



Traitement harmonisé de jeux de données multi-sites pour l'étude des sources de PM par Positive Matrix Factorization

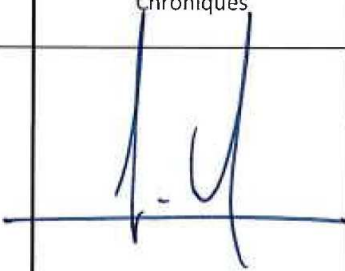
**Laboratoire Central de Surveillance
de la Qualité de l’Air**

**TRAITEMENT HARMONISE DE JEUX DE DONNEES MULTI-SITES
POUR L’ETUDE DE SOURCES DE PM PAR POSITIVE MATRIX
FACTORIZATION (PMF)**

Olivier FAVEZ (INERIS)

Dalia SALAMEH & Jean-Luc JAFFREZO (IGE)

Décembre 2017

	Rédaction	Vérification	Approbation
NOM	Olivier FAVEZ	Alexandre ALBINET	Marc DURIF
Qualité	Ingénieur de l’Unité CIME Direction des Risques Chroniques	Ingénieur de l’Unité CIME Direction des Risques Chroniques	Responsable du Pôle CARA Direction des Risques Chroniques
Visa			



LE LABORATOIRE CENTRAL DE SURVEILLANCE DE LA QUALITE DE L'AIR

Le Laboratoire Central de Surveillance de la Qualité de l'Air est constitué des laboratoires de l'IMT Lille Douai, de l'INERIS et du LNE. Il mène depuis 1991 des études et des recherches à la demande du Ministère chargé de l'environnement, et en concertation avec les Associations Agréées de Surveillance de la Qualité de l'Air (AASQA). Ces travaux en matière de pollution atmosphérique ont été financés par la Direction Générale de l'Énergie et du Climat (bureau de la qualité de l'air) du Ministère chargé de l'Environnement. Ils 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 au ministère et 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.

TABLE DES MATIERES

RESUME	7
REMERCIEMENTS ET COLLABORATIONS.....	9
1. INTRODUCTION	10
2. SITES D'ETUDE	11
3. METHODOLOGIE DE TRAITEMENT DE DONNEES	13
3.1 Principes généraux de la PMF	13
3.2 Traitements préliminaires des séries de données	14
3.3 Sélection des variables d'entrée et estimation de leurs incertitudes	15
3.4 Définition de contraintes spécifiques	16
3.5 Choix des solutions.....	17
4. RESULTATS	18
4.1 Description des profils chimiques (sources) identifiés	18
4.2 Contributions annuelles moyennes et variations saisonnières	23
5. SENSIBILITE DE LA METHODOLOGIE DE TRAITEMENT DES DONNEES.....	32
6. BILAN ET PERSPECTIVES	35
7. REFERENCES	39
8. LISTE DES ANNEXES.....	41

Ce rapport présente les résultats issus de l'utilisation de l'outil statistique *Positive Matrix Factorization* (PMF) pour une réanalyse homogène de 15 jeux de données français collectés lors de ces 6 dernières années dans le cadre du programme CARA et/ou de projets de recherche nationaux/européens. Ce travail a notamment été conduit dans le cadre du projet SOURCES (cofinancé par l'ADEME) et constitue la plus large étude de sources de PM₁₀ au niveau national par application d'outil sources-récepteur.

Cette étude inclut le développement d'une méthodologie de traitement de données harmonisée (analyse statistique du jeu de données, sélection des variables d'entrée pour l'analyse PMF, estimation de leurs incertitudes, application des contraintes chimiques spécifiques dans les profils chimiques de certains facteurs). L'application de cette méthodologie sur l'ensemble des jeux de données disponibles converge vers les principales conclusions suivantes :

- ✓ Les émissions primaires liées au transport routier influencent fortement la masse des PM₁₀ sur les sites de proximité automobile, mais également sur certains sites de fond urbain tels que Rouen, Marseille, Grenoble et Nogent (avec des contributions relatives de l'ordre de 20-30% en moyenne annuelle). Sur les autres sites, les contributions annuelles sont d'environ 10%.
- ✓ La source de combustion de la biomasse est très importante en hiver, contribuant jusqu'à 70% de la masse des PM₁₀ en moyenne saisonnière à Chamonix (contre 10-15% à Marseille, Lens et Rouen, et environ 30% sur tous les autres sites).
- ✓ Les aérosols secondaires riches en nitrate d'ammonium présentent une concentration maximale au printemps (typiquement 30% des PM₁₀), en particulier sur les sites de la moitié nord de la France.
- ✓ L'impact des aérosols secondaires riches en sulfate d'ammonium sur les niveaux de PM₁₀ est principalement observé en période estivale (en particulier dans le sud de la France).
- ✓ Les poussières minérales présentent des contributions importantes et relativement comparables (5-20%) au printemps et en été sur un grand nombre de sites.
- ✓ De nouveaux traceurs organiques - e.g., polyols et l'acide méthylsulfonique (MSA), rarement utilisés dans les études précédentes - ont permis la quantification de sources biogéniques spécifiques (e.g., émissions primaires biogéniques et aérosols organiques secondaires d'origine marine) qui présentes des contributions significatives du printemps à l'automne.

A titre de synthèse et d'illustration de ces conclusions, la Figure R1 présente les contributions moyennes des principales sources de PM à l'échelle nationale en fond urbain, obtenues en combinant les résultats issus de l'analyse PMF harmonisée pour 10 agglomérations métropolitaines (Rouen, Nogent sur Oise, Lens, Poitiers, Lyon, Grenoble, Bordeaux, Aix en Provence, Marseille, et Nice).

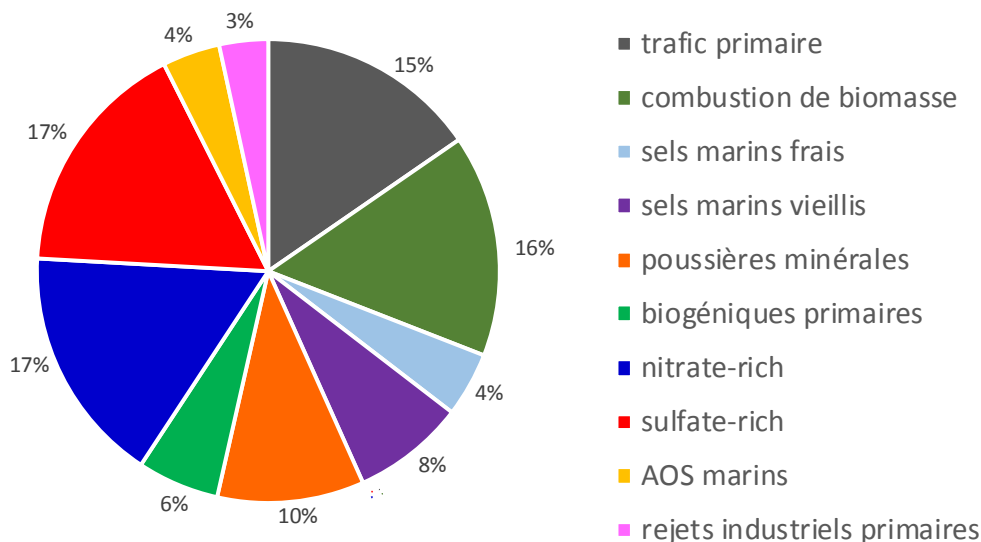


Figure R1 : Contributions annuelles moyennes des principaux facteurs constitutifs des PM_{10} en fond urbain à l'échelle nationale (Moyenne des résultats obtenus par analyse PMF harmonisée sur 10 sites de fond urbain du projet SOURCES).

Cette étude s'inscrit dans un contexte de définition de procédures d'assurance qualité pour l'application des outils de *source apportionment* au niveau européen, incluant les exercices de comparaison inter-laboratoires organisés par le JRC et les travaux de normalisation du CEN. La méthodologie mise en œuvre pourra notamment être utilisée dans de prochaines études PMF à large échelle spatiale, permettant ainsi d'améliorer la comparabilité des résultats entre les différents sites et des différentes régions. Néanmoins, une limitation inhérente à l'utilisation d'un protocole de traitement de données harmonisé réside dans l'hétérogénéité (en nombre et en qualité) du panel des espèces chimiques utilisées comme variables d'entrée. Ainsi, selon les caractéristiques spécifiques des sites et jeux de données étudiés, l'utilisation d'une approche « personnalisée » de traitement de données peut être préférée. En particulier, l'élargissement du jeu de données d'entrée avec, par exemples, de nouveaux marqueurs organiques (e.g., n-alcanes, hopanes, cellulose, oxy- et nitro-HAP dérivés, et/ou autres marqueurs de composés secondaires), des résultats de mesures isotopiques (e.g., ^{14}C , ^{15}N), et/ou des données d'analyseurs automatiques (AE33, ACSM, mesures de métaux en continu...), doit permettre d'améliorer l'identification et la quantification de certaines sources minoritaires (e.g., émissions industrielles et/ou combustion de fioul lourd) ainsi que les facteurs liés aux aérosols secondaires (organiques et inorganiques).

REMERCIEMENTS ET COLLABORATIONS

Ce rapport s'appuie largement sur les résultats obtenus par Dalia Salameh dans le cadre de son post-doctorat pour le projet SOURCES, fruit d'une collaboration entre le LCSQA/INERIS et l'Institut des Géosciences de L'Environnement (IGE, Grenoble) sur co-financement de l'ADEME.

Les jeux de données utilisés ici proviennent notamment du travail de prélèvement des échantillons par les Associations Agréées de Surveillance de la Qualité de l'Air (AASQA), en particulier : Atmo Auvergne-Rhône-Alpes, Atmo Hauts de France, Air PACA, Atmo Nouvelle Aquitaine, Atmo Grand-Est, et Atmo Normandie.

1. INTRODUCTION

L'amélioration de la qualité de l'air nécessite toujours un effort de réduction des émissions des principales sources anthropiques. Ces dernières doivent être préalablement identifiées et quantifiées aussi finement que possible, dans l'optique de l'élaboration et de l'évaluation des plans d'action.

Deux grands types de méthodologies sont principalement utilisées : (i) celles se basant sur l'utilisation de modèles numériques permettant de simuler le devenir des polluants dans l'atmosphère à partir de cadastres d'émission, de la paramétrisation des conditions météorologiques et des processus physico-chimiques de (trans-)formation des PM ; (ii) celles se basant sur la détermination de la composition chimique (et/ou de la granulométrie) des particules sur un site récepteur avec l'utilisation d'espèces spécifiques marqueurs des différentes sources de PM. Si les modèles numériques doivent permettre d'apporter des informations en tout point du territoire et selon des échelles temporelles aussi larges que souhaitées, leur validation nécessite des comparaisons avec les résultats obtenus par la mise en œuvre de méthodologies expérimentales. En raison de leur coût, ces dernières ne peuvent être que ponctuelles (dans l'espace et/ou dans le temps). Néanmoins, basées sur l'observation, elles rendent compte de situations réelles et constituent un maillon important pour une meilleure gestion de la qualité de l'air.

L'outil statistique *Positive Matrix Factorization* (PMF) est utilisé depuis 2011 par le LCSQA pour l'exploitation des données issues du programme CARA et des travaux sur les HAPs (LCSQA 2012a&b, LCSQA, 2013a&b, LCSQA 2015). Il permet la détermination de la contribution des principales sources et mécanismes de formation des polluants, à partir d'un jeu de données aussi denses que possible, en termes de nombre d'échantillons et d'espèces chimiques mesurées (Belis et al., 2014).

L'utilisation de ce type d'outil statistique revêt une part de subjectivité, notamment liée au choix des hypothèses de départ et/ou de la solution finalement retenue. Ainsi, la comparaison de résultats issus de différentes études basées sur une même approche contient une part d'incertitude liée aux inhomogénéités de traitement de données. A noter également que les travaux de normalisation devant permettre la définition de méthodologies standards et l'harmonisation des pratiques ne font que commencer (dans le cadre du CEN/TC 264 WG44 et de FAIRMODE).

Dans ce contexte, il est apparu intéressant de tester l'utilisation de l'outil PMF pour une réanalyse homogène de jeux de données collectées ces dernières années dans le cadre du programme CARA et/ou de projets de recherche nationaux/européens. Le présent rapport rend compte des résultats obtenus lors de ces travaux. Ceux-ci ont notamment été conduits dans le cadre du projet SOURCES (cofinancé par l'ADEME) et constitue la plus large étude de sources de PM₁₀ au niveau national par application de modèle statistique sources-récepteur. L'utilisation d'un traitement de données identique pour chaque site peut induire une moins bonne prise en compte de la spécificité du jeu de données et/ou du site étudié. Afin d'évaluer (qualitativement) la robustesse de ces résultats, une comparaison a été effectuée avec les solutions PMF « individualisées » issues des travaux précédemment réalisés sur certains de ces jeux de données.

2. SITES D'ETUDE

La Figure 1 présente la localisation géographique des 15 sites de mesures pour lesquels des jeux de données comparables (en termes de nombre d'échantillons, de durée d'étude, et de résultats de caractérisation chimique des PM₁₀) étaient disponibles pour la présente étude. Le Tableau 1 résume les caractéristiques (coordonnées, période de collecte, nombre d'échantillons, et typologie) des sites de mesure investigués.

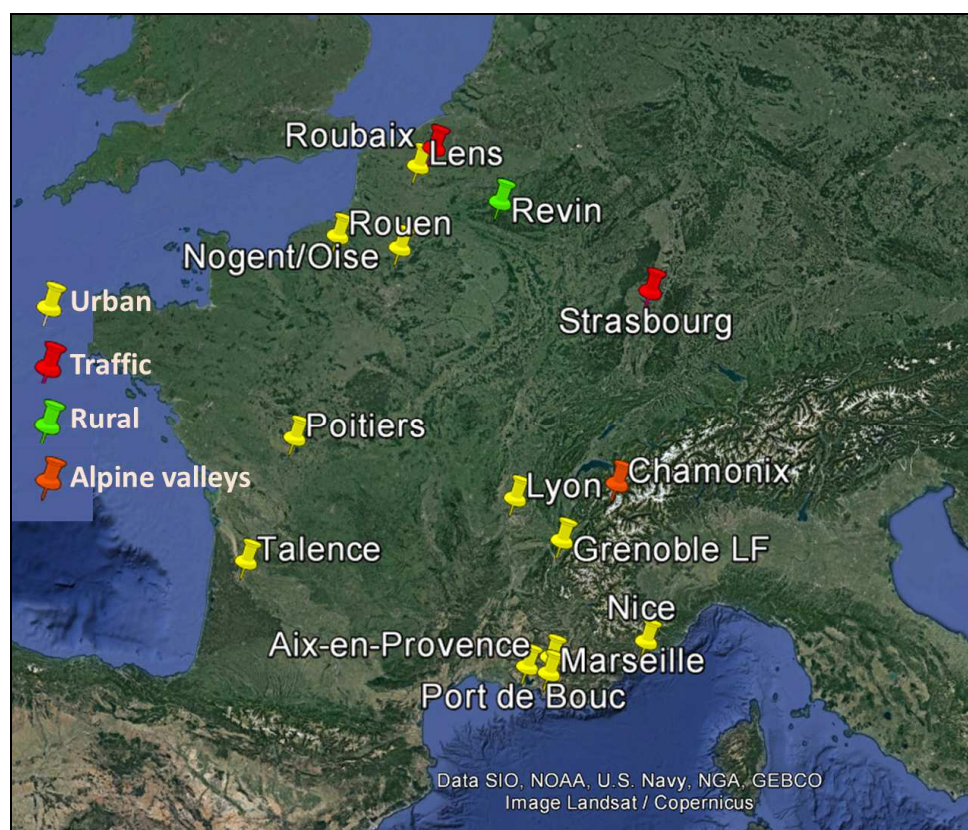


Figure 1 : Localisation des 15 sites de mesure exploités dans le cadre du projet « SOURCES ». Le code couleur indique la typologie du site : jaune pour urbain (N=11), rouge pour trafic (N=2), vert clair pour rural (N=1), et marron pour vallée Alpine (N=1).

Sur l'ensemble de ces sites, des filtres journaliers de PM₁₀ (24 h) ont été collectés par les AASQA sur des filtres en fibre de quartz et généralement à raison de 1 jour sur 3. Une caractérisation chimique détaillée a été réalisée sur l'ensemble de ces filtres. Les méthodologies mises en œuvre pour la réalisation de ces analyses chimiques sont indiquées en Annexe 1.

Ces analyses ont notamment permis la détermination des concentrations de la matière carbonée (OC et EC), des ions majeurs (Cl⁻, NO₃⁻, SO₄²⁻, Na⁺, NH₄⁺, K⁺, Mg²⁺, Ca²⁺), d'un grand nombre de métaux/éléments traces (e.g. Al, Ca, Fe, K, As, Ba, Cd, Co, Cu, La, Mn, Mo, Ni, Pb, Rb, Sb, Sr, V, Zn), et d'une large variété de marqueurs organiques spécifiques de sources, tels que les monosaccharides anhydrides (levoglucosan, mannosan, galactosan) pour la source combustion de biomasse, les polyols (arabitol, sorbitol,

mannitol)¹ pour les émissions primaires biogéniques des sols (spores fongiques), et l'acide méthylsulfonique (MSA) pour la source marine biogénique.

Tableau 1 : Présentation des caractéristiques des 15 sites d'études traités dans SOURCES

Site de mesure	Code	Coordonnées	Période de collecte	Nombre d'échantillons	Typologie
Talence	TAL	44°48'2.01"N 0°35'17.01"E	02/02/2012→ 07/04/2013	154	urbain
Lyon	LY	45°45'27.82"N 4°51'15.15"E	03/01/2012→ 31/12/2012	115	urbain
Poitiers	POI	46°34'48.80"N 0°20'25.34"E	16/11/2014→ 29/12/2015	110	urbain
Nice	NIC	43°42'7.48"N 7°17'10.53"E	04/06/2014→ 29/06/2016	184	urbain
Marseille	MRS	43°18'18.84"N 5°23'40.89"E	11/01/2015→ 27/06/2016	102	urbain
Port de Bouc	PdB	43°24'7.99"N 4°58'55.99"E	01/06/2014→ 27/06/16	185	urbain
Aix-en-Provence	PROV	43°31'49.04"N 5°26'29.00"E	18/07/2013→ 13/07/2014	56	urbain
Nogent	NGT	49°16'35.0"N 2°28'56.0"E	02/01/2013→ 02/06/2014	155	urbain
Rouen	ROU	49°25'41.4"N 1°03'29.1"E	02/01/2013→ 01/06/2014	162	urbain
Lens	LEN	50°26'12.6"N 2°49'36.7"E	02/01/2013→ 01/06/2014	167	urbain
Grenoble LF	GRE	45°9'42.84"N 5°44'8.15"E	02/01/2013→ 29/12/2014	240	urbain
Roubaix	RBX	50°42'23.6"N 3°10'50.5"E	20/01/2013→ 26/05/2014	157	trafic
Strasbourg	STRAS	48°34'24.25"N 7°45'7.6"E	02/04/2013→ 08/04/2014	78	trafic
Chamonix	CHAM	45°55'21"N 6°52'12"E	02/11/2013→ 31/10/2014	115	vallée alpine
Revin	REV	49°55'00.0"N 4°38'29.0"E	02/01/2013→ 01/06/2014	168	rural

¹ Le sorbitol n'a pas été analysé sur les sites de Lens, Nogent, Rouen, Roubaix, et Revin.

3. METHODOLOGIE DE TRAITEMENT DE DONNEES

3.1 Principes généraux de la PMF

Cet outil statistique est utilisé depuis plusieurs années dans le cadre des travaux du programme CARA [e.g., 1-2]. Il permet la détermination de la contribution des principales sources et mécanismes de formation des particules, à partir d'un jeu de données aussi dense que possible (en termes de nombre d'échantillons et d'espèces chimiques mesurées). Cette méthode d'analyse factorielle multivariée prend en compte l'évolution de la concentration des espèces chimiques mesurées et rassemble dans un même facteur les fractions des espèces dont les concentrations évoluent de la même manière dans le temps. Chaque facteur est ensuite identifié et attribué à une source (et/ou catégorie de source) en se basant sur les espèces chimiques spécifiques, marqueurs, caractérisant chaque profil de facteur résolu et leur répartition dans celui-ci.

Du point de vue mathématique, l'objectif d'une analyse PMF appliqué à la matrice de données chimiques X ($n \times m$, où n est le nombre d'échantillons et m le nombre d'espèces chimiques) est de résoudre le nombre de facteurs p , la matrice des contributions G ($n \times p$), et la matrice des profils chimiques F ($p \times m$) selon le principe de conservation de masse suivant (Paatero and Tapper, 1994) :

$$X = G \times F + E$$

Où E est la matrice résiduelle (partie de la matrice X non expliquée par le modèle)

$$x_{ij} = \sum_{k=1}^p g_{ik} \cdot f_{kj} + e_{ij}$$

Avec x_{ij} : concentration de l'espèce j le jour i ; g_{ik} : contribution du facteur k le jour i ; f_{kj} : fraction du facteur k provenant de l'espèce j ; et e_{ij} : fraction de l'espèce j le jour i non expliquée par le modèle (résidu).

Sur la base des incertitudes de mesure, le modèle vise à minimiser une fonction « Q » qui correspond au rapport du résidu (e_{ij}) de l'espèce j le jour i sur l'incertitude (σ_{ij}) associée à la concentration de cette espèce :

$$Q = \sum_{i=1}^{771} \sum_{j=1}^{32} \left(\frac{e_{ij}}{\sigma_{ij}} \right)^2$$

Toutefois, il n'existe pas de solution unique à ce système d'équations en raison de l'ambiguïté due à la liberté de rotation matricielle dans les solutions. Des changements subtils dans les profils chimiques peuvent ainsi modifier les concentrations massiques des facteurs tout en produisant des solutions mathématiquement correctes. Il existe différentes approches pour réduire cette ambiguïté rotationnelle responsable de la multitude de solutions possibles, permettant ainsi d'aboutir à une solution plus réaliste (présentant un mélange de facteurs moindre), telles que l'incorporation dans le modèle des contraintes externes supplémentaires, en se basant sur la connaissance, à priori, des sources influentes sur le site d'étude. Ces contraintes peuvent inclure des profils de sources connus, des contributions des sources connues, ou encore des variables météorologiques (même si très peu d'exemples d'utilisation de cette dernière possibilité

sont mentionnés dans la littérature scientifique).

La version du logiciel utilisé dans cette étude, PMF 5.0 de l'US-EPA, permet notamment d'implémenter deux types de contraintes dans les solutions du modèle :

- ✓ des relations numériques entre espèces (e.g. ratio entre des éléments d'un même profil de source, bilan de masse entre les espèces chimiques d'un même facteur ou de différents facteurs, équation personnalisée par l'utilisateur sur les profils ou les contributions),

- ✓ des contraintes sur les contributions d'un élément chimique dans un/ou plusieurs facteurs qui peuvent être de deux sortes : (i) le forçage fort (*hard pulling*) qui ne tient pas compte de la variation de la valeur de Q (e.g. contribution d'un élément fixée à zéro, ou à sa valeur d'origine, ou attribuée avec des limites inférieures et supérieures) ; et (ii) le forçage doux (*soft pulling*) qui présente une limite de changement autorisée dans la valeur de Q (e.g. contribution tirée au maximum, ou au minimum, ou vers une valeur cible). La valeur de %dQ prise par défaut est de 0.5 mais peut être mise à jour par l'utilisateur s'il souhaite être plus contraignant, une limite de changement plus large de la valeur de Q doit être permise.

D'une manière générale, il convient d'être prudent dans le choix du type/nombre de contraintes à imposer dans la solution PMF afin d'éviter d'introduire trop de subjectivité dans l'analyse, en forçant le modèle vers la solution souhaitée, à priori, par l'utilisateur. Dans le cadre de ce travail, une liste minimale de contraintes chimiques spécifiques et ayant un sens géochimique a été définie et imposée sur les éléments chimiques dans les profils des facteurs (cf. section 3.4).

3.2 Traitements préliminaires des séries de données

Avant de lancer l'analyse PMF sur la matrice de données chimiques, un traitement statistique des valeurs inférieures à la limite de détection (LD), des valeurs manquantes, et des valeurs extrêmes est nécessaire. Les valeurs <LD sont remplacées par LD/2, et les valeurs manquantes par la moyenne géométrique sur la période d'étude considérée. Les valeurs extrêmes liées à des événements ponctuels particuliers atypiques (ou épisodes spécifiques), et qui sont souvent mal reconstituées par la PMF, doivent être identifiées puis traitées, si possible, ou exclues de la série de données. Dans le cadre de ces travaux, l'identification des valeurs extrêmes a été réalisée en utilisant le test statistique de Grubbs qui permet de détecter les valeurs aberrantes en termes de dispersion de moyennes. Le principe de ce test, qui suppose que les données suivent une loi normale, est de comparer les valeurs absolues des écarts réduits, en d'autres termes :

$$G_i = \frac{\max |X_i - \bar{X}|}{\sigma}$$

avec X_i la valeur mesurée, $\bar{X}$ la moyenne de la série, et σ l'écart type de la série de données.

Ainsi, si la valeur G_i est supérieure à la valeur critique G_c , fixée dans ce cas à 5%, la valeur mesurée est considérée comme aberrante au niveau de signification 0,05, et doit être supprimée de la série de mesures ou traitée. Dans le cadre de ce travail, une valeur aberrante obtenue pour plusieurs espèces chimiques a été traitée en excluant

l'échantillon correspondant, alors qu'une valeur aberrante pour une espèce chimique a été traitée en invalidant cette mesure et en la remplaçant par la moyenne géométrique des valeurs de la série de donnée.

3.3 Sélection des variables d'entrée et estimation de leurs incertitudes

La sélection des variables d'entrée a été effectuée en se basant sur le pourcentage des valeurs supérieures aux LD (seuil de 60%), et sur la valeur du rapport signal sur bruit (S/N). Le Tableau 2 présente les variables d'entrée de la PMF retenues dans le cadre de ce travail. Ce jeu de données comprend : la matière carbonée (OC*, EC), les ions majeurs, les marqueurs organiques moléculaires (MSA, Σ polyols, levoglucosan, mannosan), et un grand nombre de métaux/éléments traces. Afin de ne pas doubler les informations relatives aux marqueurs organiques spécifiques utilisés comme variables d'entrée, OC* est déterminé en retranchant aux concentrations totales d'OC les concentrations de ces espèces organiques incluses dans la PMF. Pour ce faire, les concentrations en équivalents Carbone (eqC) de ces dernières espèces ont été préalablement calculées à partir de la masse molaire du carbone et du nombre d'atomes de carbone présents dans chacun de ces composés organiques, selon l'équation suivante :

$$OC^* = OC - ([MSA \times 0.12] + [\Sigma \text{polyols} \times 0.40] + [\text{levoglucosan} \times 0.44] + [\text{mannosan} \times 0.44])$$

Tableau 2 : Liste des espèces chimiques définies comme variables d'entrée pour la PMF

Input data Matrix	X échantillons et Y espèces chimiques	Coefficient "a"
	PM10	total variable
Carbonaceous fraction	OC*, EC	0.03
Ions	Cl, NO ₃ , SO ₄ , Na ⁺ , NH ₄ ⁺ , K ⁺ , Mg ²⁺ , Ca ²⁺	0.05
Organic markers	MSA	0.05
	Σ Polyols, Levoglucosan, mannosan	0.1
Metals/trace elements	Al, As, Ba, Cd, Co, Cr, Cs, Cu, Fe, La, Mn, Mo, Ni, Pb, Rb, Sb, Se, Sn, Sr, Ti, V, Zn	0.15

Pour chacune des analyses PMF effectuées, la variable PM₁₀ a été classée dans le modèle comme variable totale (incertitude augmentée d'un facteur 3 permettant de réduire son influence dans la solution) afin de pouvoir déterminer directement la contribution des facteurs identifiés (sources) aux concentrations de PM₁₀.

En se basant sur le rapport signal/bruit (S/N), les variables ont été classées dans le modèle comme « fortes » (S/N > 2), ou « faibles » (0.2 ≤ S/N ≤ 2), ou « mauvaises » (S/N < 0.2). Les variables classées « mauvaises », sur un site donné, ont été exclues de la liste de variables d'entrée du modèle PMF.

Un travail d'harmonisation de l'estimation des incertitudes a été également entrepris afin de définir une méthodologie commune de calcul des incertitudes des différentes variables d'entrée en se basant sur l'équation de calcul des incertitudes proposée par (Gianini et al., 2012) :

$$\sqrt{(DL)^2 + (x_{ij} \times CV)^2 + (x_{ij} \times a)^2}$$

avec **DL** : limite de détection (2σ_{blancs}) ; **x_{ij}** : concentration de l'espèce j le jour i ; **CV** : coefficient de variation de l'espèce j ; et **a** : coefficient de variation additionnel.

Différents tests d'erreurs ont été conduits avec pour objectif de définir un domaine de variation du coefficient « a » qui soit dépendant du type d'analyse effectuée, c'est-à-dire d'attribuer une valeur de « a » à chaque catégorie/famille d'espèces chimiques analysées. Les solutions PMF obtenues pour différentes valeurs du coefficient « a » ont été examinées de manière à optimiser la stabilité des résultats ainsi que la réalité géochimique des facteurs résolus. Les valeurs retenues *in fine* varient entre 3% et 15% selon les différentes catégories d'espèce chimique (cf. Tableau 2).

Il est également à noter que pour les valeurs < LD, une incertitude de $5/6 \cdot LD$ a été retenue ; et pour les valeurs manquantes ou remplacées (cf. section 3.2), une incertitude égale à 4 fois la moyenne géométrique a été attribuée (Polissar et al., 1998).

3.4 Définition de contraintes spécifiques

Comme mentionné précédemment, l'une des limites de la PMF est liée à une possible colinéarité de facteurs en raison de phénomènes de co-émissions ou de co-accumulation, conduisant à des ambiguïtés rotationnelles. Afin de résoudre ce problème, des contraintes chimiques spécifiques définies sur la base des connaissances géochimiques externes peuvent être injectées dans les profils chimiques (et/ou les contributions) des facteurs PMF (« soft pulling », tel que défini en section 3.1).

Tableau 3 : Liste des contraintes chimiques spécifiques appliquées de manière homogène aux facteurs PMF identifiés sur l'ensemble des 15 sites étudiés (NB : HFO = Heavy Fuel Oil).

Factor profile	Element	Type	Value
Biomass burning	Levogluconan/mannosan	Pull Up Maximally	(% dQ 0.50)
Road traffic	Levogluconan/mannosan	Set to Zero	0
Primary biogenic	Levogluconan/mannosan	Set to Zero	0
Primary biogenic	Σ polyols	Pull Up Maximally	(% dQ 0.50)
Primary biogenic	EC	Pull Down Maximally	(% dQ 0.50)
Secondary biogenic	Levogluconan/mannosan	Set to Zero	0
Secondary biogenic	Σ polyols	Pull Down Maximally	(% dQ 0.50)
Secondary biogenic	MSA	Pull Up Maximally	(% dQ 0.50)
Secondary biogenic	EC	Pull Down Maximally	(% dQ 0.50)
HFO combustion	Levogluconan/mannosan	Set to Zero	0
HFO combustion	Σ polyols	Set to Zero	0
HFO combustion	MSA	Set to Zero	0
Seasalt	Mg/Na	Set to Value	0.119 (% dQ 0.50)

Dans le cadre de cette étude, un jeu minimal et homogène de contraintes chimiques spécifiques a été défini, permettant d'obtenir des résultats optimaux pour l'ensemble des

sites étudiés. Ces contraintes sont synthétisées dans le Tableau 3. A titre d'exemple, les contributions du levoglucosan et mannosan (marqueurs de la combustion de la biomasse) ont été affectées de manière prioritaire au facteur de combustion de biomasse, et fixées à zéro dans le facteur trafic routier où leur présence n'a aucun sens géochimique.

3.5 Choix des solutions

Pour chaque site, différents types de solutions comprenant entre 6 et 11 facteurs ont été examinés afin de retenir la meilleure d'entre elles avec le nombre de facteurs optimal. Différents paramètres de performances statistiques permettant de juger de la robustesse et de la pertinence de la solution finale retenue ont été étudiés suivant les recommandations du rapport du *Joint Research Community* (JRC) (Belis et al., 2014) :

- Evolution du rapport $Q_{\text{true}}/Q_{\text{robust}}$ (< 1.5), où Q_{true} est la valeur calculée en prenant en compte tous les points de mesures, et Q_{robust} est la valeur calculée en excluant les points aberrants (*outliers*) ayant un résidu trop élevé ($e/\sigma > 4$). Les solutions retenues sur l'ensemble des 15 sites présentent un rapport $Q_{\text{true}}/Q_{\text{robust}}$ de 1, indiquant un impact nul des points aberrants sur les résultats.
- Analyse des résidus pondérés qui permet de déterminer le degré d'ajustement de chaque espèce par le modèle. En effet, de très larges résidus ou une distribution non-normale sont souvent indicateurs d'une mauvaise modélisation de l'espèce en question. Dans cette étude, les résidus pondérés présentent pour la majorité des espèces une distribution normale centrée et entre ± 4 , indiquant globalement une bonne modélisation de la plupart des variables.
- Evaluation de la stabilité de la solution à partir du test de bootstrap (BS) pour 100 itérations successives du modèle et pour un minimum de corrélation R^2 de 0.6. Le BS représente la capacité du modèle à retrouver les mêmes résultats en termes de profils chimiques et de contributions des facteurs entre le test initial (*base run*) et 100 différents tests de permutation effectués. Selon les recommandations du JRC, la valeur minimale de BS pour un facteur donné doit être $> 70\%$ et le nombre de facteurs non attribués doit être minimum. Dans cette étude, toutes les solutions présentent des valeurs de BS élevées ($> 70\%$) et le BS le plus faible était de 77%.
- Evaluation de l'ambiguïté rotationnelle et de la sensibilité de la solution aux petits changements à partir du test de Displacement (DISP) proposé par le logiciel utilisé.
- Evaluation du sens géochimique et de la réalité physique des profils de facteurs extraits en se basant sur la connaissance des empreintes chimiques des sources, de leurs traceurs spécifiques et aussi des caractéristiques du site étudié.
- Evaluation pour les solutions avec contraintes des paramètres BS et $\%dQ_{\text{robust}}$. Ce dernier paramètre indique l'augmentation de Q liée à l'application des contraintes. Sur l'ensemble des sites étudiés ici, la valeur minimale de $BS_{\text{contrainte}}$ était de 83%, et l'augmentation totale de dQ variait entre 1.4 et 6.2%. Ces résultats sont très satisfaisants, notamment en comparaison de ceux obtenus dans le cadre de récents travaux similaires (e.g., dQ= 19% dans le cas d'une étude portant sur 5 villes du bassin méditerranéen, Amato et al., 2016).

4. RESULTATS

4.1 Description des profils chimiques (sources) identifiés

Le nombre de facteurs retenus varie entre 8 et 11 selon le site étudié (Figure 2). Parmi ces facteurs, 9 ont été identifiés de façon récurrente sur la majorité des sites : combustion de la biomasse, émissions primaires liées au transport routier, *nitrate-rich*, *sulfate-rich*, biogénique primaire des sols, biogénique secondaire marin, sels marins frais et/ou vieillis, et poussières minérales. Deux autres facteurs minoritaires ont été exclusivement identifiés sur quelques sites et semblent être spécifiques à des sources locales : émissions industrielles (sur 5 sites) et résidus de combustion de fuel lourd (sur 3 sites).

Ces différents profils chimiques ont été identifiés et attribués aux différentes sources (ou catégorie de source) en se basant principalement sur la présence de marqueurs chimiques spécifiques présentés par le Tableau 4.

Tableau 4 : Marqueurs chimiques spécifiques aux différents facteurs identifiés

Facteurs identifiés	Principaux traceurs chimiques
Combustion de la biomasse	Levoglucosan, mannosan, OC, K⁺
Trafic routier primaire	EC, Ba, Cu, Fe, Sb
Nitrate-rich	NO₃⁻, NH₄⁺
Sulfate-rich	SO₄²⁻, NH₄⁺, OC
Biogénique primaire	Polyols
Biogénique secondaire marin	MSA
Poussières minérales	Ca²⁺, Al, Cu, Fe, Mn, Ti
Sels marins frais	Na⁺, Mg²⁺, Cl⁻
Sels marins vieillis	Na⁺, Mg²⁺, NO₃⁻, SO₄²⁻
Industries	As, Cd, Cr, Cs, Co, Ni, Pb, Rb, Se, V, Zn
Combustion de fuel lourd	V, Ni, La, SO₄²⁻, EC

Les principales caractéristiques chimiques des facteurs identifiés sur l'ensemble des 15 sites (Figure 2 et Tableau 5), ainsi que les particularités observées sur certains sites, sont explicitées ci-dessous :

o **Combustion de la biomasse** : ce facteur est caractérisé par la présence de levoglucosan et mannosan avec des proportions importantes d'OC (10-57% de sa masse totale, selon le site étudié), d'EC (4-71%), et de K⁺ (11-60%).

o **Trafic routier primaire** : caractérisé par la forte présence d'EC (majoritairement émis par les combustions de fuels fossiles (hydrocarbures)), jusqu'à 74% de sa masse totale sur le site urbain Rouen et 71% sur le site trafic de Roubaix. Ce facteur contient également, et en fonction du site étudié, d'importantes contributions de différents

éléments métalliques tels Ba (3-80% de sa masse totale), Cu (7-85%), Fe (4-83%), Mo (3-78%), Pb (8-64%), Sb (5-69%), Sn (3-71%), et Zn (2-60%). Ces éléments métalliques peuvent être émis par l'usure de certaines pièces mécaniques des véhicules, ou des pneus et des freins. Ce facteur attribué au trafic routier combine donc des émissions véhiculaires à l'échappement ainsi qu'une partie des émissions hors-échappement.

o **Nitrate-rich (aérosols secondaires enrichis en nitrate d'ammonium)** : caractérisé par la forte présence de NO_3^- (43-79% de sa masse totale) et NH_4^+ (26-74%) indiquant la présence du nitrate sous forme de nitrate d'ammonium (NH_4NO_3). Ce facteur contient également de proportions variables et significatives de sulfate (9-32% de sa masse totale). Une faible présence d'OC est également observée sur certains sites (9% de sa masse totale), représentant ainsi la condensation des composés organiques semi-volatils, potentiellement en mélange interne avec les sels inorganiques secondaires.

o **Sulfate-rich (aérosols secondaires enrichis en sulfate d'ammonium)** : caractérisé essentiellement par la présence de SO_4^{2-} (39-59% de sa masse) et NH_4^+ (21-58%), confirmant la présence de sulfate sous forme de sulfate d'ammonium. Il est à noter que sur les sites urbains d'Aix en Provence et Talence, ce facteur est associé à d'importantes contributions en V et Ni, généralement indicatifs de résidus de combustion de fuel lourd. Des concentrations significatives d'OC (5-22% de sa masse totale) sont également observées dans ce facteur sur la majorité des sites (exceptés Talence et Revin), probablement liées à la formation d'aérosols organiques secondaires (AOS) par des processus de réaction photochimiques, ayant lieu dans l'atmosphère, et de conversion gaz-particules assez similaires à ceux conduisant à la formation du sulfate d'ammonium.

o **Biogénique primaire des sols** : dans l'état actuel des connaissances, ce facteur est préférentiellement relié à la microflore fongique des sols et à la surface de la végétation. Il est dominé par la présence des polyols avec des proportions importantes d'OC (7-22% de sa masse totale, excepté Rouen avec 1%). La contribution d'OC à la masse de ce facteur varie entre 30 et 50%. Malgré la contrainte sur EC, ce facteur présente des fractions importantes d'EC sur les sites de Roubaix (26% de sa masse totale) et Port-de-Bouc (22%), probablement liées à des processus de remise en suspension des sols, favorisant le mélange entre les émissions biogéniques primaires et les sources primaires de combustion (e.g. transport routier pour le site trafic de Strasbourg) ou les particules carbonées déposées au sol. Soulignons que ce facteur peut également contenir une part d'aérosols organiques secondaires, comme indiqué par Waked et al. (2014), rendant difficile la comparaison directe des contributions obtenues sur les différents sites étudiés.

o **Biogénique secondaire marin** : ce facteur est caractérisé par la présence du MSA avec des proportions importantes de SO_4^{2-} , ces deux espèces pouvant provenir de l'oxydation du DMS émis par le phytoplancton marin. La contribution du sulfate à la masse de ce facteur varie entre 4% (Chamonix) et 93% (Nogent). Il est à noter que sur le site urbain de Nice, ce facteur présente également des proportions importantes d'éléments tels que le V et le Ni (55 et 48% de la masse totale, respectivement) indiquant un mélange avec la source combustion de fuel lourd (probablement issu du transport maritime). Ainsi, ce facteur source est plutôt à identifier comme un mélange « biogénique secondaire/HFO » sur ce dernier site.

O Poussières minérales : ce facteur est caractérisé par la forte présence d'éléments minéraux ou métalliques, tels que Ca^{2+} (14-70% de sa masse totale), Al (25-76%), Ti (36-78%), Fe (21-85%), et présente, en fonction du site étudié, des proportions importantes et variées des éléments tels Ba (7-76%), Co (12-78%), Cu (2-81%), Mn (16-77%), et Zn (3-78%) qui sont émis par l'usure des pneus et des freins. Cette signature caractérise ainsi les processus de remise en suspension des poussières à partir des sources telles l'érosion des sols, les chantiers de construction, la circulation routière, les aires de stationnement. Pour les sites de Grenoble, Poitiers, Lyon, Talence et Revin, ce facteur semble essentiellement d'origine terrigène avec une faible présence des éléments typiques des émissions issues de l'abrasion des freins et des pneus. A l'inverse, le site trafic de Strasbourg contient de très fortes proportions de ces marqueurs des émissions véhiculaires hors-échappement. A noter que ce facteur contient également, et de façon récurrente, des proportions significatives d'espèces secondaires inorganiques (nitrate et sulfate) ainsi que de la matière organique (potentiellement aussi d'origine secondaire), provenant de la condensation de précurseurs gazeux sur les particules terrigènes par réactions acido-basiques.

O Sels marins frais : ce facteur est caractérisé par la prédominance de Cl^- (74-92% de sa masse totale), Na^+ (24-71%), et Mg^{2+} (15-67%) qui sont des composés majeurs de l'eau de mer. Le rapport Cl/Na dans ce profil varie entre 1,3 (Nice) et 2,1 (Roubaix), mais reste proche de celui classiquement mesuré en milieu marin (1,8), indiquant une très faible transformation de ces particules au cours de leur transport.

O Sels marins vieillis : caractérisé par la présence de Na^+ (28-77% de sa masse totale), Mg^{2+} (5-57%), et également du NO_3^- (1-35%) et du SO_4^{2-} (1-27%) en proportions variables selon le site étudié. La déplétion en chlorure dans ce facteur résulte des réactions entre NaCl et H_2SO_4 et/ou HNO_3 conduisant à la formation de Na_2SO_4 ou NaNO_3 et à sa perte sous forme de $\text{HCl}_{(\text{gazeux})}$. Sur le site de Chamonix, ce facteur est plus vraisemblablement lié aux pratiques de salage des routes en hiver qu'à une influence marine compte tenu de la distance par rapport aux côtes maritimes.

O Emissions industrielles : ce facteur a été identifié uniquement sur les sites de Talence, Lyon, Poitiers, Aix en Provence et Port-de-Bouc et se caractérise principalement par la présence importante des éléments métalliques ou métalloïdes tels que As (20-82% de sa masse totale), Cd (14-36%), Pb (14-62%), Se (8-56%), et Zn (2-45%). Ces éléments sont fréquemment observés au sein de diverses émissions industrielles, telles que la métallurgie, la production de céramique, de verre, de ciment et la combustion de charbon (Alleman et al., 2010; Amato et al., 2016; Sylvestre et al., 2017).

O Combustion de fuel lourd (HFO) : ce facteur n'a été identifié sans ambiguïté sur seulement 3 sites : Poitiers, Marseille et Port-de-Bouc, alors qu'il est mélangé avec le sulfate-rich à Talence et Aix en Provence, et avec le biogénique secondaire marin à Nice. Ce facteur est caractérisé par la présence prédominante des éléments V (26-79% de sa masse totale) et Ni (55-68%), liés à la combustion de fioul lourd par les raffineries et centrales thermiques et liés aux émissions du transport maritime. Ce facteur contient également des quantités significatives d'EC (17-24% de sa masse totale) et de SO_4^{2-} (21% de sa masse à Marseille et 8% à Port-de-Bouc).

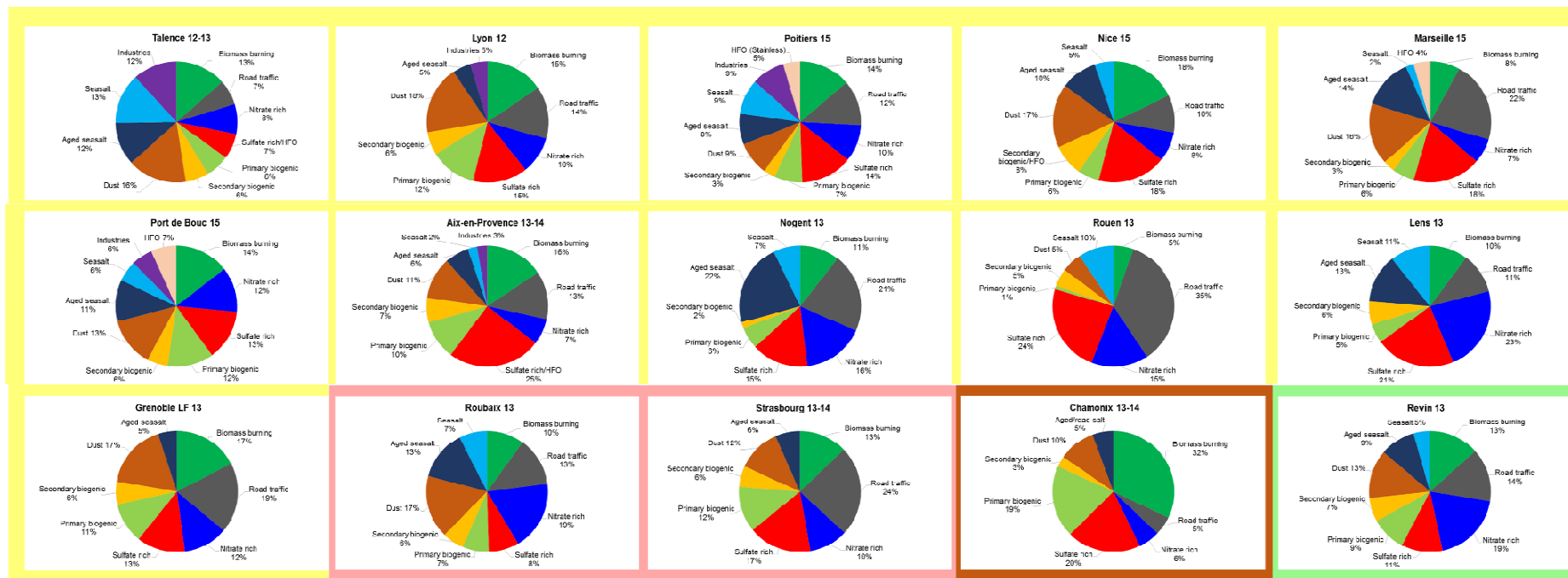


Figure 2 : Contributions relatives moyennes annuelles aux PM_{10} des facteurs identifiés pour chacun des 15 sites. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

Tableau 5 : Synthèse des concentrations annuelles moyennes (en $\mu\text{g.m}^{-3}$) obtenues pour chaque site et chaque facteur issu de l'analyse PMF harmonisée.

Facteurs ($\mu\text{g/m}^3$)	Talence 12-13	Lyon 12	Poitiers 15	Nice 15	Marseille 15	Port de Bouc 15	Aix-en-Provence 13-14	Nogent 13	Rouen 13	Lens 13	Grenoble LF 13	Roubaix 13	Strasbourg 13-14	Chamonix 13-14	Revin 13
Biomass burning	3.6	3.2	2.7	3.1	1.4	2.7	3.6	3.4	1.0	2.5	4.5	2.5	2.5	7.3	1.6
Road traffic	1.6	2.0	1.8	1.5	3.7		2.2	4.4	5.6	2.1	3.9	2.4	3.8	0.9	2.0
Nitrate rich	2.2	2.5	2.5	1.5	1.4	2.5	1.5	4.7	4.9	5.9	3.9	6.5	2.5	1.3	3.4
Sulfate rich	1.2	2.6	2.1	3.0	3.3	2.6	3.6	4.0	4.8	4.1	2.5	2.2	3.2	2.0	1.7
Primary biogenic	1.0	1.6	1.0	0.9	1.0	1.8	1.7	1.0	0.1	0.7	1.7	1.3	1.8	1.9	0.9
Secondary biogenic	0.9	0.7	0.4	1.5	0.5	0.9	0.8	0.3	0.7	0.8	0.9	1.0	0.8	0.3	0.7
Dust	3.2	2.7	1.4	2.8	2.9	2.1	1.9		0.9		3.4	4.2	1.8	1.4	1.8
Aged seasalt	2.0	0.6	1.0	1.6	2.3	1.8	1.0	3.8		1.6	0.8	2.3	0.8	1.2	0.9
Seasalt	1.9		1.2	0.8	0.3	0.8	0.4	1.2	1.3	1.3		1.4			0.4
Industries	2.5	0.8	1.3			0.8	0.5								
Heavy fuel oil (HFO)			0.5		0.8	1.2									

4.2 Contributions annuelles moyennes et variations saisonnières

Les contributions relatives et concentrations annuelles moyennes obtenues pour chaque site sont présentées respectivement sur les Figures 2 et le Tableau 5. Une synthèse de ces résultats, ainsi que de la variabilité temporelle de chacun des facteurs identifiés, est proposée dans cette section.

4.2.a. Combustion de la biomasse : Cette source a été identifiée pour chacun des sites étudiés, où elle représente une contribution significative aux PM_{10} en moyenne annuelle comprise entre 5 et 32% (correspondant à des concentrations atmosphériques de 1,0 à 7,5 $\mu g.m^{-3}$).

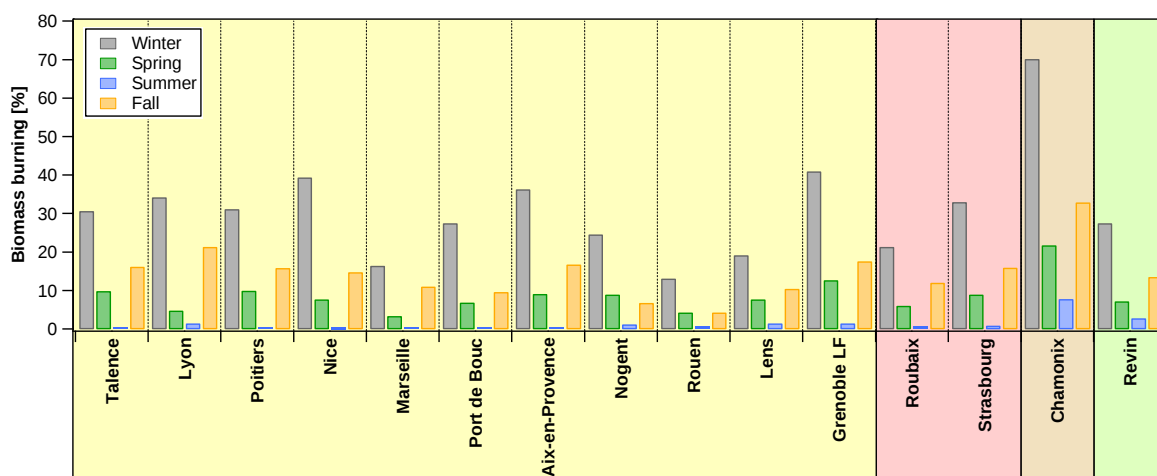


Figure 3 : Variabilités saisonnières et spatiales des contributions relatives moyenne de la combustion de la biomasse aux PM_{10} sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

Comme attendu et illustré par la Figure 3, on observe une variabilité saisonnière très marquée avec des concentrations moyennes plus élevées en hiver sur l'ensemble des sites. La concentration maximale hivernale est obtenue sur le site de Chamonix (21 $\mu g.m^{-3}$), où la combustion de la biomasse est largement prédominante avec une contribution d'environ 70% à la masse des PM_{10} . Ce résultat est en accord avec la localisation de ce site dans la vallée de l'Arve fortement impactée par le chauffage résidentiel au bois et soumis à conditions météorologiques spécifiques (fréquentes et fortes inversions thermiques) (Chevrier et al., 2016). Pour les autres sites de fond urbain, la concentration moyenne hivernale des particules liées à la combustion de biomasse est également très importante à Grenoble, Talence, Lyon, Poitiers, Nice, Aix en Provence, et Nogent (entre 6 et 10 $\mu g.m^{-3}$) correspondant à 24-41% des concentrations en PM_{10} . Les sites de Marseille, Port-de-Bouc, Rouen et Lens présentent des concentrations moyennes plus faibles (allant de 2.1 à 5.1 $\mu g.m^{-3}$) contribuant entre 13 et 27% des concentrations en PM_{10} . L'ensemble de ces résultats est en bon accord avec ceux obtenus précédemment sur les sites du programme CARA (e.g., LCSQA, 2015b). Les deux sites de proximité automobile de Roubaix et Strasbourg présentent des concentrations moyennes hivernales de 4,6 $\mu g.m^{-3}$ (21%) et 7,2 $\mu g.m^{-3}$ (33%), respectivement. Enfin, sur le site de fond rural de Revin, cette source contribue à 27% des PM_{10} (2,7 $\mu g.m^{-3}$) en hiver.

4.2.b. Trafic routier primaire : Ce facteur, combinant potentiellement les émissions directes de particules à l'échappement et hors-échappement, impacte significativement l'ensemble des sites urbains étudiés (Figure 4) et plus particulièrement les sites de Rouen, Nogent, Grenoble, Strasbourg et Marseille, avec des contributions moyennes de $5,6 \mu\text{g.m}^{-3}$ (35% des PM_{10}), $4,4 \mu\text{g.m}^{-3}$ (21%), $3,9 \mu\text{g.m}^{-3}$ (19%), $3,8 \mu\text{g.m}^{-3}$ (24%), $3,7 \mu\text{g.m}^{-3}$ (22%), respectivement. Par comparaison avec les autres sites, l'importance des valeurs maximales obtenues pour Rouen peut paraître surprenante. Comme discuté plus loin (section 5), ces valeurs peuvent être surestimées d'environ 30% (pour une contribution annuelle moyenne restant tout de même de l'ordre de 25%). A l'inverse, l'absence de ce facteur dans les solutions obtenues pour Port-de-Bouc peut résulter d'un biais de l'analyse harmonisée, et d'interférences avec les activités industrielles. Les contributions les plus faibles ont été obtenues pour Bordeaux et pour le site alpin de Chamonix (respectivement 7 et 5% des PM_{10}), avec un minimum de concentrations (environ $1 \mu\text{g.m}^{-3}$) à Chamonix. Les autres sites (Lyon, Nice, Roubaix, Poitiers, Aix en Provence, Revin, et Lens) présentent des contributions relatives globalement similaires, de l'ordre de 10-15% des PM_{10} en moyenne annuelle.

Les variations saisonnières de la contribution de cette source sont relativement peu marquées, avec toutefois des maximums observés plus fréquemment en été (Lyon, Poitiers, Strasbourg, Roubaix, Revin) et/ou surtout en automne (Lyon, Marseille, Aix, Nogent, Rouen, Lens, Grenoble, Roubaix, Revin) (Figure 4).

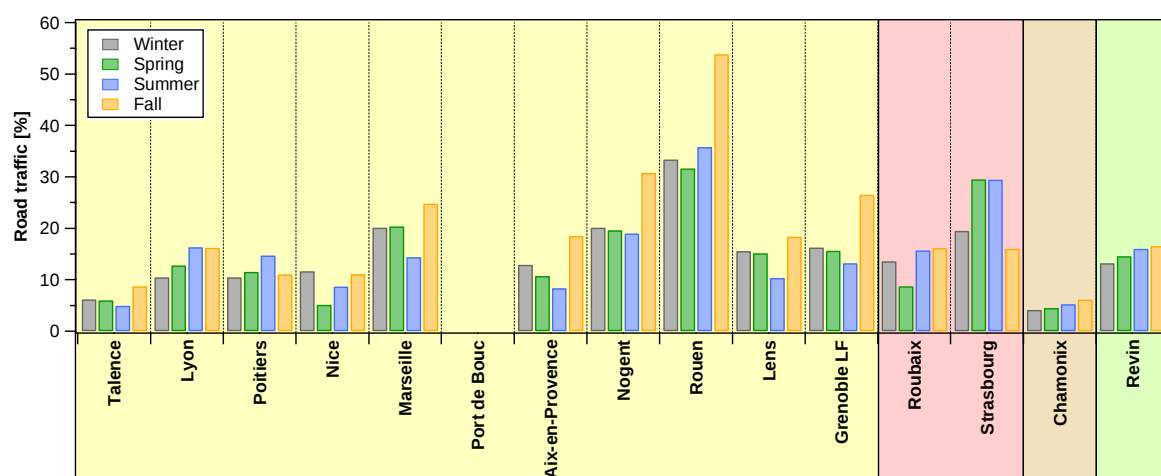


Figure 4 : Variabilités saisonnières et spatiales des contributions relatives moyennes du trafic routier primaire aux PM_{10} sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

4.2.c. Nitrate-rich (aérosols secondaires enrichis en nitrate d'ammonium) : ce facteur présente des concentrations annuelles moyennes plus élevées ($3,4\text{-}6,5 \mu\text{g.m}^{-3}$) sur les sites les plus au nord (Roubaix, Lens, Rouen, Nogent et Revin) ainsi qu'à Grenoble, où il représente entre 15 et 23% de la masse des PM_{10} (Figure 5). Sur les autres sites, les concentrations annuelles varient entre $1,3$ et $2,5 \mu\text{g.m}^{-3}$, pour une contribution à la masse des PM_{10} de l'ordre de 10%. Cette variabilité spatiale s'explique par les sources d'influence et les conditions météorologiques incluant l'importance des concentrations de précurseurs gazeux (NO_x et NH_3) en raison de la circulation routière et des activités agricoles et d'élevages, et/ou par des conditions météorologiques plus favorables aux processus photochimiques (température, humidité relative, inversions thermiques). Dans

la moitié sud de la France, la caractère semi-volatile du nitrate d'ammonium, et les températures généralement plus élevées, limite sa condensation en phase particulaire. Une variabilité saisonnière très marquée, avec des maximums au printemps et/ou en hiver selon les sites étudiés est observée (Figure 5). Pour exemple, sur le site trafic de Roubaix, ce facteur présente une concentration moyenne de $11,5 \mu\text{g.m}^{-3}$ au printemps, contribuant à 30% de la masse des PM_{10} . A Lens, Rouen, Nogent, Revin et Talence, les concentrations printanières sont de l'ordre de 7 à $9 \mu\text{g.m}^{-3}$, avec des contributions relatives de 23-34%. Sur les sites urbains de Lyon, Grenoble, Marseille et Port-de-Bouc, les concentrations les plus élevées sont mesurées en période hivernale ($2,9$ - $5,9 \mu\text{g.m}^{-3}$) contribuant à environ 15% à la masse des PM_{10} . Enfin, sur le site de Chamonix, les contributions de ce facteur aux PM_{10} restent en dessous de 10% tout au long de l'année (pour une concentration annuelle de $1,3 \mu\text{g.m}^{-3}$).

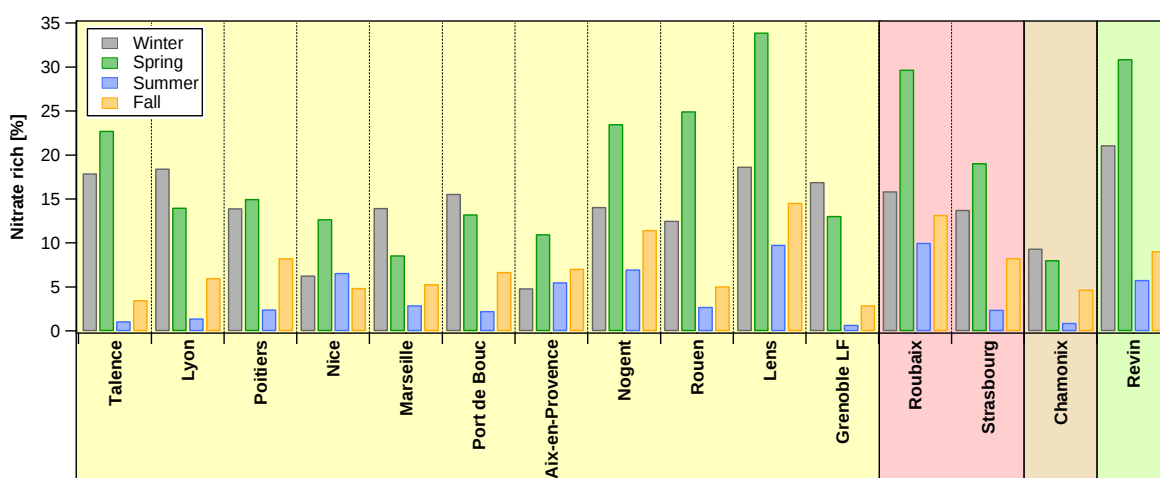


Figure 5 : Variabilités saisonnières et spatiales des contributions relatives moyennes du facteur « nitrate-rich » aux PM_{10} sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

4.2.d. Sulfate-rich (aérosols secondaires enrichis en sulfate d'ammonium) : Les concentrations annuelles moyennes de ce facteur varient entre $1,2$ et $4,8 \mu\text{g.m}^{-3}$ et présentent les valeurs les plus élevées sur les sites du nord de la France (Lens, Rouen et Nogent, $> 4 \mu\text{g.m}^{-3}$) ainsi qu'en PACA (de l'ordre de $3 \mu\text{g.m}^{-3}$). Les contributions moyennes à la masse des PM_{10} varient entre 7% (Talence) et 25% (Aix en Provence), et les aérosols secondaires enrichis en sulfate d'ammonium constituent la source majoritaire des PM_{10} pour ce dernier site. La variabilité spatiale de ce facteur est notamment liée à l'influence des sources de SO_2 (e.g., industries pétrochimiques, transport maritime, comme indiqué par l'importance des traceurs V et Ni). Les contributions naturelles aux concentrations de sulfate, via par exemple l'oxydation du diméthylsulfure (DMS) émis par le phytoplancton, doit sans doute rester faible dans ces milieux urbains.

Une variabilité saisonnière marquée avec des valeurs plus importantes en été et au printemps sur l'ensemble des sites (à l'exception de Lyon) peut être observée (Figure 6). Ce cycle saisonnier est à relier à une activité photochimique plus intense liée au fort ensoleillement favorisant l'oxydation du SO_2 . Les concentrations estivales les plus élevées sont observées sur les sites de Rouen ($5,5 \mu\text{g.m}^{-3}$), Aix en Provence ($4,8 \mu\text{g.m}^{-3}$), Lens Marseille ($4,4 \mu\text{g.m}^{-3}$), Strasbourg ($4,3 \mu\text{g.m}^{-3}$), Nice ($3,5 \mu\text{g.m}^{-3}$), et Port-de-Bouc ($2,9 \mu\text{g.m}^{-3}$), contribuant de 19 à 38% à la masse des PM_{10} . La variabilité saisonnière de cette

source est moins marquée sur les sites de Talence, Poitiers, Grenoble et Revin, qui présentent des concentrations printanières, estivales, et automnales relativement stables et légèrement plus importantes qu'en hiver. En termes de contributions relatives, des valeurs relativement comparables ont été observées sur l'ensemble des saisons, excepté au cours de la période estivale à Lyon (13-18%), Port-de-Bouc (11-14%), Nogent (11-13%), Rouen (12-19%) et Lens (13-19%), et la période hivernale à Talence (7-9%). Les sites de Roubaix et Revin présentent des contributions comparables sur l'ensemble de la période d'étude (6-8% et 8-11%, respectivement).

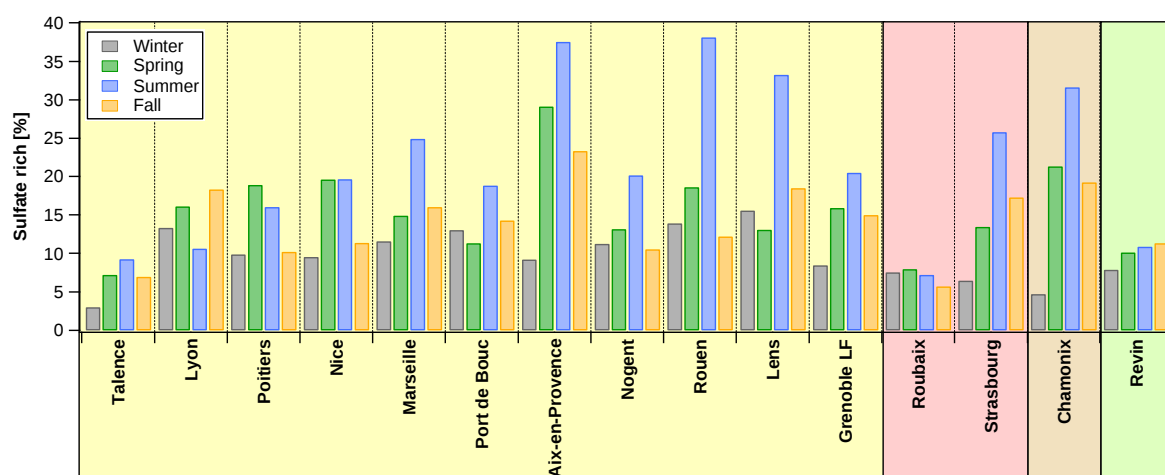


Figure 6 : Variabilités saisonnières et spatiales des contributions relatives moyennes du facteur « sulfate-rich » aux PM₁₀ sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

4.2.e. Biogénique primaire des sols : ce facteur présente les concentrations moyennes les plus élevées (1,6-1,9 $\mu\text{g}\cdot\text{m}^{-3}$) sur les sites de Chamonix, Strasbourg, Port-de-Bouc, Aix en Provence, Grenoble et Lyon, avec des contributions moyennes de 10-19% à la masse des PM₁₀ (Figure 7). Sur les autres sites, les contributions moyennes de ce facteur sont comparables (5-9%) à l'exception de Rouen où elle ne contribue qu'à une fraction négligeable des PM₁₀ (1%).

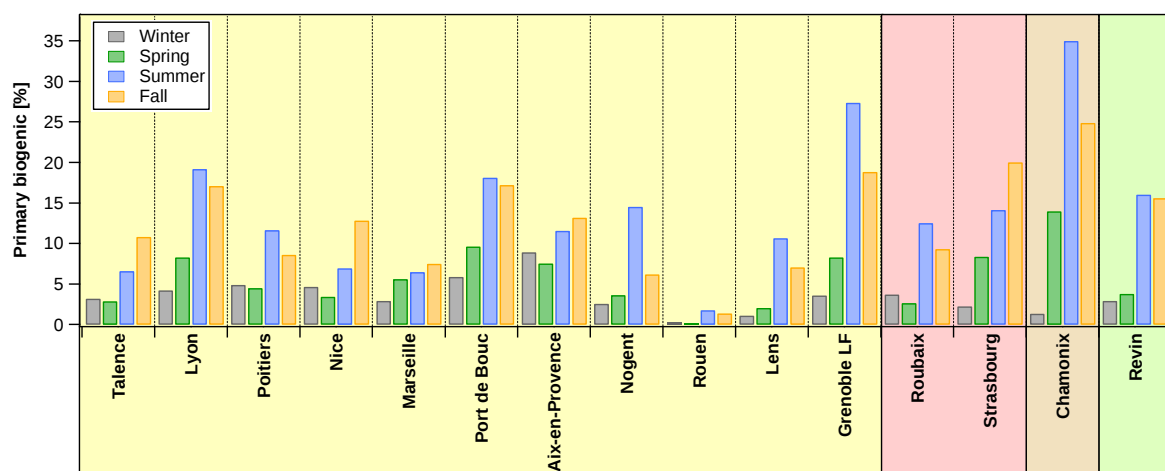


Figure 7 : Variabilités saisonnières et spatiales des contributions relatives moyennes du facteur « Biogénique primaire des sols » aux PM₁₀ sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

Les émissions de cette source sont liées aux spores fongiques et émissions de pollens, se caractérisant par une variabilité saisonnière bien marquée avec de plus fortes concentrations observées en été et en automne. Les niveaux élevés en période estivale sont liés à l'activité biologique accrue durant cette saison, favorisée par le rayonnement solaire intense et les températures élevées, alors que ceux observés en automne semblent être liés à la décomposition de la végétation. On retiendra tout d'abord que les contributions estivales et automnales de cette source sont relativement comparables (environ $2 \mu\text{g.m}^{-3}$) sur les sites de Lyon, Poitiers, Port-de-Bouc, Lens, Roubaix, et Revin. Les concentrations estivales les plus élevées ont été mesurées sur les sites de Grenoble, Chamonix, Port-de-Bouc, Nogent, Lyon, Roubaix et Strasbourg qui présentent des valeurs relativement comparables ($1,9\text{-}2,9 \mu\text{g.m}^{-3}$), contribuant entre 13% (Roubaix) et 35% (Chamonix) à la masse des PM_{10} en été (Figure 7). En automne, des concentrations comparables ($1,8\text{-}2,8 \mu\text{g.m}^{-3}$) ont été observées sur les sites de Talence, Lyon, Nice, Port-de-Bouc, Aix en Provence, Grenoble, Roubaix et Chamonix, avec un maximum constaté à Strasbourg ($3,3 \mu\text{g.m}^{-3}$), contribuant de 9% (Roubaix) à 25% (Chamonix) à la masse des PM_{10} .

4.2.f. Biogénique secondaire marin (AOS marin) : les concentrations moyennes de ce facteur varient entre $0,3 \mu\text{g.m}^{-3}$ (Chamonix) et $1,5 \mu\text{g.m}^{-3}$ (Nice) représentant des contributions moyennes relativement faibles sur l'ensemble des sites (2-8% de la masse des PM_{10}) (Figure 8). Un cycle saisonnier marqué avec des concentrations maximales en fin de printemps - début d'été, liées à l'importance des émissions biogéniques par le phytoplancton à cette période de l'année, peut-être observé. Le rayonnement solaire favorise les émissions et l'oxydation photochimique du DMS expliquant donc les fortes concentrations observées lors de ces saisons (voir section 4.1).

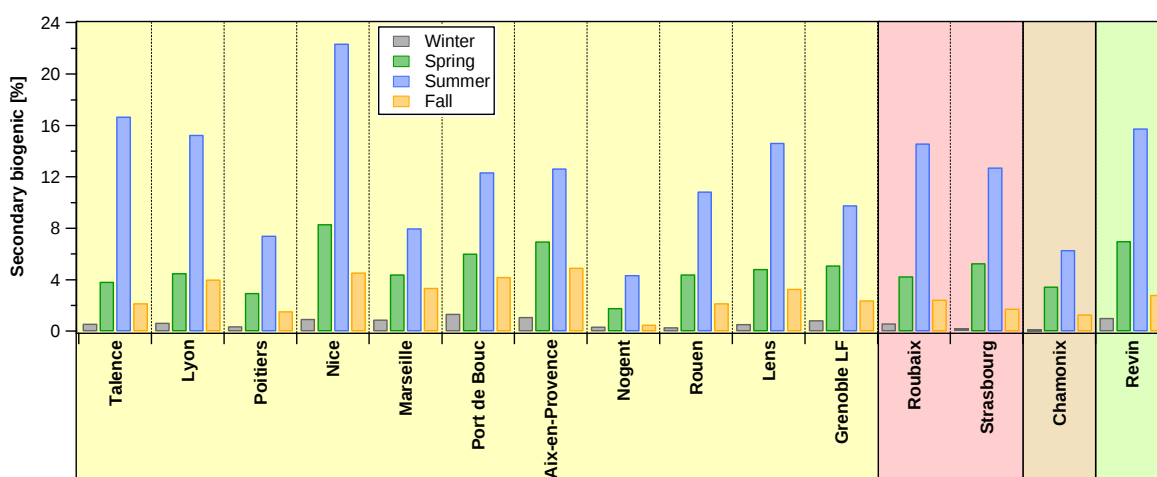


Figure 8 : Variabilités saisonnières et spatiales des contributions relatives moyennes du facteur « Biogénique secondaire marin » aux PM_{10} sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

Le maximum de concentration estival est observé pour Nice ($3,5 \mu\text{g.m}^{-3}$) avec une contribution de 22% à la masse des PM_{10} , et les plus faibles contributions sont observées sur les sites non côtiers de Poitiers ($0,9 \mu\text{g.m}^{-3}$; 7%), Nogent ($0,7 \mu\text{g.m}^{-3}$; 4%), et Chamonix ($0,5 \mu\text{g.m}^{-3}$; 6%). Sur les autres sites, les concentrations estivales sont relativement comparables ($1,2\text{-}2,2 \mu\text{g.m}^{-3}$), et contribuent entre 8% (Marseille) et 17% (Talence) à la masse des PM_{10} en été (Figure 8).

A noter également que, sur le site de Nice, ce facteur explique entre 55% et 48% de la masse totale des éléments V et Ni, traduisant un probable mélange avec les émissions de combustion de fioul lourd. Cette contribution additionnelle, avec une influence attendue de l'AOS marin, explique l'importance de ce facteur à Nice.

4.2.g. Poussières minérales : ce facteur regroupe l'import sur le site récepteur d'émissions naturelles (érosion des sols, poussières sahariennes) mais également de la remise en suspension des particules par le transport routier. Les résultats obtenus montrent des concentrations annuelles moyennes variant entre 0,9 et 4,2 $\mu\text{g.m}^{-3}$, avec les niveaux les plus élevés observés sur le site trafic de Roubaix (4,2 $\mu\text{g.m}^{-3}$) et les sites urbains de Grenoble (3,4 $\mu\text{g.m}^{-3}$), et Talence (3,2 $\mu\text{g.m}^{-3}$). En termes de contributions relatives, cette source représente entre 5% (Rouen) et 18% (Lyon) des PM₁₀ en moyenne annuelle (Figure 9).

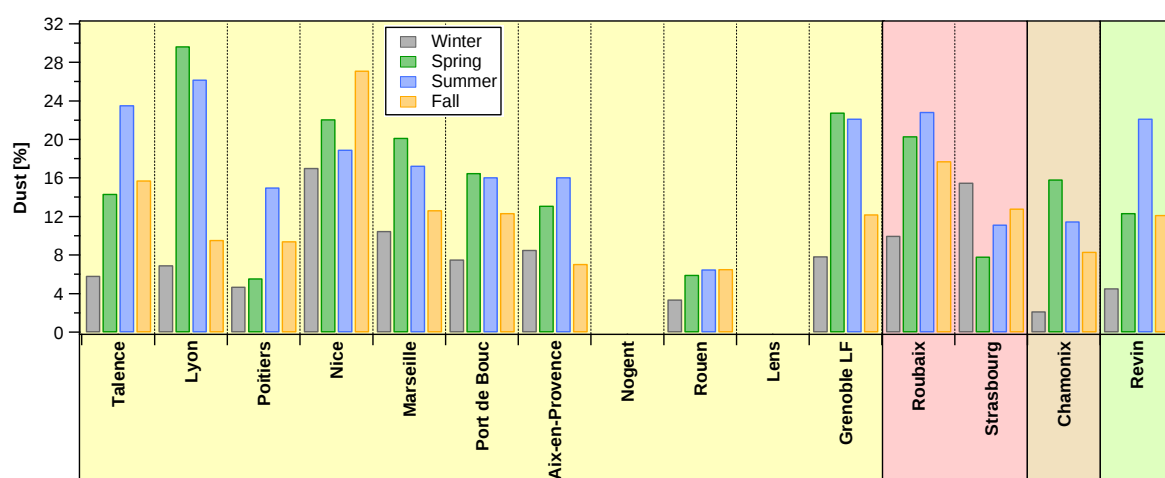


Figure 9 : Variabilités saisonnières et spatiales des contributions relatives moyennes du facteur « poussières minérales » aux PM₁₀ sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

Cette source présente une variation spatiale et saisonnière significative avec de plus faibles contributions observées en hiver, excepté à Nice (3,3 $\mu\text{g.m}^{-3}$; 17%) et Strasbourg (2,4 $\mu\text{g.m}^{-3}$; 16%), et des contributions plus importantes et relativement comparables au printemps et en été sur un grand nombre de sites, comme Lyon (3,4-4,4 $\mu\text{g.m}^{-3}$; 26-30%), Nice (2,9-3,5 $\mu\text{g.m}^{-3}$; 19-22%), Marseille (2,8-3,9 $\mu\text{g.m}^{-3}$; 17-20%), Port-de-Bouc (2,4 $\mu\text{g.m}^{-3}$; 16-17%), Aix en Provence (2,0-2,6 $\mu\text{g.m}^{-3}$; 13-16%), Grenoble (3,5-4,3 $\mu\text{g.m}^{-3}$; 22-23%), et Roubaix (4,5-5,9 $\mu\text{g.m}^{-3}$; 20-23%). A Nice, la contribution la plus importante est observée en automne (4,8 $\mu\text{g.m}^{-3}$) contribuant à 27% de la masse des PM₁₀ avec des valeurs relativement comparables pour les autres saisons (2,9-3,5 $\mu\text{g.m}^{-3}$; 17-22%). A Talence, Poitiers, et Revin, des valeurs plus élevées ont été mesurées en été, contribuant à 24%, 15% et 22% de la masse des PM₁₀, respectivement. Les niveaux élevés observés en été et au printemps sont principalement liés à des conditions météorologiques plus sèches, favorisant la remise en suspension des poussières locales, et probablement aussi aux épisodes de poussières crustales terrigènes (import de particules sahariennes notamment) qui ont lieu généralement à ces périodes de l'année.

4.2.h. Sels marins frais : cette source présente des concentrations annuelles moyennes qui varient entre $0,3 \mu\text{g.m}^{-3}$ (Marseille) et $1,9 \mu\text{g.m}^{-3}$ (Talence), et contribuent entre 2% et 13% à la masse des PM_{10} . On note une variation saisonnière relativement marquée avec de faibles contributions en été ($0,1\text{-}0,7 \mu\text{g.m}^{-3}$; 1-6%), excepté à Talence ($2,4 \mu\text{g.m}^{-3}$; 18%) (Figure 10). Des contributions plus importantes et comparables sont obtenues en hiver sur les sites de Rouen ($2,4 \mu\text{g.m}^{-3}$; 23%), Talence ($2,2 \mu\text{g.m}^{-3}$; 16%), Nogent ($2,1 \mu\text{g.m}^{-3}$; 13%), Roubaix ($2,1 \mu\text{g.m}^{-3}$; 12%), Poitiers ($2,1 \mu\text{g.m}^{-3}$; 15%), et Lens ($2,0 \mu\text{g.m}^{-3}$; 16%). Les contributions automnales sont également importantes et relativement comparables ($1,1\text{-}1,8 \mu\text{g.m}^{-3}$; 7-15%) sur les sites de Roubaix, Lens, Rouen, Nogent, Talence, Poitiers, et Port-de-Bouc. De même, les contributions printanières sont relativement comparables ($1,0\text{-}1,5 \mu\text{g.m}^{-3}$; 7-11%) sur un grand nombre de sites (Talence, Poitiers, Nice, Port-de-Bouc, Nogent, Rouen, Lens et Roubaix). Cette similarité dans l'évolution spatiale des contributions saisonnières sur la majorité des sites reflète un impact régional relativement constant des épisodes des sels de mers qui ont lieu généralement en hiver et au printemps en raison notamment des vitesses de vents plus importantes. Les occurrences de cette source coïncident également avec une faible influence des processus photochimiques entraînant une transformation des embruns marins sous l'effet de mélange avec des émissions anthropiques (favorisant l'émergence de sels marins vieillis en été, cf. 4.2.i).

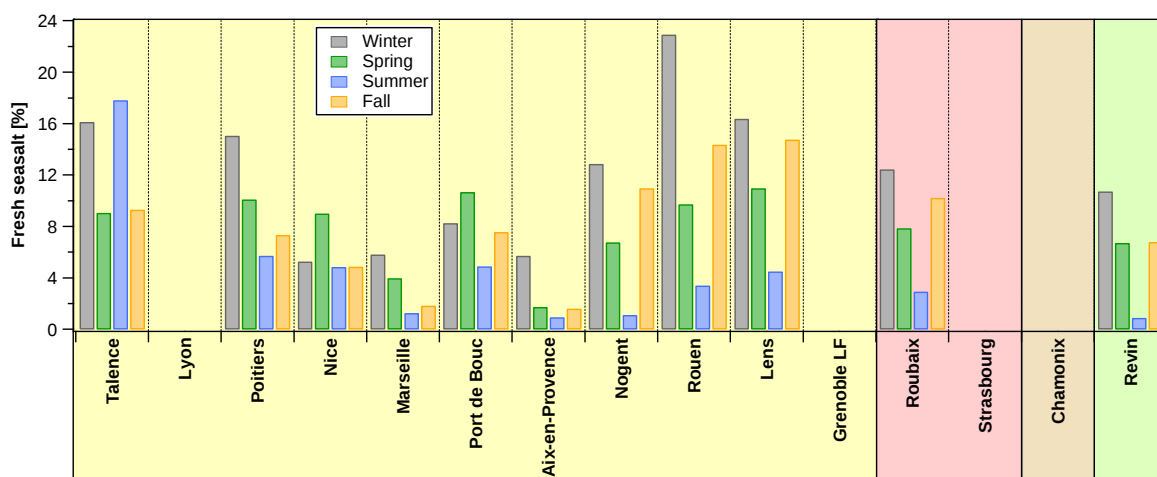


Figure 10 : Variabilités saisonnières et spatiales des contributions relatives moyennes du facteur « sels marins frais » aux PM_{10} sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

4.2.i. Sels marins vieillis : les concentrations moyennes des sels marins vieillis varient entre $0,6 \mu\text{g.m}^{-3}$ (Lyon) et $3,8 \mu\text{g.m}^{-3}$ (Nogent), et représentent entre 5 à 22% de la masse des PM_{10} en moyennes annuelles sur l'ensemble des sites (Figure 11). Cette source ne présente pas de variabilité saisonnière très significative sur la plupart des sites. Cependant, des contributions plus importantes en période chaude (été notamment) sont observées sur certains sites. En particulier, les résultats obtenus pour Nogent indiquent des contributions plus élevées en été ($5,2 \mu\text{g.m}^{-3}$; 33%), au printemps ($4,1 \mu\text{g.m}^{-3}$; 22%), et en automne ($3,6 \mu\text{g.m}^{-3}$; 23%) par rapport à l'hiver ($2,3 \mu\text{g.m}^{-3}$; 14%) (Figure 11). De même, des contributions plus élevées ont été observées à Nice et Poitiers au printemps et en été (11-12%), comparé à l'hiver (5%). A l'inverse, le site alpin de Chamonix montre des contributions plus importantes au printemps ($2,5 \mu\text{g.m}^{-3}$; 11%) et en hiver ($1,7 \mu\text{g.m}^{-3}$

; 8%) comparé aux autres saisons (0,1-0,3 $\mu\text{g.m}^{-3}$; 2%). Dans ce cas, le salage des routes explique certainement cette saisonnalité spécifique (section 4.1). Les autres sites présentent des contributions relativement comparables sur les 4 saisons, tels que Lyon (0,4-0,8 $\mu\text{g.m}^{-3}$; 4-5%), Marseille (1,9-2,7 $\mu\text{g.m}^{-3}$; 14-17%), Port-de-Bouc (1,4-2,0 $\mu\text{g.m}^{-3}$; 9-14%), Lens (1,4-1,9 $\mu\text{g.m}^{-3}$; 12-15%), Grenoble (0,6-0,8 $\mu\text{g.m}^{-3}$; 4-6%), Roubaix (2,0-2,5 $\mu\text{g.m}^{-3}$; 12-15%), et Revin (0,8-1,0 $\mu\text{g.m}^{-3}$; 7-12%).

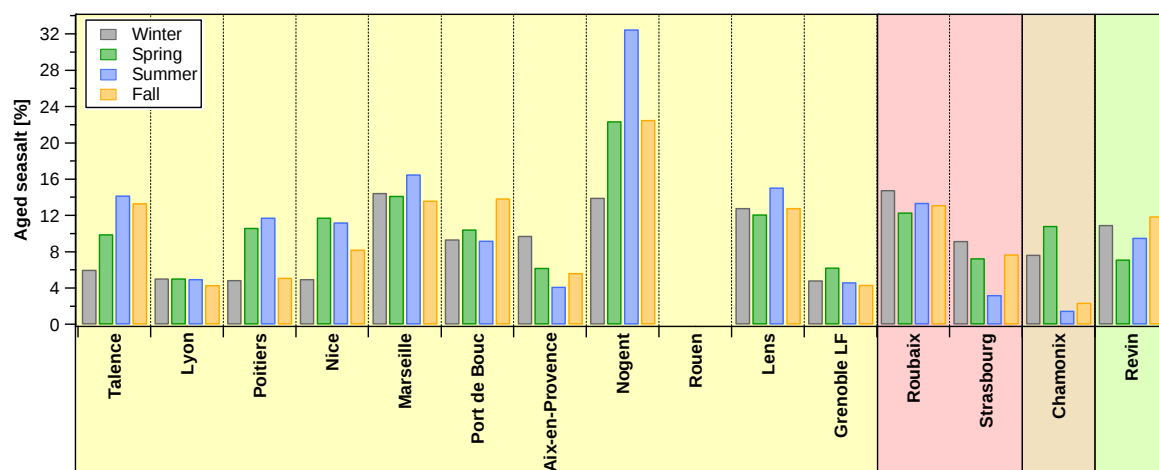


Figure 11 : Variabilités saisonnières et spatiales des contributions relatives moyennes du facteur « sels marins vieillis » aux PM_{10} sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

4.2.j. Emissions liées aux activités industrielles : ces émissions regroupent ici les facteurs liés aux rejets industriels primaires et ceux issus des produits de combustion de fioul lourd. Les émissions primaires ne sont significativement observées que sur les sites de Talence, Lyon, Poitiers, Port-de-Bouc, et Aix en Provence (Tableau 5 et Figure 12). Les résultats obtenus indiquent des concentrations annuelles moyennes variant entre 0,5 $\mu\text{g.m}^{-3}$ (Aix en Provence) et 2,5 $\mu\text{g.m}^{-3}$ (Talence), et des contributions relatives comprises entre 3% et 12% à la masse des PM_{10} en moyenne annuelle. Ce facteur se caractérise également par une variabilité saisonnière plus marquée sur les sites de Poitiers et Talence (Figure 12).

A Poitiers, le maximum de concentration est observé en automne (3,3 $\mu\text{g.m}^{-3}$), contribuant à 21% de la masse des PM_{10} pendant cette saison. A Talence, la plus faible contribution est observée en été (0,9 $\mu\text{g.m}^{-3}$; 5%) avec toutefois des contributions plus élevées et comparables pendant les autres saisons (2,5-3 $\mu\text{g.m}^{-3}$; 10-14%). A Lyon, une concentration plus élevée est observée en hiver (1,7 $\mu\text{g.m}^{-3}$) comparé aux autres saisons (0,4-0,7 $\mu\text{g.m}^{-3}$) avec toutefois des contributions relatives assez stables (3-6%) d'une saison à l'autre. De même, Port-de-Bouc présente une concentration plus élevée en hiver (1,2 $\mu\text{g.m}^{-3}$) comparé aux autres saisons (0,5-0,9 $\mu\text{g.m}^{-3}$) avec des contributions relativement comparables (6-8%, avec une valeur de 3% en automne). Enfin, à Aix en Provence, les concentrations varient de 0,3 à 0,9 $\mu\text{g.m}^{-3}$ et présentent des contributions relatives faibles et plutôt stables tout au long de l'année (2-4% des PM_{10}).

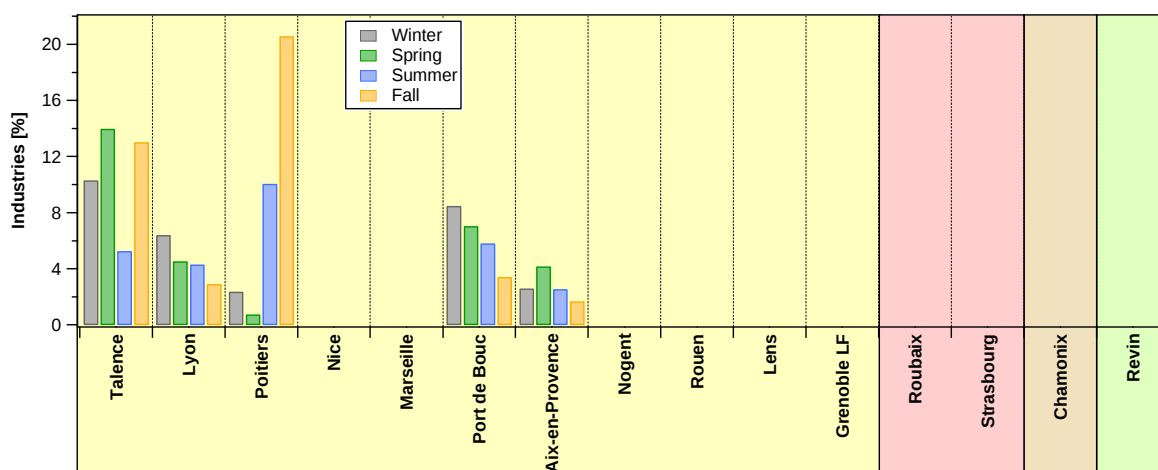


Figure 12 : Variabilités saisonnières et spatiales des contributions relatives moyennes du facteur « activités industrielles » aux PM₁₀ sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

Le facteur lié à la combustion de fioul lourd n'est spécifiquement identifié que pour les sites de Poitiers, Marseille et Port-de-Bouc (Tableau 5 et Figure 13). Il présente des concentrations annuelles moyennes comprises entre 0,5 (Poitiers) et 1,2 $\mu\text{g.m}^{-3}$ (Port-de-Bouc), ne contribuant que faiblement à la masse des PM₁₀ (4-7%). Des contributions plus élevées en été, en automne et au printemps sont observées à Port-de-Bouc (1,3-1,9 $\mu\text{g.m}^{-3}$), contribuant de 8 à 12% à la masse des PM₁₀ (Figure 13). A Marseille, la contribution la plus élevée est observée en été (1,3 $\mu\text{g.m}^{-3}$; 8%) avec des valeurs comparables pour le reste des saisons (0,6-0,8 $\mu\text{g.m}^{-3}$; 3-4%). A Poitiers, la contribution la plus importante est observée au printemps (1,0 $\mu\text{g.m}^{-3}$; 10%) avec des valeurs relativement comparables sur les autres saisons (0,3-0,5 $\mu\text{g.m}^{-3}$; 2-5%).

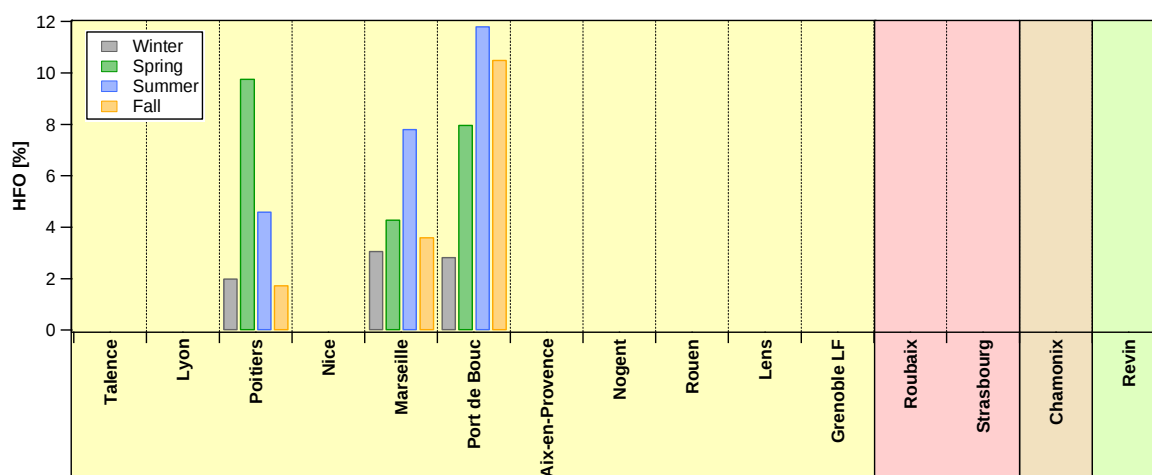


Figure 13 : Variabilités saisonnières et spatiales des contributions relatives moyennes du facteur « combustion de fioul lourd » aux PM₁₀ sur les 15 sites étudiés. La couleur du fond indique la typologie du site : jaune pour urbain, rose pour trafic, marron pour vallée alpine, et vert pour rural.

Les fortes concentrations liées aux activités industrielles et aux combustions de fioul lourd obtenues pour les sites de fond urbain de Bordeaux et Poitiers, respectivement 2,5 et 1,8 $\mu\text{g.m}^{-3}$ en moyenne annuelle, au minimum deux fois plus importantes que sur tous les

autres sites excepté Port-de-Bouc, peut paraître surprenante. En effet, de telles différences entre ces villes et d'autres agglomérations plus industrialisées (e.g., Rouen et Lyon) ne sont pas attendues, et pourraient partiellement s'expliquer par une différence de séparation entre ces facteurs et ceux liés aux aérosols secondaires d'une solution PMF à une autre. Ainsi, une étude de sensibilité de la méthodologie appliquée a été réalisée en comparant les résultats obtenus avec la méthodologie harmonisée et les résultats précédemment obtenus avec les mêmes jeux de données d'origine (voir section 5).

5. SENSIBILITE DE LA METHODOLOGIE DE TRAITEMENT DES DONNEES

Afin d'évaluer la robustesse des résultats obtenus à l'aide du traitement harmonisé des jeux de données, ces résultats ont été comparés à ceux obtenus lors d'études précédentes pour certains des mêmes jeux de données d'origine. Ces dernières ont été réalisées selon des procédures de traitement de données individualisées selon les caractéristiques propres de sites étudiés (e.g. interactions avec les acteurs de terrain apportant une connaissance ciblée des principales sources locales potentielles), et en utilisant les jeux de données, des outils et méthodologie disponibles à cette époque. Pour ce faire, deux cas ont été considérés :

- 1) Comparaison avec les résultats d'analyses PMF réalisées à partir d'un panel globalement similaire d'espèces chimiques. C'est le cas notamment des jeux de données relatifs aux sites de Bordeaux et Lyon, dont une analyse préalable avait été réalisée par le LCSQA et Atmo Aquitaine en 2013, à l'aide de du logiciel PMF3 de l'US EPA (version ne permettant pas d'appliquer de contraintes sur les profils chimiques) [LCSQA, 2013a].

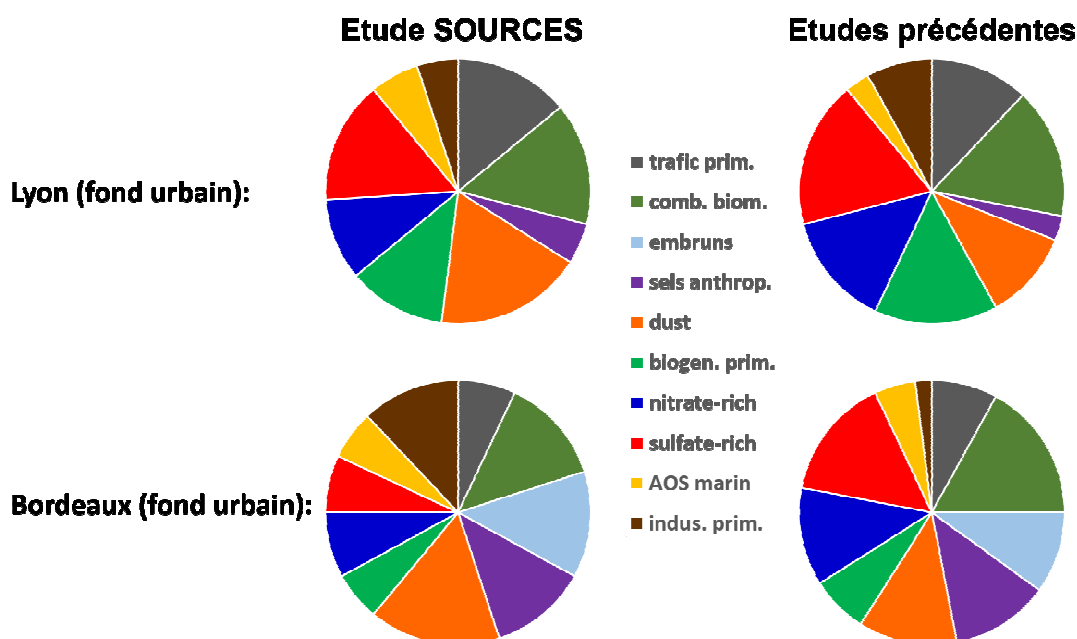


Figure 14 : Comparaison des résultats obtenus dans le cadre du projet SOURCES pour les jeux de données de Lyon (2012) et Bordeaux (2012-2013) aux résultats obtenus précédemment pour ces mêmes jeux de données à l'aide de l'outil US EPA PMF3 (pas de contraintes sur les profils chimiques).

- 2) Comparaison avec les résultats d'analyses PMF réalisées à partir d'un panel d'espèces chimiques significativement différent. C'est le cas notamment des sites de Rouen (résultats issus de la thèse de doctorat de Diogo Oliveira (Oliveira, 2017), Strasbourg (étude LCSQA) et Poitiers (Atmo Nouvelle Aquitaine, 2016). Pour chacun de ces sites, les études précédentes prenaient en compte l'oxalate comme variable chimique d'entrée. Ce composé, considéré comme un marqueur non équivoque de différents processus de formation d'AOS, a été exclu de la liste des variables d'entrée pour le traitement de données harmonisé en raison d'une influence négative sur la stabilité des solutions (forte variation des profils chimiques obtenus d'un site à l'autre).

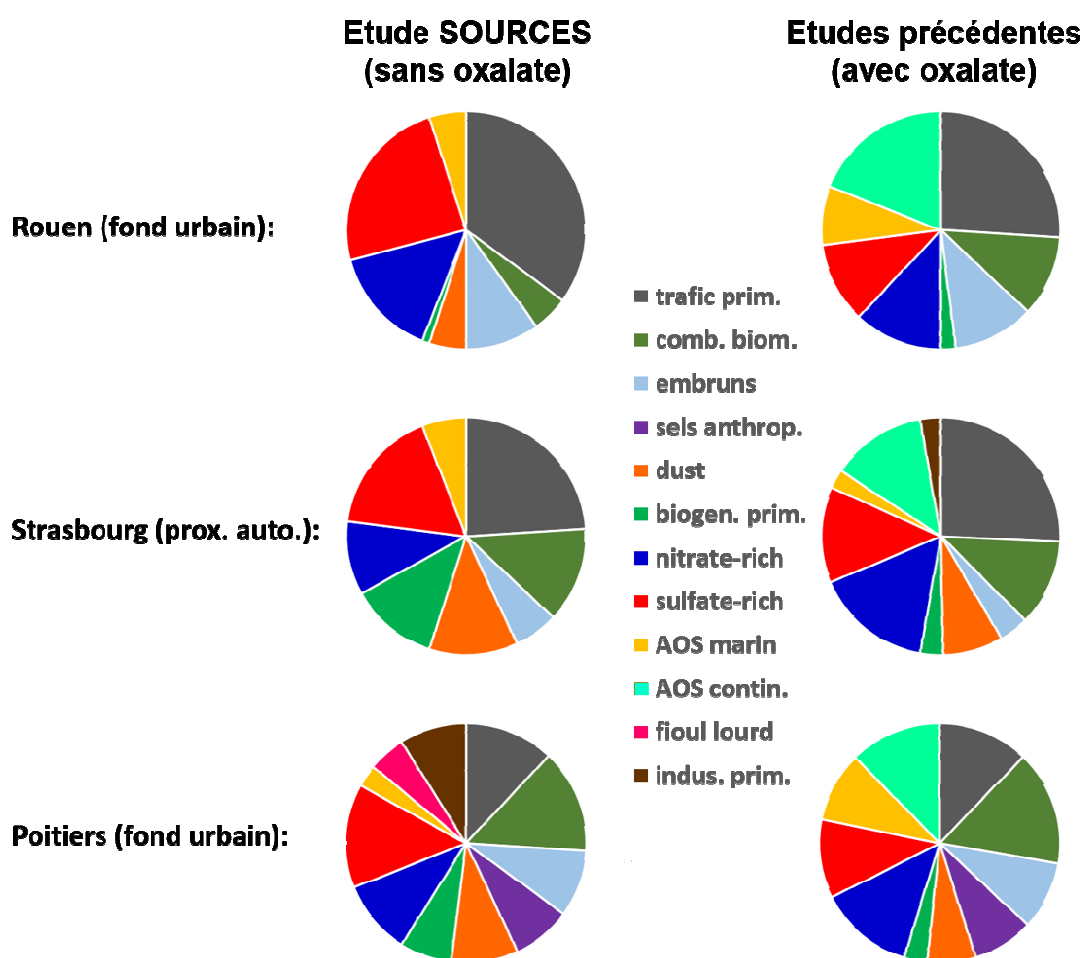


Figure 15 : Comparaison des résultats obtenus dans le cadre du projet SOURCES pour les jeux de données de Rouen (2013), Strasbourg (2013-2014) et Poitiers (2015) aux résultats obtenus précédemment pour ces mêmes sites (respectivement avec/sans prise en compte de l'oxalate dans le panel d'espèces chimiques utilisées).

Lorsque les jeux de données de départ sont similaires, une très bonne correspondance entre les deux approches pour le nombre et le type de facteurs mis en évidence est observée (Figure 14). Ainsi, l'introduction de contraintes sur les profils chimiques ne semble pas avoir d'influence significative sur ces paramètres. Ceci s'explique par le fait que l'application de ces contraintes, n'intervient qu'après que le nombre optimum de

facteurs soit déterminé. Ces résultats montrent donc la cohérence dans le choix du nombre de profils de sources significatifs et dans leur interprétation par différents opérateurs utilisant des variables d'entrée similaires.

Cependant, quelques différences dans les contributions relatives attribuées à certains facteurs, en particulier ceux étant directement attribués aux aérosols secondaires (nitrate-rich, sulfate-rich, secondaires marins), aux poussières terrigènes (pouvant également être associées à des composés secondaires), ainsi qu'aux facteurs pouvant être liés aux émissions industrielles et/ou à la combustion de fioul lourd, peuvent être observées. Les différences les plus significatives sont observées pour Bordeaux avec des contributions des facteurs sulfate-rich et émissions industrielles (incluant ici également les émissions de combustion de fioul lourd) passant respectivement de 15% à 7% et de 2% à 12% entre les résultats des études précédentes et ceux obtenus ici. Ces variations peuvent s'expliquer par l'utilisation de contraintes sur les profils chimiques dans le cas du projet SOURCES (cette étude) (Tableau 3). Elles traduisent également la part de subjectivité dans le choix de la solution PMF retenue par différents opérateurs. En fonction de leur connaissance, a priori, de l'impact attendu des sources d'émission locales potentielles, la solution finale peut donc être affectée. Ces différences illustrent par ailleurs la difficulté à séparer par analyse PMF les familles d'aérosols secondaires (e.g., sulfate-rich et nitrate-rich) des facteurs liés à des sources d'émission de certains de leurs précurseurs gazeux (e.g., rejets industriels et/ou de combustion de fioul lourd).

En outre, l'identification des facteurs de facteurs est d'autant plus problématique qu'elle nécessite une bonne prise en compte des différentes familles d'AOS, dont les marqueurs chimiques ne sont pas forcément bien connus ni systématiquement mesurés (en raison de la difficulté et des coûts d'analyse de certains de ces marqueurs). Ainsi, l'utilisation de l'oxalate comme variable d'entrée de la PMF semble permettre de distinguer un facteur directement lié aux AOS « continentaux » en complément des autres facteurs secondaires (Figure 15). Dans le cas de Strasbourg, par exemple, l'identification de ce nouveau facteur lié aux AOS (avec une contribution aux PM_{10} d'environ 17% en moyenne annuelle) induit une diminution notable de la contribution des facteurs sulfate-rich (de 17% à 13%) ainsi que des facteurs poussières minérales et biogéniques primaires (respectivement de 12% à 8% et de 12% à 3%). En effet, ces deux derniers facteurs peuvent également contenir des AOS (voir section 3). Comme dans le cas des analyses basées sur des jeux de données similaires, les différences dans les résultats obtenus entre les deux analyses avec des jeux de données plus ou moins étendus, sont également imputables à la sensibilité de l'opérateur aux spécificités du site d'étude (Figure 15). En particulier, l'attribution de contributions significativement plus faibles pour les émissions véhiculaires primaires à Rouen et pour les émissions industrielles à Poitiers résultent de la bonne adéquation entre les solutions proposées par l'outil PMF et la connaissance de l'influence réelle attendue pour ces sources d'émission locales.

On retiendra de ces comparaisons que les différences observées ne concernent que marginalement les contributions obtenues pour les sources primaires de combustion de biomasse, d'émissions véhiculaires primaires, et les sels de mer. Ainsi, l'introduction de contraintes sur ces facteurs dans le traitement de données harmonisé n'induit pas de forte variation sur les résultats obtenus pour ces types de source. En revanche, les modalités de traitement et d'utilisation des variables d'entrée à disposition, ainsi que l'influence de l'opérateur dans le choix de la solution PMF la plus plausible, peuvent

conduire à des différences non négligeables dans l'estimation des contributions des facteurs pouvant contenir une forte proportion d'aérosols secondaires. L'utilisation d'un panel élargi de marqueurs organiques, primaires et secondaires, permettrait d'affiner l'identification et la quantification des différentes familles d'aérosols secondaires (rapport LCSQA 2015a). En outre, une approche basée sur la mesure à plus haute résolution temporelle (e.g., mesures automatiques, dont analyses des métaux sur un pas de temps horaire) pourrait à l'avenir également permettre de mieux documenter l'influence réelle des sources d'émission primaires les moins influentes (typiquement, émissions industrielles et/ou combustion de fioul lourd).

6. BILAN ET PERSPECTIVES

Le présent rapport rend compte (i) du développement d'une méthodologie de traitement de données harmonisée pour l'attribution des principales sources de PM par analyse de type PMF, et (ii) de l'application de cette méthodologie sur un ensemble des jeux de données relatifs à la spéciation chimique des PM₁₀ en France au cours des 6 dernières années. Pour ce faire, il a été entrepris de regrouper l'ensemble des données nationales disponibles depuis 2011 sur une quinzaine de sites de typologies différentes à travers différents programmes et projets de recherche. Les principales conclusions de cette étude sont synthétisées dans cette section, tout d'abord en termes d'aspects méthodologiques, puis sous l'angle des résultats obtenus.

Aspects méthodologiques :

Ce travail d'harmonisation a inclus différentes étapes allant du traitement statistique de base de chaque jeu de données (e.g., valeurs inférieures aux limites de détection analytiques, points aberrants), puis la sélection des variables d'entrée à la PMF avec l'estimation de leurs incertitudes, jusqu'à l'application des contraintes chimiques spécifiques dans les profils chimiques de certains facteurs. Différents paramètres de performances statistiques (e.g., $Q_{\text{true}}/Q_{\text{robust}}$, distribution des résidus, Bootstrap) ont été pris en compte pour le choix de la meilleure solution sur chaque site étudié.

De nouveaux traceurs organiques (e.g., polyols et MSA) rarement utilisés dans les études précédentes ont permis la quantification de sources biogéniques spécifiques (e.g., émissions primaires et AOS d'origine marine).

A notre connaissance, il s'agit de la première étude PMF harmonisée et multi-sites des sources de PM. Ce travail pourra être valorisé à l'échelle européenne, notamment dans le cadre des travaux de normalisation du groupe de travail WG 44 du CEN/TC 264. De plus, cette méthode pourrait être utilisée dans les futures études PMF à large échelle spatiale, permettant ainsi d'améliorer la comparabilité des résultats (e.g. profils chimiques) entre les différents sites et des différentes régions.

Les perspectives d'optimisation de ce traitement harmonisé des jeux de données incluent :

- L'élargissement du nombre de contraintes appliquées (e.g. sur les facteurs secondaires) avec une évaluation de l'impact de ces contraintes sur les profils chimiques et les contributions des sources.

➤ L'évaluation des incertitudes sur les profils chimiques en exploitant de nouvelles fonctionnalités disponibles dans la version du logiciel PMF5 de l'US-EPA.

➤ L'utilisation de l'outil DeltaSA mis en place par le JRC pour l'évaluation de la similarité des profils chimiques permettant de comparer un profil dérivant de la PMF avec ceux disponibles dans les bases de données de SPECIATE (US-EPA) et SPECIEUROPE (EC-JRC). Cet outil permet également d'évaluer les performances du modèle utilisé.

➤ Lorsque les données sont disponibles, étendre la liste des variables d'entrée de la PMF avec les mesures de nouveaux traceurs organiques (e.g. n-alcanes, hopanes, cellulose oxy- et nitro-HAP dérivés, HULIS, 3-MBTCA, et autres marqueurs de composés secondaires), ou avec les mesures par Aethalomètre, ou des mesures isotopiques (e.g., ^{14}C , ^{15}N), ce qui pourraient permettre d'améliorer la résolution de certaines sources, y compris les profils chimiques constitutifs de processus secondaires.

Concernant ce dernier point, il apparaît qu'une limitation inhérente à l'utilisation d'un protocole harmonisé réside dans la sélection du panel des espèces chimiques utilisées comme variables d'entrée de la PMF. Selon ce panel de variables, et étant donné la typologie et les caractéristiques spécifiques des sites étudiés, l'utilisation d'une approche de traitement de données harmonisée peut ainsi conduire à des imprécisions dans les résultats obtenus, en particulier pour certaines sources primaires d'influence relativement limitée (e.g., émissions industrielles et/ou combustion de fioul lourd) et/ou des facteurs liés aux aérosols secondaires. A ce titre, dans le cas d'étude mono-site, il est recommandé de s'affranchir des contraintes induites par l'harmonisation des procédures de traitement de données.

En conclusion de ces aspects méthodologiques, il semble également important de rappeler que :

1) Les outils statistiques de type sources-récepteur (telle que l'analyse PMF) ne permettent pas d'étudier directement l'origine géographique des principales sources d'émissions identifiées.

2) Compte-tenu de la non-linéarité des mécanismes atmosphériques, les estimations de contributions de sources réalisées à l'aide des méthodologies expérimentales ne peuvent pas être utilisées simplement afin d'évaluer l'efficacité de scénarii de gestion de qualité de l'air ou pour l'étude d'impacts de l'efficacité mesures spécifiques de réduction des émissions.

Sur ces deux derniers points, le recours à la modélisation numérique - basée sur les inventaires d'émission, la paramétrisation des conditions météorologiques et de dynamique atmosphérique, ainsi que des phénomènes de (trans-)formation des aérosols - est primordial. La validation de ces modèles de chimie-transport passe cependant par la comparaison des résultats obtenus avec les données de composition chimique et de contributions de sources primaires déterminées à l'aide des méthodologies expérimentales, telle que celle utilisée dans le cadre de cette étude.

Contributions moyennes des sources aux concentrations de PM₁₀ :

De 8 à 11 facteurs ont été identifiés sur les différents sites d'étude. Les principales conclusions sur l'influence de l'impact des sources identifiées sont résumées ci-après :

✓ Les émissions primaires liées au transport routier influencent très fortement la masse des PM₁₀ sur les sites de proximité automobile, mais également sur certains sites de fond

urbain tels que Rouen, Marseille, Grenoble et Nogent (avec des contributions relatives de l'ordre de 20-30% en moyenne annuelle). Sur les autres sites, les contributions annuelles sont d'environ 10%.

- ✓ La source de combustion de la biomasse est très importante en hiver, contribuant jusqu'à 70% de la masse des PM_{10} en moyenne saisonnière à Chamonix (contre 10-15% à Marseille, Lens et Rouen, et environ 30% sur les autres sites).
- ✓ Les aérosols secondaires riches en nitrate d'ammonium présentent une concentration maximale au printemps (typiquement 30% des PM_{10}), en particulier sur les sites de la moitié nord de la France.
- ✓ L'impact des aérosols secondaires riches en sulfate d'ammonium sur les niveaux de PM_{10} est observé plus spécifiquement en période estivale (en particulier dans le sud de la France). Ce facteur constitue notamment la source majoritaire des PM_{10} à Aix en Provence, avec une contribution relative de 25%, en moyenne annuelle.
- ✓ Les poussières minérales présentent des contributions importantes et relativement comparables au printemps et en été sur un grand nombre de sites.
- ✓ Les émissions biogéniques primaires continentales (e.g., spores fongiques) sont significatives entre le début de l'été et la fin de l'automne.
- ✓ La contribution des aérosols secondaires biogéniques marins (identifiés à l'aide du MSA) est également significative au printemps – été, notamment en PACA.
- ✓ Les sels marins vieillis contribuent significativement aux niveaux de PM_{10} sur la plupart des sites étudiés, représentant jusqu'à environ 20% des PM_{10} en moyenne annuelle (à Nogent), et avec des contributions saisonnières relativement stables (typiquement de l'ordre de 10%).
- ✓ Le facteur industriel identifié sur seulement 5 des 15 sites étudiés présente une variabilité spatiale inhomogène. Il reste cependant difficile de bien isoler cette source. Il en est de même pour les résidus de combustion de fioul lourd (incluant notamment les émissions du transport maritime).

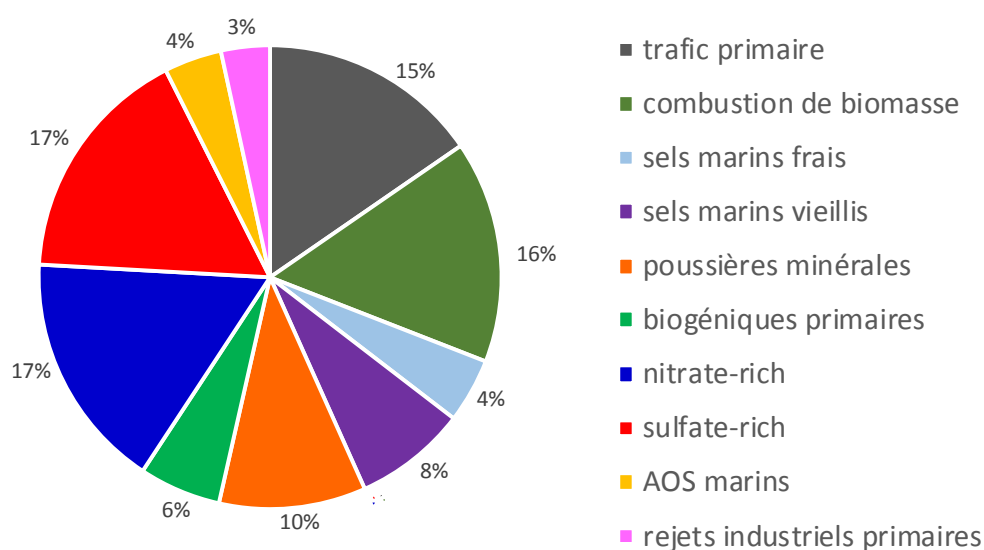


Figure 16 : Contributions annuelles moyennes des principaux facteurs constitutifs des PM_{10} en fond urbain à l'échelle nationale (Moyenne des résultats obtenus par analyse PMF harmonisée sur 10 sites de fond urbain du projet SOURCES).

A titre de synthèse et d'illustration de ces conclusions, la Figure 16 présente les contributions moyennes des principales sources de PM à l'échelle nationale en fond urbain, obtenues en combinant les résultats issus de l'analyse PMF harmonisée pour les 10 agglomérations suivantes : Rouen, Nogent sur Oise, Lens, Poitiers, Lyon, Grenoble, Bordeaux, Aix en Provence, Marseille, et Nice.

Les perspectives d'utilisation des résultats scientifiques obtenus dans le cadre de cette étude incluent notamment :

- Le couplage des sorties PMF avec des données de rétro-trajectoires de masses d'air, permettant à terme d'identifier les zones géographiques potentielles d'émissions des sources de PM.
- La comparaison des résultats obtenus avec les inventaires d'émission et des sorties de modèles numériques de type CTM (*Chemical Transport Model*). Cette étape est essentielle pour la mise en place des stratégies de gestion de la qualité de l'air.
- La mise en regard des sources de PM identifiées avec leur possibles impacts sanitaires en utilisant des indicateurs pertinents, tel, par exemple, que le potentiel oxydant des particules (et/ou de tests cellulaires) comme nouvelle métrique d'évaluation de la qualité de l'air.

7. REFERENCES

- Alleman, L. Y., Lamaison, L., Perdrix, E., Robache, A. and Galloo, J.-C.: PM₁₀ metal concentrations and source identification using positive matrix factorization and wind sectoring in a French industrial zone, *Atmos. Res.*, 96(4), 612–625, doi:<http://dx.doi.org/10.1016/j.atmosres.2010.02.008>, 2010.
- Amato, F., Alastuey, A., Karanasiou, A., Lucarelli, F., Nava, S., Calzolari, G., Severi, M., Becagli, S., Gianelle, V. L., Colombi, C., Alves, C., Custódio, D., Nunes, T., Cerqueira, M., Pio, C., Eleftheriadis, K., Diapouli, E., Reche, C., Minguillón, M. C., Manousakas, M.-I., Maggos, T., Vratolis, S., Harrison, R. M. and Querol, X.: AIRUSE-LIFE+: a harmonized PM speciation and source apportionment in five southern European cities, *Atmos. Chem. Phys.*, 16(5), 3289–3309, doi:10.5194/acp-16-3289-2016, 2016.
- Atmo Nouvelle Aquitaine, 2016. Etude de la composition chimique et des sources de particules sur le centre ville de Poitiers. Rapprt d'étude 2015-2016. http://www.atmo-nouvelleaquitaine.org/sites/aq/files/atoms/files/rapportatmona_fix_int_14_022_caracterisationparticulespoitiers_2015_2016_versionfinale_2017_01_04.pdf
- Belis, C. A., Larsen, B. R., Amato, F., Haddad, I. El, Favez, O., Harrison, R. M., Hopke, P. K., Nava, S., Paatero, P., Prévôt, A., Quass, U., Vecchi, R. and Viana, M.: European Guide on Air Pollution Source Apportionment with Receptor Models, Publ. Off. [online] Available from: http://source-apportionment.jrc.ec.europa.eu/Docu/EU_guide_on_SA.pdf, 2014.
- Cavalli, F., Viana, M., Yttri, K. E., Genberg, J. and Putaud, J. P.: Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the EUSAAR protocol, *Atmos. Meas. Tech.*, 3(1), 79–89, doi:10.5194/amt-3-79-2010, 2010.
- Chevrier, F., Brulfert, G., Jezek, I., Mocnik, G., Marchand, N., Jaffrezo, J.-L. and Besombes J.-L. : DECOMBIO - Contribution de la combustion de la biomasse aux PM₁₀ en vallée de l'Arve : mise en place et qualification d'un dispositif de suivi. *Poll Atmos.*, 2016. <http://odel.irevues.inist.fr/pollution-atmospherique/index.php?id=5952>
- Gianini, M. F. D., Fischer, A., Gehrig, R., Ulrich, A., Wichser, A., Piot, C., Besombes, J. L. and Hueglin, C.: Comparative source apportionment of PM₁₀ in Switzerland for 2008/2009 and 1998/1999 by Positive Matrix Factorisation, *Atmos. Environ.*, 54(0), 149–158, doi:<http://dx.doi.org/10.1016/j.atmosenv.2012.02.036>, 2012.
- Jaffrezo, J. L., Aymoz, G., Delaval, C. and Cozic, J.: Seasonal variations of the water soluble organic carbon mass fraction of aerosol in two valleys of the French Alps, *Atmos. Chem. Phys.*, 5(10), 2809–2821, doi:10.5194/acp-5-2809-2005, 2005.
- LCSQA, 2012a. Synthèse des travaux 2012 du programme CARA. Rapport LCSQA 2012. <http://www.lcsqa.org/rapport/2012/ineris/metrologie-particules-pm10-pm25-caracterisation-chimique-particules-synthese-tra>
- LCSQA, 2012b. Mise en oeuvre d'un « modèle récepteur » de type PMF (« Positive Matrix Factorization ») pour l'identification et la quantification des principales sources de PM₁₀ sur le site urbain de fond de Lens (Atmo Nord Pas de Calais). Rapport LCSQA 2012. <http://www.lcsqa.org/rapport/2012/ineris/mise-oeuvre-modele-recepteur-type-pmf-positive-matrix-factorization-identificati>

- LCSQA, 2013a. Synthèse des travaux 2013 du programme CARA. Rapport LCSQA 2013. <http://www.lcsqa.org/rapport/2013/ineris/synthese-travaux-2013-programme-cara>
- LCSQA, 2013b. Mise en oeuvre d'une méthodologie d'estimation des sources de HAP par modèle récepteur. Application de la Positive Matrix Factorization (PMF). Rapport LCSQA 2013. <http://www.lcsqa.org/rapport/2013/ineris/mise-oeuvre-methodologie-estimation-sources-hap-modele-recepteur-application-pos>
- LCSQA 2015a. Bilan des travaux 2014-2015 du programme CARA. Rapport LCSQA 2015. <http://www.lcsqa.org/rapport/2013/ineris/synthese-travaux-2013-programme-cara>
- LCSQA, 2015b. Impact de la combustion de biomasse sur les concentrations de PM₁₀ dans 10 agglomérations du programme CARA au cours de l'hiver 2014-15. Rapport LCSQA 2015. <https://www.lcsqa.org/rapport/2015/ineris/impact-combustion-biomasse-concentrations-pm10-programme-cara-hiver-2014-2015>
- Oliveira D.: Identification of the main sources and geographical origins of PM₁₀ in the Northern part of France. Manuscrit de Thèse de Doctorat de l'université de Lille 1. Thèse soutenue en janvier 2017. <https://www.theses.fr/201433516>
- Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, *Environmetrics*, 5(2), 111–126, doi:10.1002/env.3170050203, 1994.
- Polissar, A. V., Hopke, P. K., Paatero, P., Malm, W. C. and Sisler, J. F.: Atmospheric aerosol over Alaska: 2. Elemental composition and sources, *J. Geophys. Res. Atmos.*, 103(D15), 19045–19057, doi:10.1029/98JD01212, 1998.
- Sciare, J., d'Argouges, O., Sarda-Estève, R., Gaimoz, C., Dolgorouky, C., Bonnaire, N., Favez, O., Bonsang, B. and Gros, V.: Large contribution of water-insoluble secondary organic aerosols in the region of Paris (France) during wintertime, *J. Geophys. Res. Atmos.*, 116(D22), n/a-n/a, doi:10.1029/2011JD015756, 2011.
- Sylvestre, A., Mizzi, A., Mathiot, S., Masson, F., Jaffrezo, J.-L., Dron, J., Mesbah, B., Wortham, H. and Marchand, N.: Comprehensive chemical characterization of industrial PM_{2.5} from steel industry activities. *Atmos. Environ.*, 152, 180-190, 2017.
- Waked, A., Favez, O., Alleman, L. Y., Piot, C., Petit, J. E., Delaunay, T., Verlinden, E., Golly, B., Besombes, J. L., Jaffrezo, J. L. and Leoz-Garziandia, E.: Source apportionment of PM₁₀ in a north-western Europe regional urban background site (Lens, France) using positive matrix factorization and including primary biogenic emissions, *Atmos. Chem. Phys.*, 14(7), 3325–3346, doi:10.5194/acp-14-3325-2014, 2014.

8. LISTE DES ANNEXES

Annexes	titre
Annexe 1	Méthodologies de caractérisation chimique des PM ₁₀

ANNEXE 1

Méthodologies de caractérisation chimique des PM₁₀

Measured species	Analytical method	References
<u>Carbone organique et élémentaire</u> OC, EC	TOT (analyse thermo-optique en transmission) à l'aide du Sunset lab analyzer (EUSAAR2 protocol)	(Cavalli et al., 2010)
<u>Espèces solubles majeures</u> SO ₄ ²⁻ , NO ₃ ⁻ , Cl ⁻ , NH ₄ ⁺ , Na ⁺ , K ⁺ , Ca ²⁺ , Mg ²⁺ , MSA, oxalate	IC (Chromatographie ionique)	(Jaffrezo et al., 2005)
<u>Métaux et metalloïdes</u> Al, As, Ba, Ca, Cd, Ce, Co, Cr, Cs, Cu, Fe, K, La, Li, Mg, Mn, Mo, Na, Ni, Pb, Pd, Pt, Rb, Sb, Sc, Se, Sn, Sr, Ti, Tl, V, Zn, Zr	ICP-MS (inductively coupled plasma mass spectrometry) et ICP-AES (inductively coupled atomic emission spectrometry) après digestion acide	(Alleman et al., 2010; Waked et al., 2014)
<u>Sucres</u> - levoglucosan, mannosan, galactosan - arabitool, mannitol, sorbitol	HPLC-PAD (high performance liquid chromatography coupled to a pulsed amperometric detection)	(Sciare et al., 2011; Waked et al., 2014)



direction et secrétariat du LCSQA

INERIS - parc technologique Alata - BP 2 - F60550 Verneuil-en-Halatte
tél. 03 44 55 64 04 - www.lcsqa.org