



## Combining geostatistics and process-based water quality model to improve estimation along a stream network. Example of dissolved oxygen in the Seine River from the Greater Paris to its estuary.

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

Geosciences Mines ParisTech Fontainebleau France

Models that estimate pollutant concentrations in streams can roughly be classified into two categories: physically-based and stochastic models. The first ones tend to reproduce physical processes that occur in streams and provide dense information. But that information mostly does not suit the measurements, due to necessary approximations within modelling such complex natural systems. The stochastic models - and particularly the geostatistical ones this work is interested in - are only based on observations. They consider concentrations as random variables and provide uncertainty quantification together with the estimation: the estimation variance. But the environmental monitoring network usually provides too few measurements to obtain a relevant estimation or even to infer the geostatistical model.

This paper aims at combining both approaches to improve the water quality characterization.

First a comparison of measurements and model outputs is performed to ensure that the process-based model can be admitted as replica of reality. If so, its outputs are used to better understand the behaviour and variability of the variable of interest, and to choose and validate an appropriate class of geostatistical models. Then according to the consistency of data sets (measurements and model outputs), a geostatistical method that allows for combining different kind of data is chosen and finally the multivariate estimation provides a measurement interpolation based on process-based model outputs.

The reasoning is applied to dissolved oxygen concentrations for the modelled area of the process-based ProSe model, i.e. the downstream part of the Seine River, France. This stretch of the River Seine is sampled monthly in 25 sites (National Control Network named RCS). Oxygen is chosen because it is a global indicator of the river water quality, and because usually its time and space variations are neither well-known neither negligible, especially during summer when the levels are likely to become critical.

For summer concentrations, dams and waste water treatment plants appear to be significant singularities that act like geological faults in mining applications of geostatistics: concentrations are translated but their variability does not vary significantly from one side to the other. A co-kriging calculation including these « faults » provides an estimation of dissolved oxygen summer concentrations that goes through measurements and that presents the variability of ProSe outputs between measurements, together with a joint uncertainty quantification.