ESTIMATING THE GRADES OF POLLUTED INDUSTRIAL SITES: USE OF CATEGORICAL INFORMATION AND COMPARISON WITH THRESHOLD VALUES

N. Jeanne1,2 and Ch. De Fouquet2
1Centre National de Recherche sur les Sites et Sols Pollués, 930 Bd Lahure, BP 537, 59505 Douai Cedex, France, currently at Geovariances, 49bis Av. Franklin Roosevelt, 77212 Avon, Cedex, France. jeannee@geovariances.fr; 2Centre de Géostatistique, Ecole des Mines de Paris, 35 rue Saint-Honoré, 77300 Fontainebleau Cedex, France. fouquet@cg.ensmp.fr

Abstract: Sampling of polluted sites often leads to inaccurate estimates, particularly because of the small number of samples, the importance of sampling errors and the high spatial variability at small distance. Auxiliary information like the history of the site or qualitative measurements are of interest to improve the quality of the grade estimates. The relationship between grades and soils information (presence of coal tar, smell, clay...) are examined on a former coking plant, polluted by PAH (Polycyclic Aromatic Hydrocarbons). A sensitivity analysis shows the utility of this auxiliary information known at additional points compared to the univariate kriging of the grades. Delineation of the zones to remediate is frequently carried out by selecting the areas where the estimated grades exceed the chosen remediation grade. If the estimation is subject to large uncertainties, this selection may generate bias. Estimation of the probability that the true grade is greater than the remediation value makes it possible to take into account the uncertainties associated to the estimated grades. Moreover, it is necessary to specify the support to be retained for this selection, which differs generally from the support of the samples. Neglecting this support effect leads to bias in the calculation of the soil volumes to remediate, as the proportion of the values exceeding a given grade varies with the support size (samples, blocks of various sizes). In this paper, conditional expectation and disjunctive kriging are compared for the estimation of the probability to exceed a threshold on blocks. The evaluation of polluted sites is then improved using a consistent methodology.
1. INTRODUCTION

Nowadays, the stake of site investigation and remediation is high, due to the number of polluted sites, the sanitary risks they might represent and the high remediation costs. In France, site investigation is based on the estimation of the pollutants grades and the delineation of polluted areas. At this level, any error may lead to serious consequences, in sanitary and/or financial terms. However, despite the technical and financial investments, these estimations are often empirical. The knowledge of the pollution is usually derived from the historical information (often incomplete), which is occasionally verified by a few samples. Consequently, the number of soil samples and the resulting analyses of pollutants grades are usually scarce. In addition, a large grade variability is classically observed, even at small scale.

So, in order to improve the knowledge of the pollution without increasing too much the costs, there is an important need of additional information: historical information, organoleptical measures and soils information, or their combination. Besides, other pollutants easier to sample and analyze, or semi-quantitative in situ measures, might be of interest. The choice of a relevant auxiliary variable is of importance. Then, a modelling of the spatial relationship between the pollutant and the auxiliary variables is necessary.

Before site remediation, it is important to know which areas have to be treated depending on the chosen level of intervention. To achieve this task, a method consists in selecting the estimated values exceeding the intervention level by thresholding the grade estimate. In the case of an inaccurate estimate, it is well known that this kriged map excessively smoothes the always existing local variability that scarce data do not allow to reproduce. Therefore, using this kriged map to delineate polluted areas would potentially lead to an important bias. In order to take into account the lack of precision of kriging, it is useful to add to this estimate the probability that the true (unknown) grade exceed the intervention level. This probability will give access to the selection of areas where the pollutant grade exceed the intervention level, while knowing the risk to leave in place grades that are above the level (a risk which always exists).

Although frequently used, we will not discuss non parametric methods such as indicator kriging, due to the loss of information they imply, and the lack of consistency when considering successively several indicators. Indicator cokriging aims at minimizing the previous drawbacks by considering simultaneously the indicators at several cut-off values (Goovaerts, 1997, p.297). The larger the number of indicators, the smaller the loss of information, but also the heavier the modelling effort, as we need to model the covariance and cross-covariance functions of all the indicators.