



**HAL**  
open science

## Classification des profils gazeux odorants : Utilisation d'un système de capteurs multi-gaz associé aux techniques de reconnaissance de forme

Chaouki Regoui, Olivier Ramalho, Séverine Kirchner

### ► To cite this version:

Chaouki Regoui, Olivier Ramalho, Séverine Kirchner. Classification des profils gazeux odorants : Utilisation d'un système de capteurs multi-gaz associé aux techniques de reconnaissance de forme. 5ème Rencontres de la Société Francophone de Classification, Sep 1997, Lyon, France. hal-00697210

**HAL Id: hal-00697210**

**<https://hal.science/hal-00697210>**

Submitted on 14 May 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# CLASSIFICATION DE PROFILS GAZEUX ODORANTS: UTILISATION D'UN SYSTÈME DE CAPTEURS MULTI-GAZ ASSOCIÉ AUX TECHNIQUES DE RECONNAISSANCE DE FORMES

Chaouki REGOUI <sup>①②</sup>, Olivier RAMALHO <sup>②</sup>, Séverine KIRCHNER <sup>②</sup>

- ① Laboratoire SIP, université Paris V, UFR Maths-Info, 45 rue des saints pères 75006 Paris, France, ℳ: 01.44.55.35.56, fax: 01.44.55.35.36, e-mail: regc@math-info.univ-paris5.fr  
② CSTB, service EAE, 84 avenue Jean Jaurès BP2 77421 Marne la vallée cedex 2, France

## Résumé

L'objectif de ce travail est la classification sur critère d'information olfactive (aspect odorant), d'effluents gazeux acquis par un système multi-gaz. Notre intérêt est focalisé sur, d'une part, l'amélioration du pouvoir discriminant du système utilisé pour la reconnaissance de profils de composés émis par les matériaux de construction et, d'autre part, le temps de réponse (temps réel de quelques minutes) [1].

L'analyse factorielle est utilisée pour une exploration descriptive des données. Une classification neuronale des échantillons acquis est ensuite entreprise. Les résultats de cette classification sont comparées à ceux obtenus par analyse discriminante.

**Mots-clés:** classification, reconnaissance de formes, apprentissage, signaux multidimensionnels, capteurs multi-gaz.

## Introduction

Un système multi-capteurs de gaz est utilisé pour l'acquisition des profils de certains composés émis par les matériaux de construction permettant l'évaluation de la qualité de l'air intérieur. Le système utilisé comporte 32 capteurs chimiques, à technologie polymères conducteurs, ayant des gammes de sensibilité aux molécules ciblées qui se chevauchent et qui produisent un signal multidimensionnel (figure 1). Ce signal représente les variations relatives des résistances électriques des capteurs, suite à l'exposition à différentes substances. Cette réponse est globale et comporte 32 canaux d'information [2].

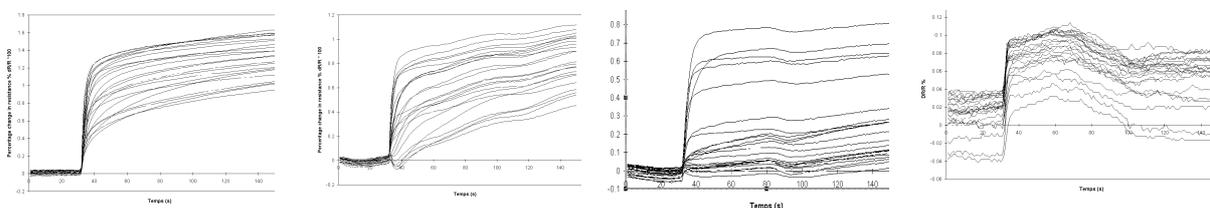


Figure 1: profils (32 capteurs) de l'acétone, le méthanol, le toluène et un air de référence neutre.

## Méthodologie de travail

Pour obtenir des réponses utilisables pour les différents gaz, un travail exploratoire a été nécessaire pour déterminer les seuils de concentration minimum engendrant une réponse des capteurs. Une fois ces seuils déterminés, l'acquisition a été effectuée. Le signal multidimensionnel acquis pour chaque échantillon de gaz comporte plusieurs phases:

- une phase de stabilisation du signal des capteurs pour obtenir une ligne de base de la réponse (de 30 secondes pour les données actuelles),

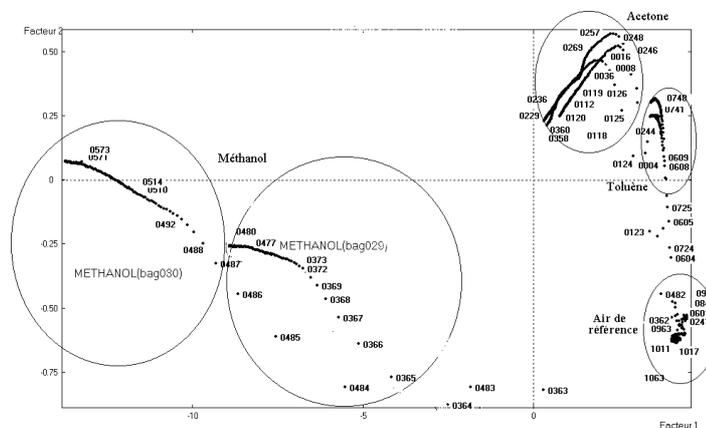
- une phase d'acquisition du signal utile, durant laquelle des échantillons de gaz sont soumis aux capteurs (durée utile du signal de 120 secondes),
- une phase de remise à la ligne de base.

La fréquence d'échantillonnage du signal d'acquisition est de 1 Hertz et la sensibilité limite de détection est de 0.01% (en variation relative de résistance).

### Le traitement des données

Une première étape, de description et d'exploration des données par des techniques factorielles est effectuée. Une seconde étape de classification proprement dite, à l'aide d'une approche neuronale a servi au traitement ultime des données [3].

La figure 2 représente le résultat de l'analyse en composantes principales de la phase utile du signal d'acquisition de gaz. Elle concerne 3 échantillons d'acétone, 2 de méthanol, 2 de toluène et 2 d'un air de référence, à raison de 120 points par échantillon. Elle montre la séparation des différentes classes d'échantillons (spécialement le méthanol). Les deux premiers facteurs cumulent 99.8% de l'inertie totale. La sensibilité des capteurs aux conditions d'acquisition a été mise en évidence par la distinction d'échantillons du même gaz pris à des instants différents.



**Figure 2: analyse en composantes principales sur toutes les données échantillonnées**

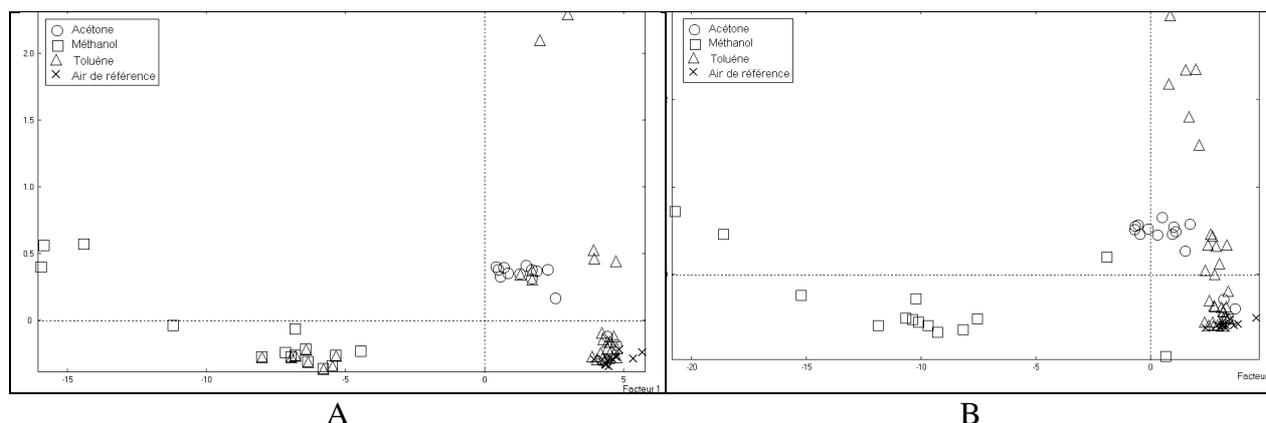
Les phénomènes ayant lieu durant la phase d'échanges chimiques (partie dynamique de la phase d'acquisition du signal utile) étant différents de ceux à l'équilibre de charges électriques une fois le signal stabilisé (partie statique de la phase d'acquisition du signal utile), une analyse des données comparant l'information de la partie dynamique et celle du plateau statique a été entreprise.

### Analyse de l'information des parties statique et dynamique

Les données qui ont été utilisées comportaient 74 échantillons représentant les quatre gaz précédemment utilisés (acétone, méthanol, toluène et air de référence neutre). La période d'acquisition du signal utile de chaque gaz est divisée en deux parties, statique et dynamique. Chaque partie est résumée par un seul paramètre. La pente de la courbe globale (intégration des 32 signaux), sur les dix premières secondes du signal utile, pour la partie dynamique et la moyenne sur les 90 dernières secondes pour le plateau statique. Les données acquises se répartissaient en acétone (14 échantillons), méthanol (15 échantillons), toluène (31 échantillons) et air de référence neutre (14 échantillons).

Une classification hiérarchique de l'ensemble des échantillons caractérisés par leurs premières coordonnées factorielles créées par analyse en composantes principales (A.C.P) et utilisant le critère d'agrégation de Ward a été effectuée. L'arbre obtenu présentait une bonne séparation des échantillons de méthanol et dans une moindre mesure d'acétone par rapport aux autres échantillons. L'analyse en composantes principales permet de juger cette séparation des échantillons de méthanol par rapport aux autres gaz notamment dans la partie dynamique (figure

3B), cependant, elle dénote la confusion entre les échantillons d'acétone et de toluène (partie statique ou dynamique).



**Figure 3: analyse en composantes principales sur les données des parties statique (A) et dynamique (B)**

Une classification neuronale, supervisée, a été effectuée sur l'ensemble des 74 échantillons acquis. La base d'apprentissage comportait 38 échantillons et la base de test 36. Les réseaux utilisés sont des réseaux multi couches utilisant l'algorithme de rétro propagation du gradient, l'initialisation des poids synaptiques étant aléatoire à valeurs normalisées [4].

Les taux de classification obtenus, après optimisation du réseau neuronal, sont résumés pour chaque gaz (tableau 1).

	Partie statique	Partie dynamique
<b>Acétone</b>	100 %	100 %
<b>Méthanol</b>	100 %	86 %
<b>Toluène</b>	73 %	93 %
<b>Air neutre</b>	86 %	100 %
<b>Taux global</b>	<b>86 %</b>	<b>94%</b>

**Tableau 1: taux de classification neuronale sur information des parties statique et dynamique**

A titre de comparaison, il a été procédé à une analyse discriminante par arbre de décision binaire. Le calcul de la fonction discriminante et son évaluation sont effectués sur les mêmes bases d'apprentissage (38 échantillons) et de test (36 échantillons) que celles utilisées par la classification neuronale.

Le calcul de la fonction discriminante tient compte des probabilités a priori d'appartenance aux différentes classes. A partir de l'arbre complet, un sous-arbre "optimal" a été sélectionné, à l'aide des échantillons de la base de test, grâce à une procédure d'élagage basée sur la suppression successive des branches les moins informatives en terme de discrimination entre les classes [5]. Les résultats obtenus sont présentés (tableau 2).

	Partie statique	Partie dynamique
<b>Acétone</b>	86 %	86 %
<b>Méthanol</b>	100 %	100 %
<b>Toluène</b>	33 %	100 %
<b>Air de référence</b>	100 %	29 %
<b>Taux global</b>	<b>69 %</b>	<b>83 %</b>

**Tableau 2: pourcentage d'échantillons bien classés par analyse discriminante des parties statique et dynamique**

### Discussion des résultats et perspectives

Les résultats obtenus sur cette base de données ont montré une bonne discrimination du système de capteurs multi-gaz pour les composés étudiés. Les gaz les mieux reconnus étant l'acétone et le méthanol.

Ces résultats sont confirmés par les deux techniques utilisées, la classification neuronale et l'analyse discriminante quoique cette dernière ait engendré des taux de classification moins élevés. Ceux-ci dénotent, cependant, la richesse du vecteur de données résultant de l'acquisition par le système multi-capteurs. Les taux de classification obtenus (reconnaissance en neuronal atteignant, sur la partie dynamique, 94 % des échantillons acquis) attestent de cette bonne discrimination.

Les performances résultant de l'utilisation de la partie dynamique du signal utile étaient supérieures à ceux obtenus avec l'information de la partie statique (et ce pour les deux techniques de classification utilisées). Les résultats obtenus grâce à l'information de la partie statique étaient cependant acceptables.

Une exploitation de l'information globale (statique et dynamique) est entreprise actuellement mais nécessite encore la résolution de quelques difficultés notamment en terme de normalisation des signaux pour éviter de perdre certains petits détails du signal, très caractéristiques de chaque gaz. Étant donné que les systèmes multi-gaz souffrent de la mauvaise répétabilité des mesures et de la dérive dans le temps de la réponse des capteurs (vieillessement), ces résultats encouragent à l'optimisation de l'outil de reconnaissance de formes ainsi développé et sa validation sur d'autres bases en cours d'acquisition. Ces nouvelles données (plusieurs autres gaz) serviront à la validation des résultats déjà obtenus notamment concernant l'importance des différentes parties du signal utile (dynamique et statique).

Une attention particulière sera cependant accordée aux conditions perturbantes d'acquisition pour vérifier leur effet sur la pertinence de l'information acquise.

### ***Bibliographie***

1. REGOUI Chaouki, RAMALHO Olivier, KIRCHNER Séverine "Discrimination of volatile organic compounds emitted by building materials using an electronic nose" Healthy Buildings/IAQ'97, Washington D.C., Septembre 97. Accepté.
2. PERSAUD K.C., PELOSI P. "Sensor arrays using conducting polymers for an artificial nose", NATO advanced research workshop 'Sensors and sensory systems for an electronic nose', Reykjavik Iceland, august 91, 237-256, Kluwer ed.
3. CRANNY A. W. J., ATKINSON. J. K. "The use of pattern recognition techniques applied to signals generated by a multi element gas sensor array as means of compensating for poor individual element response" NATO advanced research workshop 'Sensors and sensory systems for an electronic nose', Reykjavik, 197-215.
4. CHEN C.H. "Fuzzy logic & neural network handbook" Mc Graw Hill series on computer engineering, 1996.
5. LEBART L., MORINEAU A., PIRON M. "Statistique exploratoire multidimensionnelle" Dunod, 1995.