HAL open science

Apport croisé de la modélisation déterministe et géostatistique. Exemple des concentrations en nitrates de la Seine.

Chantal de Fouquet, Edwige Polus, Nicolas Flipo, Michel Poulin

To cite this version:

Chantal de Fouquet, Edwige Polus, Nicolas Flipo, Michel Poulin. Apport croisé de la modélisation déterministe et géostatistique. Exemple des concentrations en nitrates de la Seine.. 41èmes journées de statistiques, May 2009, Bordeaux, France. hal-00361354

HAL Id: hal-00361354

https://hal.science/hal-00361354v1

Submitted on 13 May 2009

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.

Manuscrit auteur, publié dans "41èmes journées de statistiques, Bordeaux : France (2009)"

hal-00361354, version 1 - 13 Feb 2009

TL - 41èmes Journées de statistiques, 25-29 mai 2009 à Bordeaux

Apport croisé de la modélisation déterministe et géostatistique. Exemple des concentrations en nitrates de la Seine

Chantal de Fouquet, Edwige Polus, Nicolas Flipo, Michel Poulin

Introduction

Les stations de mesure des concentrations (en nitrates et autres substances) sont espacées le long des cours d'eau, avec rarement plus d'une station entre confluences. Les mesures sont hétérogènes (différentes méthodes, différents organismes commanditaires), et les débits sont mesurés en d'autres stations, dites de jaugeage. Tout ceci complique le calage des modèles géostatistiques nécessaires à l'estimation des concentrations ou des flux le long des réseaux hydrographiques (Bernard-Michel 2006 ; de Fouquet et Bernard-Michel, 2006 ; Bernard-Michel et de Fouquet, 2006).

Le recours à un modèle déterministe permettrait de remédier à la rareté des mesures, en fournissant sur tout le domaine simulé (au pas de discrétisation près) débit, concentrations et flux de manière exhaustive. Un tel modèle déterministe est-il effectivement utilisable pour guider la modélisation probabiliste (Polus-Lefèbvre et al., 2008)? Nous examinons ici les résultats du modèle PROSE (Even et al., 2007) qui décrit la Seine et la Marne peu avant leur confluence jusqu'à Poses, soit un domaine d'environ 250km de l'amont à l'aval (Figure 1).

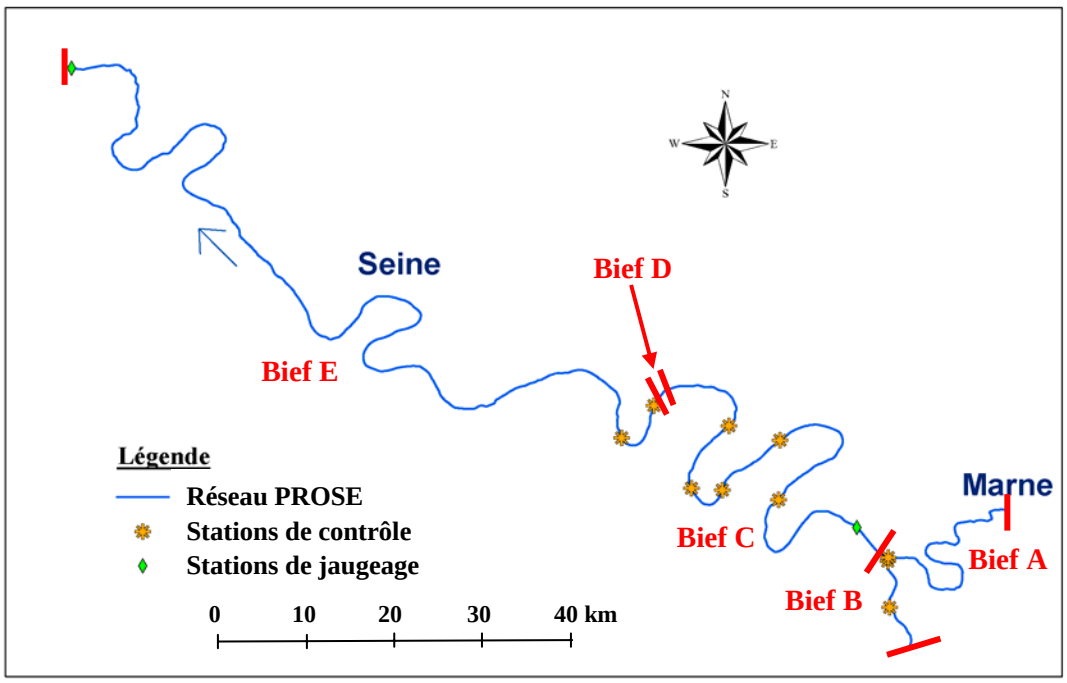

Figure 1 : Domaine modélisé par PROSE, délimité en cinq biefs.

Temporellement, les concentrations en nitrates mesurées aux stations de contrôle sont comparées à celles calculées par le modèle PROSE aux mêmes points. La stratégie d'échantillonnage temporel des mesures est reproduite sur les sorties de PROSE, de manière à assurer la pertinence des comparaisons.

Spatialement, nous nous intéressons tout d'abord aux segments entre confluences. Sur le linéaire modélisé, cinq biefs ont été délimités. Leurs extrémités sont les limites du domaine, les deux confluences majeures

hal-00361354, version 1 - 13 Feb 2009

rencontrées (Seine-Marne : limite amont du bief C ; Seine-Oise : limite amont du Bief E), ainsi que la station d'épuration Seine aval (limite amont du Bief D), en raison des discontinuités de débit et de concentrations en diverses substances qu'elle introduit.

Comparaison du modèle aux mesures : recalage des conditions aux limites

La comparaison des résultats de PROSE aux concentrations mesurées aux stations de contrôle montre des différences notables : le modèle semble être légèrement biaisé par rapport aux mesures (Figure 2a). Or les données utilisées comme conditions aux limites amont du modèle déterministe n'ont pas la même origine que les mesures de contrôle, et les protocoles de mesure respectifs diffèrent sensiblement. Le cokrigeage entre mesures de différents types permet de recaler les conditions amont, améliorant ainsi nettement la concordance entre PROSE et les mesures de contrôle. La Figure 2 présente les statistiques des écarts entre valeurs PROSE et mesures aux différentes stations de contrôle, avant et après recalage des conditions amont.

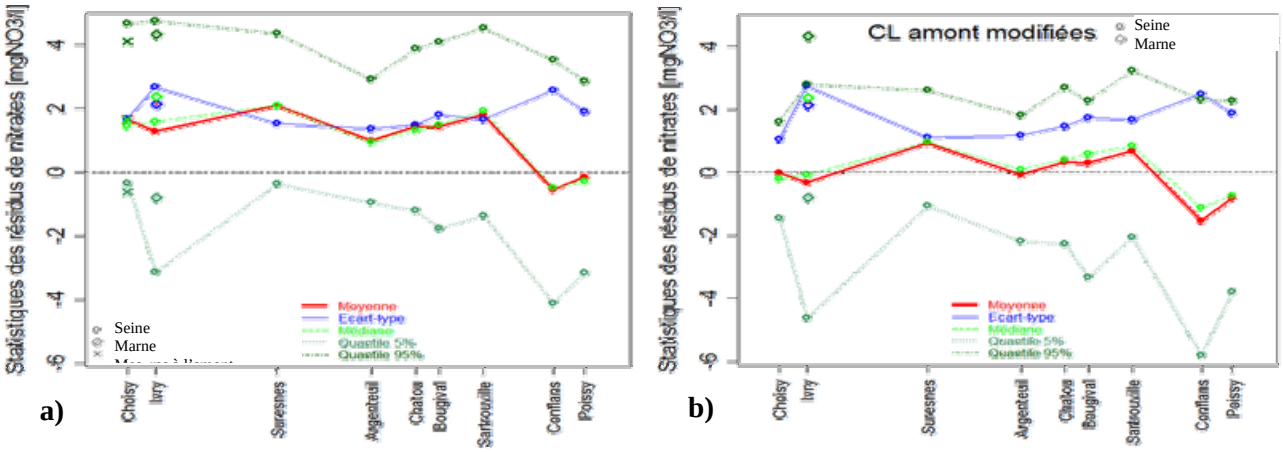

Figure 2 : Statistiques classiques des écarts entre valeurs PROSE et mesures aux stations de contrôle, avant (a) et après (b) recalage des conditions aux limites amont.

Modélisation temporelle de la liaison entre modèle et mesures

Temporellement, les variations des concentrations en nitrates sont bien reproduites par le modèle aux différentes stations. L'exemple de Suresnes, la première station à l'aval de la confluence Seine/Marne est présenté à la Figure 3.

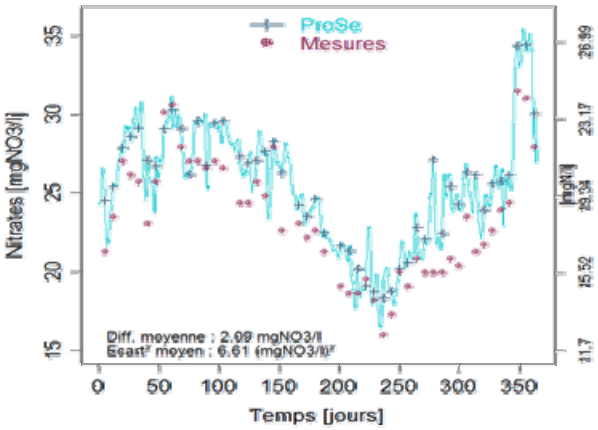

Figure 3 : Comparaison des chroniques temporelles modélisées et mesurées à Suresnes.

hal-00361354, version 1 - 13 Feb 2009

Par station, les variogrammes temporels simples et croisés entre modèle déterministe et mesures sont ajustables dans le cadre du modèle linéaire de corégionalisation (Figure 4 : l'exemple de Suresnes). Si le modèle déterministe reproduit correctement le phénomène temporel, il modifie cependant l'amplitude des fluctuations aux différentes échelles temporelles.

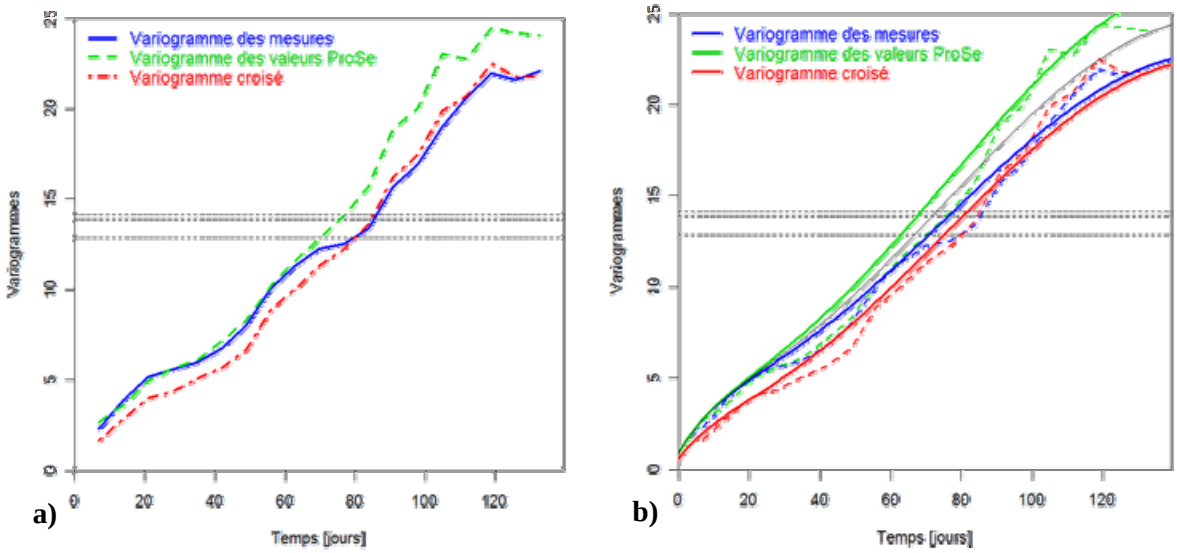

Figure 4 : Variogrammes expérimentaux simples et croisé des valeurs PROSE et mesures (a), et ajustement d'un modèle linéaire de corégionalisation (b).

Modélisation spatiale de la liaison entre modèle et mesures

Spatialement, les résultats apparaissent plus complexes. Si les variogrammes simples des résultats de PROSE et des données sont assez analogues, le variogramme croisé reste faible (Figure 4). PROSE est alors interprétable en termes de simulation non conditionnelle, puisqu'il est calé aux données uniquement à l'amont. Les mesures de contrôle permettront un recalage du modèle déterministe, analogue à la méthode classique de « repiquage des résidus » (Chilès et Delfiner, 1999).

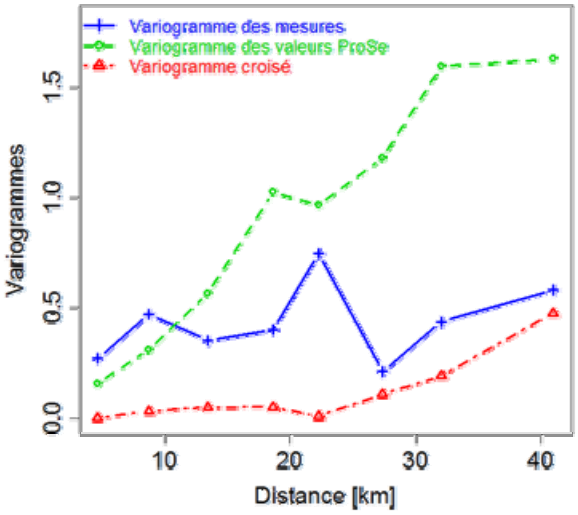

Figure 4 : Variogrammes spatiaux simples et croisé des valeurs PROSE et mesures de concentrations en nitrates.

hal-00361354, version 1 - 13 Feb 2009

Inférence des variogrammes spatiaux sur la « maquette » PROSE

Grâce à l'analogie entre les variogrammes simples des concentrations mesurées ou calculées par le modèle PROSE, ce modèle est utilisé pour l'inférence des variogrammes simples et croisés entre concentrations, débits et flux. Spatialement, ces variogrammes sont d'abord calculés par bief, afin d'éviter les discontinuités aux confluences.

Sur le domaine étudié, les débits et concentrations en nitrates sont deux grandeurs globalement peu variables, qui présentent des relations différentes selon les biefs. En général, une de ces deux variables peut être considérée comme quasiment constante, et logiquement, les variations du flux sont calquées sur les variations de l'autre variable. La Figure 5 présente l'évolution des débits, flux et concentrations en nitrates le long des biefs A et E.

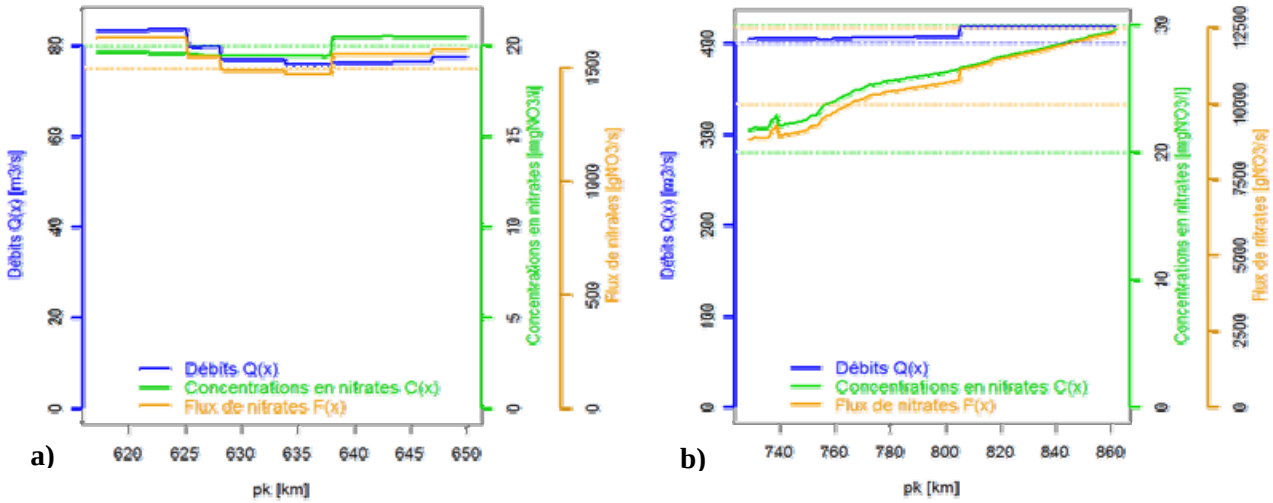

Figure 5 : Evolution des débits, flux et concentrations en nitrates le long du bief A (a : concentrations ≈ constantes), et du bief E (b : débits ≈ constants).

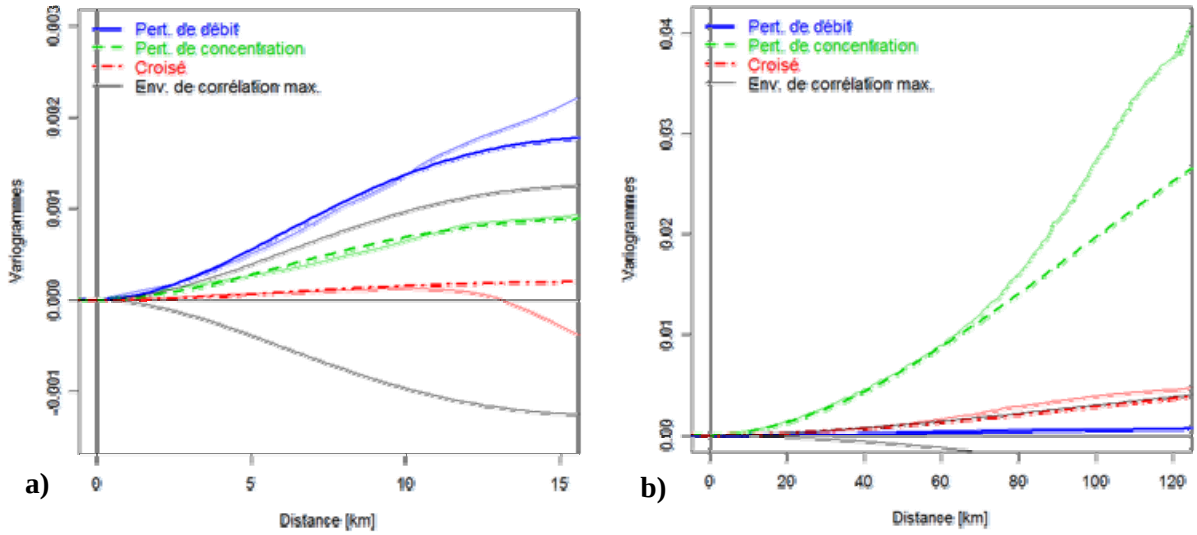

Figure 6 : Ajustement d'un modèle linéaire de corégionalisation aux perturbations de concentrations et de débits, pour les biefs A (a) et E (b).

Pour le bief A, les variations de concentration sont négligeables par rapport aux variations de débit, ces dernières étant suivies (grossièrement) par celles du flux. En revanche, le long du bief E, les variations de débit sont quasiment négligeables en regard des variations de concentrations. Ces observations sont très utiles pour guider la modélisation géostatistique.

Pour le domaine étudié, un modèle fondé sur une hypothèse de petites perturbations des débit et des concentrations est ajusté, dans le cadre du modèle linéaire de corégionalisation (Figure 6). L'ajustement trivariable des variogrammes des débits, flux et concentrations en nitrates se révèle très satisfaisant.

Perspectives

On dispose ainsi des éléments pour construire un modèle spatio-temporel en vue de l'estimation des concentrations, débits et flux en tout point du réseau hydrographique, à partir des mesures. Un krigeage bivariable entre les mesures et le modèle déterministe qui décrit les singularités du réseau permettra d'améliorer les estimations.

La décomposition des observations comme la somme de la « prédiction » donnée par le modèle, et des erreurs d'observation et de modélisation non corrélées au modèle, décomposition classique en assimilation de données, apparaît inadaptée.

Bernard-Michel C., de Fouquet C., 2006. Construction of valid covariances along a hydrographic network. Application to specific water discharge on the Moselle basin. IAMG'06: quantitative geology from multiple sources. Liège, Belgium.

Bernard-Michel C. 2006. Indicateurs géostatistiques de la pollution dans les cours d'eau. Thèse de doctorat en Géostatistique, Ecole des mines de Paris.

de Fouquet C., Bernard-Michel C. 2006. Modèles géostatistiques de concentrations ou de débits le long des cours d'eau. *Comptes-rendus Géosciences. 338(5) 307-318*

Chilès J.-P., Delfiner P. 1999. Geostatistics : modelling spatial uncertainty. N.Y. Wiley.

Even S., J.-M. Mouchel, P. Servais, N. Flipo, M. Poulin, S. Blanc, M. Chabanel et C. Paffoni (2007) Modeling the impacts of Combined Sewer Overflows on the river Seine water quality. *Sci. Total. Environ.*, n°375, p.140–51.

Polus-Lefebvre E., de Fouquet C., Bernard-Michel C., Flipo N., Poulin M. 2008. Geostatistical Model for concentrations or flow rates in streams: some results. In *Geostats 2008, proceedings of the 8th international geostatistics congress*, 1-5 dec. 2008, Santiago, Chile. Ortiz J.-M., Emery X. eds. 871-880

hal-00361354, version 1 - 13 Feb 2009