Kriging prediction of curves when spatial data are curves: an extension of cokriging to functional data

Pascal MONESTIEZ[1] and David NERINI[2]

(1) Unité Biostatistique & Processus Spatiaux, INRA Avignon, Domaine Saint-Paul, Site Agroparc 84914 AVIGNON cedex 9, France
Corresponding author : monestiez@avignon.inra.fr
Tel: (33) 4 32 72 21 87 Fax: (33) 4 32 72 21 82

(2) Laboratoire de Microbiologie, Géochimie et Ecologie Marines,
Centre d'Océanologie de Marseille, UMR 6117 CNRS, Campus de Luminy, Case 901,
13288 MARSEILLE Cedex, France

In environmental sciences, many data are curves or can be considered as curves. For instance in soil science, conductivity response curves can be associated to spatial soil samples; in atmospheric research, meteorological variables or pollutant concentration are often recorded by radiosonds giving vertical profiles and, in oceanography, surveys provide vertical profiles of temperature, salinity or other variables that are spatially dependent and sampled along the depth. In geostatistics, there is two usual way to analyse such data; the first one is to consider the profile as a third dimension which is often problematic due to strong and complex anisotropy and to non-stationarity along the vertical dimension. The second way is to discretize the curves and to model them in multivariate geostatistics as in Goulard and Voltz (1992). The latter approach suffers several drawbacks; data analysis does not consider the variable ranks along the profile and is rapidly limited when profiles are finely recorded.

We propose a method which generalizes the coregionalization approach and takes into account the functional nature of the data. Suppose that we dispose of a spatial sample $L = \{Y_{x_i},\, i = 1, ..., n\}$ such that each Y_{x_i} is a functional random variable at location x_i, valued in an Hilbert space $\mathcal{H}$ on support $[a, b] \in \mathbb{R}$. We assume that each random function Y_{x_i} belongs to a spatial stationary random field. Such hypothesis impose the definition of variance-covariance operators working in spaces of infinite dimension through a scalar product and a norm on $\mathcal{H}$. This theoretical work has been widely developed in Ramsay & Silverman, 1997 and references therein and is extended here to the spatial covariance function or variogram. Our aim is then to estimate the function Y_0 at location x_0 using the set of Y_i at location x_i by designing a kriging predictor such that

$$\widehat{Y}_0\,(t) = \sum_{i=1}^{n} \int_a^b \beta_i\,(s,t)\, Y_i\,(s)\, ds,\ \forall t \in [a,b]\,. \tag{1}$$

where the $\beta_i\,(\cdot,\cdot)$ are weight functions which verify that prediction $\widehat{Y}_0$ is unbiased and the error variance, in sense of the variance operator, is minimum.

In this paper, we show that for suitable form of weight functions $\beta_i\,(\cdot,\cdot)$ the problem of functional kriging can be reduced to an ordinary cokriging problem. When each function Y_i is decomposed in terms of linear combinations of known basis functions $\phi_1, ..., \phi_k$ such that

$$Y_i\,(t) = \sum_{k=0}^{K} \alpha_{ik}\phi_k\,(t) + \varepsilon_i\,(t)\,,$$

finding the $\beta_i\,(s,t)$ in the above problem and predicting the profile $\widehat{Y}_0(t)$ can be achieved by the cokriging of a finite number of coefficients α_{0k} at location x_0.

The method is illustrated with data in oceanography. We dispose of temperature profiles in Antarctic Ocean recorded by elephant seals equipped by Argos tracking device and data logger. The seal diving activity acts as a sampler and for technical reason only a limited number of sampled profiles can be sent to the satellite. From these data and after decomposition of reconstructed profiles on a basis of orthogonal polynomials, we fitted a Linear Model of Coregionalization (LMC) on polynomial coefficients in order to predict them by cokriging. Patterns in temperature profiles can then be reconstructed as polynomial curves anywhere in the surrounding of animal trajectory.

M. Goulard and M. Voltz (1992) Linear coregionalization model: tools for estimation and choice of multivariate variograms. *Mathematical Geology*, **24**, 269-286.

Ramsay, J. and B. Silverman (1997). *Functional data analysis.* Springer.