

Identification des sources de PM₁₀ sur un site sidérurgique par des méthodes statistiques

PM₁₀ source identification for a steelwork site using statistical methods

Anis KHLAIFI*, Anda IONESCU*, Yves CANDAU*

Résumé

L'objectif de cette étude est l'identification des principales sources contribuant aux niveaux environnementaux de poussières en suspension PM₁₀ à partir des analyses des métaux trace effectuées en un point d'échantillonnage localisé à l'intérieur d'une usine sidérurgique. Selon les informations relatives à ce site, l'une des principales sources est le parc à scories, pour laquelle nous disposons d'une analyse chimique, qu'on utilisera dans un but de validation. Une trentaine d'échantillons a été prélevée sur le site et a été analysée pour 18 métaux trace. L'utilisation de la méthode « récepteur » PMF n'a pas permis d'identifier avec certitude un seul profil qui serait attribué à la source parc à scories ; il en résulte plusieurs profils qui peuvent être associés aux différentes opérations effectuées dans le parc à scories. En dehors des résultats obtenus, ce papier se propose de mettre en évidence les difficultés engendrées par la conception de cette campagne de mesure, afin de permettre une réflexion lors de la conception de nouvelles campagnes expérimentales

Mots clés

Identification des sources. PM₁₀. Industrie sidérurgique. Signature. Modèles récepteur. PMF. ACP.

Abstract

The aim of this study is the identification of the main sources of PM₁₀, based on some analyses of trace metals performed in a sampling point inside a steelwork site. It is known that one of the main sources is the slag yard. A chemical analysis is available for this source, giving a sort of fingerprint, which has been used in a validation purpose. Twenty eight samples were sampled on the steelwork site and they were analysed for 18 trace metals. The PMF receptor model did not permit to assign with certainty a precise profile to the slag yard; there are several profiles that can be assigned to different operations in the slag yard. Moreover the presentation of the specific results, this paper focuses on the difficulties encountered consequently to the measurement campaign design, in order to be avoided for new other experimental campaigns.

Keywords

Source Identification. PM₁₀. Steelwork. Fingerprint. Receptor Models. Positive Matrix Factorisation (PMF). Principal Component Analysis (PCA).

* Centre d'études et de recherche en thermique, environnement et systèmes – Université Paris Est Créteil – 61, avenue du Général de Gaulle – 94010 Créteil Cedex – E-mail : khlafi.anis@gmail.com, ionescu@u-pec.fr, candau@u-pec.fr

1. Introduction

Cette étude a été réalisée dans le cadre d'un projet européen, « *Tracing of Airborne Dust* » [1], ayant comme objectif principal d'estimer la contribution de l'industrie sidérurgique aux niveaux globaux de PM₁₀ mesurés dans son environnement proche. Il s'agit de sources complexes, dont le profil d'émission est inconnu et qui comprend plusieurs espèces. L'identification des sources de particules en suspension PM₁₀ est dans ce cas une identification en aveugle, où on souhaite obtenir des profils ou « signatures » de ces sources.

Une campagne spéciale de collecte d'échantillons dans l'environnement des sources a été mise en place. Les échantillons ont été ensuite analysés pour déterminer les concentrations en certains éléments chimiques, dont la sélection est cruciale pour une identification pertinente des sources. Les concentrations en éléments chimiques pour tous les échantillons constituent une base de données qu'on peut analyser statistiquement, par des méthodes dites « récepteur ». Les modèles de type récepteur sont décrits dans Hopke [2-4]. La méthode spécialement conçue pour une séparation en aveugle des sources dans le domaine de l'environnement est la factorisation matricielle positive (PMF), méthode qui est appliquée dans la présente étude.

2. Présentation du site d'étude

L'usine sidérurgique présente sur le site d'étude est une aciérie électrique (figure 1). La principale source du site est le parc à scories. Les scories du four électrique (Melting) sont transférées vers le parc à scories (Slag yard) appelé aussi « slag » ci-après, où elles sont trempées (Slag Quenching) pour être refroidies, à l'aide des jets d'eau (Water Treatment). Pendant le processus de trempage des scories, l'évaporation de l'eau produit un nuage turbulent de vapeur, entraînant les particules de poussières dans l'air. Les scories humides sont ensuite séchées (Sludge drying). Les scories sont mises en tas en fonction de leur granulométrie (fines et grosses). Afin d'optimiser l'espace occupé par les scories, les tas sont quelquefois déplacés. Les opérations liées aux scories sont considérées comme des sources importantes de poussières ; de plus, ces opérations sont fréquentes. L'autre source importante du site est le four à arc.

3. Données disponibles

3.1 Mesures relatives aux émissions du parc à scories

Un échantillon en provenance du parc à scories a été analysé en utilisant un spectromètre à rayons X,

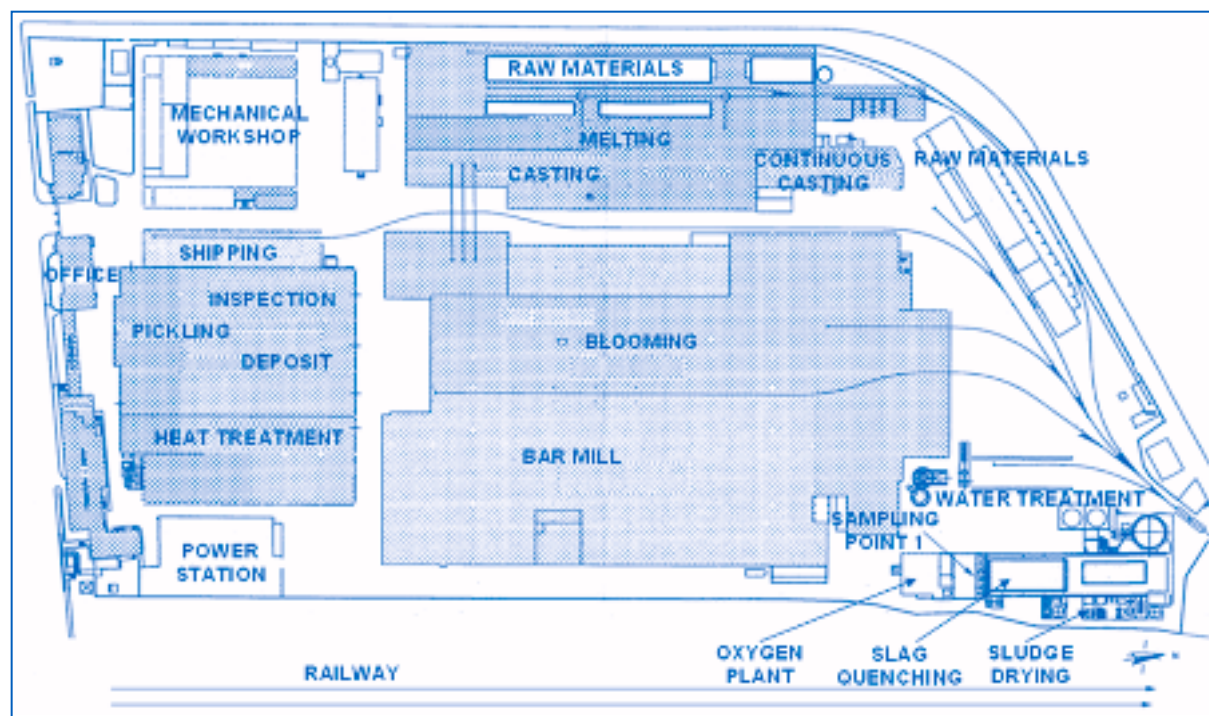


Figure 1.

Schéma du site sidérurgique [1] – *Steelwork site* [1].

Légende : *Raw materials* : matériaux bruts (ferrailles de récupération) ; *Melting* : four électrique à arc, où les ferrailles sont refondues ; *Oxygen plant* : production d'oxygène ; *Casting, Continuous casting* : coulée, coulée continue ; *Blooming* : production de billettes ; *Bar mill* : laminoir ; *Heat treatment* : traitement thermique ; *Slag quenching* : trempage des scories ; *Sludge drying* : séchage des vidanges ; *Pickling* : décapage ; *Inspection* : inspection ; *Deposit* : dépôt ; *Shipping* : transport ; *Railway* : chemin de fer ; *Mechanical workshop* : atelier mécanique ; *Laboratory* : laboratoire ; *Office* : bureaux.

Tableau 1.

Composition moyenne en métaux trace pour les scories prélevées dans le parc à scories obtenue par ICP-MS [1].
Slag mean composition (mass percentage) for the samples exposed in the slag yard, analysed by ICP-MS [1].

	Concentration µg/g																
	Fe	Cr	Mn	Ni	V	Cu	Zn	Co	As	Se	Cd	Sn	Sb	Te	Hg	Tl	Pb
Moyenne	1 785	870	951	14,7	50,2	16,9	8,1	1,2	1,2	5	< 0,1	< 0,1	< 0,1	0,3	< 0,1	< 0,1	0,2
Écart-type	229	106	83	2,0	4,2	4,1	1,0	0,6	0,2	1	—	—	—	0,1	—	—	0,1

par fluorescence. Les résultats de cette analyse mettent en évidence un contenu important de CaO (62 %), composant caractéristique de la sidérurgie secondaire (en aval du four électrique), du SiO₂ (caractéristique de la croûte terrestre (30 %), 5 % de magnésie (MgO), 2 % d'alumine (Al₂O₃), ainsi que d'autres composants, chacun en pourcentages inférieurs à 1 % : K₂O, Na₂O, TiO₂ et S.

Une deuxième analyse des échantillons de scories (préalablement mis en solution) a été effectuée par ICP-MS (*Inductively Coupled Plasma Mass Spectrometry*) (Tableau 1). Pour les métaux Cd, Sn, Sb, Hg, Tl, les concentrations sont en dessous de la limite de détection, soit 0,1 µg.g⁻¹.

3.2. Mesures dans l'environnement (au récepteur)

Des prélèvements de poussières en suspension ont été effectués en un point d'échantillonnage (*Sampling point* sur la Figure 1) situé à 12 m de hauteur et à une distance de 5 m du *slag* (de hauteur 14 m). Un total de 28 échantillons a été analysé en valeurs moyennes journalières, pendant 1 mois : la concentration de PM₁₀ et les concentrations de 18 métaux trace (Al, V, Cr, Mn, Fe, Co, Ni, Zn, Cu, As, Se, Cd, Sn, Sb, Te, Hg, Tl et Pb) ont été obtenues par ICP-MS.

Le point de mesure est supposé être sous le vent dominant des sources de l'usine sidérurgique [1]. Selon les directions de vent enregistrées par la météorologie nationale, le point de mesure ne serait que pendant deux jours sous le vent venant des sources. La station météorologique n'étant pas localisée sur le site sidérurgique, il est possible que les données enregistrées ne soient pas représentatives pour notre site d'étude. Étant donné que les distances entre les différentes sources du site et le point d'échantillonnage sont relativement faibles (le récepteur se trouve sur le site sidérurgique), il peut être contaminé par diffusion par toutes les autres sources autres que le parc à scories.

4. Résultats et interprétations

4.1. Corrélation entre le niveau global PM₁₀ et la masse totale de particules analysées

L'analyse par ICP-MS ne concernait pas la totalité des espèces chimiques présentes sur les filtres, mais seule une sélection de métaux. C'est pour cela qu'on

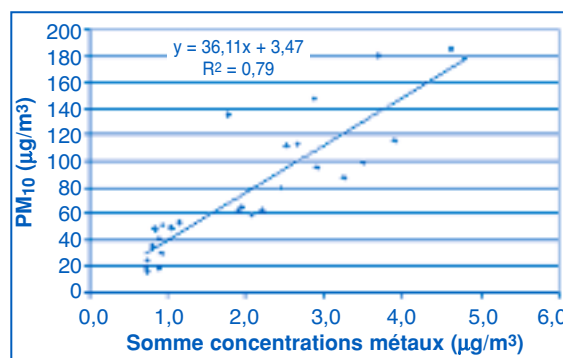


Figure 2.

Concentration de PM₁₀ mesurée et concentration massique totale de métaux trace analysés pour les 28 échantillons prélevés au récepteur.
Measured PM₁₀ concentration versus total mass concentration of analysed trace metals, for the 28 analysed samples at the receptor.

remarque des différences notables entre les PM₁₀ et la masse totale des éléments analysés, pour un même échantillon.

La Figure 2 met en évidence une corrélation linéaire significative (R² = 0,8) entre la masse totale et la masse analysée sur le filtre. La masse totale de PM₁₀ est en moyenne 36 fois plus importante que la masse analysée.

4.2. Analyse par la méthode PMF

La méthode PMF [2-5] est une méthode de factorisation matricielle qui introduit trois traitements spécifiques : (i) un poids individuel à chaque mesure représentant son incertitude ; (ii) une contrainte de non-négativité pour les matrices de profils et de contributions, et (iii) un traitement spécifique des données manquantes ou qui sont inférieures à la limite de détection. Ces traitements permettent ainsi une interprétation physique plus appropriée des résultats.

L'algorithme utilisé est le ME2 [5], cas particulier de l'algorithme général ME ("*Multilinear Engine*") qui représente une formulation multilinéaire de la méthode PMF [5]. L'algorithme spécifique pour un modèle bilinéaire correspond à la factorisation par un produit de deux matrices, et il a été appelé ME2.

4.2.1 Incertitude de mesure et limite de détection

Afin d'évaluer la précision de l'analyse ICP-MS qui a permis d'avoir les mesures au récepteur, un maté-

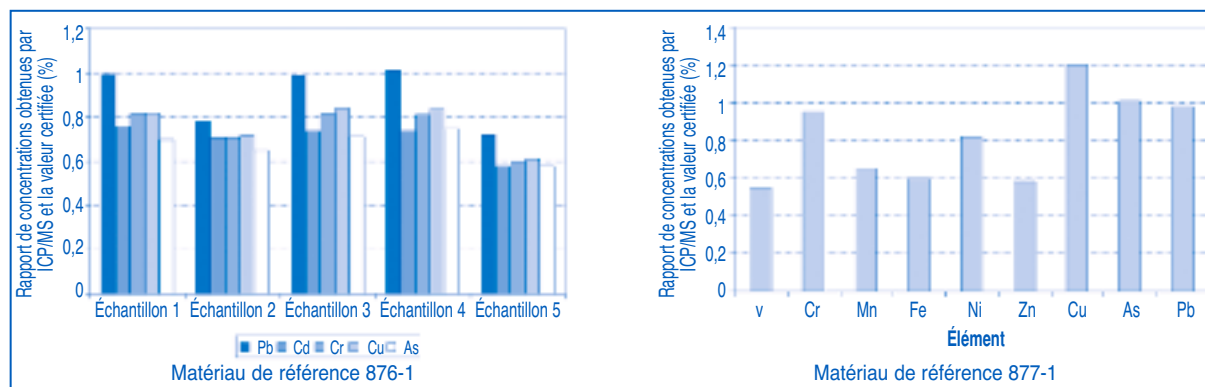


Figure 3.

Concentrations obtenues par ICP-MS pour des matériaux de référence (876-1, 877-1) comparées aux valeurs certifiées correspondant à ces matériaux [1].

Concentrations obtained for reference materials (876-1, 877-1) using the ICP-MS technique compared to the certified values for the same materials [1].

Tableau 2.

Incertitude de mesure pour les différents métaux

(écart entre la valeur obtenue par ICP-MS et la valeur certifiée de l'échantillon de référence, rapportée à la valeur certifiée).

Measurement uncertainty for different metals (difference between the value obtained via ICP-MS and the certified (reference) value of the standard sample, divided by the reference value).

	Éléments																	
	Al	V	Cr	Mn	Fe	Co	Ni	Zn	Cu	As	Se	Cd	Sn	Sb	Te	Hg	Tl	Pb
Incertitudes %	60	45	5	35	40		18	42	20	1		25	2	1				1

Matériau de référence, standard, a été analysé par la même technique, et les concentrations mesurées ont été comparées à celles certifiées [1]. Les deux matériaux standards choisis présentaient des compositions en métaux similaires à la composition des poussières du site du projet.

Les résultats de l'analyse ont été comparés aux valeurs certifiées (Figure 3). Les valeurs obtenues par ICP-MS sont généralement plus basses que les valeurs certifiées, sauf dans le cas du plomb (Pb), dont les valeurs mesurées sont très proches de celles certifiées.

Les sources de la non-conformité pourraient être une erreur systématique résultant du pesage, ou de la dissolution incomplète des échantillons. Certains éléments (V, Cr, Cd, Ni, As) sont mesurés à des concentrations très basses et donc la conformité aux valeurs de référence est difficile à réaliser.

Notons enfin que pour le fer, l'importante incertitude est due en particulier à la difficulté de la détermination de cet élément par ICP-MS.

Suite à cette analyse comparative, on estime l'incertitude de mesure au niveau de chaque métal comme la différence entre la valeur obtenue par ICP-MS et la valeur certifiée, rapportée à la valeur certifiée (Tableau 2).

C'est à partir de ces valeurs relatives d'incertitude que nous allons calculer l'incertitude de chaque mesure ponctuelle, qui sera prise en compte dans le cadre de la méthode PMF.

4.2.2 Choix du nombre de sources

La PMF demande en entrée le nombre de sources à estimer. Si la connaissance de la réalité terrain est suffisante, ceci permettra de répondre à cette question. Sinon, plusieurs tests doivent être faits, avec un nombre différent de sources.

Dans la Figure 4, on présente la variation de la fonction objectif Q en fonction du nombre de sources. Le critère Q représente la somme des carrés des écarts pondérés par les inverses des incertitudes entre les valeurs mesurées et les valeurs reconstituées par la PMF. Pour éviter les problèmes de minima locaux, on fait aussi varier le nombre de points de départ pour la recherche de l'optimum, points qui sont choisis aléatoirement. On remarque qu'en augmentant le nombre de sources, l'écart modèle-mesure diminue de la manière suivante : au début, il y a une forte variation, lorsqu'on passe de trois à quatre, cinq et six sources, ensuite on remarque une inflexion de la courbe : la variation est beaucoup moins importante. Ainsi, selon ces différentes courbes, obtenues pour différentes initialisations, un bon choix du nombre de sources serait de six.

Du point de vue de la qualité de la modélisation, on considère que l'on a une bonne reconstitution des données initiales à partir des sources retenues si la valeur de Q est approximativement égale à N , nombre de points de la matrice de mesures diminué du nombre total d'éléments dans les matrices de profils et de contributions. Ceci revient à considérer que le rapport entre l'erreur de modélisation et l'incertitude de la mesure est de l'ordre de 1.

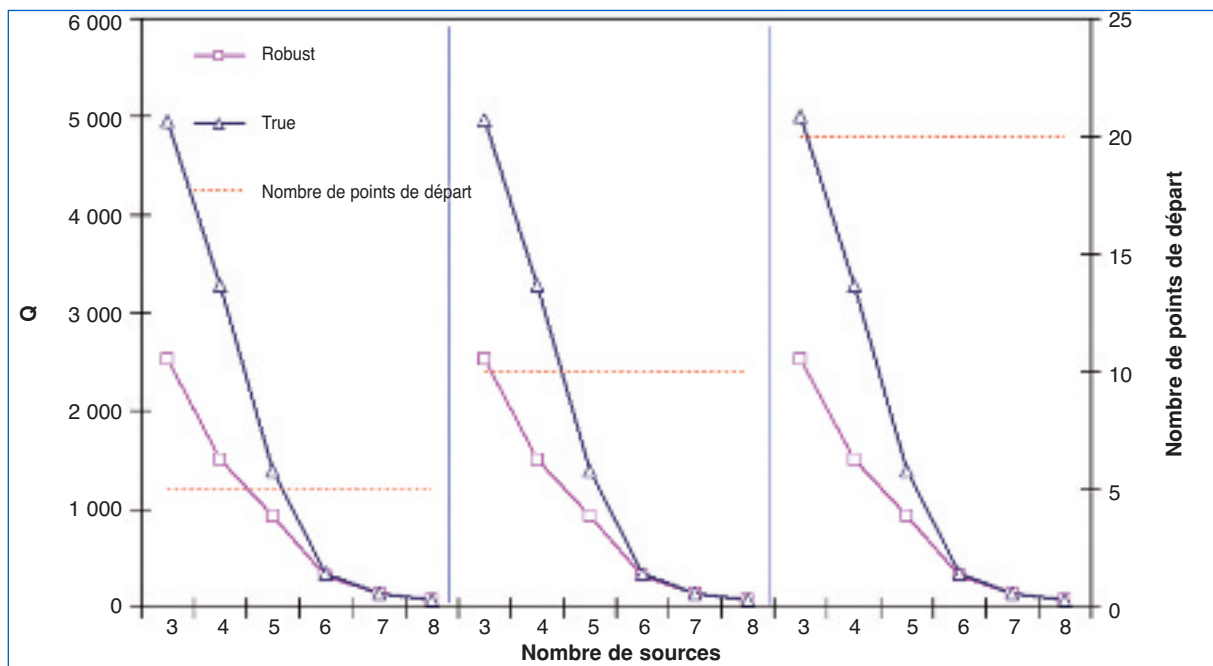


Figure 4.

Choix du nombre optimal de sources dans l'application de la PMF. En abscisse, on présente le nombre de sources et en ordonnée, sur l'échelle de gauche, le critère Q calculé de deux manières : Q true et Q robust [6], et sur l'échelle de droite on représente le nombre de points de départ choisis pour la recherche du minimum de Q.

Choice of the optimal number of sources for the PMF. The abscissa represents the number of sources and the co-ordinate, the Q-criterion, calculated in two ways (Qtrue and Qrobust [6]).

The co-ordinate axis on the right represents the number of initial points for searching the minimum of the Q-criterion.

Dans notre cas (28 échantillons, 18 éléments, 6 sources) $N = 504$, ce qui est en accord avec la valeur de Q obtenue pour le cas de six sources ($N = 444$).

On remarque que la courbe correspondant à Q_{true} se superpose à celle de Q_{robust} , à partir de six sources, où on a aussi l'inflexion de la courbe. On rappelle que la différence entre Q_{true} et Q_{robust} réside dans le fait que pour diminuer l'effet de disproportion des mesures de concentration élevée, on introduit une fonction filtre [6] dans la fonction à optimiser.

4.2.3 Profils et contributions des sources obtenues par PMF

Les profils obtenus par PMF sont normalisés par rapport à la concentration massique de PM_{10} . Ces profils sont obtenus en divisant chaque profil initial par un facteur d'ajustement déterminé en procédant à une régression multiple traduisant le fait que la concentration massique totale de l'échantillon est la somme des contributions de toutes les sources. Les profils normalisés sont adimensionnels variant entre 0 et 1.

Une comparaison des profils obtenus à celui de la seule source connue, le slag, a été tentée.

Une incertitude règne quant à la présence de l'aluminium dans la composition de cette source, car il n'y a pas spécifiquement de valeur mesurée pour ce métal (cf. section 3.2). Par ailleurs, Kuo *et al.* [7] présentent une analyse détaillée des scories et on peut remarquer que le calcium est l'élément le plus

détecté, ensuite le silicium, l'aluminium, le magnésium, le fer etc. Nous remarquons le même ordre d'importance dans le profil du slag mesuré dans notre étude, ce qui nous laisse penser qu'on a plus d'aluminium que de fer.

Dans la suite, on essaiera de comparer la signature du slag aux différents profils obtenus par PMF [5] sans prendre en compte l'aluminium, mais en considérant sa présence non négligeable et plutôt importante dans le profil de cette source.

Les coefficients de corrélation entre les profils obtenus par la PMF, et la signature du slag varient de la manière suivante : 0,82 avec la source 1 ; 0,95 avec la source 2 ; 0,85 avec la source 3 ; 0,80 avec la source 4 ; 0,74 avec la source 5 et 0,83 avec la source 6. On remarque une bonne corrélation entre le profil du slag et tous les profils de la PMF, la meilleure étant avec la source 2 ($r = 0,95$), ensuite la source 3 ($r = 0,85$). Il y a une différence importante entre les profils 2 et 3 : dans la source 2, l'aluminium est l'élément majoritaire, tandis que dans la source 3 il est complètement absent.

On complète l'identification de la source par à scories par une analyse visuelle, en superposant le profil slag mesuré aux différents profils des sources obtenues (Figure 5). Les profils sont présentés sur une échelle logarithmique permettant de mieux visualiser les éléments qui sont en faible proportion.

De manière générale, les profils sont surévalués par rapport au slag, mais il faut noter que dans le cas des profils PMF, on présente le rapport métal/masse

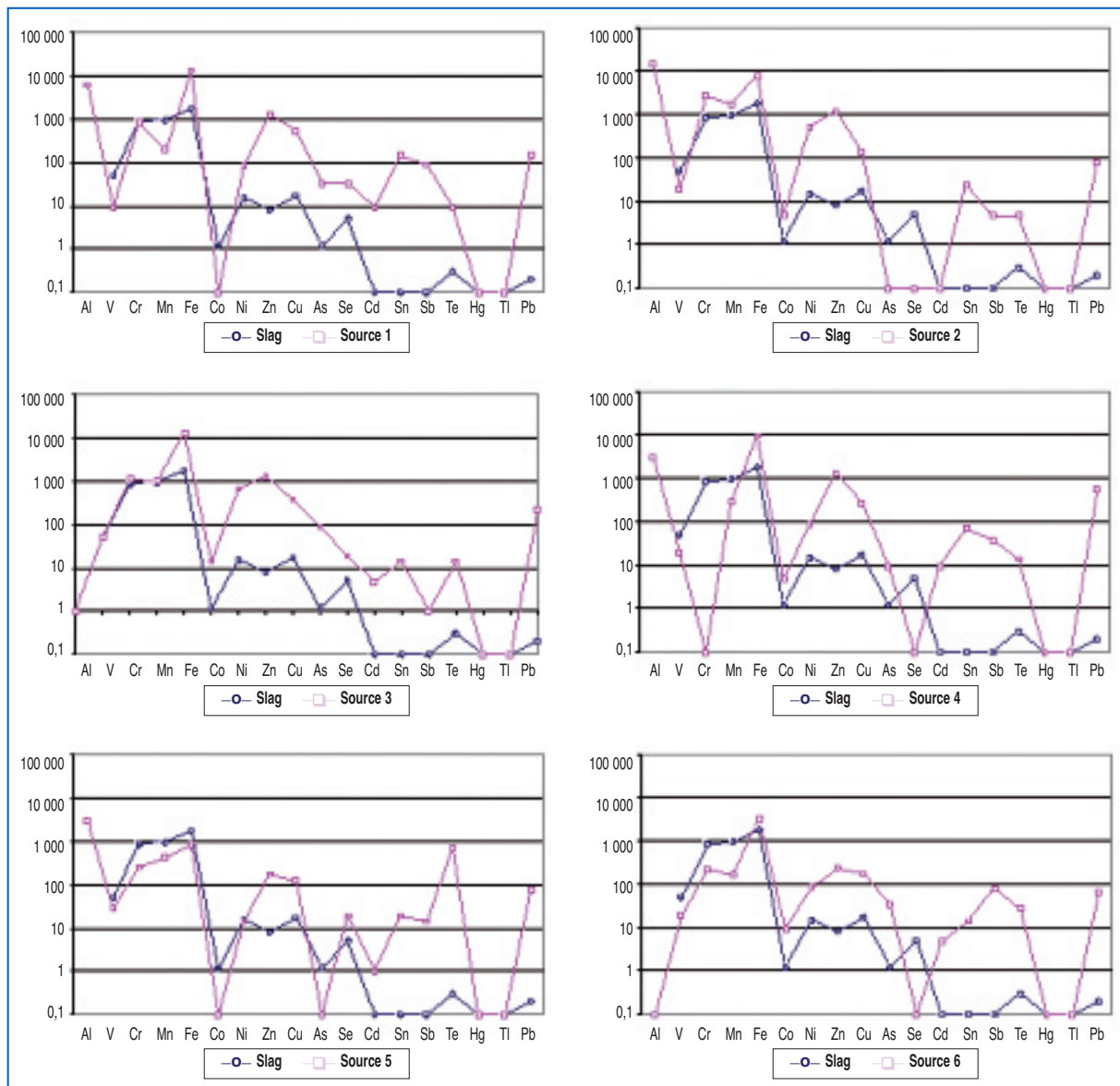


Figure 5.

Analyse comparative entre les profils (masse métal (µg)/masse PM₁₀ (g), échelle logarithmique) obtenus par la PMF après ajustement, et celui du slag mesuré (µg métal/g poussière).

Comparison between the profiles (metal mass (µg)/PM₁₀ mass (g), logarithmic scale) obtained by PMF, and the profile obtained by measurement for the slag source (µg metal/g dust).

de PM₁₀, tandis que pour le slag mesuré le rapport est vraisemblablement donné en fonction de la masse totale des particules (supérieure évidemment au PM₁₀).

Cette surestimation est systématique et d'autant plus importante pour Zn, Cu, Sn, Sb, Hg, Tl et Pb. La présence du plomb est quasi-constante dans tous les profils obtenus par la PMF, ce qui peut être expliqué par un sol contaminé en plomb.

Notons que dans l'analyse du slag mesuré, seuls le Zn, le Cu et le Pb sont en proportions non négligeables, les autres métaux (Sn, Sb, Hg, Tl) étant en dessous de la limite de détection. On note aussi une présence importante de Ni, Zn et Cu dont les proportions sont similaires dans les quatre premiers profils

et en plus, en surestimation toujours par rapport au Slag. Ceci nous laisse penser que ces profils correspondent à des sources combinées.

L'identification du slag parmi les profils obtenus est une tâche assez difficile, nous avons un seul profil slag mesuré, mais il y a plusieurs opérations effectuées sur les scories, dont la composition, de plus, peut différer suivant leur granulométrie.

Dans cette hypothèse, le profil slag mesuré ne représente au mieux qu'un profil moyen correspondant à plusieurs sources.

En analysant les contributions journalières des différentes sources obtenues par la PMF (Figure 6), on note que la source 1 est celle qui a une contribution quasi-constante et importante. Vu que le point de

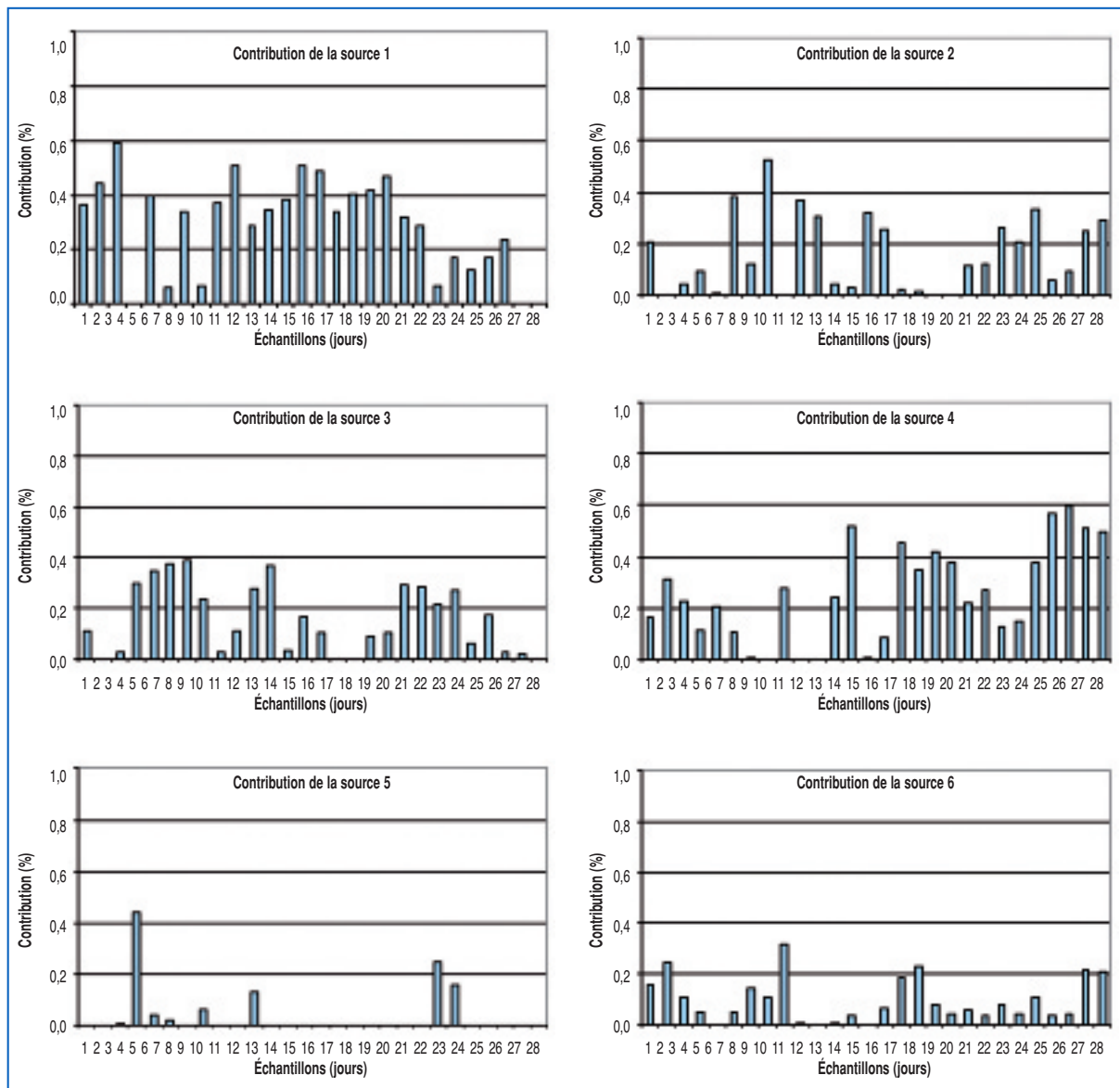


Figure 6.
Contributions journalières des sources obtenues par la PMF aux différents échantillons journaliers.
Daily source contributions obtained by PMF.

mesure est très proche du slag, on peut penser, en effet, que cette source peut être l'une des composantes importantes du slag, dans l'hypothèse où l'aluminium est un élément important de cette source. L'autre hypothèse plausible est que cette source soit le four électrique lui-même, qui fonctionne en continu.

On peut remarquer la contribution quasi-inexistante, sauf quelques épisodes particuliers, de la source 5 et celle très faible de la source 6. Ces sources correspondent à des opérations particulières, ponctuelles dans le parc à scories. On rappelle qu'il y a plusieurs opérations liées aux traitements de scories, comme la trempé des scories, leur séchage ou les déplacements des tas de scories, qui sont rangées par granulométrie (fines et grosses) ; ceci peut justi-

fier d'avoir des profils assez semblables, caractérisant, avec quelques différences, la même source principale : le slag (le parc à scories). Une connaissance précise du calendrier de fonctionnement des sources de ce site pourrait faciliter la discrimination entre les sources, voire les opérations liées à la source principale et le type de granulométrie.

4.3. Analyse par la méthode ACP

Notre hypothèse de retrouver parmi les profils des sources, des profils qui correspondent aux particules fines et des profils caractérisant les particules grosses, mais pour les mêmes sources, peut être soutenue par les résultats obtenus en appliquant une analyse en composantes principales (ACP).

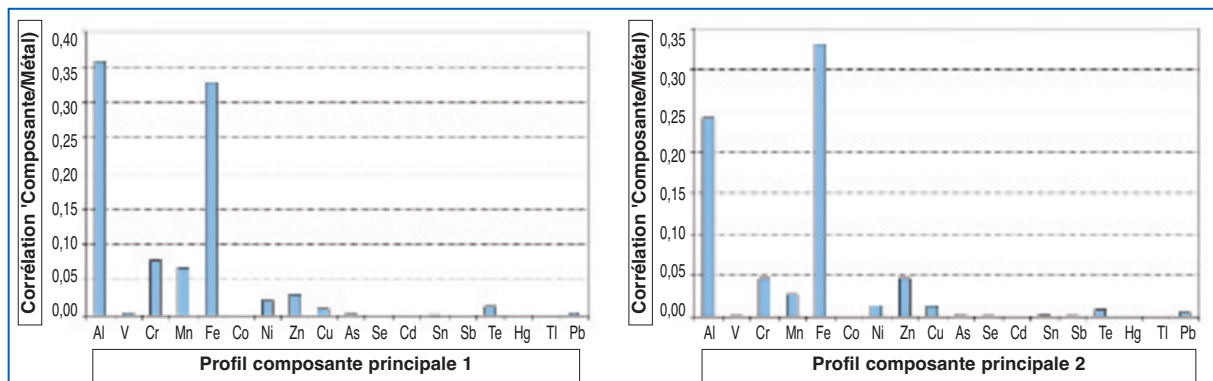


Figure 7.

ACP avec rotation VARIMAX appliquée avec des données brutes:
coefficients de corrélation entre chaque concentration de métal et la composante principale (1 et 2).
PCA/VARIMAX applied to raw data:
correlation coefficients between each metal concentration and the principal component (1 and 2).

L'ACP [8] a été appliquée aux données brutes, contrairement au cas usuel où elles sont centrées et réduites, les variables étant les concentrations des métaux mesurés sur les différents échantillons.

Les deux premières composantes obtenues expliquent 99 % de l'information dans les échantillons (la première, 90 % et la deuxième 9 %).

Les coefficients de corrélation entre chaque variable métal et la composante principale sont présentés sur la Figure 7.

Comme on pouvait s'y attendre, les métaux les plus significatifs mis en évidence par cette analyse sur des données brutes sont ceux qui se trouvent en plus grande quantité : l'aluminium et le fer.

On note aussi que les deux premières composantes, les plus significatives présentent des profils semblables.

Ces profils ne peuvent pas être comparés à l'analyse quantitative du slag, bien qu'on aurait un coefficient de corrélation de 0,93 entre la composante 1 et le slag, mais qualitativement on voit que le profil est très semblable.

Deux interprétations différentes seraient possibles. La première interprétation consiste à identifier une composante au slag (on retrouve bien les mêmes éléments, l'ordre de grandeur est respecté mais pas les proportions) et l'autre composante, à une autre source importante du site : le four à arc. La deuxième interprétation consisterait à considérer une composante comme une combinaison de sources : la composante 1 pourrait être associée aux grosses particules et la composante 2 aux fines.

Enfin, on peut faire une combinaison de ces deux interprétations, la première composante, des particules plus grosses correspondant au slag, la deuxième, des particules plus fines ayant pu s'échapper du filtre du four à arc. Dans cette lumière, les différentes sources obtenues par la PMF (plus nombreuses) deviennent, pour la plupart, des composantes du slag.

5. Conclusion

Ce papier présente l'analyse d'un cas réel d'étude d'identification des sources de pollution sur un site sidérurgique. Ce sont plutôt des composantes d'un processus industriel ayant lieu sur le site sidérurgique que des sources différentes qu'on peut identifier. Les éléments connus sur le fonctionnement du site, ainsi que la nature peu différente des sources présentes sur le site ne permettent pas une identification suffisamment fiable des sources obtenues.

L'identification du slag parmi les profils obtenus par PMF est une tâche assez difficile, nous avons un seul profil slag mesuré, mais il y a plusieurs opérations effectuées sur les scories, dont la composition, de plus, peut différer suivant leur granulométrie. Dans cette hypothèse, le profil slag mesuré ne représente au mieux qu'un profil moyen correspondant à plusieurs sources.

Par contre, une analyse rétrospective de l'expérience permet de mettre en évidence quelques éléments de réflexion à prendre en compte lors de la conception des nouvelles campagnes de mesures dans le but d'identification des sources.

La première question qu'on se pose est de savoir si l'ensemble d'éléments traceurs choisi est suffisant pour remonter aux sources de ce site. La régression linéaire entre la masse totale de PM₁₀ et la concentration totale d'éléments analysés est caractérisée par un coefficient de détermination de 0,8 qui soutient la proportionnalité entre la masse totale de PM₁₀ et celle de métaux analysés. Par contre, le choix des éléments retenus n'a pas un pouvoir vraiment discriminant, car ils sont présents, pour la plupart, dans toutes les sources du site.

Le point d'échantillonnage a été placé de façon à recevoir les particules provenant des sources sidérurgiques, par rapport au vent dominant. Or, selon les données météorologiques fournies pour ce site, le point d'échantillonnage n'a été sous le vent provenant de l'usine sidérurgique que pendant deux jours. Ceci

soulève la deuxième question, qui est celle de la représentativité de l'emplacement du point d'échantillonnage ou de la représentativité des données météorologiques, relevées sur un site météorologique, assez éloigné du site sidérurgique. En même temps, vu la proximité du point d'échantillonnage par rapport à toutes les sources du site, elles peuvent toutes y contribuer, par une simple diffusion.

La comparaison des profils obtenus par PMF et celui du slag mesuré ne permet pas une identification certaine du slag. Par contre, l'analyse des contributions pourrait donner des renseignements complémentaires très utiles si ces contributions sont corrélées au calendrier de fonctionnement des diverses sources ainsi qu'au vent mesuré sur le site.

La reconstitution de données initiales à partir des profils obtenus par PMF est bonne avec un nombre acceptable de sources. Il n'y a pas de nette différenciation entre les profils obtenus : on retrouve une certaine redondance dont on peut supposer que les raisons sont les suivantes : plusieurs opérations relatives à la même « source » ; des granulométries différentes pour des signatures très similaires. Dans ce cas il serait plus approprié d'identifier les sources à partir de leurs contributions au niveau global de poussières interprétées par rapport à la direction du vent et au calendrier (horaire) de fonctionnement des sources (informations manquantes dans notre analyse).

Il faudrait également mentionner la présence d'un certain nombre d'éléments de concentration en dessous de la limite de détection. On pourrait peut-être obtenir des meilleurs résultats en éliminant ces éléments ou trouver un traitement plus adapté.

Un dernier point important à signaler concernant les données d'origine est le nombre d'échantillons analysés (28), qui est très faible. Il est évident qu'un nombre plus important d'échantillons (de l'ordre de 100) conduirait à des résultats plus robustes statistiquement.

L'ACP appliquée sur les données brutes a mis en évidence deux profils très similaires pouvant correspondre à des profils moyens de types différents de particules. La deuxième interprétation serait de considérer la première composante le slag et la deuxième le four à arc.

Cette présentation n'a pas comme unique but de présenter les résultats obtenus pour l'identification des sources sur un site donné, mais aussi de mettre en évidence et discuter les principales difficultés rencontrées, pour que certaines d'entre elles puissent être évitées avant de concevoir l'expérience. En effet, ce travail nous a permis de mettre en évidence certains points faibles des expériences menées ou des questions qui nécessitent encore une réflexion, ce qu'on ne trouve malheureusement pas souvent dans les études publiées. On les résume ci-dessous :

- représentativité d'espèces sélectionnées (élément crucial) ;
- représentativité du point d'échantillonnage ;
- nombre d'échantillons suffisant ;
- connaissance des informations exogènes : données météorologiques (en particulier le vent, qui devrait être mesuré sur le site d'étude), des connaissances sur l'horaire de fonctionnement des sources.

Il est évident que si les données initiales n'ont pas un potentiel informatif suffisant, quel que soit le traitement appliqué les résultats ne seront pas satisfaisants.

Remerciements

Les auteurs tiennent à remercier les collègues participants du projet européen TOAD qui a financé cette étude, et tout particulièrement : P. Le Louër, F. Berho, J. Hodges et D. Harvey pour les discussions très intéressantes.

References

- Harvey D, Horne A, Hodges J, Haines A *et al.* Tracing of Airborne Dust, Technical Report EUR 22457, European Commission, Technical steel research, Brussels 2007.
- Hopke PK. Factor analysis. Receptor modelling in environmental chemistry. New York, John Wiley 1985 : 155-97.
- Hopke PK. An introduction to receptor modelling. Receptor modelling for air quality management. Amsterdam, Elsevier 1991 : 1-10.
- Hopke PK. The application of receptor modeling to air quality data. *Pollution Atmosphérique* 2010.
- Paatero P. The Multilinear Engine- A table-driven, least-squares program for solving multilinear problems, including the n-way parallel factor analysis model. *Journal of Computational and Graphical Statistics* 1999 ; 8(4) : 854-88.
- Lee E, Chan CK, Paatero P. Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. *Atmospheric Environment* 1999 ; 33 : 3201-12.
- Kuo Y, Lin T, Tsai P. Metal behavior during vitrification of incinerator ash in a coke bed furnace. *J Hazard Mater* 2004 ; 109 (1-3) : 79-84.
- Saporta G. Probabilités, analyse des données et statistique. Editions Technip 1990.

