

# MODÉLISATION GÉOSTATISTIQUE DES DÉBITS LE LONG DES COURS D'EAU.

C. Bernard-Michel ‡(actuellement à †) & C. de Fouquet ‡

† *MISTIS, INRIA Rhône-Alpes.*

*655 avenue de l'Europe, 38334 SAINT ISMIER Cedex.*

‡ *Ecole des Mines de Paris, Centre de Géosciences.*

*35, rue Saint Honoré, 77305 Fontainebleau Cedex.*

## Résumé

L'estimation géostatistique des débits ou des concentrations le long des cours d'eau nécessite des modèles spécifiques de fonctions aléatoires, les modèles usuels définis pour des espaces euclidiens n'étant plus nécessairement valables pour la topologie définie sur des graphes. De plus, ces modèles doivent prendre en compte la forte non stationnarité des débits de l'amont vers l'aval du réseau hydrographique, ainsi que la conservation des débits aux confluences. Nous proposons un modèle de fonctions aléatoires définies sur l'arbre décrivant le réseau hydrographique, en introduisant le concept de filets élémentaires définis à une dimension sur les rivières (chemins allant des sources à l'exutoire). En chaque point, le débit est alors la somme des débits élémentaires définis sur les rivières issues des sources à l'amont de ce point ; le débit spécifique ou la concentration s'obtiennent par combinaison des filets élémentaires, de façon à assurer la conservation des débits ou des flux (de Fouquet et Bernard-Michel, 2006). L'étude des propriétés d'ordre 2 du modèle est présentée ainsi qu'une application à l'étude des débits du bassin de la Moselle (100 confluences, 100 stations).

**Mots-clés:** réseau hydrographique, débit, débits, spécifiques, géostatistique, graphe, arbre, covariance, variogramme, concentration.

## Abstract

Estimating discharges along a stream network requires specific models of random function because usual covariance model are no longer valid on such structures. Moreover, variables are generally highly non stationary because of the discontinuity at each confluence but also because of relationships with soil properties, that have to be taken into account. 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 (de Fouquet and Bernard-Michel, 2006). 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. In practise, the coefficients of the linear combination are determined according to the conservation equation of discharge and flux at the forks. An application to water discharge on the Moselle Basin (north-east of France) is presented. The hydrographic network of this Basin is made of about 100 important nodes and

100 monitoring stations are available for the last 10 years.

**Key-words:** hydrographic networks, water discharge, geostatistics, covariance, streams, graph, tree.

## 1 Introduction

Estimer les débits le long d'un réseau hydrographique nécessite de prendre en compte la géométrie de ce réseau. Or les modèles géostatistiques usuels ont été développés pour des espaces euclidiens. En effet, la distance euclidienne intervient explicitement dans les théorèmes de Bochner et de Schoenberg qui donnent la caractérisation spectrale des covariances et des variogrammes (Chilès et Delfiner, 1999). Sur un arbre, la distance considérée n'étant plus euclidienne, les modèles usuels ne sont plus nécessairement admissibles, à l'exception du variogramme linéaire et de la covariance exponentielle (Ver Hoef et al., 2006 ; de Fouquet et Bernard-Michel, 2006). Nous présentons un modèle général de covariances ou de variogrammes sur un arbre pour des variables non stationnaires, avec une application à l'étude des débits du bassin de la Moselle. La modélisation est guidée par une étude exploratoire préalable. Dans la suite, la variable «débit» désigne en fait la moyenne annuelle du débit.

## 2 Non stationnarité des débits

L'analyse des données de débit du réseau hydrographique drainé par la Moselle confirme l'intuition physique : le débit augmente de l'amont vers l'aval. Pour le bassin de la Moselle, le variogramme est linéaire pour la distance «au fil de l'eau» (Fig.1(a)); le nuage de corrélation (Fig.1(b)) montre une forte croissance du débit en fonction de la surface drainée. Cette non stationnarité résulte en particulier de l'addition des débits aux confluences: à l'aval immédiat d'une confluence, le débit est la somme des débits des différents affluents à l'amont immédiat de la confluence.

Notons  $d_3$  le débit à la confluence (modélisé par une variable aléatoire  $D_3$ ),  $d_2$  et  $d_1$  les débits (positifs) immédiatement à l'amont (modélisés par des variables  $D_1$  et  $D_2$ ). La relation (2.1)

$$D_3 = D_1 + D_2 \quad (2.1)$$

implique la non stationnarité (2.2) de l'espérance et de la variance de l'amont vers l'aval, sous hypothèse d'indépendance de  $D_1$  et  $D_2$  :

$$E(D_3) = E(D_1) + E(D_2) \text{ et } Var(D_3) = Var(D_1) + Var(D_2) \quad (2.2)$$

Plus généralement, l'additivité des débits aux confluences, rend non pertinente la modélisation par une fonction aléatoire stationnaire sur l'ensemble du réseau (Bernard-Michel, 2006).

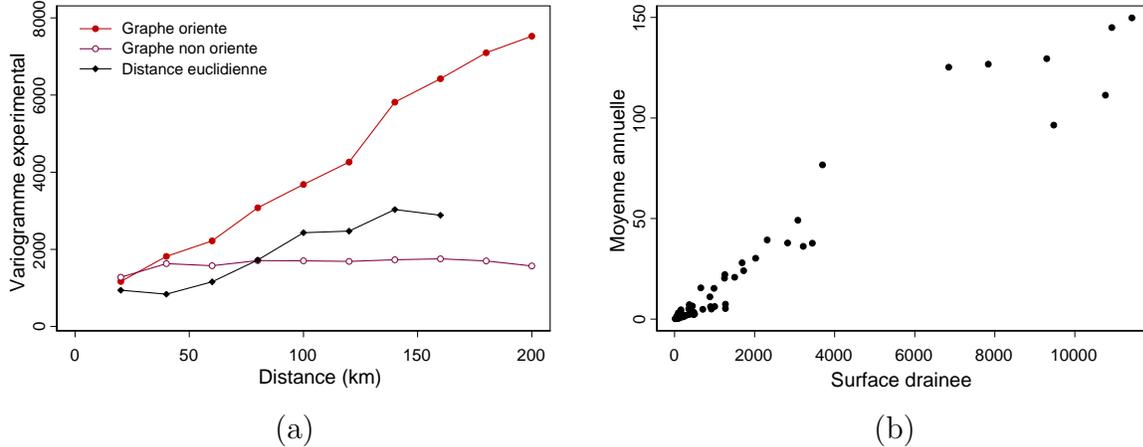


Figure 1: (a) Variogrammes des débits en 1996 sur le réseau hydrographique de la Moselle calculés pour la distance euclidienne en 2D, la distance curviligne définie sur le graphe orienté et la distance curviligne sur le graphe non orienté (b) Moyenne annuelle par station des débits en 1996 en fonction de la surface drainée

### 3 Modèle de covariance sur un arbre

Pour modéliser une fonction aléatoire sur un graphe de type arbre, nous introduisons le concept de «filets élémentaires» (voir Fig. 2).

Considérons une rivière isolée sans aucune confluence entre la source et l'exutoire. Sur cette rivière, on peut définir à une dimension une fonction aléatoire quelconque, par exemple stationnaire ou intrinsèque.

Pour un arbre à  $N$  «sources», considérons  $N$  fonctions aléatoires, indépendantes ou non, définies sur les  $N$  rivières allant des sources à l'exutoire. Ces fonctions aléatoires à une dimension, notées  $Y_j$ ,  $j \in 1, \dots, N$ , définissent des «filets élémentaires» supposés stationnaires ou intrinsèques. Le modèle consiste à écrire la variable étudiée comme une somme (pour les débits) ou une combinaison linéaire (pour les débits spécifiques ou les concentrations) des filets élémentaires. En tout point du réseau à  $n$  sources, la fonction aléatoire  $D$  décrivant le débit, le débit spécifique, ou la concentration s'écrit :

$$D(x) = \sum_{j \in 1, \dots, n} W_j(x) Y_j(x), \forall x \in R \quad (3.1)$$

où  $W_j(x)$  est le poids affecté en  $x$  au filet  $j$ . Il est nul si  $x$  n'est pas sur la rivière  $j$  ; dans le cas contraire, le poids se déduit de la conservation des débits ou des flux aux confluences. Par exemple, pour les débits,  $W_j(x) = 1$  pour toutes les rivières passant par le point  $x$  et  $W_j(x) = 0$  sinon.

## 4 Application aux débits spécifiques

Pour se ramener à des filets potentiellement stationnaires, nous proposons d'étudier le débit spécifique  $T$ , défini comme le rapport du débit  $D$  à la surface drainée  $S$  (Sauquet, 2000):

$$T = \frac{D}{S} \quad (4.1)$$

Comme le débit, le débit spécifique n'est pas modélisable par une fonction aléatoire stationnaire d'ordre 2 sur l'ensemble du réseau hydrographique. En effet, l'additivité des débits aux confluences (2.1) implique la relation (4.2) pour les débits spécifiques

$$T_3 = \frac{S_1}{S_1 + S_2} T_1 + \frac{S_2}{S_1 + S_2} T_2 \quad (4.2)$$

avec  $T_3$  et  $S_3$  le débit spécifique et la surface drainée à la confluence et  $T_1, T_2, S_1, S_2$  les débits spécifiques et les surfaces drainées immédiatement à l'amont de la confluence. L'espérance et la variance du débit spécifique sont donc discontinues aux confluences. En revanche, l'hypothèse de stationnarité des débits spécifiques entre deux confluences paraît plus réaliste que pour les débits, permettant une modélisation fondée sur des filets élémentaires stationnaires. Considérons:

- $R$  le réseau hydrographique
- $I(x)$  l'ensemble des rivières contenant  $x$
- $J(i, x)$  l'ensemble des confluences à l'amont de  $x$  et sur le filet élémentaire  $i$
- $Y_i$  les filets élémentaires. Les fonctions  $Y_i$  sont supposées stationnaires, indépendantes, d'espérance  $E(Y_i(x)) = m$  et de variance  $Var(Y_i(x)) = \sigma^2$ ,  $\forall x \in R$  constantes, et de covariance  $C_1(h)$  ou de variogramme  $\gamma_1(h)$
- $S_j^i$  la surface drainée à l'amont de la confluence  $j$  sur la branche contenant le filet élémentaire  $i$
- $S_j$  la surface drainée juste à l'aval de la confluence  $j$
- $\omega_j^i = \frac{S_j^i}{S_j}$  le poids affecté au filet élémentaire  $i$  immédiatement à l'aval de la confluence  $j$
- $W(i, x)$  le poids affecté en  $x$  au filet élémentaire  $i$
- $K(x, y) = \frac{W(i, y)}{W(i, x)}$  pour  $y$  à l'aval de  $x$

Alors, le débit spécifique  $T(x)$ ,  $x \in R$  s'écrit:

$$T(x) = \sum_{i \in I(x)} W(i, x) Y_i(x) \text{ avec } W(i, x) = \begin{cases} \prod_{j \in J(i, x)} \omega_j^i \\ 1 \text{ si } J(i, x) = \{\emptyset\} \end{cases} \quad (4.3)$$

Pour un réseau à deux confluences (voir Fig. 2(b)), on a par exemple:

$$T(x) = \frac{S_2^1 S_1^1}{S_2 S_1} Y_1(x) + \frac{S_2^2 S_1^2}{S_2 S_1} Y_2(x) + \frac{S_2^3}{S_2} Y_3(x) \quad (4.4)$$

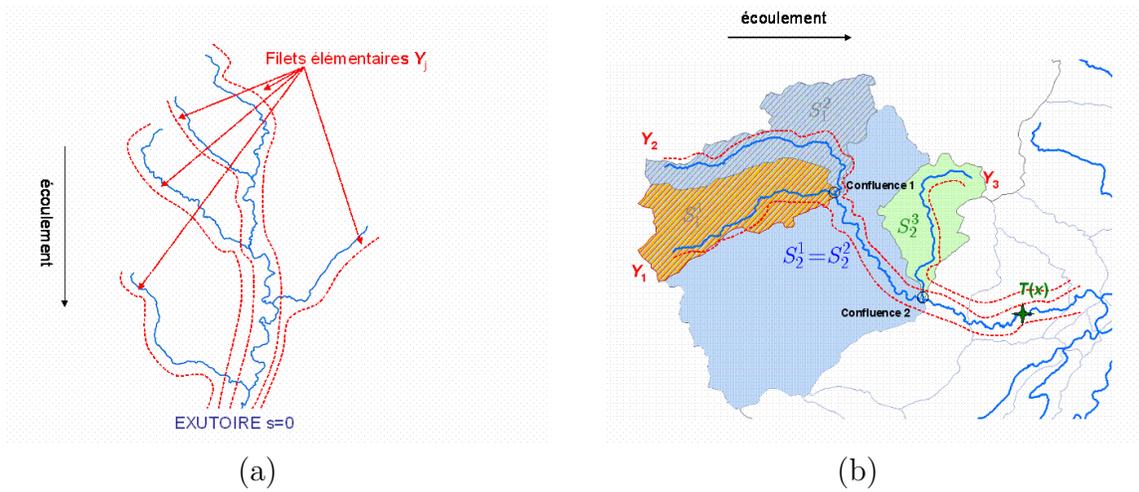


Figure 2: (a) Filets élémentaires le long d'un réseau hydrographique (b) Exemple de réseau hydrographique avec deux confluences

## 4.1 Propriétés du modèle

Dans cette construction, les propriétés statistiques d'ordre 2 du modèle résultant s'obtiennent facilement. Les paramètres de la covariance  $C$  ou du variogramme  $\gamma$  peuvent alors être inférés par maximum de vraisemblance (Ver Hoef, 2006).

### 4.1.1 Moyenne et variance

$$E(T(x)) = m \text{ et } Var(T(x)) = \sigma^2 \sum_{i \in I(x)} W(i, x)^2 \quad (4.5)$$

### 4.1.2 Covariance et variogramme

On note  $x \sim y$  si  $x$  et  $y$  sont connectés au fil de l'eau,  $x$  étant à l'amont de  $y$ .

$$C(x, y) = \begin{cases} C_1(0) \sum_{i \in I(x)} W(i, x)^2 & \text{si } x = y \\ C_1(|x - y|) \\ \quad \times \sum_{i \in I(x) \cap I(y)} W(i, x)W(i, y) & \text{si } x \sim y \\ 0 & \text{sinon,} \end{cases} \quad (4.6)$$

$$Var(T(y) - T(x)) = \begin{cases} 0 & \text{si } x = y \\ 2\gamma(y - x)[K(x, y) \sum_{i \in I(x)} W(i, x)^2] \\ + \sigma^2 [ \sum_{i \in I(y) \setminus I(x)} W(i, y)^2 + (K(x, y) - 1)^2 \sum_{i \in I(x)} W(i, x)^2 ] & \text{si } x \sim y \\ \sigma^2 ( \sum_{i \in I(y)} W(i, y)^2 + \sum_{i \in I(x)} W(i, x)^2 ) & \text{sinon,} \end{cases}$$

## 5 Conclusion

Pour le bassin de la Moselle, les hypothèses d'indépendance et de stationnarité des filets élémentaires ne sont vérifiées que pour certaines années (Bernard-Michel, 2006). De plus, l'inférence du modèle est compliquée par le nombre trop faible de stations entre deux confluences (le plus souvent, une seule station). D'autre part, il serait préférable de tenir compte de certaines variables liées aux débits, notamment la pluviométrie. Ce modèle constitue donc une première approche géostatistique à l'étude des débits, des concentrations ou des flux dans les cours d'eau. Il permet la construction d'un modèle de covariance valide pour un réseau hydrographique. A terme, il serait souhaitable d'associer modélisation géostatistique et modélisation phénoménologique.

## Références

- [1] de Fouquet C. and Bernard-Michel C. (2006), Modèles géostatistiques de concentrations ou de débits le long des cours d'eau, *Comptes rendus Géosciences* (338), 307–318.
- [2] Bernard-Michel C. (2006), Indicateurs géostatistiques de la pollution dans les cours d'eau, Thèse de doctorat, Ecole des Mines de Paris, Centre de Géosciences.
- [3] Chilès J-P. and Delfiner P. (1999), *Geostatistics: modeling spatial uncertainty*. New-York, Wiley.
- [4] Sauquet E., Gottschalk L. and Leblois E. (2000), Mapping average annual runoff: a hierarchical approach applying a stochastic interpolation scheme, *Hydrological sciences Journal* 45, 799–815.
- [5] Ver Hoef J. M., Peterson E. and Theobald D. (2006), Spatial statistical models that use flow and stream distance, *Environmental and Ecological Statistics* 13(4), 449–464.