The following supplement accompanies the article

Spatial and temporal interaction between sediment and microphytobenthos in a temperate estuarine macro-intertidal bay

F. Orvain[1,2,*], S. Lefebvre[1,3], J. Montepini[1], M. Sébire[1], A. Gangnery[4], B. Sylvand[5,6]

[1]Université de Caen Basse-Normandie – FRE3484 BIOMEA CNRS, Esplanade de la Paix, BP 5186, 14032 Caen Cedex, France
[2]CNRS, UMR 7208 BOREA, Muséum d'histoire Naturelle, CRESCO, 38 Rue du Port Blanc, 35800 Dinard, France
[3]Université de Lille1 Sciences et Technologies- CNRS, UMR 8187 Laboratoire d'Océanologie et Géosciences, Station Marine de Wimereux, 28 Avenue Foch, 62930 Wimereux, France
[4]Ifremer-Laboratoire Ressources Halieutiques, Station de Port-en-Bessin, Avenue du Général de Gaulle, 14520 Port-en-Bessin, France
[5]GEMEL-Normandie, Station Marine de Luc-sur-Mer, B.P. 49, 54 rue du Dr Charcot, 14530 Luc-sur-Mer, France
[6]M2C-Morphodynamique Continentale et Côtière, 2-4 Rue des Tilleuls, Université de Caen, 14032 Caen Cedex, France

*Email: francis.orvain@unicaen.fr

Marine Ecology Progress Series 458: 53–68 (2012)

Supplement. Detailed equations for variogram and kriging methods, and detailed kriging results.

Spatial analysis and kriging.

Several procedures can be used for kriging (ordinary kriging, universal kriging, cokriging). These methods all produce interpolated maps upon the basis of the spatial autocorrelation structure. For producing interpolated maps, a fine systematic grid between each interpolated point (or 'node') is applied and the kriging method evaluates the interpolated value at each node. The resulting kriged maps are well recognised to provide the most powerful visualizations of spatial pattern (Legendre & Legendre 1998). The interpolation grid covers the entire study area by using a systematic grid with a step of 100 m between each point (i.e the smallest sampling interval). At each node of the grid, the interpolated value of the variable ($\hat{z}_i$) is estimated by using the observed values (z_i) of the neighbours by applying a weight (w_i) to each observation point as following:

$$\hat{z}_i = \sum_n w_i z_i \qquad \text{(S1)}$$

This weight is calculated as a function of the distance between the node of the grid and the observed points (see further section on the spatial autocorrelation analysis). This calculation is based upon the semi-variance in respect to the degree of similarity between neighbours (observed spatial structure). The modelled semi-variogram is used to quantify the weights in question and for interpolation ('kriging'), which provides optimal, unbiased estimates of points not sampled. The principle is different between the kriging methods. In the ordinary kriging method, it is assumed that the studied variable *z(x,y)* is a stationary,

regionalised random function. This method thus requires that the condition of first-order and second-order stationarity is verified. This condition refers to spatial structures that are homogeneously distributed all over the studied area without displaying either a drift or a gradient. This condition is difficult to meet in coastal ecology (at least over a wide scale), because coastal ecology studies transition areas between continental and marine ecosystems, where gradients systematically occurs. In practice, it must be known that a drift (non-stationary expectation) exists within certain zones. In such a case, the universal kriging is required.

The universal kriging method consists of taking into account the drift by using a polynomial regression as a function of the coordinates of the point within the study area (x and y). By computing the residues (i.e. the observed values from which we subtracted the drift extrapolated by the polynomial regression), the drift effect can be dropped and the second-order stationarity of the residues is verified. The simple kriging method can then be used without misleading effects to produce variograms and interpolated values of the residues. The final interpolated values are obtained by adding the values of the drift (obtained from the polynomial regression) with the value of the interpolated residue (obtained from kriging) for each node.

Three parameters were analysed to study autocorrelation (correlograms and semi-variograms) and to produce interpolated maps: chl *a* content, silt content (<63 µm), and median grain size. The normality of data was verified by using Kolmogorov-Smirnov tests and a log transformation was applied if necessary.

For each variable, polynomial regression were adjusted as a function of the coordinates (x,y):

$$z_0(\boldsymbol{x},\boldsymbol{y}) = \sum_{i=1}^{k} \boldsymbol{a}_l . \boldsymbol{f}^l \qquad \text{(S2)}$$

where k is the degree of the polynomial regression (1, 2 or 3) and $f_1=1$; $f_2=x$; $f_3=y$; $f_4=x^2$; $f_5=y^2$; $f_6=x.y$. We compared the observed values (z) and the values estimated by the polynomial regression (z_0). The correlation coefficient was calculated and its significance was tested (test F). When the correlation was significant, first-order and second-order stationarity was not guaranteed and the universal kriging was preferred.

Spatial autocorrelation

The spatial autocorrelation was examined for raw data (z) and their residues ($z - z_0$) to compare the predictability of ordinary and universal kriging at the end of the analysis. Autocorrelation was mathematically described by correlograms, which are graphs of autocorrelation coefficients as a function of the distance separating points $\boldsymbol{d}$. A n*n matrix of geographical distances $\boldsymbol{D}$ was calculated among observation sites before constructing the correlogram (Legendre & Legendre 1998). The Sturge's rule was applied to decide the number of classes $\boldsymbol{N}$ for distance intervals $\boldsymbol{d}$:

$$\boldsymbol{N} = 1 + 3.3 \log_{10}\left[\frac{\boldsymbol{n}\cdot(\boldsymbol{n}-1)}{2}\right]$$

where $\boldsymbol{n}$ is the number of observation sites (and the expression $\left[\frac{\boldsymbol{n}\cdot(\boldsymbol{n}-1)}{2}\right]$ allows calculating the number of pairs of observation). Autocorrelation coefficients were calculated at each given distance class $\boldsymbol{d}$ extending from the minimal distance between observations within the

criss-cross sampling design to half of the extent of the whole sampled area. For each specified distance class, spatial autocorrelation was then calculated by the Geary's ***C*** and Moran's ***I*** spatial autocorrelation statistics (Legendre & Legendre 1998):

$$\boldsymbol{C}(\boldsymbol{d}) = \frac{\frac{1}{2\boldsymbol{W}(\boldsymbol{d})}\sum_{\boldsymbol{h}=1}^{\boldsymbol{n}}\sum_{\boldsymbol{i}=1}^{\boldsymbol{n}}\boldsymbol{w}_{\boldsymbol{hi}}(z_h - z_i)^2}{\frac{1}{(\boldsymbol{n}-1)}\sum_{i=1}^{n}(z_i - \bar{z})^2} \quad \text{for h} \neq i \qquad \text{(S3)}$$

$$\boldsymbol{I}(\boldsymbol{d}) = \frac{\frac{1}{\boldsymbol{W}(\boldsymbol{d})}\sum_{\boldsymbol{h}=1}^{\boldsymbol{n}}\sum_{\boldsymbol{i}=1}^{\boldsymbol{n}}\boldsymbol{w}_{\boldsymbol{hi}}(z_h - \bar{z})(z_i - \bar{z})}{\frac{1}{\boldsymbol{n}}\sum_{i=1}^{n}(z_i - \bar{z})^2} \quad \text{for h} \neq i \qquad \text{(S4)}$$

where $\bar{z}$ is the averaged value of the variable of interest and z_h and z_i are the observed values of this variable at sites *h* and *i* which are separated by the distance *d*. *W(d)* is the number of observation pairs separated by distance *d*. For a given distance class *d*, the weights w_{hi} are the Kronecker delta i.e. the value $w_{hi} = 1$ when sites *h* and *i* are within the distance *d* and $w_{hi} = 0$ otherwise. In this way, only the pairs of sites (*h,i*) separated by the specified distance d are taken into account in the calculation of any given coefficient.

Moran's ***I*** is usually related to Pearson's correlation and takes values within the interval [±1]. Positive aurocorrelation in the data produces positive values of ***I*** while negative autocorrelation produces negative values. The Geary's ***C*** coefficient varies from 0 to some unspecified value larger than 1. Positive correlation produces values of *C* between 0 and 1 whereas negative autocorrelation produces values larger than 1. When there is no autocorrelation, $\boldsymbol{C} = \mathbf{1}$ and $\boldsymbol{I} = \mathbf{0}$. The 2 autocorrelation coefficients were tested for significance, and confidence intervals were computed. The distance classes showing significant positive and negative autocorrelation were determined by following a statistical procedure for testing significance under the null hypothesis of random spatial distribution or absence of autocorrelation (Legendre & Legendre 1998). The significance of coefficients was represented by circled dots on the correlograms (Figs. 2 & 4 in main text).

The semi-variance values upon observed autocorrelation to provide the most suitable model that can predict semi-variance for all intermediate distances. Autocorrelation was thus calculated using the semi-variance statistic $\gamma(h)$ for a range of distance intervals *d*:

$$\gamma(d) = \frac{1}{2\boldsymbol{W}}\sum_{\boldsymbol{h}=1}^{\boldsymbol{n}}\sum_{\boldsymbol{i}=1}^{\boldsymbol{n}}\boldsymbol{w}_{\boldsymbol{hi}}(z_h - z_i)^2 \qquad \text{(S5)}$$

The equation of the model has been chosen to guarantee positive interpolated values during ordinary kriging. The spherical, exponential, Gaussian, and linear models were thus fitted to the observed semi-variogram. All equations are detailed by Legendre & Legendre (1998). The most suitable model was retained (by analysing the least square criteria). Among others, the spherical model was adjusted:

$$\gamma(d) = C_0 + C_1\left[1.5\frac{d}{a} - 0.5\left(\frac{d}{a}\right)^3\right] \text{ if } d \leq a;\ \gamma(d) = C_0 \text{ if } d > a \qquad \text{(S6)}$$

where C_0 is the nugget effect parameter and C_1 is the spatially structured component; the sill is equal to $C_0 + C_1$. The distance at which the variance levels off is referred to as the range (parameter *a*).

We used an iterative non-linear least squares regression according to Nelder-Mead simplex method to estimate parameter values (Nelder & Mead 1965). The isotropy of the spatial structure was tested by comparing the correlograms and semi-variograms in perpendicular directions (north-south and east-west axes). Eqs. (S3), (S4) & (S5) were used to make the calculations with anisotropy along the 2 main axes north-south and east-west. For instance, for calculating statistics in the north-south direction: the Kronecker delta $w_{hi} = 1$ only when sites h and i are aligned with this direction (and always when the distances between the 2 sites are within the distance d) and $w_{hi} = 0$ otherwise. When anisotropy was well developed, the best fitted models (Eq. S6) were also calculated separately in the 2 perpendicular directions when applying the Nelder-Mead simplex methods. In this case, an anisotropic ratio was deduced from the 2 ranges calculated in each perpendicular direction according to:

$$\theta = \frac{a_{N/S}}{a_{E/W}} \qquad \text{(S7)}$$

where $a_{N/S}$ is the range obtained from the best-fitted model of variogram that was calculated along the north-south direction and $a_{E/W}$ was the same but along the east-west direction.

Interpolated maps

In the ordinary kriging method, Krige (1952) has showed that the calculation of each observation point (from 1 to $\boldsymbol{n}$) depends on (1) the vector of semi-variance $\boldsymbol{d}$ between all observation points and the grid node to be estimated, and (2) the square matrix of semi-variance between each pair of observations (squared matrix of $\mathbf{C}_{ih}$). This computation results in the matrix multiplication ($\mathbf{C = W.D}$):

$$\begin{bmatrix} c_{11} & \dots & c_{1n} & 1 \\ . & \dots & . & 1 \\ . & \dots & . & 1 \\ c_{n1} & \dots & c_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ . \\ w_n \\ \lambda \end{bmatrix} = \begin{bmatrix} D_1 \\ . \\ . \\ D_n \\ 1 \end{bmatrix} \qquad \text{(S8)}$$

where λ is the lagrange parameter introduced to minimise the variance of estimates under the constraint that $\sum \boldsymbol{w}_i = 1$. This constraint allows certifying that, during the interpolation process, the estimation error at all observation points is minimised (no difference between interpolated value and observed value). The variogram model is essential to build the matrix $\mathbf{C}$ and the vector $\mathbf{D}$ and to finally provide the weighting function of the whole map by applying the following matrix inversion:

$$\mathbf{W = D \,.\, C^{-1}} \qquad \text{(S9)}$$

where D is the variogram function of distances d between all observation points and the grid node (affected by a potential anisotropy):

$$\mathrm{D} = \gamma\left(\sqrt{x_i^2 + \theta.y^2}\right) \qquad \text{(S10)}$$

where x_i is the Euclidian distance between the x-coordinates (along east-west axis) of the grid node and observation sites and y_i is the same but for y-coordinates (along the north-south axis) affected by ϕ, the anisotropy ratio (which is equal to 1, when anisotropy does not occur).

In the universal kriging method, the unbiased estimator is affected by the trend, which is characterised by the polynomial regression (Eq. S2). The universality condition can be written:

$$\sum_{i=1}^{k} \mu_l . f_i^{l_i} - f_0^{l_0} = 0 \text{ with } l_i=1,2\ldots., k \qquad \text{(S11)}$$

f_0 corresponding to the values of z_0 at the interpolation node to take into account the drift (Eq. S2). We obtain a system of k equations with k unknown coefficients (μ_i).

To obtain the minimum value of the variance of residues, the universal kriging system can be described by the following matrix form (**C = W.D**):

$$\begin{bmatrix} \boldsymbol{c}_{11} & \boldsymbol{c}_{12} & \ldots & \boldsymbol{c}_{1n} & \boldsymbol{f}_1^1 & \boldsymbol{f}_1^2 & \ldots & \boldsymbol{f}_1^k \\ \boldsymbol{c}_{21} & \boldsymbol{c}_{22} & \ldots & \boldsymbol{c}_{2n} & \boldsymbol{f}_1^2 & \boldsymbol{f}_2^2 & \ldots & \boldsymbol{f}_2^k \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{c}_{n1} & \boldsymbol{c}_{n2} & \ldots & \boldsymbol{c}_{nn} & \boldsymbol{f}_n^1 & \boldsymbol{f}_2^n & \ldots & \boldsymbol{f}_n^k \\ \boldsymbol{f}_1^1 & \boldsymbol{f}_2^1 & \ldots & \boldsymbol{f}_n^1 & 0 & 0 & 0 & 0 \\ \boldsymbol{f}_1^2 & \boldsymbol{f}_2^2 & \ldots & \boldsymbol{f}_2^n & 0 & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & 0 & 0 & 0 & 0 \\ \boldsymbol{f}_1^k & \boldsymbol{f}_2^k & \ldots & \boldsymbol{f}_n^k & 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \\ \ldots \\ \boldsymbol{w}_n \\ \mu_1 \\ \mu_2 \\ \ldots \\ \mu_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{d}_1 \\ \boldsymbol{d}_2 \\ \ldots \\ \boldsymbol{d}_n \\ \boldsymbol{f}_0^1 \\ \boldsymbol{f}_0^2 \\ \ldots \\ \boldsymbol{f}_0^n \end{bmatrix} \qquad \text{(S12)}$$

The interpolated value is equal to: $\hat{z}_0(\boldsymbol{x},\boldsymbol{y}) = \sum_{i=1}^{k} \boldsymbol{a}_l . \boldsymbol{f} + \sum_{n} \boldsymbol{w}_i \boldsymbol{z}_i$ (instead of Eq. S1 for ordinary kriging). The variogram model (of residuals) is also used to build the matrix ***C*** and the vector ***d*** and to finally provide the weighting function of the whole map by applying the matrix inversion (as in Eq. S7).

The spatially dependent predictability of the modelled semi-variogram was assessed by analysing the goodness of fit (regression methods) and other criteria derived from cross-validation (Journel & Huijbregts 1978). Cross-validation tests were performed at the end of kriging to verify the good predictability of the interpolation for both ordinary and universal kriging (Legendre & Legendre 1998) and the choice between the 2 methods was based on the coefficient of correlation between the measured value z_i at the position (x_i,y_i) and the interpolated value $\hat{z}_{-i}$ after removal of the ith datum that was measured at the position (x_i,y_i). For all variables, there was a need to disaggregate the kriging interpolation between the 3 subdomains to take into account the discontinuities due to the channels. The comparison between universal and ordinary kriging was done for each of the 3 independent sectors and the best-fit method was retained to produce maps combining universal and ordinary (after log transformation) kriging if necessary.

Kriging results

The distribution of pooled data of chl *a* biomass was unimodal and a log transformation was required to obtain a normal distribution (p = 0.63). The model of variogram was spherical with a very little nugget effect and a range of 1354 m (Fig. 2, Table 2). This value matched well with the smallest distance for which the Geary's and Moran indices exhibited a lack of autocorrelation (1350 m). The model was very well adjusted to the experimental variogram (R^2 = 0.936), and the autocorrelation function was found to be the same at all geographic directions (Fig. 2). Concerning universal kriging, the

experimental semi-variogram that was calculated after removal of the drift (Eqs. S2, S8 & S9) was also fitted to a spherical model but not as well as with the first variogram ($R^2 = 0.454$). The normality condition was also verified in this case ($p = 0.37$). The nugget effect (1.97) was also very low compared to the sill (17.5), showing the quality of the sampling strategy. The value of the range is smaller than for ordinary kriging with a value of 504 m (Table 2).

The 2 kriging methods clearly provided different spatial structures. Some patches were not revealed in Brévands by the ordinary kriging while they appeared with the universal kriging. In Géfosses, a misleading gradient was created by the universal kriging and this gradient masked a patch located at the north of the area. The spatial patterns were different between the 3 areas as confirmed by cross-validation results (Table 1). The universal kriging provided the best interpolation in Grands Veys and Brévands while the ordinary kriging was the best fitted in Géfosses (Table 1).

For mud content of sediment, the cross-validation results displayed a better fitting by using the ordinary kriging in the 3 sub-domains with anisotropy (Table 1). However, the conditions for kriging were not well verified neither using the ordinary kriging, nor using the universal kriging. Indeed, the distribution of log-transformed silt content was clearly not normal, but this should be attributed to a significant anisotropy ($p < 0.0001$). The model of variogram was spherical with anisotropy and a small nugget effect of 0.689 (Table 2, Fig. 4). Nugget effects were significantly decreased when taking into account anisotropy. Patches were more extensively extended along the north-south axis. The averaged size of the patches can be estimated by the range of the variogram model (when using the universal kriging), which is of ca. 800 m (Table 2), but the form of these patches were affected by an anisotropic ratio θ of 0.44 (see. Eq. S7). This was also confirmed by correlograms (Fig. 4A,B) that were clearly different according the direction and also indicating the lack of significant autocorrelation for a distance higher than 700 m.

For the median grain-size, there was a significant lack of stationarity and normality on the log-transformed data (also producing a linear model of variogram, Table 2). The normality and stationarity was obtained only after removal of the drift, reinforcing the idea that universal kriging was the most appropriate for sediment (Table 1). A spherical model was well adjusted to the experimental variogram ($R^2 = 0.801$, Fig. 4F), with a range of 929 m, which was a little bit higher than the range of median diameter. The nugget effect was particularly small compared to the sill, revealing the small sampling error of this variable.