I'Analyse de Données Fonctionnelles en Océanographie

Quand les données sont des courbes

David NERINI

Mediterranean Institute of Oceanography
UMR 7294 CNRS, Aix-Marseille University

david.nerini@univ-amu.fr

Courbes échantillonnées

- En oceanographie, les données sont souvent des courbes échantillonnées

Distribution de taille de zooplancton

D. Nerini, B. Ghattas / Computational Statistics & Data Analysis 51 (2007) 4984–4993 4985

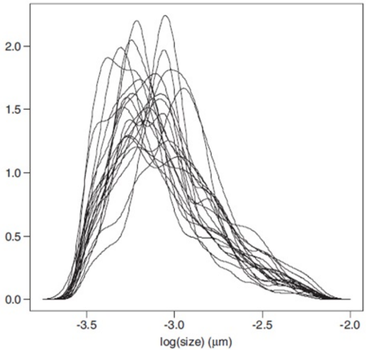
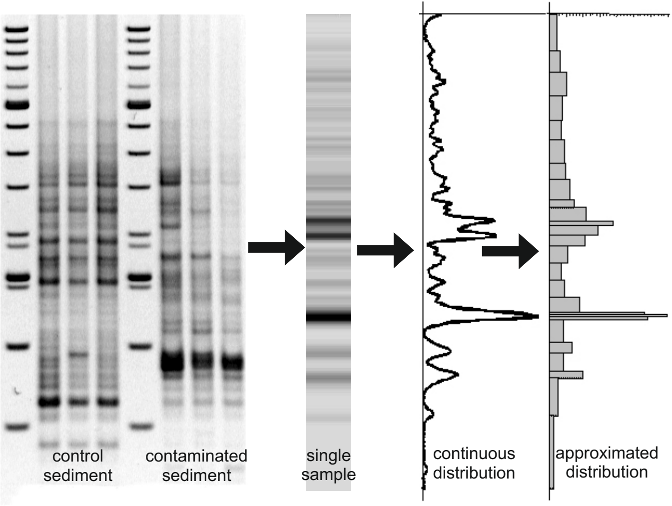


Fig. 1. Zooplankton size spectra. Each curve is a density which arises from a digital imaging system designed for estimating the size distribution of marine zooplankton net samples (Grosjean et al., 2004).

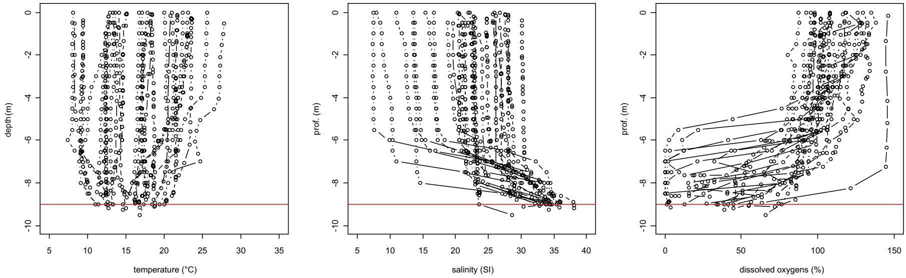
Courbes échantillonnées

Spectre d'électrophorèse



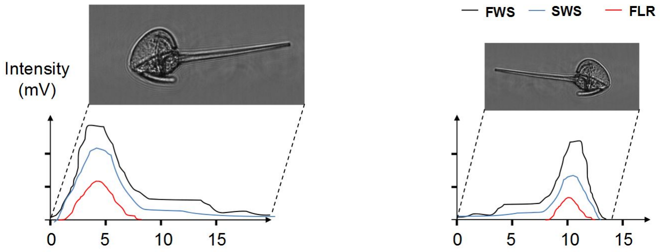
Courbes multiples échantillonnées

Profils physicochimiques



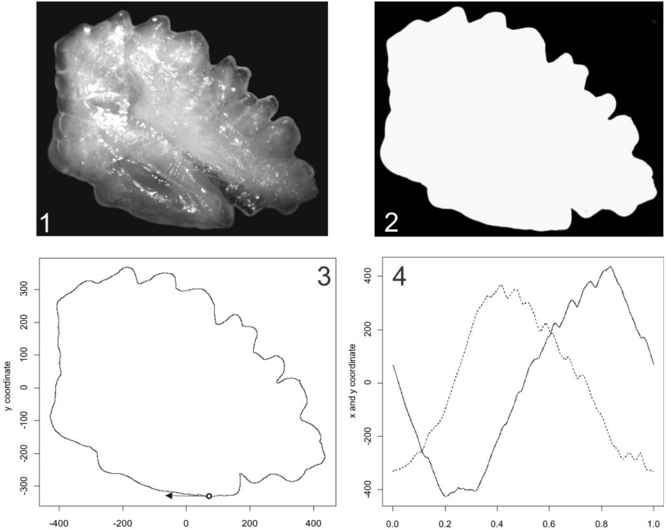
Courbes multiples échantillonnées

Emprunte optique d'un dinoflagellé Ceratium



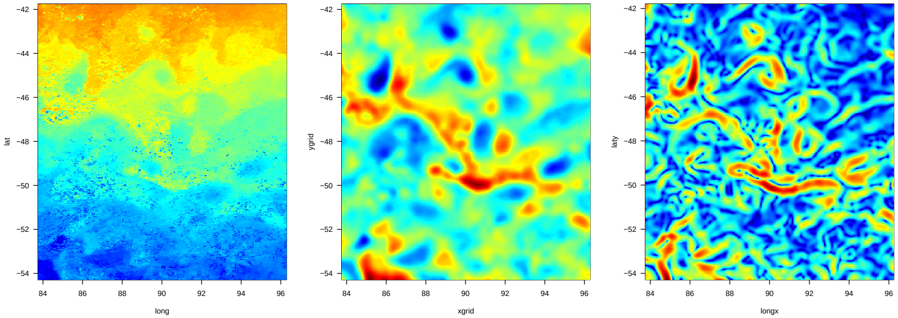
Courbes appariés échantillonnées

Contour d'otolithe



Images satellites

(SST) Brute, lissée & dérivée spatiale



Definition

Une données fonctionnelles est l'échantillon d'une variable aléatoire indicée par un argument

- Nombreuses observations bruitées

Definition

Une données fonctionnelles est l'échantillon d'une variable aléatoire indicée par un argument

- Nombreuses observations bruitées
- Beaucoup de points échantillonnés

Definition

Une données fonctionnelles est l’échantillon d’une variable aléatoire indicée par un argument

- Nombreuses observations bruitées
- Beaucoup de points échantillonnés
- La grille d’échantillonnage peut changer d’une courbe à l’autre

ATTENTION

- Impossibilité de créer un tableau de données !

Definition

Une données fonctionnelles est l'échantillon d'une variable aléatoire indicée par un argument

- Nombreuses observations bruitées
- Beaucoup de points échantillonnés
- La grille d'échantillonnage peut changer d'une courbe à l'autre

ATTENTION

- Impossibilité de créer un tableau de données !

- Hypothèse clef : **régularité**

$$Y_i\left(t_j\right) = \hat{Y}_i\left(t_j\right) + \varepsilon_i\left(t_j\right)$$

où $Y_i\left(t_j\right)$ est la donnée brute, $\hat{Y}_i\left(t\right)$ une courbe régulière et $\varepsilon_i\left(t_j\right)$ un résidu.

- A partir de l'échantillon $\mathbf{Y} = (Y(t_1), \cdots, Y(t_p))'$, on crée une combinaison linéaire de K fonctions de base fixées d'avance $\{\phi_1, \cdots, \phi_K\}$ tel que

$$Y(t) = \sum_{k=1}^{K} c_k \phi_k(t) + \varepsilon(t).$$

- A partir de l'échantillon $\mathbf{Y} = (Y(t_1), \cdots, Y(t_p))'$, on crée une combinaison linéaire de K fonctions de base fixées d'avance $\{\phi_1, \cdots, \phi_K\}$ tel que

$$Y(t) = \sum_{k=1}^{K} c_k \phi_k(t) + \varepsilon(t).$$

- L'estimation des paramètres $\mathbf{c} = (c_1, \cdots, c_K)'$ est réalisée par régression des moindres carrés en minimisant :

$$SSE(\mathbf{c}) = (\boldsymbol{Y} - \boldsymbol{\Phi}\mathbf{c})'(\boldsymbol{Y} - \boldsymbol{\Phi}\mathbf{c})$$

- Solution similaire à celle d'une régression classique :

$$\hat{\mathbf{c}} = \left(\boldsymbol{\Phi}'\boldsymbol{\Phi}\right)^{-1} \boldsymbol{\Phi}\mathbf{Y}$$

- A partir de l'échantillon $\mathbf{Y} = (Y(t_1), \cdots, Y(t_p))'$, on crée une combinaison linéaire de K fonctions de base fixées d'avance $\{\phi_1, \cdots, \phi_K\}$ tel que

$$Y(t) = \sum_{k=1}^{K} c_k \phi_k(t) + \varepsilon(t).$$

- L'estimation des paramètres $\mathbf{c} = (c_1, \cdots, c_K)'$ est réalisée par régression des moindres carrés en minimisant :

$$SSE(\mathbf{c}) = (\boldsymbol{Y} - \boldsymbol{\Phi}\mathbf{c})'(\boldsymbol{Y} - \boldsymbol{\Phi}\mathbf{c})$$

- Solution similaire à celle d'une régression classique :

$$\hat{\mathbf{c}} = \left(\boldsymbol{\Phi}'\boldsymbol{\Phi}\right)^{-1} \boldsymbol{\Phi}\mathbf{Y}$$

- Finalement la courbe est reconstituée avec :

$$\hat{Y}(t) = \sum_{k=1}^{K} \hat{c}_k \phi_k(t)$$

- A partir de l'échantillon $\mathbf{Y} = (Y(t_1), \cdots, Y(t_p))'$, on crée une combinaison linéaire de K fonctions de base fixées d'avance $\{\phi_1, \cdots, \phi_K\}$ tel que

$$Y(t) = \sum_{k=1}^{K} c_k \phi_k(t) + \varepsilon(t).$$

- L'estimation des paramètres $\mathbf{c} = (c_1, \cdots, c_K)'$ est réalisée par régression des moindres carrés en minimisant :

$$SSE(\mathbf{c}) = (\boldsymbol{Y} - \boldsymbol{\Phi}\mathbf{c})'(\boldsymbol{Y} - \boldsymbol{\Phi}\mathbf{c})$$

- Solution similaire à celle d'une régression classique :

$$\hat{\mathbf{c}} = \left(\boldsymbol{\Phi}'\boldsymbol{\Phi}\right)^{-1} \boldsymbol{\Phi}\mathbf{Y}$$

- Finalement la courbe est reconstituée avec :

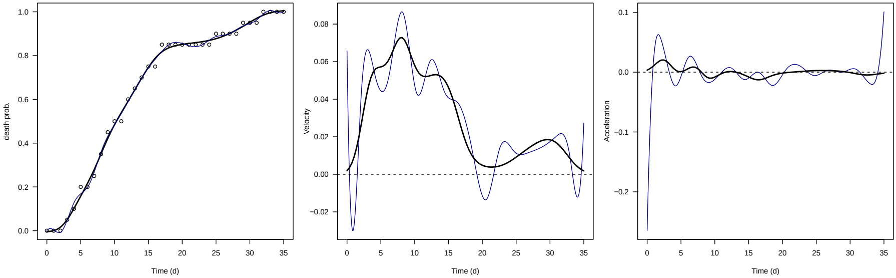
$$\hat{Y}(t) = \sum_{k=1}^{K} \hat{c}_k \phi_k(t)$$

- Dimension réduite quand $K < p$

- On peut rajouter des contraintes (régularité, positivité, convexité, monotonie, ...)

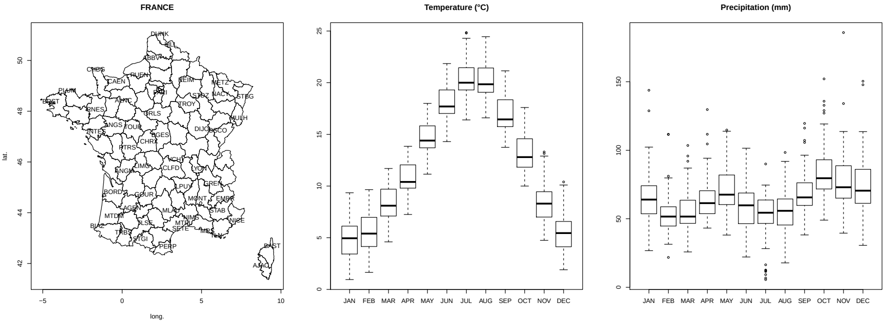
$$SSE(f) = \frac{1}{n}\sum_{i=1}^{n}(y_i - f(t_i))^2 + \lambda \times \|Lf\|^2$$

Gorgones soumises à un stress thermique



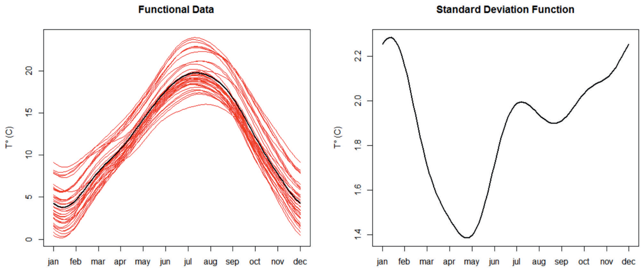
Statistiques de base

Temp. / Prec. en France (moyenne 1960-2000)



Fonction moyenne et fonction variance

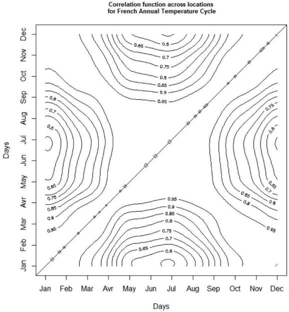
Température en France

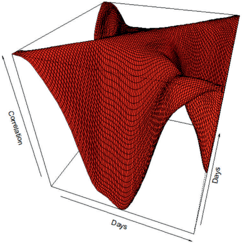


- $$\bar{Y}(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i(t), \; \sigma_Y(t) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i(t) - \bar{Y}(t)\right)^2$$

Opérateur de corrélation & de Variance

Fonction bivariée de corrélation



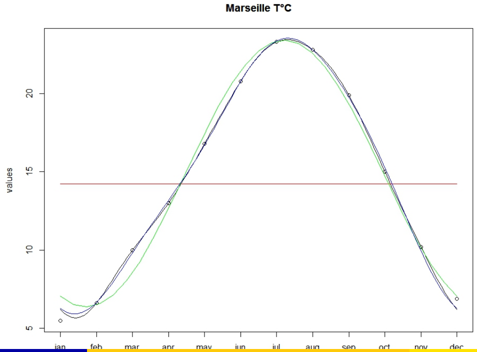


- $$\rho(s,t) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i(t)-\bar{Y}(t)}{\sigma(t)}\right)\left(\frac{Y_i(s)-\bar{Y}(s)}{\sigma(s)}\right)$$

Ajustement des données

- L'augmentation du nombre K de fonctions de base conduit à un meilleur ajustement
- Le choix de la base dépend des données (base de Fourier, B-splines, polynômes, ...)

Température ajustée à Marseille \ Base de Fourier



ACP fonctionnelles

- **Objectifs** : En utilisant les données $\left\{\hat{Y}_1, \cdots, \hat{Y}_n\right\}$ trouver la meilleure projection 2D des observations qui sont des courbes
- La position des individus dans l'espace factoriel est interprétée comme un changement dans la forme des courbes

ACP fonctionnelles

- **Objectifs** : En utilisant les données $\left\{\hat{Y}_1, \cdots, \hat{Y}_n\right\}$ trouver la meilleure projection 2D des observations qui sont des courbes
- La position des individus dans l'espace factoriel est interprétée comme un changement dans la forme des courbes
- Le problème de projection se résume à la décomposition spectrale de l'opérateur de variance-covariance V_Y (décomposition aux valeurs propres)

$$V_Y \xi_k = \lambda_k \xi_k$$

ce qui est équivalent à

$$\int_{\mathcal{D}} v_Y(s,t)\, \xi_k(s)\, ds = \lambda_k \xi_k(t)\,,\ t \in \mathcal{D}$$

ACP fonctionnelles

- **Objectifs** : En utilisant les données $\left\{\hat{Y}_1, \cdots, \hat{Y}_n\right\}$ trouver la meilleure projection 2D des observations qui sont des courbes
- La position des individus dans l'espace factoriel est interprétée comme un changement dans la forme des courbes
- Le problème de projection se résume à la décomposition spectrale de l'opérateur de variance-covariance V_Y (décomposition aux valeurs propres)

  $$V_Y \xi_k = \lambda_k \xi_k$$

  ce qui est équivalent à

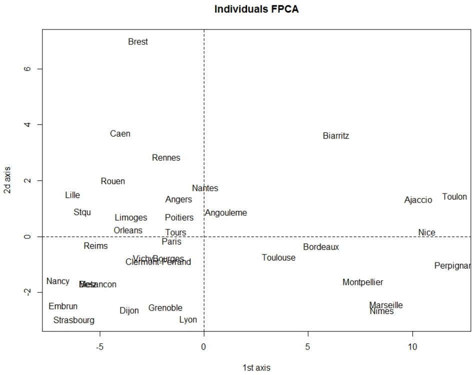
  $$\int_{\mathcal{D}} v_Y(s,t)\, \xi_k(s)\, ds = \lambda_k \xi_k(t)\,,\ t \in \mathcal{D}$$

- Fonctions $\xi_k(t)$ associées aux valeurs propres λ_k.
- Décomposition aux v. p. des courbes équivalente à celle réalisée sur les coefs de la décomposition à une métrique $\mathbf{\Phi}$ prêt.

Représentations de l'ACPF

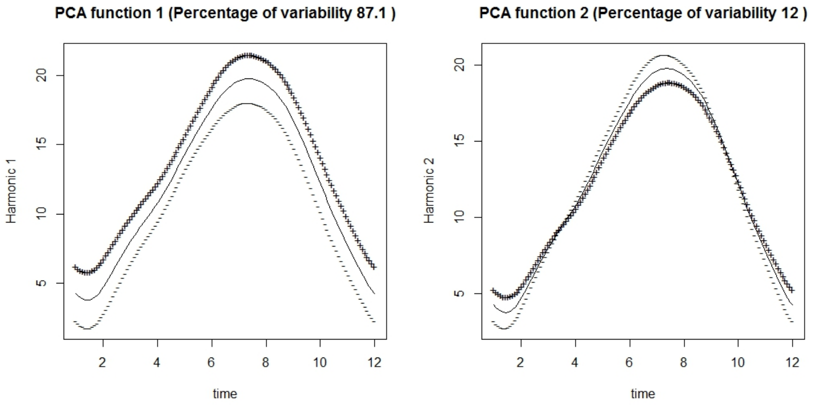
- La représentation 2D des individus est classique

Représentation factorielle 2D



Représentation de l'ACPF

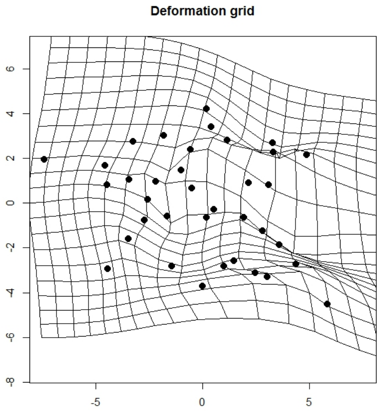
- Représentation 2D des variables comme une perturbation de la courbe moyenne $\bar{Y}(t) \pm \sqrt{\lambda_k}\xi_k(t)$ lors d'un déplacement le long d'un axe factoriel

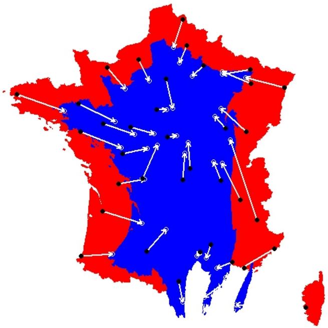


Autre representation

- Analyse procustéenne + SPM entre espace géographique et espace projectif

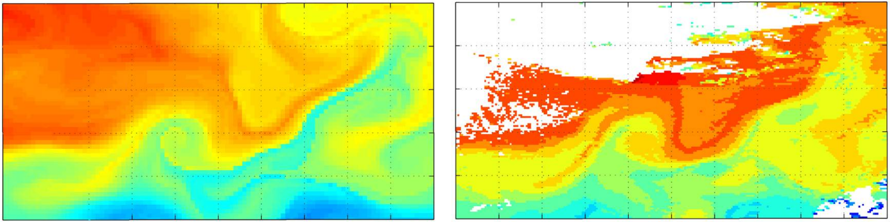
De la position GPS à l'espace factoriel





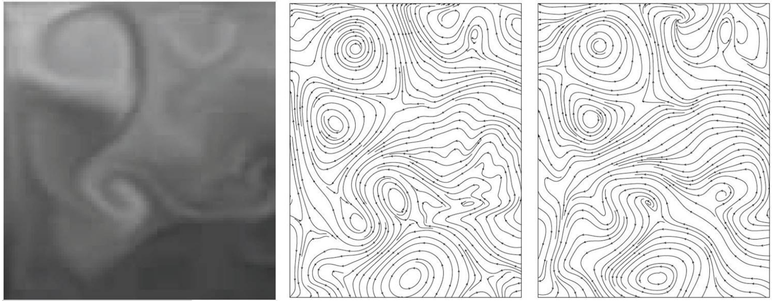
Extension : splines vectorielles

Comparaison sorties de modèles / images MODIS



- Permet d'évaluer spatialement l'efficacité d'un modèle
- Trouver une mesure de distance entre deux surfaces
- Nécessité de régulariser les deux objets

Modèle OPA, champ observé & approximé

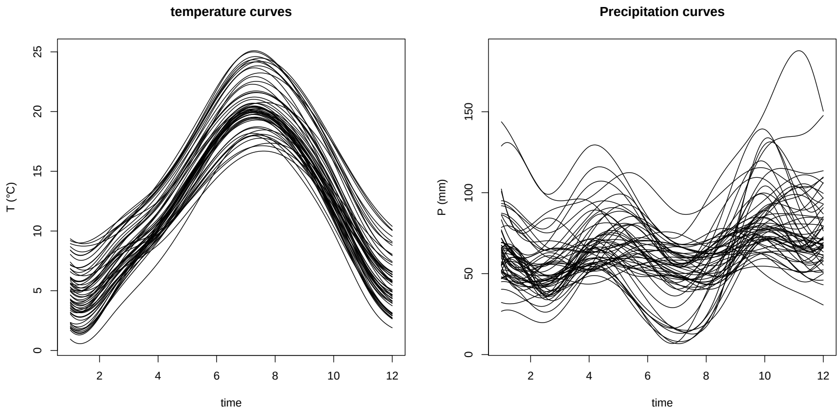


$$\min \left\{ \sum_{i=1}^{n} \|\mathbf{f}(\mathbf{x}_i) - \mathbf{f}_i\|^2 + \lambda \times \int \alpha \|\nabla \mathrm{div}\mathbf{f}\|^2 + \beta \|\nabla \mathrm{rot}\mathbf{f}\|^2 \right\}$$

- $\alpha >> \beta$ pour un système hautement turbulent (variabilité forte du rotationnel)
- $\alpha << \beta$ pour une activité verticale intense (flot localement divergent)

Prec. profile / Temp. profile

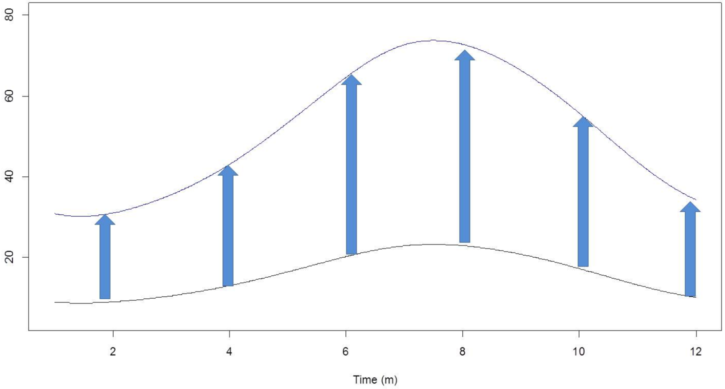
- Starting from a sample $(P_i, T_i)_{i=1,\cdots,n}$, we want to explain variations of profile P using functional variable T as predictor.



- The simplest form of the model is

$$P_i(t) = \alpha + \beta T_i(t) + \varepsilon_i(t)$$

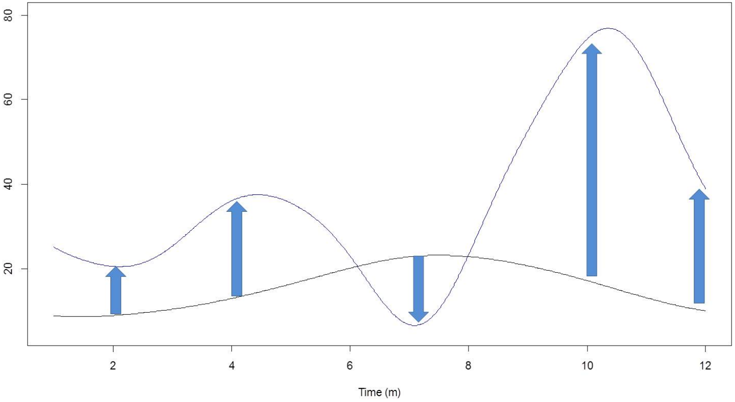
Example : $P(t) = 3 \times T(t) + 4$



- OK but . . . limited flexibility
- and curve shape cannot be modified. Try a more complex version :

- OK but . . . limited flexibility
- and curve shape cannot be modified. Try a more complex version :

Example : $P(t) = \beta(t) \times T(t) + \alpha(t)$



- In both cases, solutions are those of standard linear regression
- Interactions at same time t

- In both cases, solutions are those of standard linear regression
- Interactions at same time t
- Full Functional Linear Model :

$$P_i(t) = \alpha(t) + B(T_i)(t) + \varepsilon_i(t)$$

where

$$B(T_i)(t) = \int \beta(s,t) T_i(s) ds$$

is a linear operator of finite trace (HS) with kernel function β.

- Functional coefficients α and β are solution of the normal equations :

$$\begin{cases} \hat{B}\hat{V}_T &= \hat{V}_{TP} \\ \hat{\alpha} &= \hat{\mu}_P - \hat{B}(\hat{\mu}_T) \end{cases}$$

- PROBLEM : In infinite dimension, no solution can be found when computing :

$$\hat{B} = \hat{V}_{TP} \hat{V}_T^{-1}$$

- PROBLEM : In infinite dimension, no solution can be found when computing :

$$\hat{B} = \hat{V}_{TP}\hat{V}_T^{-1}$$

- SOLUTION : the use of regularization method
  1. Tikhonov

$$\hat{B} = \hat{V}_{TP}\left(\kappa I + \hat{V}_T\right)^{-1}$$

  2. Spectral cut regularized inverse

$$\hat{B} = \hat{V}_{TP}\hat{V}_T^{\dagger}$$

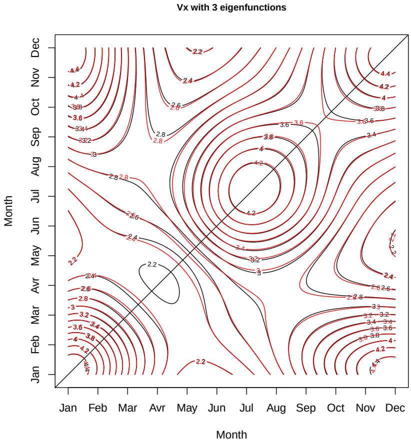
  with

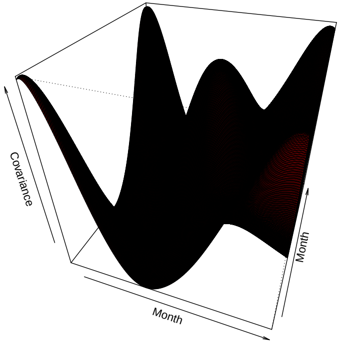
$$\hat{V}_T^{\dagger} = \sum_{j=1}^{k} \lambda_j^{-1} \left(\xi_j \otimes \xi_j\right)$$

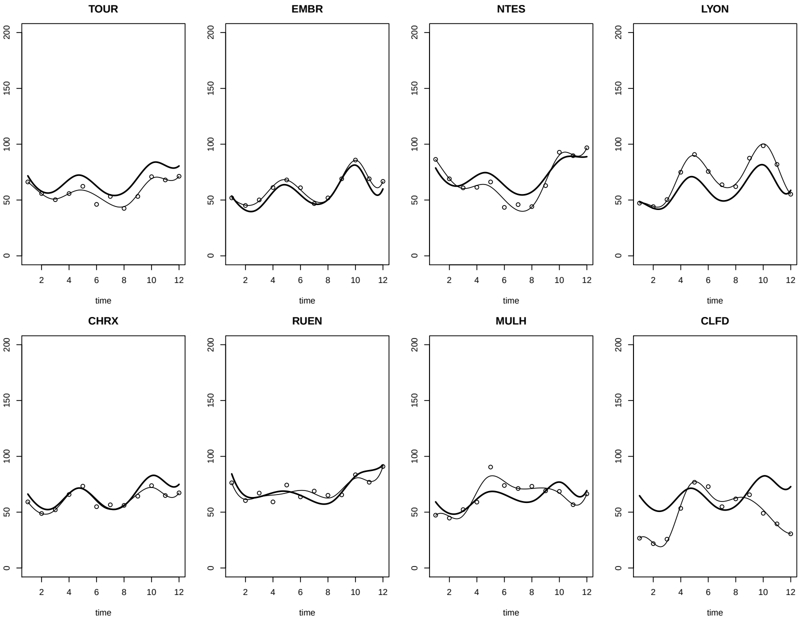
- Precipitation curve in new town $n+1$ is predicted with

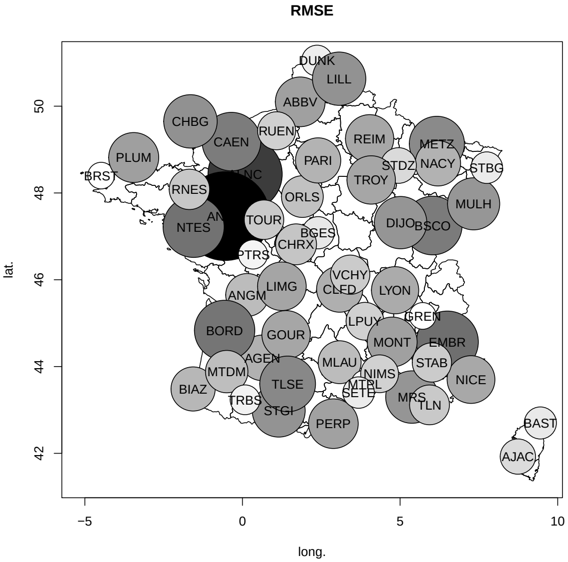
$$\hat{P}_{n+1} = \hat{\alpha}(t) + \hat{B}(T_{n+1})$$

Example : Regularized V_T





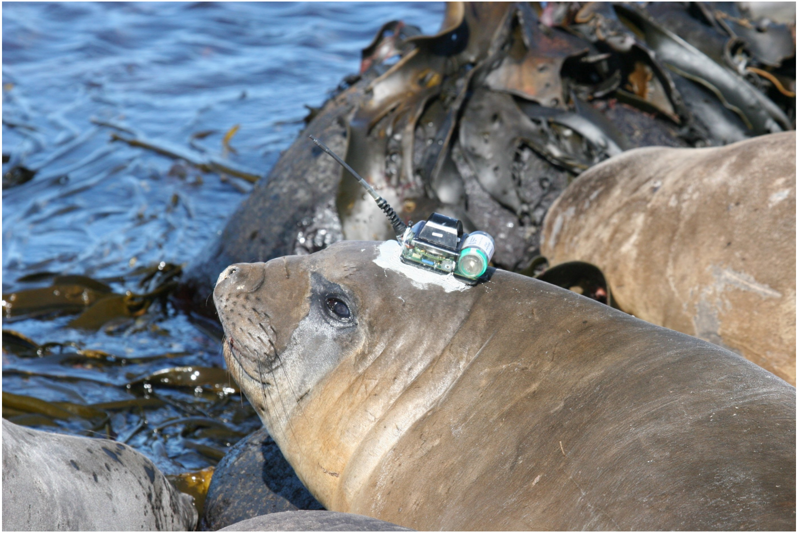

TOUR
EMBR
NTES
LYON
CHRX
RUEN
MULH
CLFD
0
50
100
150
200
2
4
6
8
10
12
time

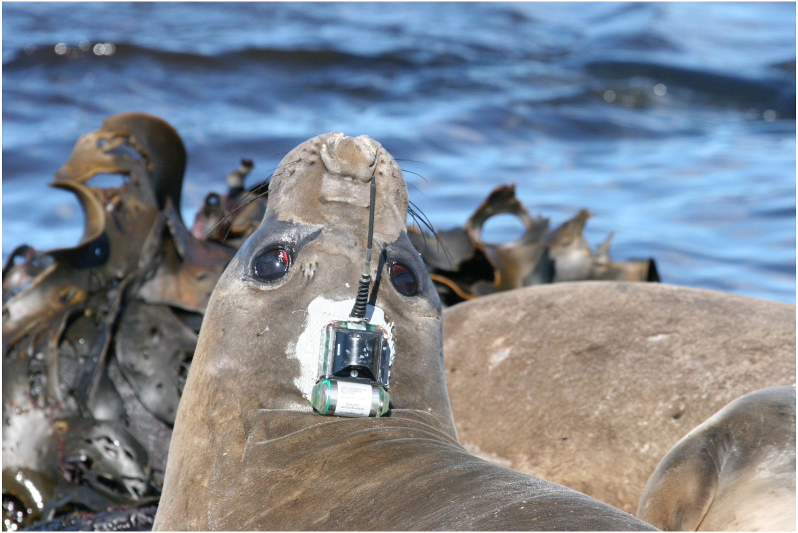

RMSE
DUNK
LILL
ABBV
CHBG
RUEN
CAEN
REIM
METZ
PLUM
PARI
STDZ
NACY
STBG
BRST
TROY
RNES
ORLS
MULH
NTES
TOUR
DIJO
BSCO
BGES
CHRX
PTRS
VCHY
LIMG
CLFD
LYON
ANGM
GREN
LPUY
BORD
GOUR
MONT
EMBR
AGEN
MLAU
STAB
MTDM
NIMS
NICE
MTPL
BIAZ
TLSE
SETE
MRS
TRBS
TLN
STGI
PERP
BAST
AJAC
lat.
long.
50
48
46
44
42
−5
0
5
10

- Since a decade, marine mammals constitute valuable auxiliaries for operational oceanography

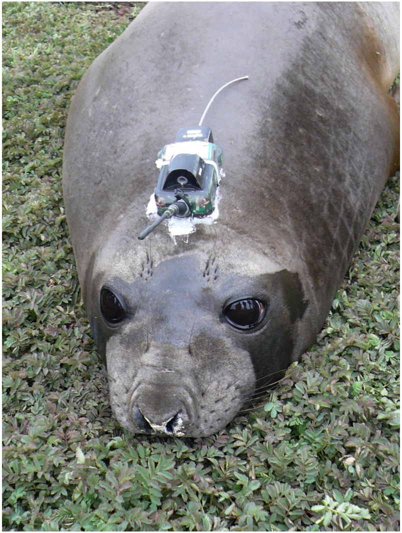
- Since a decade, marine mammals constitute valuable auxiliaries for operational oceanography

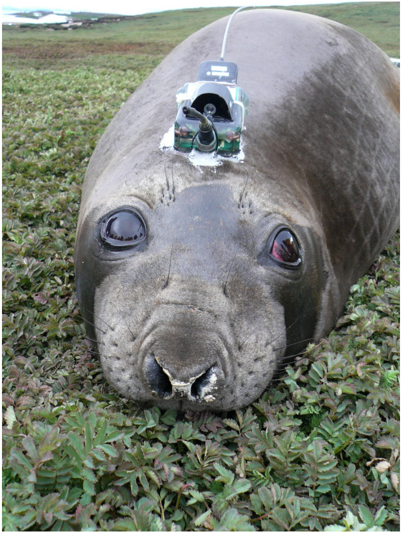
Elephant seals dataset



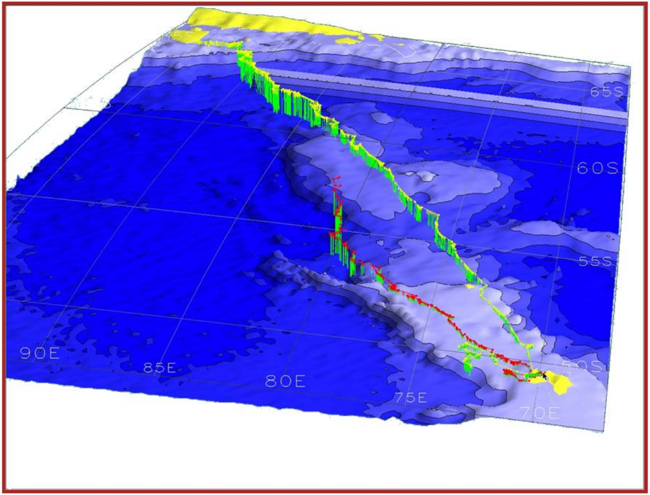



© C. Guinet, CEBC CNRS, 2013

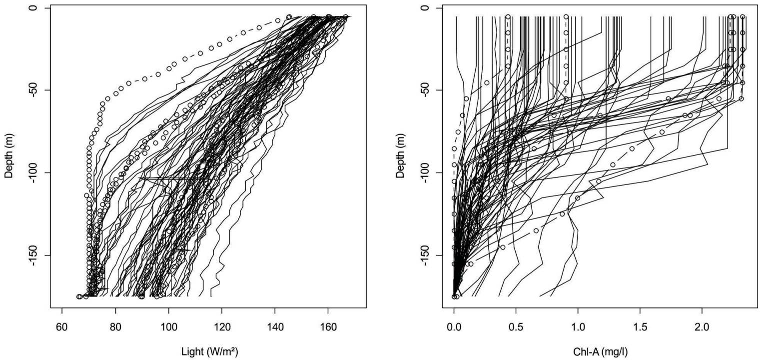





TURNER

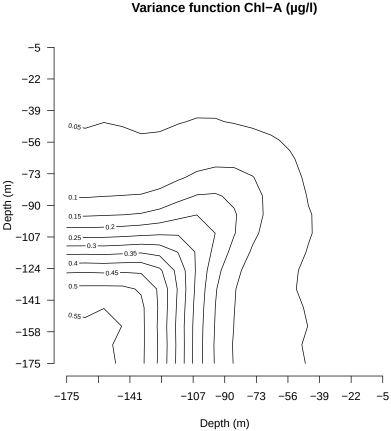

65S
60S
55S
50S
90E
85E
80E
75E
70E

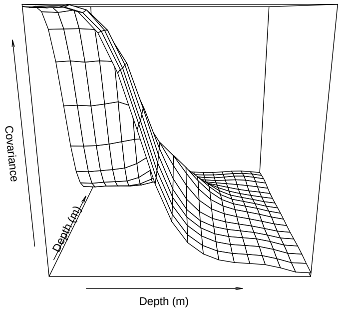
Sampled data



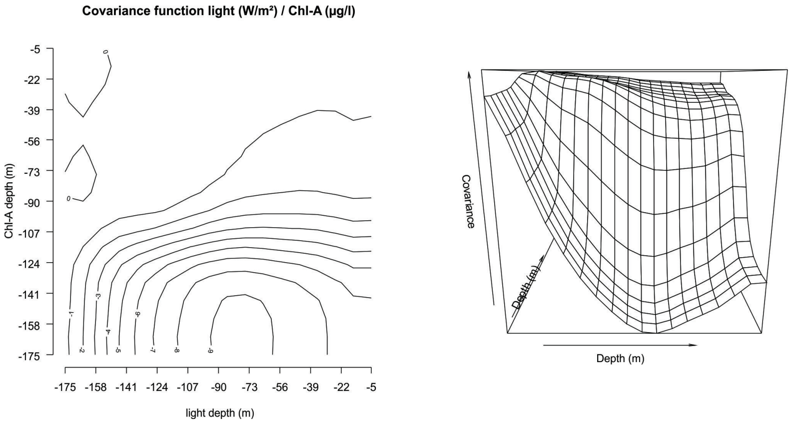
$$C(p) = \int_0^{P_{max}} \beta(s,t)L'(s)ds + \varepsilon(t)$$

ChlA Variance function



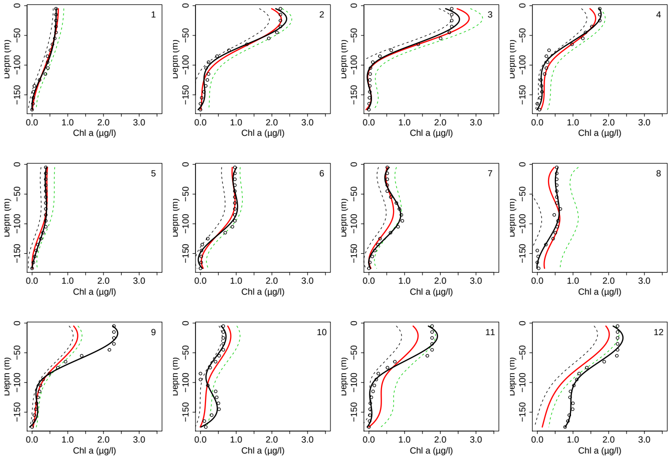


Cross-covariance function

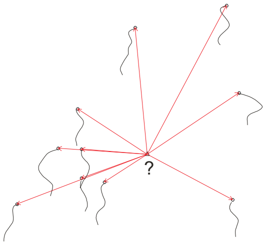


Using the model

Prediction



- Consider a collection of curves $E = \{Y_i, i = 1, \ldots, n\}$ sampled at n random spatial position inside a domain $\mathcal{D}$
- Each observation is a realization of a functional random variable Y_i at position $\mathbf{x}_i \in \mathcal{D}$, where $\mathcal{H}$ is an Hilbert space equipped with standard inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and the associated norm $\|\cdot\|_{\mathcal{H}}$.



- If we admit second order stationarity assumptions, the mean function $\mu(t)$ is the same at any point of the domain

$$\mathbb{E}(Y_i) = \mu, \ \forall \mathbf{x}_i \in \mathcal{D}$$

- If we admit second order stationarity assumptions, the mean function $\mu(t)$ is the same at any point of the domain

$$\mathbb{E}(Y_i) = \mu, \ \forall \mathbf{x}_i \in \mathcal{D}$$

- The covariance operator C_{ij} such that

$$C_{ij} = \mathbb{E}\left[(Y_i - \mu) \otimes \left(Y_j - \mu\right)\right]$$

- If we admit second order stationarity assumptions, the mean function $\mu(t)$ is the same at any point of the domain

$$\mathbb{E}(Y_i) = \mu, \ \forall \mathbf{x}_i \in \mathcal{D}$$

- The covariance operator C_{ij} such that

$$C_{ij} = \mathbb{E}\left[(Y_i - \mu) \otimes (Y_j - \mu)\right]$$

only depends on the vector $\mathbf{h}$ between a function Y_i at position $\mathbf{x}_i$ and Y_j at position $\mathbf{x}_j = \mathbf{x}_i - \mathbf{h}$

$$\begin{aligned} C_{ij}(f) &= \mathbb{E}\left[(Y_i - \mu) \otimes (Y_j - \mu)(f)\right] \\ &= \mathbb{E}\left[\langle f, Y_i - \mu\rangle_{\mathcal{H}} (Y_j - \mu)\right], \ f \in \mathcal{H} \end{aligned}$$

- We seek to estimate Y_0, the curve at unknown location $\mathbf{x}_0$, with the linear model

$$\widehat{Y}_0 = \sum_{i=1}^{n} B_i\,(Y_i)$$

with $B_i : \mathcal{H} \to \mathcal{H}$ being linear operators such that

$$B_i\,(f)\,(t) = \int_{\tau} \beta_i\,(s,t)\,f\,(s)\,ds,\ \forall f \in \mathcal{H},\ \forall t \in \tau.$$

- We seek to estimate Y_0, the curve at unknown location $\mathbf{x}_0$, with the linear model

$$\widehat{Y}_0 = \sum_{i=1}^{n} B_i\left(Y_i\right)$$

with $B_i : \mathcal{H} \rightarrow \mathcal{H}$ being linear operators such that

$$B_i\left(f\right)\left(t\right) = \int_{\tau} \beta_i\left(s,t\right) f\left(s\right) ds, \ \forall f \in \mathcal{H}, \ \forall t \in \tau.$$

- The $B_i's \in \mathcal{B}\left(\mathcal{H}\right)$ the space of HS operators equipped with the inner product

$$\forall\left(A,B\right) \in \mathcal{B}^2\left(\mathcal{H}\right), \ \langle A,B\rangle_{\mathcal{B}} = \sum_{k,l} \langle A\left(u_k\right), u_l\rangle_{\mathcal{H}} \langle B\left(u_k\right), u_l\rangle_{\mathcal{H}}$$

where the functions u_k form an orthonormal basis of $\mathcal{H}$.

- Looking for an unbiased estimator leads to develop the following equation

$$\mathbb{E}\left(\widehat{Y}_0 - Y_0\right) = 0,$$

which, under second order stationarity assumptions, gives the unbiasedness condition

$$\left[\sum_{i=1}^{n} B_i\right](\mu) = \mu.$$

- Looking for an unbiased estimator leads to develop the following equation

$$\mathbb{E}\left(\widehat{Y}_0 - Y_0\right) = 0,$$

which, under second order stationarity assumptions, gives the unbiasedness condition

$$\left[\sum_{i=1}^{n} B_i\right](\mu) = \mu.$$

- Interesting, but the mean function μ is unknown !

- Trying to extend the previous condition to the more general constraint

$$\left[\sum_{i=1}^{n} B_i\right](f) = f,\ \forall f \in \mathcal{H}$$

- Trying to extend the previous condition to the more general constraint

$$\left[\sum_{i=1}^{n} B_i\right](f) = f,\ \forall f \in \mathcal{H}$$

leads to find operators B_i checking

$$\sum_{i=1}^{n} B_i = I.$$

- Trying to extend the previous condition to the more general constraint

$$\left[\sum_{i=1}^{n} B_i\right](f) = f,\ \forall f \in \mathcal{H}$$

leads to find operators B_i checking

$$\sum_{i=1}^{n} B_i = I.$$

- Damn it ! The Identity operator is NOT an integral operator over $\mathcal{H}$

- Reconsider the last condition with

$$\sum_{i=1}^{n} B_i = K$$

- Reconsider the last condition with

$$\sum_{i=1}^{n} B_i = K$$

where K would be a linear operator which should verify

$$\begin{aligned} K(f) &= f, \ \forall f \in \mathcal{H} \\ K \in \mathcal{B}(\mathcal{H}) &\Leftrightarrow \|K\|_{\mathcal{B}}^2 < \infty \end{aligned}$$

- Can we found a Hilbert space where such an operator exists ? The answer is YES.

- Let now $\mathcal{H}$ be a RKHS. This space owns a reproducing kernel function κ defined as

$$
\begin{array}{rcl}
\kappa & : & \tau \times \tau \to \mathbb{R} \\
(s,t) & \longmapsto & \kappa(s,t)
\end{array}
$$

which checks the following conditions

$$
\begin{array}{rcl}
\forall t & \in & \tau,\ \kappa(\cdot,t) \in \mathcal{H} \\
\forall t & \in & \tau,\ \forall f \in \mathcal{H},\ \langle \kappa(.,t), f \rangle_{\mathcal{H}} = f(t).
\end{array}
$$

- Let now $\mathcal{H}$ be a RKHS. This space owns a reproducing kernel function κ defined as

$$\begin{array}{rcl} \kappa & : & \tau \times \tau \to \mathbb{R} \\ (s,t) & \longmapsto & \kappa(s,t) \end{array}$$

which checks the following conditions

$$\begin{array}{rcl} \forall t & \in & \tau,\ \kappa(\cdot,t) \in \mathcal{H} \\ \forall t & \in & \tau,\ \forall f \in \mathcal{H},\ \langle \kappa(.,t), f \rangle_{\mathcal{H}} = f(t). \end{array}$$

- The operator K attached to the kernel function κ verifies $K(f) = f$
- It is a HS operator $\Leftrightarrow \|K\|_{\mathcal{B}}^2 < \infty$
- The form of K is entirely attached to the choice of $\langle \cdot, \cdot \rangle_{\mathcal{H}}$.

- We now consider the problem of finding an estimator

$$\widehat{Y}_0 = \sum_{i=1}^{n} B_i\left(Y_i\right)$$

for which $B_i\left(Y_i\right)(t) = \langle \beta_i\left(.,t\right), Y_i \rangle_{\mathcal{H}}$

- We now consider the problem of finding an estimator

$$\widehat{Y}_0 = \sum_{i=1}^{n} B_i\left(Y_i\right)$$

for which $B_i\left(Y_i\right)(t) = \langle \beta_i\left(., t\right), Y_i \rangle_{\mathcal{H}}$

- The constraint for unbiasedness can be extended to

$$\sum_{i=1}^{n} B_i = K$$

for suitable choice of a RKHS $\mathcal{H}$ possessing a kernel operator K which plays the role of an identity operator

- We now consider the problem of finding an estimator

$$\widehat{Y}_0 = \sum_{i=1}^{n} B_i\left(Y_i\right)$$

for which $B_i\left(Y_i\right)(t) = \langle \beta_i\left(.,t\right), Y_i \rangle_{\mathcal{H}}$

- The constraint for unbiasedness can be extended to

$$\sum_{i=1}^{n} B_i = K$$

for suitable choice of a RKHS $\mathcal{H}$ possessing a kernel operator K which plays the role of an identity operator

- The objective is now to find estimates of B_i, $i = 1, ..., n$ by constrained minimization of

$$E\left\|\widehat{Y}_0 - Y_0\right\|_{\mathcal{H}}^2$$

- Seeking the minimizer of $E\left\|\widehat{Y}_0 - Y_0\right\|_{\mathcal{H}}^2$ under the constraint $\sum_{i=1}^n B_i = K$ leads to solve the constrained problem

$$\min_{(B_1,\ldots,B_n)' \in \mathcal{B}^n(\mathcal{H})} \sum_i \sum_j \left\langle B_j, C_{ij} B_i \right\rangle_{\mathcal{B}} + \langle K, C_{00} \rangle_{\mathcal{B}} - 2 \sum_i \langle C_{i0}, B_i \rangle_{\mathcal{B}}$$

- Seeking the minimizer of $E\left\|\widehat{Y}_0 - Y_0\right\|^2_{\mathcal{H}}$ under the constraint $\sum_{i=1}^n B_i = K$ leads to solve the constrained problem

$$\min_{(B_1,\dots,B_n)' \in \mathcal{B}^n(\mathcal{H})} \sum_i \sum_j \left\langle B_j, C_{ij} B_i \right\rangle_{\mathcal{B}} + \langle K, C_{00} \rangle_{\mathcal{B}} - 2 \sum_i \langle C_{i0}, B_i \rangle_{\mathcal{B}}$$

- The constraint can be included into the minimization problem by constructing the functional F such that

$$\begin{aligned} F(B_1, \dots, B_n, \Lambda) &= \sum_i \sum_j \left\langle B_j, C_{ij} B_i \right\rangle_{\mathcal{B}} + \langle K, C_{00} \rangle_{\mathcal{B}} - 2 \sum_i \langle C_{i0}, B_i \rangle_{\mathcal{B}} \\ &\quad + 2 \times \left\langle \Lambda, \sum_i B_i - K \right\rangle_{\mathcal{B}} \end{aligned}$$

where $\Lambda \in \mathcal{B}(\mathcal{H})$ acts as a Lagrange multiplier

- Let $\delta_{B_i}^{\Delta} F$ be the Gateaux differential of F at B_i in direction Δ such that

$$\delta_{B_i}^{\Delta} F\left(B_1, ..., B_i, ..., B_n, \Lambda\right) =$$
$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[F\left(B_1, ..., B_i + \varepsilon\Delta, ..., B_n, \Lambda\right) - F\left(B_1, ..., B_i, ..., B_n, \Lambda\right)\right]$$

- Let $\delta_{B_i}^{\Delta} F$ be the Gateaux differential of F at B_i in direction Δ such that

$$\delta_{B_i}^{\Delta} F\left(B_1, ..., B_i, ..., B_n, \Lambda\right) =$$
$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[F\left(B_1, ..., B_i + \varepsilon\Delta, ..., B_n, \Lambda\right) - F\left(B_1, ..., B_i, ..., B_n, \Lambda\right)\right]$$

- This limit exists for all Δ and, thanks to the self-adjoint property of C_{ij}, can easily be computed for all $i = 1, ..., n$ as

$$\delta_{B_i}^{\Delta} F\left(B_1, ..., B_i, ..., B_n, \Lambda\right) = 2\left\langle \sum_j C_{ij} B_j - C_{i0} + \Lambda, \Delta \right\rangle_{\mathcal{B}}$$
$$\delta_{\Lambda}^{\Delta} F\left(B_1, ..., B_i, ..., B_n, \Lambda\right) = \left\langle \sum_i B_i - K, \Delta \right\rangle_{\mathcal{B}}$$

- Let $\delta_{B_i}^{\Delta} F$ be the Gateaux differential of F at B_i in direction Δ such that

$$\delta_{B_i}^{\Delta} F\left(B_1, ..., B_i, ..., B_n, \Lambda\right) =$$
$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[F\left(B_1, ..., B_i + \varepsilon\Delta, ..., B_n, \Lambda\right) - F\left(B_1, ..., B_i, ..., B_n, \Lambda\right)\right]$$

- This limit exists for all Δ and, thanks to the self-adjoint property of C_{ij}, can easily be computed for all $i = 1, ..., n$ as

$$\begin{aligned}
\delta_{B_i}^{\Delta} F\left(B_1, ..., B_i, ..., B_n, \Lambda\right) &= 2\left\langle \sum_j C_{ij} B_j - C_{i0} + \Lambda, \Delta \right\rangle_{\mathcal{B}} \\
\delta_{\Lambda}^{\Delta} F\left(B_1, ..., B_i, ..., B_n, \Lambda\right) &= \left\langle \sum_i B_i - K, \Delta \right\rangle_{\mathcal{B}}
\end{aligned}$$

- A sufficient condition required to find the mimimum of F is to make all its $(n+1)$ Gateaux directional derivatives equal to 0

- The solution to the previous minimization problem provides a system of $(n+1)$ equations

$$\begin{aligned} \sum_j C_{ij} B_j + \Lambda &= C_{i0},\ i = 1, ..., n \\ \sum_i B_i &= K \end{aligned}$$

which forms the functional kriging system.

- PROBLEM : computers do not love infinite dimensional spaces

- PROBLEM : computers do not love infinite dimensional spaces
- SOLUTION : the use of basis functions

- PROBLEM : computers do not love infinite dimensional spaces
- SOLUTION : the use of basis functions
- We consider that each profile arrives as a linear combination of known orthonormal basis functions ϕ_k

$$Y_i(t) = \sum_{k=1}^{p} \alpha_k(\mathbf{x}_i)\, \phi_k(t)$$

- The space $\mathcal{H}$ that is spanned by functions Y_i is a finite dimensional Hilbert space with inner product

$$\forall (f, g) \in \mathcal{H}^2,\ \langle f, g \rangle_{\mathcal{H}} = \int_{\mathcal{T}} f(t)\, g(t)\, dt$$

- PROBLEM : computers do not love infinite dimensional spaces
- SOLUTION : the use of basis functions
- We consider that each profile arrives as a linear combination of known orthonormal basis functions ϕ_k

$$Y_i(t) = \sum_{k=1}^{p} \alpha_k(\mathbf{x}_i)\,\phi_k(t)$$

- The space $\mathcal{H}$ that is spanned by functions Y_i is a finite dimensional Hilbert space with inner product

$$\forall (f,g) \in \mathcal{H}^2,\ \langle f,g\rangle_{\mathcal{H}} = \int_{\mathcal{T}} f(t)\,g(t)\,dt$$

- This space owns a reproducing kernel function for any set $(\phi_1, ..., \phi_p)$ defining an orthonormal basis in $\mathcal{H}$

$$\kappa(s,t) = \sum_{k=1}^{p} \phi_k(s)\,\phi_k(t)$$

- The matrix form of the operator K reduces to the identity $\mathbf{I}_p$

- In that case, the functional kriging problem boils down to a standard cokriging problem working with the coefficients of the basis expansion of each Y_i
- The search for the weighting function β_i of the kriging estimator

$$\widehat{Y}_0(t) = \sum_{i=1}^{n} \int_{\tau} \beta_i(s,t)\, Y_i(s)\, ds, \ \forall t \in \tau$$

reduces to solve the following kriging matrix system

$$\begin{pmatrix} \mathbf{C}_{11} & \cdots & \mathbf{C}_{n1} & \mathbf{I} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{C}_{1n} & \cdots & \mathbf{C}_{nn} & \mathbf{I} \\ \mathbf{I} & \cdots & \mathbf{I} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_n \\ \mathbf{\Lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{01} \\ \vdots \\ \mathbf{C}_{0n} \\ \mathbf{I} \end{pmatrix}$$

- In that case, the functional kriging problem boils down to a standard cokriging problem working with the coefficients of the basis expansion of each Y_i
- The search for the weighting function β_i of the kriging estimator

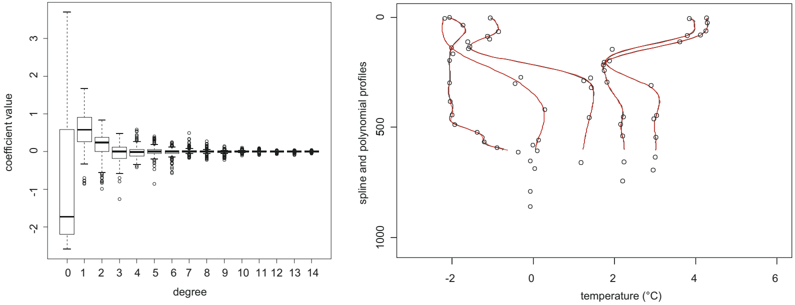
$$\widehat{Y}_0(t) = \sum_{i=1}^{n} \int_{\tau} \beta_i(s,t)\, Y_i(s)\, ds, \ \forall t \in \tau$$

reduces to solve the following kriging matrix system

$$\begin{pmatrix} \mathbf{C}_{11} & \cdots & \mathbf{C}_{n1} & \mathbf{I} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{C}_{1n} & \cdots & \mathbf{C}_{nn} & \mathbf{I} \\ \mathbf{I} & \cdots & \mathbf{I} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_n \\ \mathbf{\Lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{01} \\ \vdots \\ \mathbf{C}_{0n} \\ \mathbf{I} \end{pmatrix}$$

- The problem is now to find admissible estimators for the C_{ij}

- Each profile is expanded into a Legendre polynomial basis where coefficients are estimated by quadrature using a spline regularization procedure

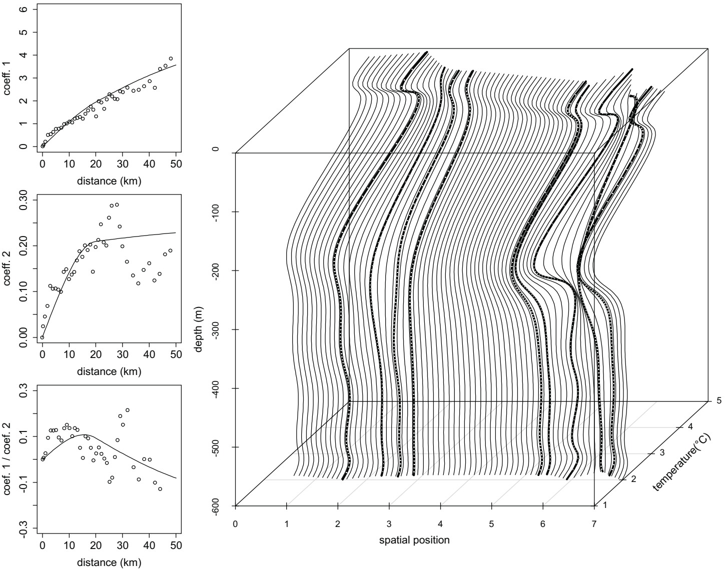


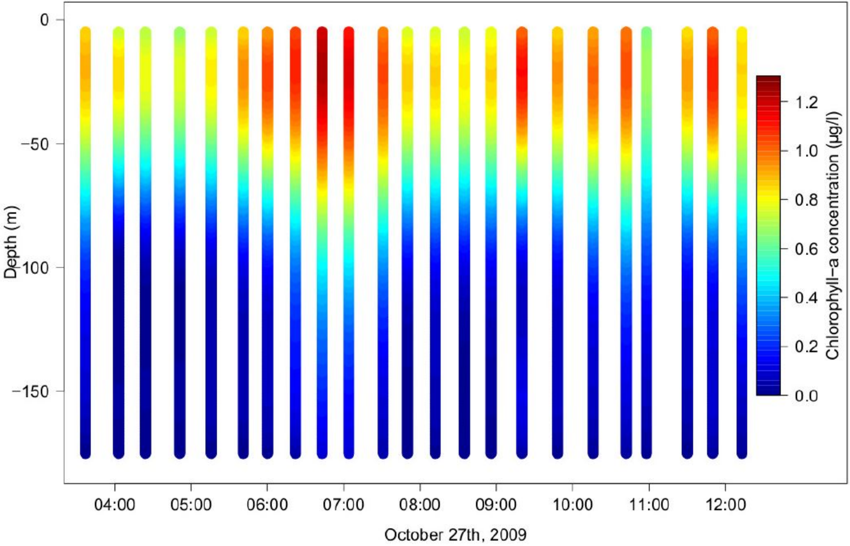
- One can use more polynomial coefficients than the number of sampled points and still achieve a nice polynomial fit

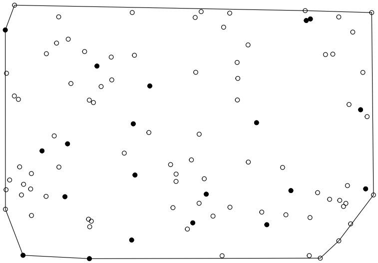
- The estimation of the covariance structures is realized through the fit of a linear coregionalization model
- We suppose that the matrix $\mathbf{C}(\mathbf{h})$ of cross-covariances between coefficients is proportional to a number s of correlation real functions $\rho_u(\mathbf{h})$ such that

$$\mathbf{C}(\mathbf{h}) = \sum_{u=1}^{s} \mathbf{P}_u \rho_u(\mathbf{h})$$

  where $\mathbf{P}_u$ are positive semi-definite coregionalization matrices
- In practice, the estimation of the matrices $\mathbf{P}_u$ is carried out through weighted least squares fitting of variogram models to experimental data
- Here, $s = 2$, the covariogram models are exponential with 2 scales


coeff. 1
distance (km)
coeff. 2
distance (km)
coef. 1 / coef. 2
distance (km)
depth (m)
spatial position
temperature(°C)


0
−50
−100
−150
Depth (m)
04:00
05:00
06:00
07:00
08:00
09:00
10:00
11:00
12:00
October 27th, 2009
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Chlorophyll−a concentration (µg/l)



- We dispose of a small sized sample $\{(X_i, Y_i), i = 1, \cdots, n\}$
- How can we use cross-covariance information to improve prediction on X or Y?

1. Nerini, D., Monestiez, P. & C., Mante (2010) Cokriging for functional data, *J. Mult. Anal.* **101**, 409-418

1. Nerini, D., Monestiez, P. & C., Mante (2010) Cokriging for functional data, *J. Mult. Anal.* **101**, 409-418
2. Bayle, S., Monestiez, P., Guinet, C.& Nerini, D. (2015) Moving toward finer scales in oceanography : predictive linear functional model of Chlorophyll a profiles from light data, *Progress In Oceanography*, **134**, 221-231

1. Nerini, D., Monestiez, P. & C., Mante (2010) Cokriging for functional data, *J. Mult. Anal.* **101**, 409-418
2. Bayle, S., Monestiez, P., Guinet, C.& Nerini, D. (2015) Moving toward finer scales in oceanography : predictive linear functional model of Chlorophyll a profiles from light data, *Progress In Oceanography*, **134**, 221-231
3. Goulard, M. and M., Voltz (1992) Linear coregionalization model : tools for estimation and choice of multivariate variograms, *Mathematical Geology*, **24**, 269-286

1. Nerini, D., Monestiez, P. & C., Mante (2010) Cokriging for functional data, *J. Mult. Anal.* **101**, 409-418
2. Bayle, S., Monestiez, P., Guinet, C.& Nerini, D. (2015) Moving toward finer scales in oceanography : predictive linear functional model of Chlorophyll a profiles from light data, *Progress In Oceanography*, **134**, 221-231
3. Goulard, M. and M., Voltz (1992) Linear coregionalization model : tools for estimation and choice of multivariate variograms, *Mathematical Geology*, **24**, 269-286
4. Berlinet, A. and C. Thomas-Agnan (2004) Reproducing Kernel Hilbert Spaces in Probability and Statistics, Kluwer Academic Publishers

1. Nerini, D., Monestiez, P. & C., Mante (2010) Cokriging for functional data, *J. Mult. Anal.* **101**, 409-418
2. Bayle, S., Monestiez, P., Guinet, C.& Nerini, D. (2015) Moving toward finer scales in oceanography : predictive linear functional model of Chlorophyll a profiles from light data, *Progress In Oceanography*, **134**, 221-231
3. Goulard, M. and M., Voltz (1992) Linear coregionalization model : tools for estimation and choice of multivariate variograms, *Mathematical Geology*, **24**, 269-286
4. Berlinet, A. and C. Thomas-Agnan (2004) Reproducing Kernel Hilbert Spaces in Probability and Statistics, Kluwer Academic Publishers
5. Besse, Ph. and J., O., Ramsay (1986) Principal Components Analysis of Sampled Functions, *Psychometrika*, **51**, 285-311

1. Nerini, D., Monestiez, P. & C., Mante (2010) Cokriging for functional data, *J. Mult. Anal.* **101**, 409-418
2. Bayle, S., Monestiez, P., Guinet, C.& Nerini, D. (2015) Moving toward finer scales in oceanography : predictive linear functional model of Chlorophyll a profiles from light data, *Progress In Oceanography*, **134**, 221-231
3. Goulard, M. and M., Voltz (1992) Linear coregionalization model : tools for estimation and choice of multivariate variograms, *Mathematical Geology*, **24**, 269-286
4. Berlinet, A. and C. Thomas-Agnan (2004) Reproducing Kernel Hilbert Spaces in Probability and Statistics, Kluwer Academic Publishers
5. Besse, Ph. and J., O., Ramsay (1986) Principal Components Analysis of Sampled Functions, *Psychometrika*, **51**, 285-311
6. Cardot, H., Ferraty, F., Sarda, P. (1999) Functional Linear Model, *Statistics and Probability Letters*, **45**, 11-22

1. Nerini, D., Monestiez, P. & C., Mante (2010) Cokriging for functional data, *J. Mult. Anal.* **101**, 409-418
2. Bayle, S., Monestiez, P., Guinet, C.& Nerini, D. (2015) Moving toward finer scales in oceanography : predictive linear functional model of Chlorophyll a profiles from light data, *Progress In Oceanography*, **134**, 221-231
3. Goulard, M. and M., Voltz (1992) Linear coregionalization model : tools for estimation and choice of multivariate variograms, *Mathematical Geology*, **24**, 269-286
4. Berlinet, A. and C. Thomas-Agnan (2004) Reproducing Kernel Hilbert Spaces in Probability and Statistics, Kluwer Academic Publishers
5. Besse, Ph. and J., O., Ramsay (1986) Principal Components Analysis of Sampled Functions, *Psychometrika*, **51**, 285-311
6. Cardot, H., Ferraty, F., Sarda, P. (1999) Functional Linear Model, *Statistics and Probability Letters*, **45**, 11-22
7. Craven, P. and G.,Wahba (1979) Smoothing Noisy Data with Splines Functions, *Numer. Math.*, **31**, 377-403

1. Nerini, D., Monestiez, P. & C., Mante (2010) Cokriging for functional data, *J. Mult. Anal.* **101**, 409-418
2. Bayle, S., Monestiez, P., Guinet, C.& Nerini, D. (2015) Moving toward finer scales in oceanography : predictive linear functional model of Chlorophyll a profiles from light data, *Progress In Oceanography*, **134**, 221-231
3. Goulard, M. and M., Voltz (1992) Linear coregionalization model : tools for estimation and choice of multivariate variograms, *Mathematical Geology*, **24**, 269-286
4. Berlinet, A. and C. Thomas-Agnan (2004) Reproducing Kernel Hilbert Spaces in Probability and Statistics, Kluwer Academic Publishers
5. Besse, Ph. and J., O., Ramsay (1986) Principal Components Analysis of Sampled Functions, *Psychometrika*, **51**, 285-311
6. Cardot, H., Ferraty, F., Sarda, P. (1999) Functional Linear Model, *Statistics and Probability Letters*, **45**, 11-22
7. Craven, P. and G.,Wahba (1979) Smoothing Noisy Data with Splines Functions, *Numer. Math.*, **31**, 377-403

1. Cuevas, A., Febrero, M., Fraiman, R. (2002) Linear functional regression : the case of fixed design and functional response, *The Canadian Journal of Statistics*, **30**, 285-300

1. Cuevas, A., Febrero, M., Fraiman, R. (2002) Linear functional regression : the case of fixed design and functional response, *The Canadian Journal of Statistics*, **30**, 285-300
2. Dabo-Niang, S. and A. -F., Yao (2007) Kernel regression estimation for continuous spatial processes, *Mathematical Methods for Statistics*, **16**, 298-317

1. Cuevas, A., Febrero, M., Fraiman, R. (2002) Linear functional regression : the case of fixed design and functional response, *The Canadian Journal of Statistics*, **30**, 285-300
2. Dabo-Niang, S. and A. -F., Yao (2007) Kernel regression estimation for continuous spatial processes, *Mathematical Methods for Statistics*, **16**, 298-317
3. Delicado,P., Giraldo, R., Comas, C. and Mateu, J. (2010) Statistics for spatial functional data : some recent contributions *Environmetrics*, **21**, 224-239

1. Cuevas, A., Febrero, M., Fraiman, R. (2002) Linear functional regression : the case of fixed design and functional response, *The Canadian Journal of Statistics*, **30**, 285-300
2. Dabo-Niang, S. and A. -F., Yao (2007) Kernel regression estimation for continuous spatial processes, *Mathematical Methods for Statistics*, **16**, 298-317
3. Delicado,P., Giraldo, R., Comas, C. and Mateu, J. (2010) Statistics for spatial functional data : some recent contributions *Environmetrics*, **21**, 224-239
4. Menafoglio, A., Secchi, P.C. and Dalla Rosa, M. (2013) A Universal Kriging predictor for spatially dependent functional data of a Hilbert Space, *Electro. J. Stat.* **7**,2209-2240