

# Schéma d'échantillonnage pour les campagnes de mesures de la qualité de l'air : une approche par optimisation

Thomas Romary, Chantal de Fouquet, Laure Malherbe

► **To cite this version:**

Thomas Romary, Chantal de Fouquet, Laure Malherbe. Schéma d'échantillonnage pour les campagnes de mesures de la qualité de l'air : une approche par optimisation. 42èmes Journées de Statistique, 2010, Marseille, France, France. inria-00494818

**HAL Id: inria-00494818**

**<https://hal.inria.fr/inria-00494818>**

Submitted on 24 Jun 2010

**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.

# SCHÉMA D'ÉCHANTILLONNAGE POUR LES CAMPAGNES DE MESURES DE LA QUALITÉ DE L'AIR : UNE APPROCHE PAR OPTIMISATION

Thomas Romary<sup>1</sup>, Chantal de Fouquet<sup>1</sup> & Laure Malherbe<sup>2</sup>

<sup>1</sup> *Equipe géostatistique, centre de géosciences, Mines ParisTech,  
35 rue Saint Honoré, 77305 Fontainebleau, France*

<sup>2</sup> *Institut National de l'Environnement Industriel et des Risques (INERIS), Direction  
des risques chroniques, Parc Technologique Alata, 60550 Verneuil-en-Halatte, France*

## RESUME

Dans ce travail, nous présentons une méthode d'échantillonnage optimal dans un contexte spatial, pour une application aux campagnes de mesure de la concentration de benzène dans l'air ambiant, à l'échelle de l'agglomération.

Dans un premier temps, nous nous fondons sur l'analyse des données de campagnes antérieures, réalisée sur deux agglomérations différentes, pour définir une modélisation a priori. Précisément, la modélisation retenue est une modélisation en dérive externe comprenant une dérive et un résidu corrélé spatialement (voir par exemple l'ouvrage de Chilès et Delfiner (1999)). L'analyse conduite nous permet de choisir les variables auxiliaires pertinentes et de déterminer un modèle de variogramme a priori pour le résidu.

Dans un second temps, nous nous proposons d'optimiser l'implantation des appareils de mesure de la concentration de benzène sur une troisième agglomération en suivant la méthodologie proposée dans un article de Brus et Heuvelink (2007). Pratiquement, nous cherchons à minimiser la moyenne sur l'agglomération de la variance de krigeage universel, dont les paramètres sont fondés sur la modélisation a priori. Cette optimisation est effectuée par recuit simulé, décrit dans Kirkpatrick et al. (1983). Nous présentons les résultats obtenus sur trois agglomérations différentes. Enfin, nous discutons des avantages et inconvénients de cette méthodologie en la comparant notamment à une démarche déterministe mise en œuvre dans le rapport de De Fouquet et Faucheux (2009).

## Mots-clefs

Statistique spatiale - Collecte de données

Ce travail a été réalisé pour le Laboratoire Central de Surveillance de la Qualité de l'Air. Il a été financé par le Ministère de l'Énergie, de l'Écologie, du Développement durable et de la Mer. Les données ont été fournies par les associations AIRAQ, ATMO Nord-Pas-de-Calais et ATMO Champagne-Ardenne.

## ABSTRACT

In this work, we present a spatial statistical methodology to design benzene air concentration measurement surveys, at the urban scale.

In a first step, we define an a priori modeling based on an analysis of data coming from previous campaigns on different agglomerations. More precisely, we retain a modeling with an external drift which consists of a drift plus a spatially correlated residual (see for instance the book by Chilès and Delfiner (1999)). The analysis performed leads us to choose the most relevant auxiliary variables and to determine an a priori variogram model for the residual.

In a second step, we optimize the measuring devices positioning on a third agglomeration following the methodology proposed in Brus and Heuvelink (2007). Practically, we aim at finding the design that minimizes the mean over the urban domain of the universal kriging variance, whose parameters are based on the a priori modeling. This optimization is performed by simulated annealing, described e.g. in Kirkpatrick (1983). We present the results obtained on three different agglomerations. Finally, we discuss the pros and cons of this methodology comparing it to a deterministic approach carried out in a previous work (De Fouquet and Faucheux (2009)).

### Keywords

Spatial statistics - Data collection

## Références

- [1] Chilès, J. P. and Delfiner, P. (1999) *Geostatistics, Modeling Spatial Uncertainty*. John Wiley & Sons.
- [2] Brus, D. J. and Heuvelink, G. B. M. (2007) Optimization of sample patterns for universal kriging of environmental variables. *Geoderma*, 138, 86–95.
- [3] Kirkpatrick, S., Gelatt Jr., C. D., and Vecchi, M. P. (1983) Optimization by Simulated Annealing. *Science*, 220, 671–680.
- [4] Faucheux, C., Malherbe, L., De Fouquet, C. and Cárdenas, G. (2009) How to build a sampling scheme : recommendations for measurement surveys of air quality. *Submitted to Environmetrics*.