Utilisation de méthodes multivariées pour caractériser et optimiser un réseau de surveillance de la qualité de l'air

Use of multivariate analyses for the characterization and optimization of air quality monitoring network

Hervé PLAISANCE* (Docteur ès Sciences), Jean-Claude GALLOO (Docteur ès Sciences, René GUILLERMO (Professeur)

RÉSUMÉ

Deux techniques multivariées, la classification ascendante hiérarchique et le graphique en étoile, ont été utilisées pour examiner les niveaux en NO et NO_2 obtenus aux 14 stations du réseau de surveillance de la qualité de l'air de l'agglomération lilloise. Cette approche statistique a permis d'identifier, par la construction d'une classification, deux groupes homogènes de distribution de NO séparant les sites exposés à une pollution de proximité due aux transports de ceux non influencés par ce même type de pollution. Les distributions de NO_2 se répartissent dans cette même analyse en trois classes, correspondant à des degrés différents d'exposition des stations vis-à-vis de la pollution de fond induite par les transports.

Ce classement des stations obtenu par l'analyse des distributions de NO et de NO_2 est globalement en accord avec nos connaissances sur l'environnement des sites de mesures. Il a permis de bien définir le type d'influence « transport » subi par les différentes stations du réseau et de reconsidérer, pour certaines d'entre elles, leurs objectifs de mesures. Avec l'appui d'un outil de prospection, tel qu'un laboratoire mobile, cette analyse par classification peut également servir de support à l'optimisation du dispositif de surveillance du réseau.

ABSTRACT

Two multivariate analyses, hierarchical clustering analysis and star symbol plot, have been applied to examine the NO and NO_2 levels on the fourteen stations of air quality network of Lille. Two groups of NO distributions have been identified by this statistical approach : first group, the sites exposed to the the pollution of proximity generated by the transports and second one, the sites non influenced by this same type of pollution. The NO distributions are divided among three classes, corresponding to three different degrees of exposure of stations beside the background pollution induced by the transports.

This classification of stations obtained by the analysis of their NO and NO_2 distributions are globally in agreement with our knowledges on the environment of sites. It has allowed to well define the type of "transport" influence undergone by the different stations and to reconsider, for some stations, their measurement target. With the support of a mobile laboratory, these analysis can be also useful to the optimization of network design. A method of optimization is proposed in this article.

La répétition récente, en été et automne 1997, des pics d'ozone et d'oxydes d'azote enregistrés dans plusieurs villes de l'hexagone désigne le transport comme cause principale de la pollution atmosphérique en zone urbaine. Avec l'adoption récente de la loi sur l'air et l'utilisation rationnelle de l'énergie [1] et les futures directives européennes, on s'oriente vers plus de surveillance et donc vers un accroissement du parc des analyseurs. Ces moyens supplémentaires amènent les gestionnaires de réseaux à se poser des questions sur la résolution spatiale optimale du dispositif de mesures, c'est-à-dire, le nombre et la localisation des sites adéquats pour estimer le plus justement possible la pollution atmosphérique sur une agglomération. La conception même d'un réseau dépend bien évidemment d'une stratégie de surveillance fixée par le gestionnaire qui peut répondre à divers objectifs de mesures. Les principaux sont :

– l'évaluation en continu de la qualité de l'air dans l'agglomération ;

* *École des Mines de Douai, Département Chimie et Environnement, 941 rue Charles Bourseul, BP 838, 59508 Douai, France.*

– la mesure de la pollution de fond sur l'ensemble de la zone concernée ;

– l'estimation de la part des différents types de source de pollution ;

– l'obtention de données pour valider et calibrer des modèles en vue de réaliser une prévision de la pollution ;

– la surveillance des émetteurs industriels présents dans la zone à surveiller ;

– l'évaluation des risques d'exposition des habitants et plus spécifiquement des zones sensibles (écoles, hôpitaux...).

Les coûts importants d'acquisition, de fonctionnement et de maintenance du matériel de mesure rendent difficile pour un gestionnaire de concevoir un réseau qui puisse satisfaire à l'ensemble de ces objectifs. Sa stratégie de surveillance va donc prendre en compte certains de ces objectifs jugés prioritaires au regard des caractéristiques propres de la zone à surveiller (présence ou non de forts émetteurs, topographie particulière, conditions météorologiques...).

Dans les réseaux, le choix des sites de mesure se fait à partir d'une connaissance approfondie du terrain et sur le concept de représentativité d'un espace ou d'une population exposée. Quelques travaux [2-4] développent une voie complémentaire en proposant l'utilisation d'outils d'exploitation des données dans le but d'optimiser l'emplacement des sites de mesures préexistants. Cette démarche porte pour l'essentiel sur l'analyse des séries antérieures de données obtenues aux différentes stations [3,5]. Divers paramètres statistiques, tels que le coefficient de corrélation linéaire, le pourcentage de valeurs identiques ou proches dans les séries de données, et les tests de comparaison des distributions sont utilisés dans ces études afin de bien caractériser les niveaux de pollution des stations et d'évaluer leur degré de similitude.

Malgré la diversité et le volume important des données à traiter, peu d'analyses multidimensionnelles sont utilisées dans ces méthodologies d'optimisation. Les travaux menés dans cette étude proposent justement de classer les niveaux de pollution en NO et NO_2 des stations du réseau lillois géré par l'association AREMA-LRT, à l'aide de deux techniques multivariées complémentaires : le graphique en étoile et une classification ascendante hiérarchique. À partir du calcul des différents percentiles obtenus sur les données de trois années de mesures, un profil de pollution en NO et NO_2 est établi pour chaque station. Une image multivariée, appelée graphique en étoile, va permettre de visualiser ces différents profils de pollution. Une classification ascendante hiérarchique va ensuite répartir les sites en différentes classes de profils de pollution ressemblants. Les caractéristiques de ces classes sont discutées de manière à révéler les liens dominants existant entre l'environnement du site et son appartenance à une classe de profils de pollution.

Sur les bases de ce classement des sites, sera proposée, en conclusion, une démarche générale d'optimisation du réseau de mesures.

Caractéristiques du lieu d'étude

Le réseau AREMA-LRT est censé couvrir le vaste territoire de la communauté urbaine de Lille (Figure 1) qui se définit comme une entité multipolaire de 87 communes et de 1,1 millions d'habitants [6]. Elle s'étend sur 611 km^2, avec une densité moyenne de population de 1 746 habitants/km^2. Son tissu urbain est surtout caractérisé par un habitat dense au niveau des centres-ville des trois principales agglomérations de la zone : Lille, Roubaix et Tourcoing. Ces trois ensembles resserrés sont entourés par un tissu urbain moins dense comprenant de nombreuses voies rapides (grandes avenues, rocades...) et de grands axes routiers à forte circulation (autoroutes et nationales), reliant Lille aux principales villes d'Europe du Nord. On compte six grandes voies routières de liaison, cinq autoroutes (A1, A25, A23, A22 et A27) et une nationale (N41) qui totalisaient plus de 250 000 véhicules par jour en 1996 avec une forte proportion de poids lourds (environ 10 %), dénotant l'importance du transport routier de passage dans cette zone urbaine. Notons également que ces grandes voies de circulation sont surtout situées au sud de la communauté urbaine de Lille ; l'autoroute A1 en est la principale puisqu'elle draine à elle seule près de 90 000 véhicules par jour. La circulation correspondant aux trajets travail-domicile est aussi en pleine progression dans cette zone en raison du fort développement périurbain. Le nombre de véhicules aux principales entrées et sorties de la seule ville de Lille est évalué approximativement à 370 000 véhicules par jour [7].

La région présente néanmoins une topographie peu accidentée et des vents très fréquents de secteurs ouest à sud-ouest qui confèrent à l'atmosphère de cette zone une bonne faculté de dispersion des polluants [8].

Description des sites de mesure et base des données

Durant la période d'étude considérée, allant du 1er janvier 1994 au 31 décembre 1996, le réseau AREMA-LRT comprenait 14 stations de mesures dotées d'analyseurs à chimiluminescence mesurant en continu les NO_x. La figure 1 présente également le positionnement géographique de ces 14 points de mesures. Ces stations ont des degrés d'exposition différents vis-à-vis des trois grands types de sources pouvant influencer les mesures de NO_x, à savoir l'habitat, le trans-

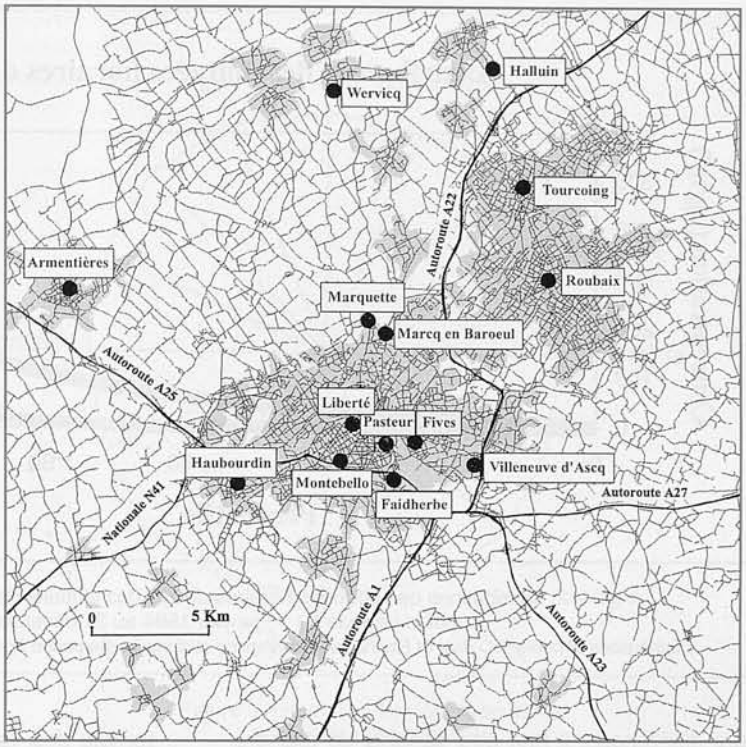


Figure 1. Implantation des 14 stations de mesure des NO_x du réseau AREMA-LRT et principales voies de circulation de la communauté urbaine de Lille.

Locations of fourteen NO_x stations of AREMA-LRT network and main roads of Urban Community of Lille.

port et l'industrie. Sur les bases d'une enquête menée par J.L. Houdret [9] auprès des réseaux français, il a été possible d'établir un premier classement des stations en plusieurs groupes : urbain, transport, industriel ou mixte. La définition de ces groupes, ainsi que le résultat du classement pour les stations AREMA-LRT sont présentés dans le tableau 1. Ce classement reflète à la fois la connaissance que les membres du réseau ont de l'environnement des sites et les choix stratégiques qui ont conduit initialement à l'implantation de ces stations. D'après ce classement, il semble que le réseau AREMA-LRT soit composé de stations possédant des caractéristiques environnementales très diverses, ce qui témoigne de la multiplicité des objectifs de mesure de ce réseau. Ce classement des stations sera comparé à celui obtenu par l'analyse statistique des données.

L'investigation réalisée dans cette étude porte sur l'analyse comparative des distributions calculées à partir des moyennes horaires des concentrations de NO et NO_2 mesurées à ces 14 stations, du 1er janvier 1994 au 31 décembre 1996. Une distribution est formée d'un ensemble de percentiles successifs qui caractérise le niveau de concentration d'un polluant mesuré sur un site. La figure 2 présente, à titre d'exemple, les distributions des données horaires de NO obtenues pour deux stations du réseau AREMA-LRT : Faidherbe et Wervicq. Pour apprécier le plus justement possible

Tableau 1. Description de l'environnement des stations NO_x.

Description of the environment of NO_x stations.

Station	Type d'influence
1 - Marquette	IU
2 - Roubaix	UF
5 - Liberté	TUP
11 - Haubourdin	IU
13 - Villeneuve d'Ascq	UF
16 - Marcq-en-Barœul	TUPI
17 - Wervicq	IR
19 - Pasteur	TUP
22 - Armentières	UF
23 - Montebello	UF
24 - Faidherbe	TUP
25 - Fives	UF
26 - Halluin	IR
28 - Tourcoing	UF

IR : Industrielle Rurale - Site rural influencé par une pollution d'origine industrielle.
IU : Industrielle Urbaine - Site urbain influencé par une pollution venant d'émetteurs industriels.
TUP : Transport, Urbain de Pointe - Site subissant des influences de proximité liées aux transports et à l'habitat.
TUPI : Transport, Urbain de Pointe, Industriel - Site influencé à la fois par une pollution de proximité due aux transports et à l'habitat et par des apports venant de l'industrie.
UF : Urbaine de Fond - Site influencé uniquement par une pollution de fond venant de l'habitat.

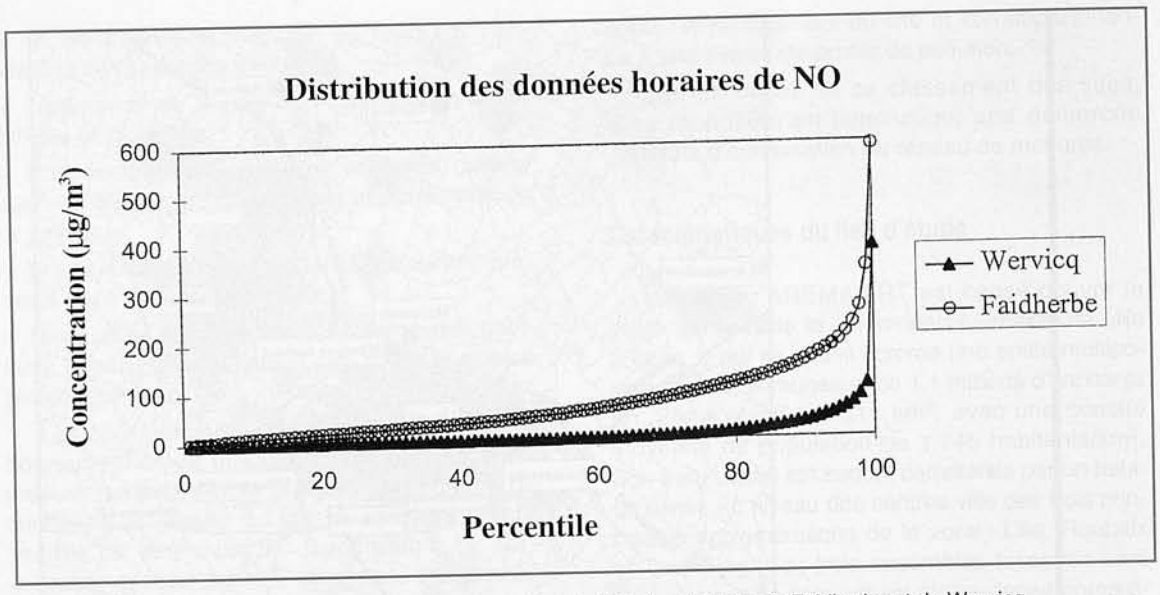


Figure 2. Distributions des données horaires de NO des stations de Faidherbe et de Wervicq pour la période du 1er janvier 1994 au 31 décembre 1996.
Distributions of hourly NO data of Faidherbe and Wervicq stations for the period from january 1st to december 31th,1996.

leurs différences, il faut être capable de prendre en compte les multiples écarts de percentiles trouvés entre ces distributions.

Les techniques multivariées utilisées dans ces travaux permettent la comparaison d'éléments multidimensionnels, tels que des distributions décrites par plusieurs percentiles. Dans notre analyse statistique, nous avons choisi de retenir pour chaque distribution les valeurs de 40 percentiles allant par pas de 2 du 20e au 98e percentile. La partie de la courbe de distribution située en dessous du 20e percentile n'est pas prise en compte dans l'analyse (Figure 2), car elle est jugée trop dépendante de mesures proches de la limite de quantification de l'appareil et présente en outre peu d'intérêt pour distinguer des différences entre les distributions.

Méthodes statistiques utilisées

Graphique en étoile

Le graphique en étoile est un outil de visualisation des tendances et ressemblances entre les différents individus, décrit par plusieurs variables (Figure 3). Dans notre cas, une étoile représente « le profil de pollution » d'une station, pour un composé donné (NO ou NO_2), formé à partir des 40 percentiles retenus de la courbe de distribution (compris entre le 20e et le 98e). Chaque étoile est constituée d'une série de raies partant d'un point central. Une raie est la représentation d'une variable (percentile) pour un individu donné (station), dans une échelle rapportée à la valeur la plus grande de la variable considérée. Pour calculer la longueur d'une raie, on considère l'ensemble des valeurs du percentile obtenues pour les 14 stations. La valeur maximale du percentile p (y^p_{max}) va correspondre à la raie la plus longue pour les 14 représentations de ce percentile. La longueur de cette raie est notée L^p_{max}. La valeur minimale, pour ce même percentile, est représentée par une raie de longueur égale à : $L^p_{min} = 0,1 \times L^p_{max}$. Les 12 autres valeurs du percentile seront associées à des raies de longueurs intermédiaires calculées à partir de la relation suivante :

$$L_{ip} = 0,9 \times \left(\frac{Y_{ip}}{y^p_{max}} \right) \times L^p_{max} + 0,1 \times L^p_{max}$$

y_{ip} : valeur du percentile p pour la station i ;

y^p_{max} : valeur la plus grande du percentile p obtenue pour les 14 stations ;

L_{ip} : longueur de la raie correspondant à la valeur du percentile p pour la station i ;

L^p_{max} : longueur de la raie correspondant à la valeur la plus grande du percentile p.

Pour notre application, ce type de graphe constitue « l'empreinte » du profil d'un polluant mesuré à une station. Cette représentation sera un support graphique très utile pour interpréter les résultats de la classification automatique, réalisée à partir des mêmes percentiles retenus pour chaque station.

Classification ascendante hiérarchique dans un espace métrique

La classification ascendante hiérarchique est un outil d'analyse des données qui va permettre,

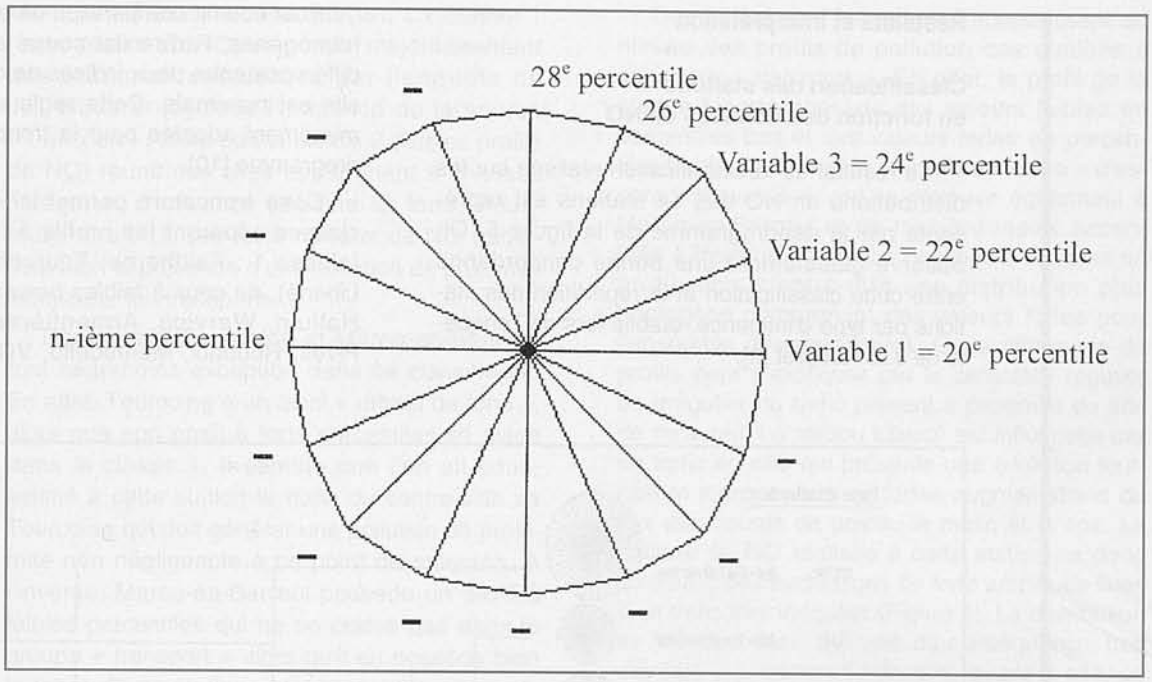


Figure 3. Exemple d'un graphe en étoile.
Example of a star symbol plot.

par la construction de partitions hiérarchisées, de mesurer les proximités entre les profils de pollution des 14 stations. Les 40 percentiles retenus sur chaque courbe de distribution sont considérés dans cette analyse comme des variables. Chacune de ces 40 variables définit une dimension dans un hyperespace. Les 14 profils de station forment donc un nuage de 14 points dans cet espace à 40 dimensions. Les différences entre les profils de station vont être mesurées par le calcul des distances entre ces points.

Le résultat de la classification ascendante hiérarchique se présente sous la forme d'un arbre « couché » communément appelé dendrogramme (Figure 4). La méthode de classification est basée sur l'algorithme de Ward. Ce dendrogramme est construit par itérations successives en partant de la partition initiale formée des 14 points correspondant aux profils de station, munis chacun d'une masse égale à 1. À chaque itération, on calcule pour chaque couple de points un indice de distance, δ (i,i'), qui se définit de la façon suivante :

$$\delta(i,i') = \frac{m_i \times m_{i'}}{m_i + m_{i'}} \times \left\| Y^i - Y^{i'} \right\|^2$$

Y^i : vecteur contenant les coordonnées du profil de la station i dans l'espace des percentiles à 40 dimensions ;

m_i : masse du profil de la station i ;

$Y^{i'}$: vecteur contenant les coordonnées du profil de la station i' dans l'espace des percentiles à 40 dimensions ;

$m_{i'}$: masse du profil de la station i'.

Les deux points dont l'indice de distance est le plus faible sont agrégés, c'est-à-dire qu'ils vont être remplacés dans le nuage des points par leur centre de gravité, affecté d'une masse égale à la somme des masses de ces deux points. Il est aisément démontré que l'indice de distance de ces deux points correspond également à la perte d'inertie intraclasses provoquée dans le nuage par leur agrégation. Ce procédé est réitéré (ici 13 fois) jusqu'à ce que tous les points soient réunis au centre de gravité du nuage. Les agrégats de points appelés également des « nœuds », sont placés sur le dendrogramme à des hauteurs correspondant aux indices de distance associés à ces différentes agrégations (Figure 4).

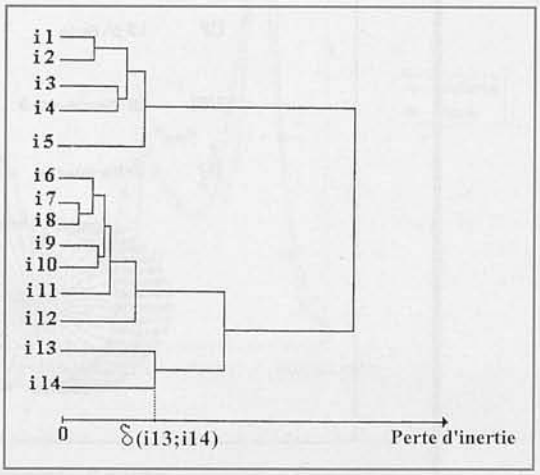


Figure 4. Exemple d'un dendrogramme à 14 individus.
Example of a fourteen individuals dendrogram.

Résultats et interprétation

Classification des stations en fonction des données de NO

Le résultat de la classification réalisée sur les distributions en NO des 14 stations est représenté par le dendrogramme de la figure 5. On observe globalement une bonne concordance entre cette classification et la répartition des stations par type d'influence, établie lors de l'enquête de J.L. Houdret [9].

Pour obtenir une partition de profils en classes homogènes, l'arbre est coupé au niveau où la différence entre deux indices de distance successifs est maximale. Cette règle est la plus communément adoptée pour la troncature d'un dendrogramme [10].

Cette troncature permet ici de former deux classes séparant les profils à forts percentiles (classe 1 : Faidherbe, Tourcoing, Pasteur, et Liberté), de ceux à faibles percentiles (classe 2 : Halluin, Wervicq, Armentières, Haubourdin, Fives, Roubaix, Montebello, Villeneuve d'Ascq,

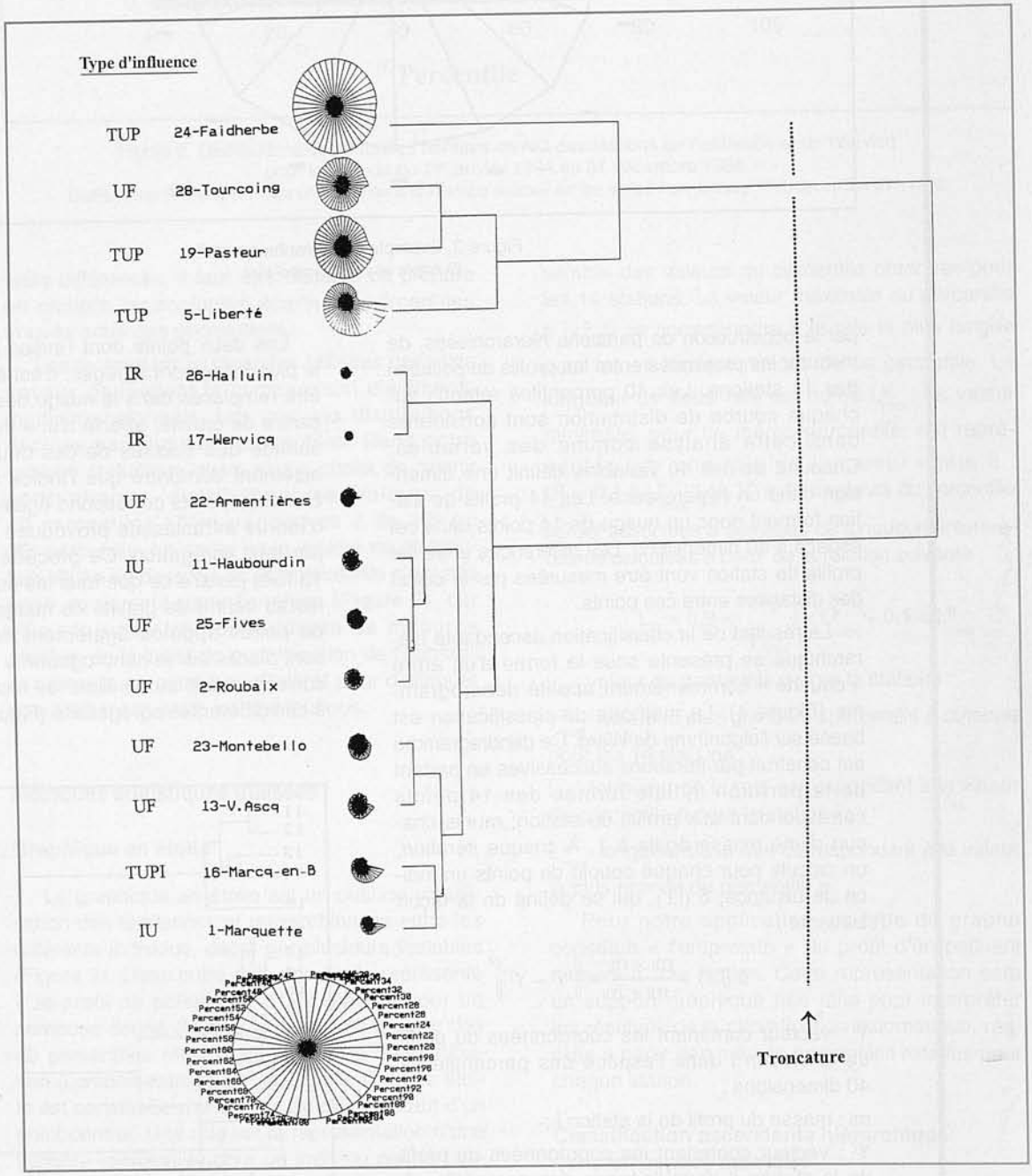


Figure 5. Graphiques en étoile et dendrogramme obtenus pour les profils de NO des 14 stations du réseau AREMA-LRT.

Star symbol plots and dendrogram obtained on the NO profiles of fourteen AREMA-LRT stations.

Marcq-en-Barœul et Marquette). La classe 1 (forts profils de NO) regroupe majoritairement des stations considérées par l'enquête de J.L. Houdret [9] sous l'influence de la source « transport », alors que la classe 2 (faibles profils de NO) réunit des sites appartenant à diverses catégories d'influence (urbaine de fond, industrielle rurale...). Les forts niveaux de NO traduisent bien la présence d'une pollution de proximité induite par les transports.

Deux sites (Tourcoing et Marcq-en-Barœul) font néanmoins exception dans ce classement. En effet, Tourcoing a un label « urbain de fond », alors que son profil à forts percentiles se place dans la classe 1. Il semble que l'on ait sous-estimé à cette station le trafic du centre-ville de Tourcoing qui doit générer une pollution de proximité non négligeable à ce point de mesures. À l'inverse, Marcq-en-Barœul possède un profil à faibles percentiles qui ne se classe pas dans le groupe « transport » alors qu'il en possède bien le label. Cette station est certes située à proximité d'un axe routier de relative importance, la N17 (près de 5 500 véhicules par jour en 1996), mais des bâtiments de grandes dimensions, présents entre la station et la N17, semblent jouer un rôle d'écran à la pollution en NO et expliquer les faibles niveaux enregistrés sur ce site. D'après les résultats de cette classification, il apparaît nécessaire de reconsidérer les labels des stations de Tourcoing et de Marcq-en-Barœul.

Des différences importantes apparaissent au niveau des profils de pollution des stations à caractère « transport ». En effet, le profil de la station Liberté possède des valeurs faibles en percentiles bas et des valeurs fortes en percentiles hauts. Cela produit un profil en forme « d'escargot renversé », qui se retrouve également à Marcq-en-Barœul et de manière moins accentuée à Pasteur. Faidherbe possède à l'inverse un profil rond, c'est-à-dire une distribution plus homogène comprenant des valeurs fortes pour l'ensemble des percentiles. Cette différence de profils peut s'expliquer par le caractère régulier ou irrégulier du trafic présent à proximité du site de mesures. La station Liberté est influencée par un trafic en ville qui présente une évolution journalière marquée par de fortes augmentations du flux aux heures de pointe, le matin et le soir. La mesure du NO réalisée à cette station va donc connaître des fluctuations de forte amplitude liées à ce trafic très irrégulier (Figure 6). La distribution de NO sera donc, par voie de conséquence, très étendue. La station Faidherbe, quant à elle, se situe près de l'autoroute A1. Elle est sous l'influence directe d'une circulation automobile dense et régulière. La mesure de NO à Faidherbe va donc présenter des niveaux élevés avec de faibles variations journalières (Figure 6). Il est donc logique d'observer pour ce site une distribution moins étendue de valeurs fortes de NO.

On peut également remarquer que les stations de type industriel rural (Wervicq et Halluin),

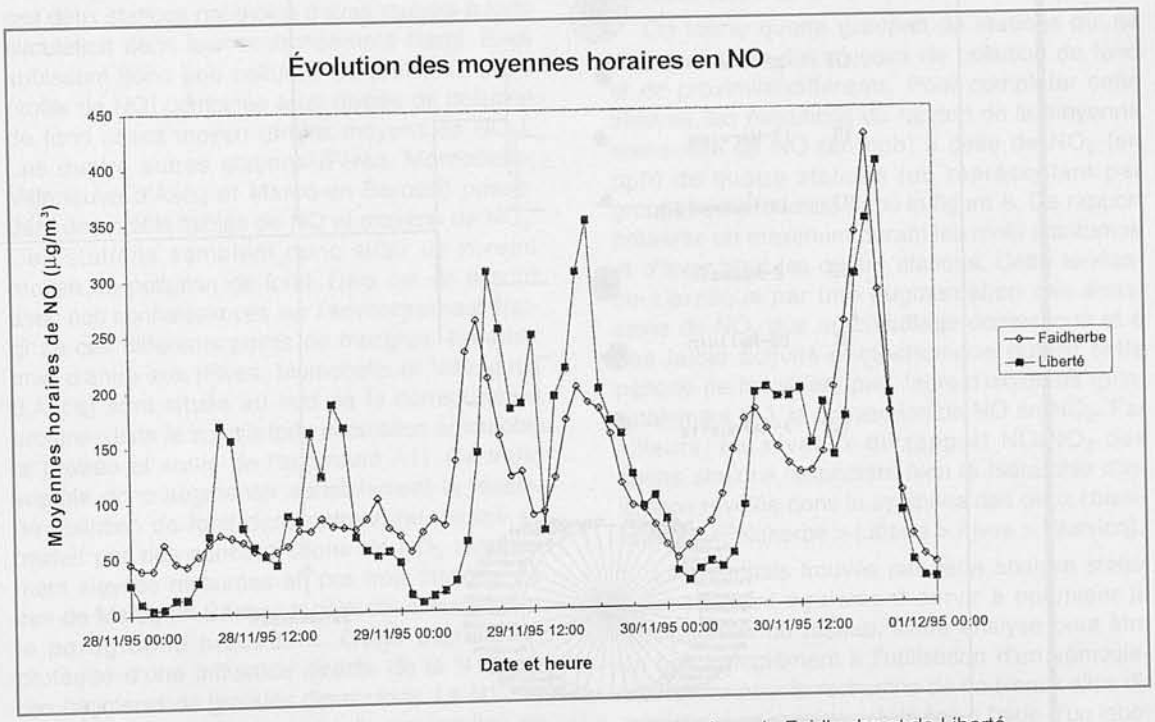


Figure 6. Évolution des moyennes horaires de NO aux stations de Faidherbe et de Liberté, pour la période du 28-11-95 à 0 h au 1-12-95 à 0 h.

Variations of hourly NO values at Faidherbe and Liberté stations for the period from 28-11-95 0 h to 1-12-95 0 h.

situées dans la classe 2, forment le premier agrégat de l'arbre de classification. Les profils très ressemblants de ces deux stations sont marqués par de très faibles valeurs pour tous les percentiles. Le caractère industriel de ces deux sites n'a donc aucune influence sur la mesure du NO. Les autres stations de cette classe 2 (Armentières, Haubourdin, Fives, Roubaix, Montebello, Villeneuve d'Ascq, Marcq-en-Barœul, et Marquette) présentent des profils à faibles percentiles. Elles ont en commun d'appartenir au tissu urbain et de ne pas être directement influencées par la source « transport ».

Classification des stations en fonction des données de NO_2

Le dendrogramme obtenu pour les profils de NO_2 des 14 stations est présenté dans la figure 7. L'examen des graphiques en étoile révèle que les écarts entre les profils de NO_2 sont globalement plus faibles que ceux obtenus pour NO. Cela se

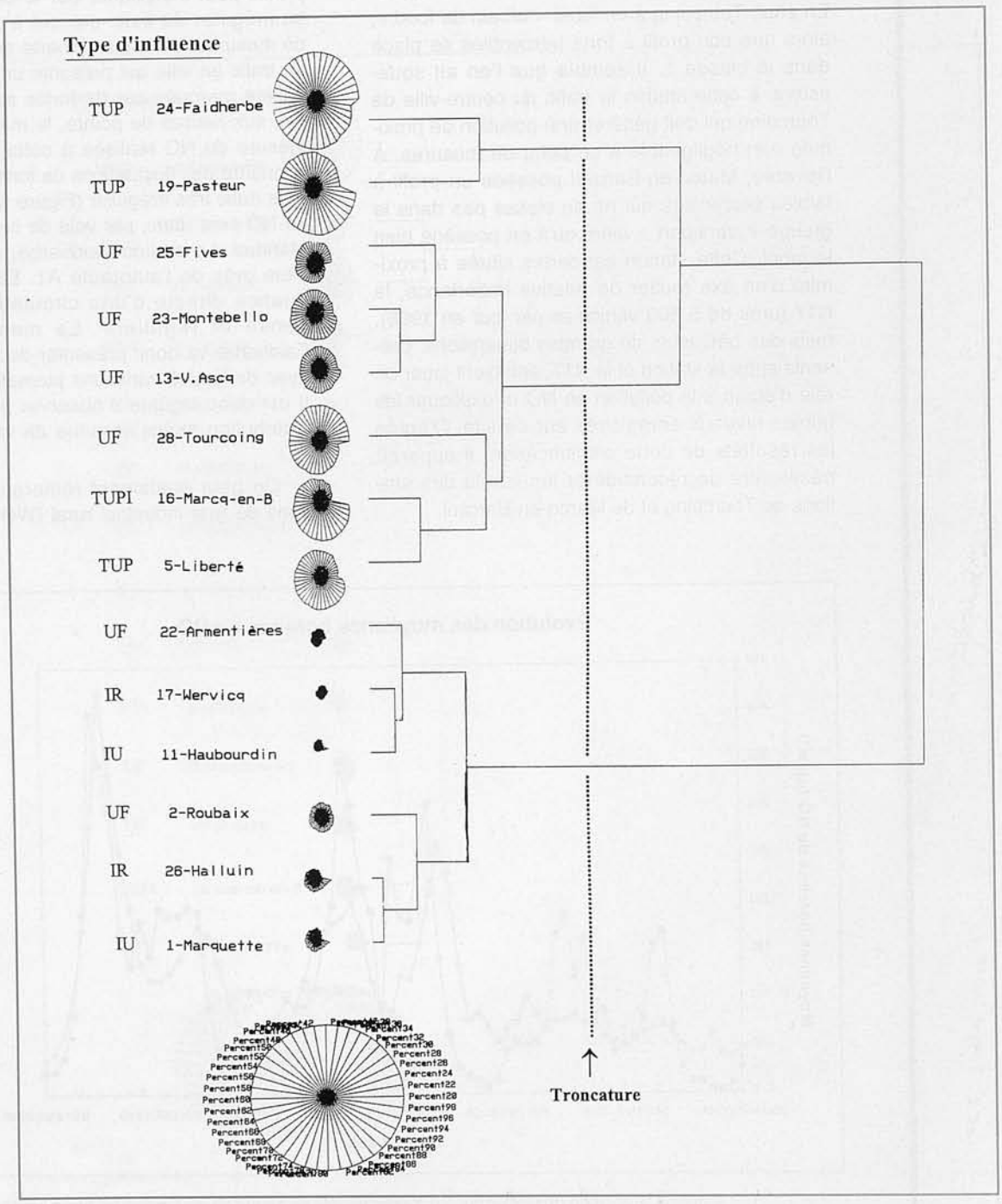


Figure 7. Graghiques en étoile et dendrogramme des profils de NO_2 de 14 stations du réseau AREMA-LRT.
Star symbol plots and dendrogram obtained on the NO_2 profiles of fourteen AREMA-LRT stations.

traduit sur le dendrogramme par un étalement plus important des branches de l'arbre de classification. Le caractère plus passif de la pollution par NO_2 peut expliquer ce resserrement des profils. NO_2 est un polluant secondaire qui présente une réactivité beaucoup moins importante dans l'atmosphère que NO. Ce dernier est rapidement oxydé autour des sources d'émissions, alors que NO_2 va diffuser sur plusieurs centaines de mètres, voire plusieurs kilomètres. Les concentrations de NO_2 mesurées à l'échelle d'une ville sont donc logiquement plus homogènes que celles de NO.

En adoptant la même règle de troncature que pour le dendrogramme de NO, nous formons ici trois classes de profils de NO_2.

Un premier groupe est composé de deux stations (Faidherbe et Pasteur) qui possèdent de forts profils de NO_2. Ces stations ont en commun d'appartenir à la classe « transport » pour le dendrogramme de NO et d'être situées au sud de la communauté urbaine, dans une zone présentant une densité importante de grands axes routiers à forte circulation. Cette classe semble donc regrouper des sites marqués à la fois par une pollution de proximité due aux transports (expliquant les forts profils de NO) et une pollution de fond venant des transports (expliquant les forts profils de NO_2).

La deuxième classe regroupe six stations présentant des profils moyens de NO_2. Deux sites (Tourcoing et Liberté) marqués par une pollution de proximité due aux transports (forts profils de NO) font partie de cette classe « intermédiaire ». Situées plutôt au nord de la communauté urbaine, ces deux stations ont moins d'axes routiers à forte circulation dans leur environnement élargi. Elles subissent donc une pollution de proximité (forts profils de NO) combinée à un niveau de pollution de fond assez moyen (profils moyens de NO_2). Les quatre autres stations (Fives, Montebello, Villeneuve d'Ascq et Marcq-en-Barœul) possèdent des profils faibles de NO et moyens de NO_2. Ces stations semblent donc subir un niveau moyen de pollution de fond. Cela est en accord avec nos connaissances sur l'environnement élargi de ces différents points de mesures. En effet, trois d'entre eux (Fives, Montebello et Villeneuve d'Ascq) sont situés au sud de la communauté urbaine, dans la zone à forte circulation automobile (entrée et sortie de l'autoroute A1). Ce trafic semble donc augmenter sensiblement le niveau de pollution de fond dans cette zone, ce qui se traduit par des concentrations de NO_2 relativement élevées mesurées en ces trois stations. Le cas de Marcq-en-Barœul a déjà été évoqué dans le paragraphe précédent. Cette station est protégée d'une influence directe de la N17 par des bâtiments de grandes dimensions. La N17 va donc contribuer uniquement à l'augmentation de la pollution de fond mesurée en cette station, c'est-à-dire à l'élévation sensible du niveau de NO_2. Cette deuxième classe regroupe donc des stations de proximité (forts profils de NO) présentant une pollution de fond assez faible (profils moyens de NO_2) et des stations urbaines exposées uniquement à un niveau moyen de pollution de fond (profils faibles de NO et moyens de NO_2).

La troisième classe regroupe six stations présentant des profils faibles de NO_2. Ces sites sont soit de type rural c'est-à-dire placés hors du tissu urbain (Halluin et Wervicq), soit de type urbain mais exposés à un faible niveau de pollution de fond (Haubourdin, Roubaix, Marquette, et Armentières). Notons que ces six stations, faiblement influencées par la source « transport », sont situées au nord et à l'ouest de la communauté urbaine. Il semble donc y avoir une distribution spatiale cohérente de la pollution de fond générée par les transports. En effet, cette pollution, représentée par le niveau de NO_2, diminue à mesure que l'on s'éloigne des grands axes routiers situés au sud de la communauté urbaine. Les forts profils de NO_2 traduisent bien la présence d'un niveau élevé de pollution de fond induit par les transports.

Synthèse des deux classifications et démarche d'optimisation du réseau de mesures

Le recoupement des informations obtenues par ces deux classifications permet de bien caractériser et quantifier la pollution liée aux transports mesurée par ces différentes stations. Sur la base de ces résultats, un nouveau classement est établi, qui est présenté dans le tableau 2.

On forme quatre groupes de stations qui se distinguent par des niveaux de pollution de fond et de proximité différents. Pour compléter cette analyse, les évolutions du rapport de la moyenne mensuelle de NO (en ppb) à celle de NO_2 (en ppb) de quatre stations (un représentant par groupe) sont fournies dans la figure 8. Ce rapport présente un maximum durant les mois d'automne et d'hiver pour les quatre stations. Cette tendance s'explique par une augmentation des émissions de NO_x due au chauffage domestique et à une faible activité photochimique durant cette période ne favorisant pas, faute d'oxydants (principalement O_3), la conversion de NO en NO_2. Par ailleurs, les niveaux du rapport NO/NO_2 des quatre stations respectent bien la hiérarchie d'influence révélée dans la synthèse des deux classifications (Faidherbe > Liberté > Fives > Wervicq).

Les résultats trouvés par cette analyse statistique peuvent également servir à optimiser la configuration du réseau. Cette analyse peut être un bon complément à l'utilisation d'un véhicule-laboratoire pour la recherche de nouveaux sites de mesures. Les mesures réalisées à l'aide d'un laboratoire mobile en différents points d'une zone peuvent fournir de nouveaux profils de pollution.

Tableau 2. Synthèse des deux classifications des stations NO_x.
Synthesis on the two dendograms of NOx stations.

	Groupe 1	Groupe 2	Groupe 3	Groupe 4
Type d'influence : Pollution de proximité (classification des niveaux de NO) : Pollution de fond (classification des niveaux de NO_2) :	Forte Forte	Forte Moyenne	Faible Moyenne	Faible Faible
Stations	Faidherbe Pasteur*	Liberté* (Tourcoing)	Fives Montebello Villeneuve d'Ascq (Marcq-en-Barœul*)	Haubourdin Roubaix Marquette* Armentières Wervicq Halluin

* Station influencée plus ou moins fortement par une circulation routière irrégulière.
( ) : les objectifs de mesures de ces stations sont à reconsidérer.

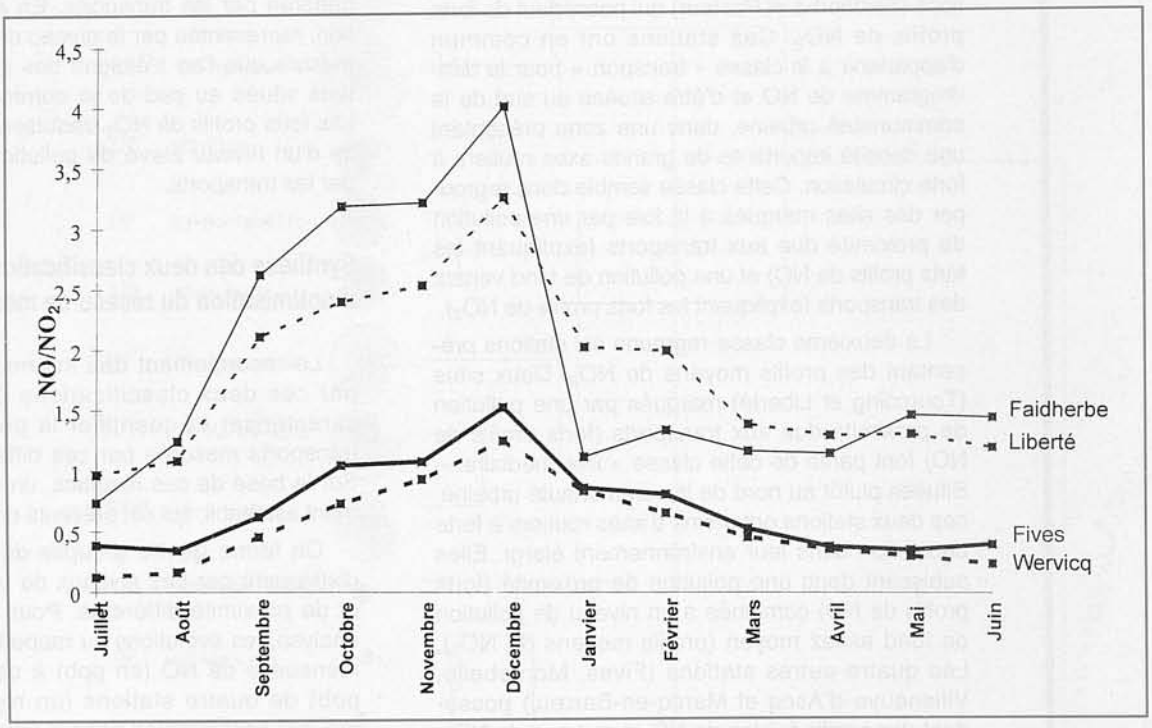


Figure 8. Variations saisonnières des moyennes mensuelles du rapport NO/NO_2 dans quatre stations de mesure du réseau AREMA-LRT.
Seasonal variations of the monthly mean of ratio NO/NO_2 at four AREMA-LRT stations.

Introduits dans les classifications, ces nouveaux profils pourront être comparés à ceux des stations déjà existantes du réseau. On pourra ainsi sélectionner les sites de mesures qui répondent le mieux à la stratégie de développement du réseau ou qui apportent des informations supplémentaires sur la pollution présente dans la zone à surveiller.

Conclusion et perspectives

L'analyse statistique réalisée dans cette étude révèle que les niveaux de NO et de NO_2 des stations du réseau AREMA-LRT sont essentiellement déterminés par la pollution émise par les transports. Les distributions de NO se répartissent en deux classes suivant que la station est influencée ou non par une pollution de proximité venant des transports. Les distributions en NO_2 forment trois classes homogènes qui correspondent à des degrés d'exposition différents des stations vis-à-vis de la pollution de fond liée aux transports. Des différences de profils de NO ont été observées entre les stations de proximité placées au centre-ville et celle située près d'une autoroute. La synthèse des résultats obtenus a permis de bien définir le type d'influence « transport » subi par les différentes stations du réseau et de reconsidérer, pour certaines d'entre elles, leurs objectifs de mesures.

Ces résultats peuvent également servir de support à l'optimisation du dispositif de surveillance du réseau.

Cette démarche peut bien évidemment s'appliquer aux autres polluants (SO_2, O_3, poussières...) qui font également l'objet d'une surveillance en réseau. Elle nécessite néanmoins un nombre suffisant de stations à intégrer dans l'analyse statistique pour que la classification et les groupes de profils formés puissent avoir une réelle signification.

Remerciements

Nous présentons nos très sincères remerciements à M. Jean-Yves Saison, responsable du réseau AREMA-LRT, pour nous avoir permis d'exploiter les données des stations de mesures de l'agglomération lilloise. Ces travaux ont été menés dans le cadre des activités du Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA) et ont bénéficié du soutien financier du ministère de l'Environnement et de l'ADEME.

Mots clés

Pollution de l'air. Milieu urbain. Réseau. NO_x. Statistiques. Classification. Optimisation.

Keywords

Air pollution. Urban area. Network. NO_x. Statistics. Classification. Optimization.

Bibliographie

1. Loi n° 96-1236 du 30 décembre 1996 sur l'air et l'utilisation rationnelle de l'énergie. Journal officiel de la République française, 1er janvier 1997 : 11-8.
2. Nakamori Y, Ikeda S, Sawaragi Y. Design of air pollutant monitoring system by spatial sample stratification. Atmospheric Environment 1979 ; 13 (1) : 97-103.
3. Kainuma Y, Shiozawa K, Okamoto S. Study of the optimal allocation of ambient air monitoring stations. Atmospheric Environment 1990 ; 24B (3) : 395-406.
4. Hwang JS, Chan CC. Redundant measurements of urban air monitoring networks in air quality reporting. Journal Air and Waste Management Association 1997 ; 47 : 614-9.
5. Handscombe CM, Elsom DM. Rationalization of the national survey of air pollution monitoring network of the United Kingdom using spatial correlation analysis: a case-study of the Greater London area. Atmospheric Environment 1982 ; 16 (5) : 1061-70.
6. Gottard S. Indice national de la qualité de l'air ATMO : Application au cas de la métropole lilloise. Rapport du Réseau AREMA-LRT, Lille 1995.
7. Évolution de la circulation à Lille intra-muros de 1978 à 1993. Rapport de la Direction de la circulation et voirie, Communauté Urbaine de Lille 1994.
8. Bessemoulin J. Atlas climatique de la France. Direction de la météorologie nationale, Paris 1989.
9. Houdret JL. Étude des sites de mesures en France. Rapport de l'École des Mines, Douai 1997.
10. Volle M. Analyse des données. Édition Economica, Paris 1993.