Indicateurs géostatistiques de la pollution dans les cours d'eau

Caroline BERNARD-MICHEL

Thèse de doct. en géostatistique, Ecole des Mines de Paris, 2006

Résumé

La qualité chimique des cours d'eau est mesurée par un réseau de stations constitué progressivement depuis une quarantaine d'années, qui fournit une base de données très riche, mais très hétérogène.

Le système d'évaluation de la qualité de l'eau préconise actuellement de synthétiser les mesures par station par des indicateurs statistiques tels que la moyenne arithmétique et le quantile 90. Ces calculs reposent sur deux hypothèses implicites mais erronées : l'indépendance des mesures et la stationnarité des concentrations durant l'année. En effet, les concentrations en nutriments (nitrates, nitrites et orthophosphates) présentent généralement des variations saisonnières et les variogrammes expérimentaux mettent en évidence une corrélation temporelle. Dans une première partie, nous examinons les biais et les incertitudes des indicateurs actuels. Le krigeage prend en compte l'irrégularité de l'échantillonnage et les corrélations temporelles dans l'estimation et dans le calcul d'incertitude associé. Une simplification par segments d'influence est proposée. Le biais du quantile empirique est réduit par une simple interpolation linéaire. L'apport de ces méthodes est étudié théoriquement et par simulation sur le bassin Loire Bretagne qui présente des stratégies d'échantillonnage très différentes selon la station et l'année.

Pour interpoler ces indicateurs le long des cours d'eau, la question se pose ensuite de modéliser leur corrélation spatiale. Or les modèles usuels de covariance, développés pour des espaces euclidiens, ne sont plus nécessairement valables sur une structure arborescente. Un modèle général de fonctions aléatoires le long d'un réseau hydrographique a donc été développé. En tout point du réseau, les cours d'eau sont considérés comme la combinaison de filets « élémentaires » définis par les chemins de l'ensemble des sources à l'exutoire. Les questions d'indépendance des fonctions aléatoires entre filets et de leur stationnarité sont discutées et l'inférence de ce modèle est examinée sur le bassin de la Moselle.

Abstract

In order to assess river quality, different parameters such as nutrients concentrations are measured in different monitoring stations, setting up a very important but heterogeneous database.

The French evaluation system of water quality recommends to summarize the information contained in these measurements by a few statistical indicators such as the annual mean of concentrations or the 90% quantile. They are estimated using the classical statistical inference based on hypothesis proved to be incorrect: time correlations and seasonal variations are ignored. Actually, in France, nitrate concentrations are generally higher in winter. Biases and confidence intervals can be reduced by kriging or segments of influence and a linear interpolation of the empirical quantile is proposed. Methods are analyzed theoretically and experimentally on the Loire Bretagne basin.

Estimating indicators along a stream network then requires specific models of random functions because usual covariance models are no longer valid on such structures. We propose a global model of random functions along a tree graph introducing the concept of "elementary thin streams", defined by the whole set of paths between sources and outlet. At each point of the network, the river is considered to be the linear combination of these streams on which one dimensional stationary random functions are defined. An application to water discharge on the Moselle Basin (north-east of France) is presented