, à l'initiative du Ministère chargé de l'Environnement et de l'ADEME (Agence De l'Environnement et de la Maîtrise de l'Énergie). Ce programme comprend des mesures systématiques à long terme des espèces impliquées dans la pollution photo-oxydante comme les HCNM qui sont mesurés sur trois des dix sites du dispositif : Donon depuis 1992, Peyrusse-Vieille depuis 2000 et La Tardière depuis 2002.

Par ailleurs, les inventaires d'émissions de HCNM sont des outils importants pour évaluer l'efficacité des politiques de réduction des émissions mises en place. Ils sont également essentiels en tant que données d'entrée des modèles de prévision. Cependant, compte tenu de la multiplicité des HCNM, de la variété des sources et de la variabilité spatio-temporelle des teneurs, les inventaires d'émission ne sont spécifiés que sur quelques zones d'échelle réduite. Les incertitudes sur les quantités émises sont très élevées, variant de 30 à 100 % pour les COVNM [1].

Il est donc complémentaire de confronter les résultats d'inventaires à des données d'observations notamment en s'appuyant sur les concentrations mesurées sur un site récepteur. Il s'agit de relier ces mesures aux sources par l'utilisation de modèles sources-récepteur. Outre la validation des inventaires d'émissions, la démarche va permettre d'identifier les signatures de sources ayant un impact sur les niveaux de fond en HCNM, d'évaluer et d'analyser leur contribution temporellement et géographiquement.

Suivant la présentation comparée des modèles ci-avant, le PMF (*Positive Matrix Factorization*) permet

une bonne identification de signatures de sources sans connaissance *a priori* des profils d'émissions. Il nécessite un nombre d'observations important mais est cependant peu sensible aux valeurs atypiques parfois mesurées sur un site de fond. En outre, ce modèle est intéressant par les possibilités de paramétrage permettant ainsi de consolider les résultats lors d'une étude de sensibilité. Ce modèle s'appuie sur le principe de conservation de la masse pour relier de manière linéaire les concentrations mesurées sur le site récepteur aux quantités d'espèces émises sans connaissance *a priori* des sources. Ce principe est discuté lorsqu'il s'agit de mesures en zones rurales éloignées des sources. En effet, les polluants émis sont impliqués dans plusieurs processus de transformation, de transport et de mélange qui rendent complexes les relations sources-récepteur. Néanmoins, les associations d'espèces mises en évidence par ce type d'approche peuvent révéler la signature de sources ou de catégorie de sources.

À notre connaissance, concernant les COVNM, le PMF n'a été appliqué qu'à des mesures réalisées en zone urbaine [2] voire urbaine de fond [3]. Fort d'une base de données conséquente de 10 ans de mesures et pour 46 espèces hydrocarbonées, l'application du PMF à des observations sur des sites récepteurs ruraux constitue une approche nouvelle. Le modèle intégrant l'incertitude des données d'entrée, il sera possible de pondérer la prise en compte des concentrations mesurées suivant la qualité de la mesure mais aussi suivant le potentiel de réactivité propre à chacune des espèces. Il s'agira également d'être prudent dans l'interprétation des résultats et de tester la sensibilité du modèle en fonction des données d'entrée et de son paramétrage.

Ce travail a pour objectif la mise en évidence d'associations d'espèces hydrocarbonées pour identifier des profils de sources à l'origine des concentrations mesurées sur des sites de fond. Il s'agira alors, d'une part, d'analyser les contributions des différents facteurs pour une confrontation aux inventaires d'émissions et, d'autre part, d'identifier les secteurs géographiques à l'origine des niveaux de fond observés.



Figure 1.

Situation géographique des sites du dispositif MERA (o) et sites concernés par les mesures d'HCNM (\*).

## 2. Matériel et méthode

### 2.1 Concentrations en HCNM mesurées

L'ensemble de la base de données correspondant aux mesures validées des 46 espèces de 2 à 9 atomes de carbone a été pris en compte pour les prélèvements des périodes de 1997 à 2006 au Donon (N = 977), de 2000 à 2006 à Peyrusse-Vieille (N = 505) et de 2002 à 2006 (N = 485) à La Tardière (Figure 1). L'analyse descriptive de cette base a montré des tendances significatives à la baisse des concentrations d'espèces comme le benzène, le toluène ou l'isopentane. En revanche, les espèces biogéniques (isoprène) et à longue durée de vie (éthane) ont tendance à stagner voire à augmenter [4].

Afin de comparer les contributions des différents profils identifiés aux données d'émission exprimées en masse, les concentrations sont exprimées en  $\mu\text{g}/\text{m}^3$ . Les mesures inférieures à la Limite de Détection (LD) et les données manquantes sont traitées suivant les recommandations de Hopke [5].

### 2.2 Incertitudes sur les concentrations mesurées

Le PMF est un modèle linéaire qui détermine une solution par itération en minimisant les résidus des valeurs modélisées en tenant compte de l'incertitude sur les données d'entrée. À chaque mesure et pour chaque espèce est associée une incertitude qui traduit la qualité de la donnée. L'incertitude de base considérée, est l'incertitude de la méthode de mesure déterminée pour chaque espèce. 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, l'incertitude associée sera égale aux 5/6<sup>e</sup> de la LD [5].

L'incertitude sur les valeurs mesurées dans le cas de l'application du modèle sources-récepteur à des espèces réactives n'est pas seulement composée de l'incertitude sur la méthode de mesure [6]. En effet, la réactivité photochimique de l'espèce peut être prise en compte de manière à pondérer davantage les observations des espèces pour lesquelles la relation sources-récepteur est très dépendante de la photochimie. Latella *et al.* [6] ont proposé de diviser l'incertitude de mesure par un terme de stabilité de l'espèce. Ce terme est lié à la constante de réaction de l'espèce avec le radical OH, au temps de transport entre la source et le récepteur et à une estimation empirique de la concentration en radicaux OH à partir de la concentration en ozone mesurée sur le site. Ce mode de calcul conduit à des incertitudes très élevées relativement aux concentrations mesurées amenant des difficultés de résolution pour le PMF. Pour prendre en compte cette composante de manière adaptée aux faibles niveaux de concen-

trations mesurées en zone rurale, la méthode appliquée dans ce cas consiste à estimer une erreur potentielle  $E_j$  pour chaque espèce déterminée suivant l'équation (1) en considérant que la principale voie chimique de déplétion des HCNM est la réaction de pseudo 1<sup>er</sup> ordre d'oxydation avec OH\*.

$$E_j = x_{ij} (1 - e^{-k_j [^*OH]\tau}) \quad (1)$$

$x_{ij}$  est la concentration mesurée en espèce  $j$  pour l'échantillon  $i$ .

$k_j$  est la constante cinétique d'oxydation de l'espèce  $j$  par les radicaux OH [7].

[\*OH] est la concentration moyenne en radicaux OH.

$\tau$  est le temps de transport entre la source et le récepteur. Il sera pris égal au temps de remontée de la rétrotrajectoire (72 h ou 120 h)

L'incertitude élargie finale  $S_{ij}$  de l'espèce  $j$  pour l'échantillon  $i$  est composée de l'incertitude de mesure  $U_{ij}$  et de l'incertitude  $E_j$  liée à la déplétion potentielle de l'espèce  $j$ . Elle est déterminée en suivant les recommandations du guide d'expression de l'incertitude [8] par la relation suivante :

$$S_{ij} = 2x_{ij} \sqrt{\left(\frac{U_{ij}}{2}\right)^2 + \left(\frac{E_j}{\sqrt{3}}\right)^2} \quad (2)$$

Le choix de la composante d'incertitude peut avoir un impact sur les résultats du modèle lorsque des scénarios extrêmes sont considérés. Ne pas prendre en compte la déplétion photochimique n'est pas recevable dans notre cas de figure puisque les sites de mesure sont éloignés des sources. À l'inverse considérer un temps de transport constant de cinq jours pour toutes les espèces est excessif au regard de leur durée de vie dans l'atmosphère. Un scénario intermédiaire (temps de transport de trois jours) est plus réaliste et le choix des variantes (influence de la [\*OH]) n'influence pas significativement les résultats du modèle.

### 2.3 Protocole d'application du modèle PMF

Le PMF a été appliqué de la même façon sur les trois sites du dispositif MERA. 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. Il s'appuie sur les éléments de base du modèle définis dans la partie méthodologie. Le nombre de facteurs n'étant pas connu *a priori*, il est nécessaire d'exécuter le modèle pour un nombre variable de facteurs afin de déterminer la solution idéale. La validité de la solution retenue est ensuite confirmée en étudiant les résultats obtenus pour un nombre de facteurs variant de plus ou moins un. La stabilité du résultat est également examinée en fonction de la rotation des facteurs. Le résultat final est complété par un ensemble d'indicateurs qualité des sorties du modèle.

La méthodologie d'utilisation a été appliquée en imposant les choix discutés en début de ce chapitre

concernant les données d'entrée et notamment le modèle d'incertitude. Une étude de sensibilité de la méthode a donc été menée de manière à consolider les résultats obtenus. Les éléments testés et les résultats obtenus ont été :

- très peu de sensibilité aux valeurs extrêmes ;
- des sorties de modèles cohérentes et comparables lorsque les données hivernales et estivales sont ou ne sont pas différenciées ;
- un impact notable du modèle d'incertitude ayant conduit au choix décrit précédemment.

### 2.4 Analyse par classe de rétrotrajectoires de masses d'air

De nombreux travaux ont montré l'intérêt de prendre en considération l'histoire de la masse d'air pour l'identification de zones sources. Les concentrations en HCNM et les ratios spécifiques vont donc être examinés en fonction des secteurs de provenance des masses d'air. Les prélèvements d'HCNM des trois sites ont pu être associés avec les rétrotrajectoires finissantes à 12 h TU sur ce même site. Les rétrotrajectoires ont été obtenues auprès du BADC (*British Atmospheric Data Center*). Une rétrotrajectoire comprend un point toutes les trois heures de coordonnées longitude, latitude et altitude. La pression du point finissant est fixée à 950 hPa.

Afin d'améliorer la lisibilité des flux des différentes masses d'air, les rétrotrajectoires ont été regroupées en utilisant la méthode de classification non hiérarchique (K-means) décrite dans la partie méthodologie. Le temps de remontée a été choisi à trois jours. Le choix du nombre de classes a été fait par examen de la variation de la distance interclasse en fonction du nombre de classes [9].

## 3. Modélisation des profils de sources et de leur contribution pour les trois sites MERA

Les résultats du modèle expliquent 86,7 %, 84,5 % et 88,3 % de la concentration totale en HCNM mesurée respectivement pour les sites du Donon, de La Tardière et de Peyrusse-Vieille. Entre 77 et 80 % de la variabilité des concentrations des espèces sont expliqués par six facteurs modélisés. Les concentrations totales modélisées sont bien corrélées avec les concentrations totales mesurées puisque les coefficients de corrélation calculés sur l'ensemble des échantillons varient de 0,89 pour Peyrusse-Vieille à 0,97 pour La Tardière.

### 3.1 Identification des facteurs modélisés

Les facteurs modélisés ont été comparés aux profils de sources connus issus de la littérature de manière à identifier des signatures de sources tout en tenant compte de la durée de vie des espèces. L'examen du profil saisonnier a permis de consolider

l'identification. Les facteurs modélisés ainsi définis et les profils de source d'émission sont représentés sur des mêmes graphes. Cinq des six facteurs modélisés sont communs aux trois sites.

### 3.1.1 Sources « évaporation de carburants »

Le premier facteur commun aux trois sites est marqué principalement par l'éthane et le propane associés avec le butane, l'isobutane, le pentane et l'isopentane. Ces espèces sont généralement associées aux sources « évaporation de carburants » comme le montre le profil moyen d'émission « évaporation liée au stockage, l'extraction et la distribution de carburants » (Figure 2). Le facteur ne fait pas apparaître de produits de combustion ou d'alcane plus lourds qui s'évaporent plus lentement [10], en revanche, l'éthane et le propane apparaissent en complément comme la signature du profil d'émission « fuites de gaz » ou « raffineries de pétrole » [11]. Les coefficients de corrélation calculés à partir des poids des 46 espèces (Tableau 1) entre les trois sites sont supérieurs à 0,98, montrant le lien élevé entre les facteurs.

La représentation des boîtes à moustaches mensuelles (Figure 2) permet de visualiser les profils saisonniers des contributions modélisées. Ce profil présente des contributions plus élevées en hiver. L'été est marqué par une activité photochimique plus soutenue et davantage de dispersion des polluants qui conduisent à des concentrations observées moins élevées même si le processus d'évaporation est renforcé avec la température.

### 3.1.2 Chauffage résidentiel

L'éthane, le propane, l'acétylène (éthyne), le butane, le benzène, le toluène et l'éthène présentent des poids significatifs pour les seconds facteurs communs modélisés (Figure 3). Contrairement à un profil « échappement automobile », la contribution du benzène pour ce profil est deux fois supérieure à celle du toluène. 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 et issus de la base de données GENEMIS. L'éthène présente une contribution assez faible comparativement au profil d'émission probablement due à sa faible durée de vie.

Une plus faible corrélation est obtenue entre le profil associé à Peyrusse-Vieille et ceux des sites de La Tardière et du Donon (Tableau 1). Le site de Peyrusse-Vieille est situé dans une zone plus tempérée et reste le plus éloigné des zones urbaines. En outre, les variations saisonnières des contributions des profils du Donon et de La Tardière (Figure 3) sont très marquées avec des valeurs très faibles en été et plus élevées en hiver.

### 3.1.3 Échappement automobile

Le troisième facteur commun est dominé par l'éthène, le toluène, les xylènes, le propylène et l'acétylène (Figure 4). Les facteurs des trois sites sont assez bien corrélés (Tableau 1). En outre, les contributions associées au facteur sont bien corrélées avec les concentrations moyennes en dioxyde d'azote

Tableau 1.  
Comparaison des six facteurs modélisés par le PMF pour les trois sites MERA.  
Le coefficient de corrélation (Pearson) est calculé en considérant les poids des 46 espèces pour chaque facteur.

	Évaporation de carburants	Chauffage résidentiel	Échappement automobile	Sources lointaines	Biogénique	Profil mixte
(Peyrusse) vs (Tardière)	0,98	0,56	0,79	0,97	0,89	0,55
(Peyrusse) vs (Donon)	0,99	0,67	0,62	1,00	1,00	0,13
(Tardières) vs (Donon)	0,98	0,97	0,89	0,98	0,89	0,14

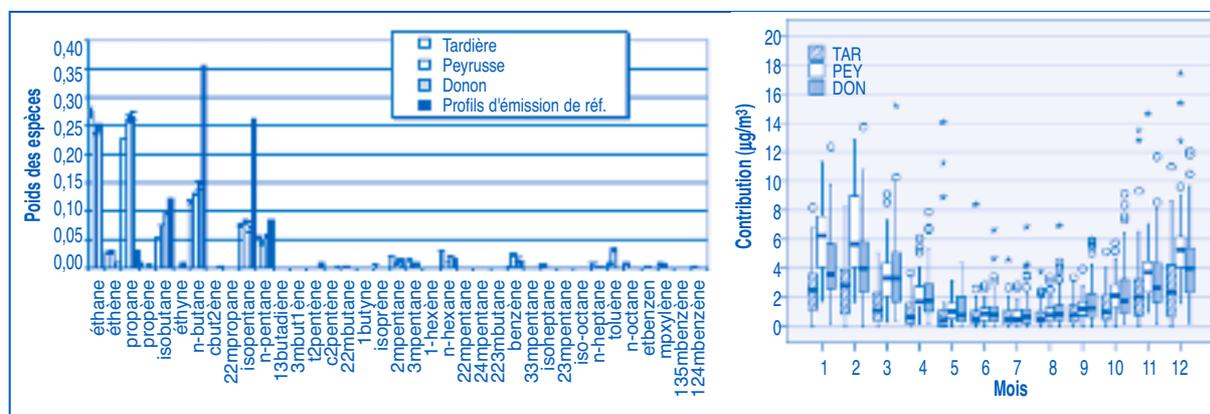


Figure 2.  
(À gauche) Composition normalisée du profil de source « **Évaporation de carburants** » modélisé pour les trois sites et profil de source d'émission associé.  
(À droite) Boîtes à moustaches mensuelles des contributions des facteurs modélisés (en  $\mu\text{g}\cdot\text{m}^{-3}$ ).

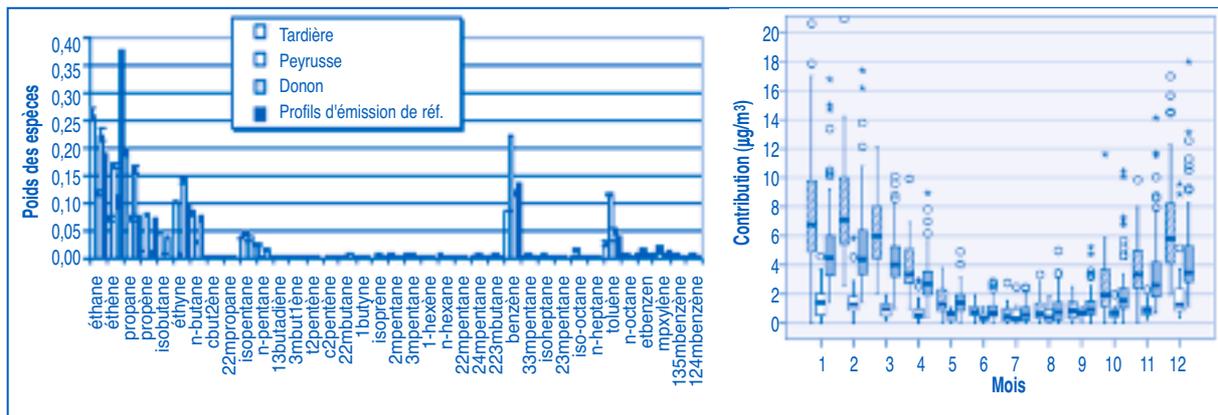


Figure 3.

(à gauche) Composition normalisée du profil de source « **Chauffage résidentiel** » modélisé pour les trois sites et profil de source d'émission associé. (à droite) Boîtes à moustaches mensuelles des contributions des facteurs modélisés (en  $\mu\text{g}\cdot\text{m}^{-3}$ ).

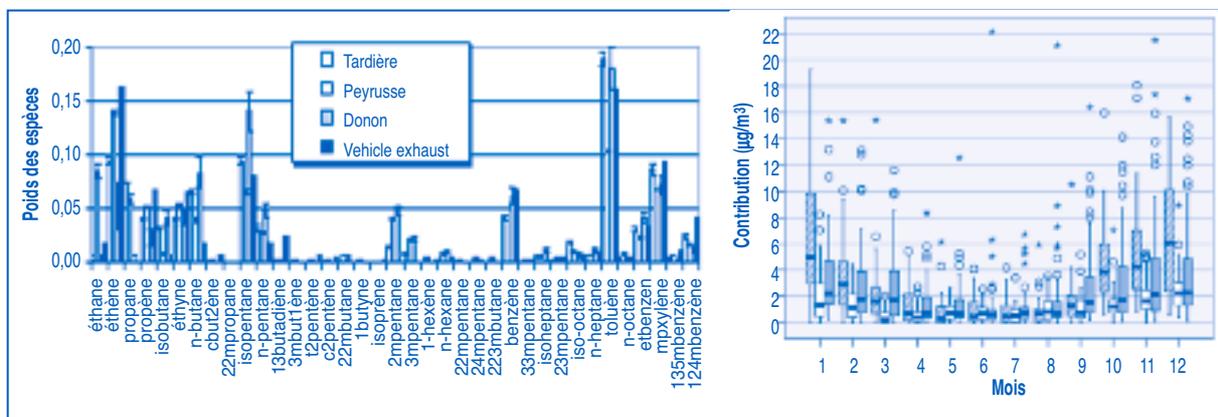


Figure 4.

(À gauche) Composition normalisée du profil de source « **Échappement automobile** » modélisé pour les trois sites et profil de source d'émission associé. (À droite) Boîtes à moustaches mensuelles des contributions des facteurs modélisés (en  $\mu\text{g}\cdot\text{m}^{-3}$ ).

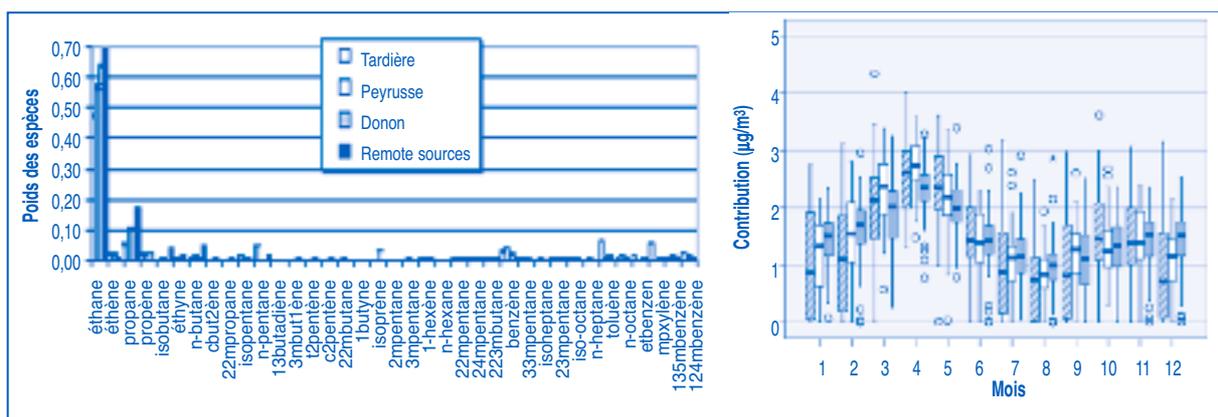


Figure 5.

(À gauche) Composition normalisée du profil de source « **Sources lointaines** » modélisé pour les trois sites et profil de source d'émission associé. (À droite) Boîtes à moustaches mensuelles des contributions des facteurs modélisés (en  $\mu\text{g}\cdot\text{m}^{-3}$ ).

( $\text{NO}_2$ ) mesurées aux heures de prélèvement d'HCNM ( $R = 0,64$  pour Donon ;  $R = 0,75$  pour La Tardière). Ces facteurs sont identifiés comme des profils d'échappement automobile par comparaison avec les profils d'émission obtenus pour un parc automobile représentatif du parc français [12]. Les contributions

du facteur « échappement automobile » sont plus importantes en période hivernale compte tenu de la faible durée de vie des espèces concernées en été, des émissions renforcées lors du démarrage à froid des véhicules, et des conditions atmosphériques plus favorables à l'accumulation des polluants en hiver.

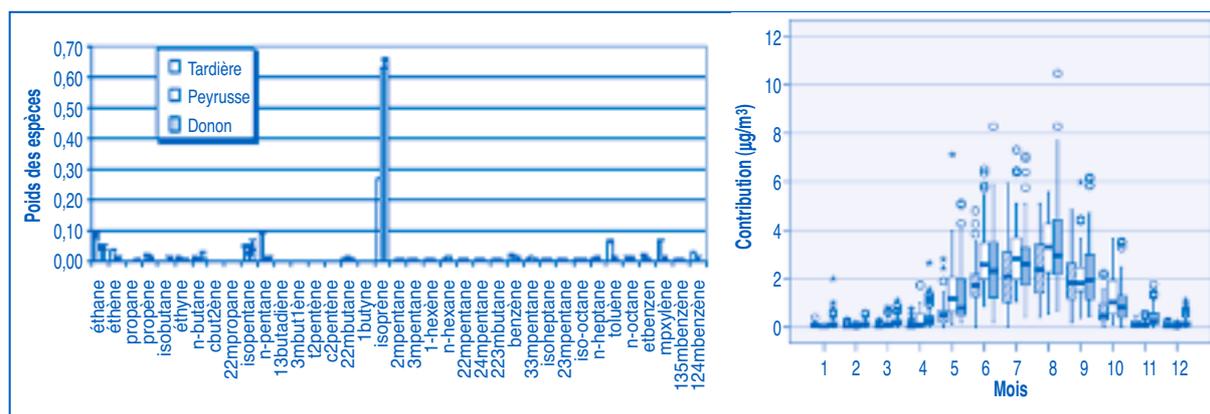


Figure 6.

(À gauche) Composition normalisée du profil de source « **Biogénique** » modélisé pour les trois sites et profil de source d'émission associé. (À droite) Boîtes à moustaches mensuelles des contributions des facteurs modélisés (en  $\mu\text{g}\cdot\text{m}^{-3}$ ).

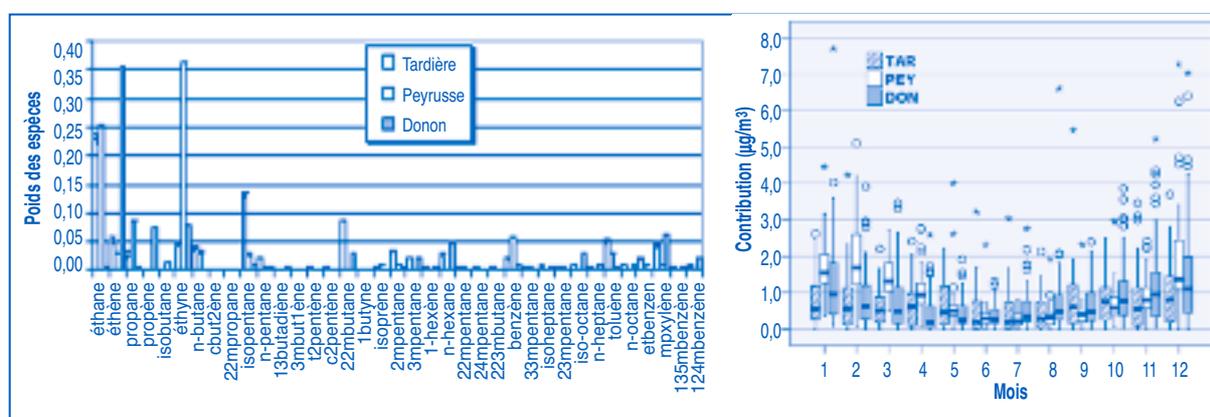


Figure 7.

(À gauche) Composition normalisée des **sixièmes profils** de source modélisés pour les trois sites.  
(À droite) Boîtes à moustaches mensuelles des contributions des facteurs (en  $\mu\text{g}\cdot\text{m}^{-3}$ ).  
Comparaison intersites des contributions.

### 3.1.4 Sources lointaines

L'éthane et le propane sont les composés dominant le quatrième facteur (Figure 5). En raison de leur faible réactivité, ces espèces sont souvent associées à des masses d'air anciennes. Elles peuvent également être émises par des sources proches du type « fuite de gaz naturel ». En considérant que ces sources sont plutôt représentées par le profil « évaporation de carburant » et compte tenu de la typologie rurale des sites, ce profil est attribué aux niveaux de fond dus à des sources éloignées. Ce profil est communément appelé « Sources lointaines » [13, 14]. De plus, les contributions évoluent typiquement au cours des saisons montrant une augmentation progressive de novembre à avril liée à l'accumulation des espèces avant une baisse prononcée à partir de mai en raison de la déplétion photochimique et de la meilleure dispersion.

### 3.1.5 Sources biogéniques

Le facteur 5 est labélisé « Sources biogéniques » avec l'isoprène comme seul composé dominant (Figure 6). La contribution de la source automobile peut être écartée en zone rurale compte tenu de la

durée de vie très limitée de cette espèce. Comme les émissions sont reliées à la température et à l'ensoleillement [15], la contribution de ce facteur est maximale en été et quasi nulle en hiver.

### 3.1.6 Profils mixtes

Contrairement aux cinq profils communs présentés précédemment, les sixièmes profils sont assez différents pour les trois sites (Figure 7). L'éthane est l'espèce prépondérante pour le profil du Donon. Compte tenu de la durée de vie de cette espèce, il s'agit probablement d'une source proche qui de plus, d'après le suivi chronologique des contributions, est apparue significativement à partir de fin 2003.

Ce sont les profils de La Tardière et de Peyrusse-Vieille qui se ressemblent le plus ( $R = 0,55$ ). Ils sont marqués par des espèces comme l'éthane, l'acétylène ou l'isopentane qui peuvent être attribuées à un profil de source de combustion incomplète comme l'échappement automobile ou le chauffage résidentiel. Les contributions sont assez faibles comparativement aux autres facteurs et sont plus importantes en hiver surtout pour le site de Peyrusse-Vieille. Elles ne montrent pas d'évolution annuelle.

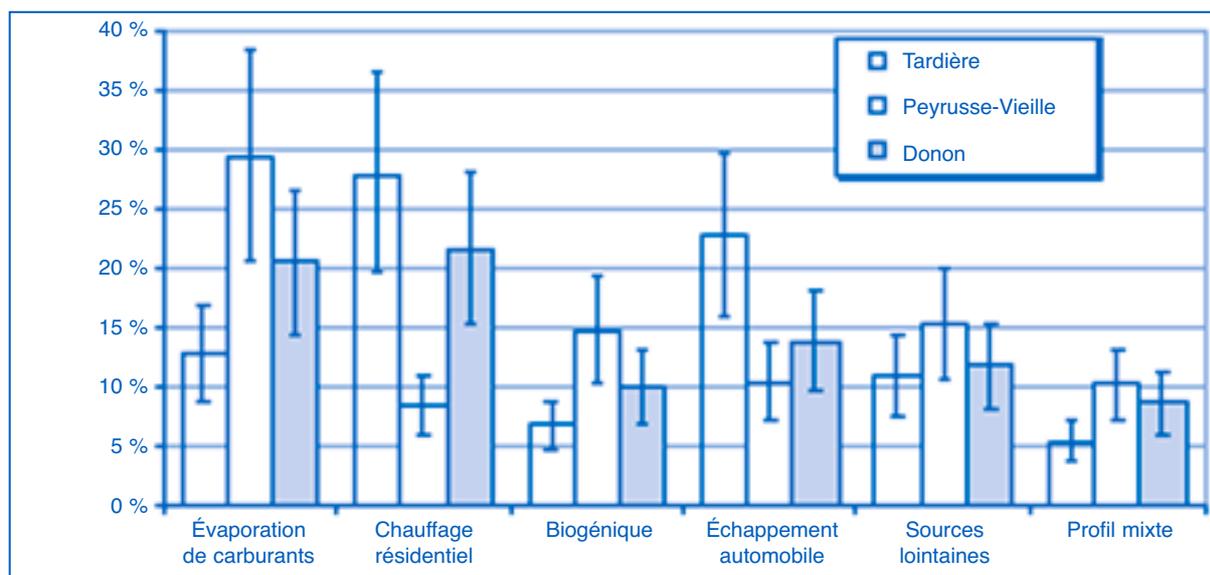


Figure 8.  
Contributions relatives moyennes des facteurs modélisés et incertitudes associées, pour les trois sites MERA pour la période 2002-2006.

### 3.2 Analyse des contributions des facteurs modélisés

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

$$C_k = \frac{\sum_{i=1}^n g_{ik}}{\sum_{k=1}^p \sum_{i=1}^n g_{ik}} \times 100 \quad (3)$$

Les résultats comparatifs entre les trois sites pour chaque facteur modélisé et pour la période commune 2002-2006 sont représentés par des histogrammes, avec des barres représentant les incertitudes des contributions (Figure 8). En effet, le modèle fournit une estimation des incertitudes sur les poids des espèces par facteur (F) et sur les contributions modélisées (G).

Le site de La Tardière est marqué par de plus grandes contributions du « chauffage résidentiel » (28,0 %) et du facteur « échappement automobile » (22,8 %). La proximité de la commune de La Chataignerie au sud du site peut en être la cause. Les profils « chauffage résidentiel » (21,8 %) et « évaporation de carburants » (20,3 %) contribuent majoritairement aux concentrations en HCNM mesurées sur le site du Donon. Ces éléments confirment le caractère plus influencé de ces deux sites par des sources anthropiques plutôt urbaines. Pour Peyrusse-Vieille, c'est la contribution du profil « évaporation de carburants » (29,5 %) qui prédomine largement. Le site de Lacq situé à 60 km à l'est et les températures plus chaudes peuvent expliquer cette influence. En revanche ce site est moins influencé par des profils de sources de combustion de type urbain, la deuxième plus forte

contribution étant celle du profil « sources lointaines ». Le profil biogénique est significativement plus important à Peyrusse-Vieille et au Donon avec des contributions moyennes de 14,8 % et 10,0 % respectivement contre 6,7 % pour La Tardière.

#### 3.2.1 Comparaison des contributions de facteurs aux inventaires d'émission

Les contributions relatives des différents facteurs peuvent être confrontées à l'inventaire d'émissions national sachant que cet inventaire concerne sans distinction tous les COVNM alors que notre étude se limite aux HCNM. Les profils d'émission modélisés sont rapprochés des secteurs d'émission pour lesquels les parts relatives aux émissions totales ont été déterminées.

Le profil modélisé « **chauffage résidentiel** » contribue de  $8,4 \pm 2,6$  % à  $28,0 \pm 8,5$  % (moyenne de  $19,4 \pm 6,4$  %) aux concentrations mesurées sur les trois sites MERA. Ces résultats sont cohérents avec le secteur « combustions non industrielles » qui représente 13,7 % des émissions totales sur la période 2002-2006. Une autre étude menée en vallée Alpine a estimé à 22 % la contribution du chauffage résidentiel aux teneurs en COVNM en zone rurale [16].

Dans les inventaires d'émission, le secteur SNAP « Transport routier » inclut les émissions liées à l'évaporation des carburants des véhicules et à la combustion des moteurs. En conséquence, les contributions relatives des profils modélisés « **évaporation de carburants** » et « **échappement automobile** » ont été additionnées pour être comparées à la somme des secteurs d'émission « transport routier » et « extraction et distribution de fuels fossiles » qui représentent 31,6 %. Les contributions relatives modélisées variant de  $34,4 \pm 7,6$  % à  $39,8 \pm 9,5$  % (moyenne  $36,6 \pm 8,4$  %) sont comparables. Les

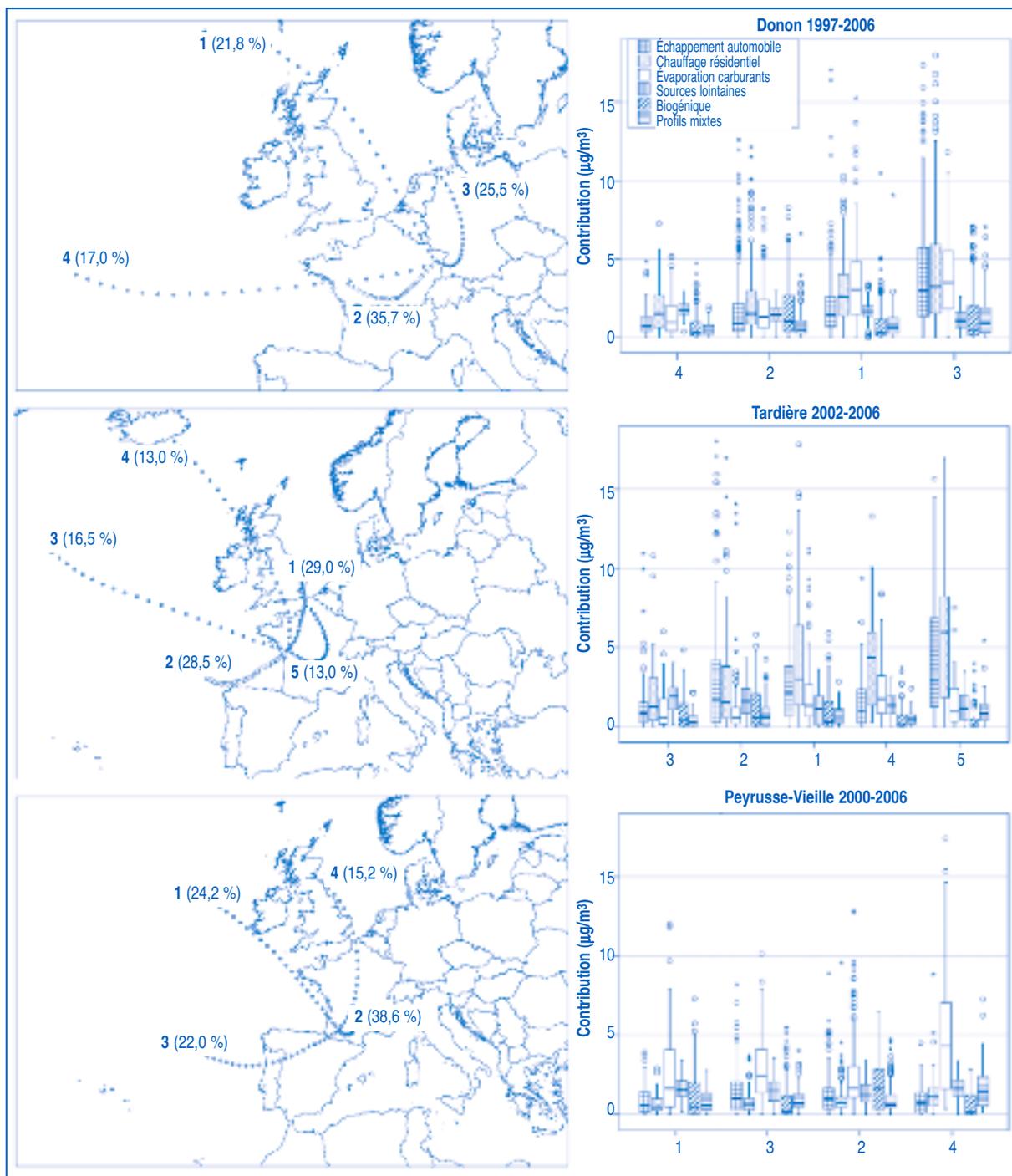


Figure 9.

(Colonne gauche) Rétrotrajectoires moyennes représentant les classes de rétrotrajectoires arrivant sur le site.  
(Colonne droite) Distribution des contributions de profils de sources modélisés par le PMF suivant les classes de rétrotrajectoires.

secteurs « autres sources mobiles » (5,0 %) et en partie le secteur « agriculture » (8,6 %) comprennent également des émissions du type moteur thermique qui pourraient être associées au profil modélisé « échappement automobile ».

Le profil « **biogénique** » contribue de  $6,7 \pm 2,1$  % à  $14,8 \pm 4,7$  % avec une moyenne pour les trois sites de  $10,5 \pm 3,5$  %. Une étude réalisée dans le cadre du programme ESCOMPTE [17] avait estimé que les

émissions d'isoprène représentent environ 15 % des émissions en COVNM en France. La contribution moyenne de  $10,5 \pm 3,5$  % obtenue à partir des trois sites MERA donne une bonne représentation des émissions biogéniques.

Finalement, la contribution du profil « **sources lointaines** » est en moyenne de  $12,7 \pm 3,9$  % pour les trois sites. Ce résultat est assez cohérent avec les résultats d'autres études : cette contribution a été

estimée à 17 % dans la vallée de Chamonix en hiver [16], de 14 % à 34 % sur les sites urbains de Marseille, Grenoble et Strasbourg [18]. Cette composante a également été évaluée à 17 % en zone urbaine au Texas [14].

Dans les inventaires d'émission, le secteur « usage de solvants » est important puisqu'il contribue à 32,4 % des émissions des COVNM en France, même si les estimations de ce secteur sont remises en cause [19]. Cependant, les résultats du PMF ne mettent pas en évidence un profil de ce type. Pour les inventaires d'émission, ce profil inclut, outre le toluène, une grande part de composés oxygénés et chlorés qui ne sont pas pris en compte dans cette étude. Néanmoins, il est possible que la contribution du toluène pour le profil modélisé « échappement automobile » soit surestimée intégrant une part associée au secteur « usage de solvants ». Géographiquement, ces deux sources peuvent être corrélées car toutes deux de type urbain.

### 3.2.2 Analyse des contributions par classe de rétrotrajectoires

Pour poursuivre l'interprétation des résultats, il est intéressant d'étudier les contributions des profils modélisés par classe de rétrotrajectoires. Pour chaque site et pour chaque classe, la distribution des contributions d'un facteur est représentée par des boîtes à moustaches (Figure 9).

Cette analyse met en évidence pour les trois sites une origine océanique (Classe 4 pour Donon, 3 pour La Tardière et 1 pour Peyrusse-Vieille) pour laquelle les contributions moyennes des profils « sources lointaines » sont majoritaires alors que celles des profils associés aux sources « chauffage résidentiel » et « échappement automobile » sont assez basses comparativement aux autres classes. Ces classes sont caractérisées par des vitesses de vent élevées limitant l'impact des émissions de sources proches. L'analyse des ratios d'espèces menée par ailleurs a montré que les masses d'air correspondantes sont les plus anciennes.

À l'inverse, les classes 1 et 3 du Donon et 4 et 5 de La Tardière sont marquées par des contributions moyennes significativement plus élevées pour les profils de sources « chauffage résidentiel », « échappement automobile » et « évaporation carburants », montrant l'influence de zones urbaines continentales. À noter que ces classes sont composées majoritairement de rétrotrajectoires hivernales provenant du Nord et traversant des zones de fortes émissions comme le Sud de l'Angleterre, le Nord de la France, le Benelux ou encore le Sud de l'Allemagne. Les classes de rétrotrajectoires courtes (classe 3 du Donon et classe 5 de La Tardière), pour des masses d'air qualifiées de jeunes, soulignent l'impact des zones les plus proches des sites de mesures.

Le site de Peyrusse-Vieille se distingue car il apparaît comme moins influencé localement. C'est le profil « évaporation de carburants » qui montre la plus

grande variabilité de sa contribution suivant les classes de rétrotrajectoires. La classe 4, représentant les masses d'air provenant du secteur Nord, est caractérisée par des contributions significativement plus élevées pour ce profil. C'est aussi une classe principalement hivernale, regroupant des rétrotrajectoires de masses d'air froides favorisant l'accumulation d'espèces peu réactives. La classe 1 est marquée par une forte amplitude de variation des contributions et la classe 3 par des valeurs assez fortes, ce qui peut être lié à l'influence du site de Lacq dans ces secteurs de provenance.

Les contributions du profil « biogénique » sont plus élevées pour les classes de masses d'air en période estivale, provenant du secteur Sud et caractérisées par des vents faibles (les classes 2 des trois sites). C'est particulièrement le cas pour Peyrusse-Vieille, site pour lequel la classe 2 affiche la plus grande occurrence (38,6 %) avec les températures les plus élevées et par conséquent, les niveaux en isoprène les plus importants.

## 4. Conclusion

Le modèle sources-récepteur PMF a été appliqué à la base de données comprenant 46 espèces hydrocarbonées mesurées sur les trois sites du dispositif MERA. Ce modèle a permis de mettre en évidence des associations entre les espèces et d'identifier des profils de sources à l'origine des concentrations mesurées sur le site récepteur. L'application de ce type de modèle sur des sites ruraux, éloignés des sources ainsi que l'application de ce type de modèle à des espèces réactives comme les HCNM constituent une approche nouvelle. L'hypothèse de la conservation des masses sur lequel s'appuie le modèle peut ne pas être vérifiée compte tenu de la réactivité différente des espèces. La réactivité est prise en compte par l'ajout d'une composante d'incertitude comme données d'entrée du modèle de manière à pondérer les espèces suivant leur réactivité potentielle. Une étude de sensibilité du modèle a permis de consolider les résultats obtenus par rapport aux choix méthodologiques. Les résultats ont montré la robustesse de la méthode.

Les résultats du PMF ont mis en évidence cinq facteurs communs aux trois sites MERA. La comparaison de ces facteurs aux profils d'émissions issus de la littérature permet d'identifier assez distinctement les signatures de sources suivantes : « évaporation de carburants » ; « chauffage résidentiel », « échappement automobile », « sources lointaines » et « biogéniques ».

L'analyse de ces contributions par classe de rétrotrajectoires de masses d'air met en évidence la part d'influence locale, la part continentale où les profils de sources urbaines contribuent davantage et la part d'influence océanique dominée par le profil « sources lointaines ».

Les comparaisons des contributions de profils modélisés avec les données d'inventaire d'émission ou avec les données d'autres études montrent la cohérence des résultats obtenus. Toutefois, le profil « usage de solvants » très significatif dans les inventaires n'a pas été identifié dans cette analyse.

Suite à l'identification de profils, l'étude s'est poursuivie par l'application de modèles sources-récepteur orientés dans le but de localiser des zones sources potentielles contribuant aux observations. Avec quelques apports méthodologiques, les résultats apparaissent cohérents avec les inventaires d'émission

spécialisés, avec les résultats d'autres études ou avec les sorties du modèle déterministe Unified EMEP [20]. Ces travaux sont en cours de publication par ailleurs.

## Remerciements

Le dispositif MERA est financé par l'ADEME et le Ministère de l'Environnement. Les auteurs remercient l'ensemble des gestionnaires des sites MERA pour leur précieuse collaboration.

## References

1. OMINEA. National inventories of air emissions in France: Organisation and methodology. CITEPA 2008 : 1026 p.
2. Xie Y, Berkowitz CM. The use of positive matrix factorization with conditional probability functions in air quality studies: An application to hydrocarbon emissions in Houston, Texas. *Atmospheric Environment* 2006 ; 40 (17) : 3070-91.
3. Lanz VA, Hueglin C, Buchmann B *et al.* 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.* 2008 ; 8 (9) : 2313-32.
4. Sauvage S, Plaisance H, Locoge N, Wroblewski A, Coddeville P, Galloo JC. Long term measurement and source apportionment of non-methane hydrocarbons in three French rural areas. *Atmospheric Environment* 2009, In Press, Corrected Proof.
5. Hopke. A guide to positive Matrix Factorization. EPA Workshop proceedings materials from the workshop on UNMIX and PMF as applied to PM<sub>2.5</sub>. 2000.
6. Latella A, Stani G, Cobelli L *et al.* Semicontinuous GC analysis and receptor modelling for source apportionment of ozone precursor hydrocarbons in Bresso, Milan, 2003. *Journal Of Chromatography A* 2005 ; 1071 (1-2) : 29-39.
7. Atkinson R, Arey J. Atmospheric degradation of volatile organic compounds. *Chemical Reviews* 2003 ; 103 (12) : 4605-38.
8. ISO13005. Guide to the expression of uncertainty in Analytical measurement. Geneva 1995 (ISBN 92-67-10188-9).
9. Dorling SR, Davies TD, Pierce CE. Cluster analysis: A technique for estimating the synoptic meteorological controls on air and precipitation chemistry – Method and applications. Part A. General Topics. *Atmospheric Environment* 1992 ; 26 (14) : 2575.
10. Watson JG, Chow JC, Fujita EM. Review of volatile organic compound source apportionment by chemical mass balance. *Atmospheric Environment* 2001 ; 35 (9) : 1567.
11. Passant NR. Speciation of UK emissions of non-methane volatile organic compounds. AEA Technology 2002. Report No. AEAT/ENV/R/0545 Issue 1.
12. Fontaine H, Galloo JC. Volatile organic compound emission factors in the exhaust gases of light duty vehicles: uncertainties, influence of fuel formulation and new data. Technical report 2002, Department of Chemistry and Environment, École des Mines de Douai, Douai, France. Final report PRIMEQUAL-PREDIT 1999, March 2002.
13. Hellen H, Hakola H, Laurila T. Determination of source contributions of NMHCs in Helsinki (60 degrees N, 25 degrees E) using chemical mass balance and the Unmix multivariate receptor models. *Atmospheric Environment* 2003 ; 37 (11) : 1413-24.
14. Buzcu B, Fraser MP. Source identification and apportionment of volatile organic compounds in Houston, TX. *Atmospheric Environment* 2006 ; 40 (13) : 2385-400.
15. Fuentes JD, Wang D. On the seasonality of isoprene emissions from a mixed temperate forest. *Ecological Applications* 1999 ; 9 (4) : 1118-31.
16. Rouvière A, Brulfert G, Baussand P, Chollet JP. Monoterpene source emissions from Chamonix in the Alpine Valleys. *Atmospheric Environment* 2006 ; 40 (19) : 3613-20.

17. Simon V, Dumergues L, Ponche JL, Torres L. The biogenic volatile organic compounds emission inventory in France: Application to plant ecosystems in the Berre-Marseilles area (France). *Science of The Total Environment* 2006 ; 372 (1) : 164-82.
18. Troussier F. Caractérisation du comportement spatio-temporel des composés organiques volatils en atmosphère urbaine et périurbaine. PhD Thesis. École des Mines de Douai 2006. Université des Sciences et Technologies de Lille.
19. Niedojadlo A, Becker KH, Kurtenbach R, Wiesen P. The contribution of traffic and solvent use to the total NMVOC emission in a German city derived from measurements and CMB modelling. *Atmospheric Environment* 2007 ; 41 (33) : 7108-26.
20. Simpson D, Fagerli H, Jonson JE, Tsyro S, Wind P, Tuovinen JP. Transboundary acidification and eutrophication and ground level ozone in Europe: Unified EMEP Model Description, EMEP Status Report 1/2003 Part I, EMEP/MSC-W Report. The Norwegian Meteorological Institute, Oslo, Norway.