Optimal spatial design for air quality measurement surveys: what criteria ?

Thomas Romary[1], Chantal de Fouquet[1]
1 Geostatistics team, Geosciences centre, Mines ParisTech, 35 rue Saint Honoré, 77305 Fontainebleau, France, thomas.romary@mines-paristech.fr

Laure Malherbe[2]
2 Institut National de l'Environnement Industriel et des Risques (INERIS), Direction des risques chroniques, Parc Technologique Alata, 60550 Verneuil-en-Halatte, France

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 two different agglomerations. More precisely, we retain a modeling with an external drift which consists of a drift plus a spatially correlated residual. The statistical analysis performed leads us to choose the most relevant auxiliary variables and to determine an a priori variogram model for the residual. An a priori distribution is also defined for the variogram parameters, whose values appear to vary from a campaign to another. In a second step, we optimize the positioning of the measuring devices on a third agglomeration according to a Bayesian criterion. 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, while accounting for the prior distribution over the variogram parameters. Two optimization algorithms are then compared: simulated annealing and a particle filter based algorithm.

Keywords: Optimal Design, Geostatistics, External Drift Kriging

1 Introduction

Mapping air pollution as precisely as possible is a major issue for French Local Air Quality Monitoring Agencies (the AASQAs) both for regulatory and information purposes and for public health concerns. Seasonal or annual average concentration maps can be obtained from passive sampling data collected at a large number of sites across the area of interest. The AASQAs regularly carry out such sampling surveys over various areas at various scales. Given those considerations, they have to design sampling schemes so that resulting concentration maps will fulfill precision criteria.
The interpolation is performed by kriging, see *e.g.* [1], and [2] for an application in atmospheric sciences. With its internal quantification of spatial variability through the covariance function (or variogram), kriging methodology can produce maps of optimal

predictions and associated prediction error variance from incomplete and possibly noisy spatial data. Kriging also provides a prediction error variance that can be seen as a criterion quality of the resulting maps. As it only depends on the spatial repartition of the points over the domain, it is a straightforward criterion for the quality of a sampling design, once a geostatistical model has been fitted on the phenomenon under study. This work completes and extend [3].

2 Materials and Methods

The proposed methodology consists of three steps: an estimation step, the definition of a quality criterion for the sampling design and an optimization step.

2.1 Estimation

A geostatistical analysis of the data collected during previous surveys is performed in order to set up the model (covariates and covariance) that will be used when applying the optimization method to another area. Data from benzene sampling surveys conducted in two French cities (Lille and Reims) have been used to fit the geostatistical model, which is made of a drift plus a spatially correlated residual:

$$Z(x) = \beta_0 + Y'(x)\beta + S(x), \tag{1}$$

where Z is the benzene concentration variable, $x \in \mathcal{X} \subset \mathbb{R}^2$ is the spatial coordinate, Y is the matrix of auxiliary variables exhaustively known on $\mathcal{X}$, $'$ is the transpose operator, β is a vector of parameters and $S(x)$ is a centered, spatially correlated residual.

2.2 Criterion building

The criterion to optimize is defined from the set up model by the integral over the domain under study of the weighted prediction error variance:

$$O(\eta) = \frac{1}{|\mathcal{X}|} \int_{\mathcal{X}} \mathbb{V}(Z(x) - \widehat{Z}(x)) w(x) dx, \tag{2}$$

where $|\mathcal{X}|$ is the area of $\mathcal{X}$. Practically, this integral is evaluated on a grid discretizing the domain $\mathcal{X}$. A non uniform weight function $w(x)$ can be designed to obtain a more accurate mapping in some areas, for instance in function of auxiliary variables values. When some parameters of the model cannot be fitted accurately, we can associate them an a priori distribution, then a Bayesian version of (2) can be considered:

$$O_{Bayes}(\eta) = \frac{1}{|\mathcal{X}|} \int_{\times} \int_{\mathcal{X}} \mathbb{V}(Z(x) - \widehat{Z}(x)|\theta) w(x) p(\theta) dx d\theta, \tag{3}$$

where $\theta \in \Theta$ is the set of uncertain parameters and $p(\theta)$ is its a priori distribution.

2.3 Optimization

Once the model to use has been clearly identified, we have to optimize (3) on a discretization of $\mathcal{X}$. Optimizing (3) on a large grid is a hard combinatorial problem. Therefore, we rely on heuristics algorithm to perform the global optimization: a simulated annealing scheme and a interacting particle algorithm.

3 Results

A third French agglomeration (Bordeaux) is taken as application case. The performances of both algorithms of are compared, in terms of optimization quality and computing time. We also show how the method can be used to dimension the network of passive samplers.

4 Concluding remarks

The current work has been carried out with the aim of supplying scientific and technical support to the French local air quality monitoring agencies. For the moment it has been applied to benzene sampling over urban areas but it can be extended to other pollutants such as NO2 and to larger spatial domains like regions.

References

[1] Chilès, J. P., and Delfiner, P. *Geostatistics, Modeling Spatial Uncertainty.* John Wiley & Sons, New-York, 1999.

[2] de Fouquet, C., Gallois, D., and Perron, G. Geostatistical characterization of the nitrogen dioxyde concentration in an urban area. part 1: spatial variability and cartography of the annual concentration. *Atmospheric Environment 41* (2007), 6701–6714.

[3] Romary, T., de Fouquet, C., and Malherbe, L. Sampling design for air quality measurement surveys: An optimization approach. *Atmospheric Environment In Press, Corrected Proof* (2011), –.