Progress in Oceanography 134 (2015) 221–231

Contents lists available at ScienceDirect

Progress in Oceanography

journal homepage: www.elsevier.com/locate/pocean

Moving toward finer scales in oceanography: Predictive linear functional model of Chlorophyll *a* profile from light data

Séverine Bayle [a], Pascal Monestiez [a], Christophe Guinet [b], David Nerini [c,*]

[a] INRA, UR546 Biostatistique et Processus Spatiaux (BioSP), F-84914 Avignon, France
[b] CEBC-CNRS, 79360 Villiers en Bois, France
[c] Mediterranean Institute of Oceanography (MIO), UMR 7294, Pytheas Institute (OSU), Aix-Marseille University, Campus de Luminy, Case 901, 13288 Marseille Cedex 09, France

ARTICLE INFO

Article history:
Received 16 July 2013
Received in revised form 2 January 2015
Accepted 10 February 2015
Available online 19 February 2015

ABSTRACT

The Southern Ocean plays a key role in ocean–atmosphere carbon dioxide fluxes. Estimation of carbon exchanges between ocean and atmosphere must rely on accurate estimations of primary productivity which require measurements of phytoplankton concentration within the water column. In this paper, we are interested in relationships between primary productivity and light in the Antarctic ocean. The originality of this work is twofold. Starting from physical hypothesis, a statistical model is constructed for the prediction of Chlorophyll *a* (Chl *a*) profiles where light profiles are used as a covariate. Taking into account of the functional nature of the data, solutions are proposed to estimate continuous vertical profiles from discrete data sampled by elephant seals equipped with a new generation of oceanographic tags. Bootstrapped prediction intervals show a good quality of prediction of Chl *a* profiles, giving access to the shape of the profiles along depth and to the submesoscale structure of phytoplankton within the euphotic layer of the Southern Ocean.

Introduction

Marine phytoplankton contributes to roughly half of the biosphere's primary production and therefore represents a fundamental level between living and inorganic stocks of carbon (Behrenfeld et al., 2006). But there is conflicting evidence on how this biological productivity will respond to global warming and climate change, particularly in the Southern Ocean, which plays an essential role in the carbon cycle. In that context, measurement of Chl *a* concentration in the waters of the Southern Ocean is an important indicator of the spatial and temporal variability of primary productivity (Behrenfeld and Falkowski, 1997) and must enable a better quantification of CO_2 fluxes.

However, the understanding of both the primary production variability and its spatial structure at submesoscale is hampered by the lack of *in situ* observations. Furthermore, the degree of confidence for observations of primary production derived from satellite-based estimates of phytoplankton biomass is still open to debate, especially in the Southern ocean (Guinet et al., 2013a). There is evidence of major limitations regarding the use of satellite assessment of primary production within the Southern Ocean. Satellites scan the sea surface, while deep fluorescence maxima can be found at depths of 40 m and 75 m within the frontal zone of the Antarctic Circumpolar Current (Queguiner and Brzezinski, 2002). Primary production cannot be properly assessed due to persistent cloud cover that precludes satellite detection of ephemeral phytoplankton blooms (Arrigo et al., 1997; Buesseler and Boyd, 2003). A better description of spatial (horizontal and vertical) and temporal (seasonal, inter-annual) distribution of phytoplankton is therefore essential to understand how primary production within the Southern Ocean and therefore CO_2 fluxes will respond to climate change.

Phytoplankton concentration is generally quantified through active measurement of Chl *a* fluorescence. However few autonomous platforms such as ARGO floats or gliders are used to collect this data in the Antarctic ocean (Fedak, 2013).

In recent years, a number of broad ranging deep diving marine predators have been equipped with electronic tags to investigate their foraging ecology and sample *in situ* oceanographic variables over broad areas of the ocean (Charrassin et al., 2008; Boehlert et al., 2001; Fedak et al., 2002; Block et al., 2002; McMahon et al., 2005; Biuw et al., 2007; Charrassin et al., 2008). Oceanographic variables sampled with animal-borne electronic tags include usually temperature and salinity. More recently fluorometers were integrated in order to obtain simultaneous estimations of Chl *a* in the water column (Xing et al., 2012; Guinet et al., 2013a).

* Corresponding author.
E-mail addresses: severine.bayle@avignon.inra.fr (S. Bayle), pascal.monestiez@avignon.inra.fr (P. Monestiez), christophe.guinet@cebc.cnrs.fr (C. Guinet), david.nerini@univ-amu.fr (D. Nerini).

http://dx.doi.org/10.1016/j.pocean.2015.02.001
0079-6611/

However the high energy requirement of these measurement devices only allows daily sampling of few profiles. Recently, using Southern elephant seals simultaneously equipped with a fluorometer and a light logger, Jaud et al. (2012) showed that light attenuation was strongly correlated with Chl *a* concentration measured by the fluorometer within the euphotic layer. In a pioneering work, Teo et al. (2009) using the bio-optical model of Morel (1988) demonstrated that *in situ* Chl *a* concentration profiles estimated from fluorescence measurements could be estimated from light and depth data collected by electronic tags.

The relationship between Chl *a* and fluorescence is known to be modulated by the taxonomic composition of phytoplankton and physiological acclimation mechanisms essentially related to light. Among the physiological acclimation mechanisms affecting relationships between Chl *a* and light, the depression of the fluorescence signal in surface waters during daylight is the most obvious one, especially at maximum solar elevation (Marra, 1997; Holm-Hansen et al., 2000; Xing et al., 2012). This so-called fluorescence quenching does indeed represent a collection of different photoprotective mechanisms to avoid photodamage under excessive sunlight energy (Kiefer, 1973; Maxwell and Johnson, 2000).

To our knowledge, the fluorescence profiles collected in Teo et al. (2009) were not quenching-corrected and this could have resulted in an underestimation of Chl *a* concentration within the first 30–60 m of the water column from the fluorescence measurements. This, in addition to other factors such as the presence of non-phytoplanktonic particles (zooplankton, inorganic particles, etc.) which might contribute to light attenuation, could explain the poor performance of the prediction of the Chl *a* profiles from light data under given conditions, and Teo et al. (2009) recognized the need to improve their method.

Following these two studies, we propose to construct a statistical model in order to predict *in situ* Chl *a* concentration profiles using light profiles as a predictive variable. Compared to the previous studies, the originality of this work is to include into the model the functional nature of the data *i.e.* to consider sampled data as observed curves. Indeed, the dataset used to construct the model has been sampled by elephant seals equipped with tags. In the course of an elephant seal trajectory, each data profile arrives as a discrete set of observations of light and fluorescence sampled at varying depths. The first section of this paper is devoted to problems of data sampled with marine mammals for light and Chl *a* variables. As for the bio-optical model described in Teo et al. (2009), the construction of the linear functional model starts from the Beer–Lambert relationship defined by Bouguer (1729) which connects the light absorption to properties of an homogeneous environment. This is the objective of the second section: bring the construction of the statistical model by linking Chl *a* profiles to light profiles under more realistic assumptions than those of Beer–Lambert. We will focus more particularly on the fact that the model is constructed using the derivative of log light profiles. This section also includes technical solutions for parameters estimation, construction of observed profiles from elephant seals dataset, including constraints over the shape of profiles, construction of bootstrap prediction intervals and measures of model accuracy. The third section present the main results. The paper ends with a discussion about the modeling choices and advantages of the proposed statistical method.

Elephant seal dataset

In October 2009, at the beginning of the austral spring, 3 post-breeding Southern elephant seal females from Kerguelen Island were anesthetized by intravenous injection of tiletamine and zolazepam 1:1. These animals have been equipped with a Satellite Relayed Data Logger which samples pressure, temperature, salinity and fluorescence (CTD-Fluo SRDL) developed by the Sea Mammal Research Unit, St Andrews University, Scotland, in collaboration with the Centre d'Etudes Biologiques de Chizé. A MK9 time depth recorder (TDR)(Wildlife Computer, USA) glued on the back of the CTD-Fluo SRDL has also been added. It is fixed next to the fluorometer with both the fluorometer and light sensor facing backward. Previous studies have shown that location of the light sensor at the back or at the front of the satellite does not change the relationship between light and fluorescence (Jaud et al., 2012). The package was then glued on the fur of the Southern elephant seal's head using a two component industrial epoxy (Araldite AW 2101).

Through the whole post breeding foraging trip (Fig. 1), fluorescence was generally measured twice a day from the CTD-Fluo tags while light and pressure was monitored continuously at 2 Hz by the logger. The CTD-Fluo SRDL included a Keller type pressure sensor (series PA7, 0 dbar to 2000 dbar ± 1 dbar), a fast response Platinum Resistance Thermometer (PRT) (−5 °C to 35 °C ± 0.005°C, 0.7 s response time), an induction conductivity sensor (Valport, UK, range: 0 mS cm^{-1} to 80 mS cm^{-1}, accuracy: better than 0.02 mS cm^{-1}), and a Cyclops 7 fluorometer from Turner Design with a dynamic range set between 0 µg to 5 µg of Chl *a* (µg l^{-1}). The MK9 TDR loggers were set to sample depth (0–1500 m ± 1 m), water temperature (−40 °C to +60 °C ± 0.1 °C) and light ($5 \cdot 10^{-2}$ W cm^{-2} to $5 \cdot 10^{-12}$ W cm^{-2} in blue Wavelength) every 2 s. The MK9 integrated light sensor was tested under laboratory condition and was found to exhibit its highest sensitivity at 465 nm, with a mean sensitivity range of 405–480 nm (Vacquié-Garcia, 2014). The blue light wavelength is the least water attenuated wavelength, and therefore the blue light is able to reach greater depth encompassing the whole euphotic layer. A complete description of these tags is available in Boehme et al. (2009). Light values are converted inboard via a log treatment reducing light measurements to 3 digit values.

Before deployment on Southern elephant seals, each CTD-Fluo SRDL was calibrated at sea during the BOUSSOLE campaign (Guinet et al., 2013a) by comparison with *in situ* measurements

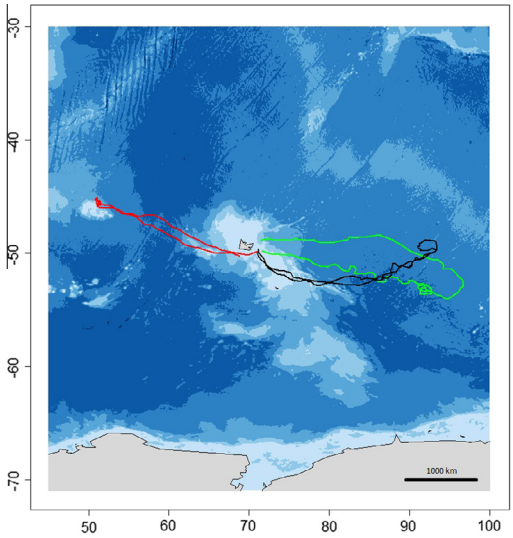

Fig. 1. Paths of the three elephant seals. Starting from Kerguelen Islands (center), animals turn back to land after a journey of 3 months, with Antarctic continent at the bottom (green: path 1, black: path 2, red: path 3). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

from Niskin bottles. A coefficient was calculated for each tag to convert the fluorescence values to an actual Chl *a* concentration.

When CTD-Fluo SRDL were deployed on elephant seals, Chl *a* concentration ($\mu g\, l^{-1}$) derivated from fluorescence measurements was assessed continuously at a two-second sampling rate for the last 180 m of the ascent phase of the dive. The depth of 180 m was selected as threshold because it encompasses the euphotic layer, which is generally close to 150 m. Each profile, transmitted via ARGOS, consists of a maximum of eighteen sections of ten meters long. The average fluorescence value is associated with the median depth of each segment (−5 m to −175 m). About two fluorescence profiles are sampled and ARGOS transmitted daily by these tags. Temperature and salinity were treated similarly for the first 180 m, and six measurements were made at depths exceeding 180 m to ensure that the best reconstruction of the high resolution temperature and salinity profile might also be transmitted.

Daylight fluorescence profiles are affected by quenching, defined as photo-inhibition due to an excess of light, resulting in an artificial deep maximum Chl *a* concentration. In well mixed waters representing about 84% of available profiles, the fluorescence profiles obtained during daylight hours were post-processed to correct the quenching effect according to the procedure proposed in Xing et al. (2012). Daylight profiles obtained in well stratified water (less than 20%) were excluded from the analysis as quenching could not be corrected accurately under these conditions (Xing et al., 2012). Processed Chl *a* data including the 3 individuals (tags 11259, 11260 and 11263) used in this study are freely downloadable at http://dx.doi.org/10.7491/MEMO.1 (Guinet et al., 2013a).

The 3 tags were recovered in January 2010 at the beginning of the austral summer when Southern elephant seal females come back to Kerguelen Island to molt and MK9 pressure and light data were downloaded.

In this study, only complete Chl *a* profiles composed with 18 observations were used for the construction of Chl *a* profiles. So, among the 436 Chl *a* profiles sampled, 407 were included into the statistical sample. Among them we selected those sampled only during daylight hours with a sun angle greater than 20° above the horizon to match them to the corresponding light data profiles sampled during ascending phase. A complete dataset is composed of $n = 208$ pairs of Chl *a* and light data profiles (Fig. 2, Table 1).

Statistical methods

Constructing the functional linear model

Let $C(z)$ be the Chl *a* variable and $L(z)$ the light variable. These variables are functions with argument z ranging from subsurface $Z_m = 5$ m to maximum depth $Z_M = 175$ m in the interval $[Z_m; Z_M]$. We wish to investigate to what extent Chl *a* profiles can be predicted from information contained in light profiles using a collection of pairwise functions (L_i, C_i), $i = 1, \ldots, n$. This sample of observed profiles must be constructed from pointwise data sampled by the elephant seals.

Table 1
Different elephant seals with number of available profiles. The last column is the number of Chl *a* profiles with 18 observations sampled during daylight hours.

ARGOS number	Number of elephant seal	Total number of registered Chl *a* profiles with 18 observations	Number of considered Chl *a* profiles
11260	1	142	73
11263	2	148	73
11259	3	117	62

Usually, a light profile in a homogeneous and weakly concentrated liquid medium is constructed using the Beer–lambert equation (Bouguer, 1729). The intensity of light is supposed to decrease exponentially in accordance with depth z following the relationship

$$L(z, \lambda) = L_0(\lambda) \exp(-q(\lambda) z),$$

where L is the light intensity, z is depth, λ is a wavelength, L_0 represents the light intensity at the surface, and q is the light attenuation coefficient. Remind that in our case, only the blue wavelength has been sampled. Then, considering a fixed value of λ (blue wavelength in this study), this last equation is the solution of the following ordinary differential equation

$$\frac{dL}{dz}(z) = -qL(z), \quad L(z = Z_m) = L_0.$$

The Beer–Lambert assumptions suggest that the profile of Chl *a* concentration is constant alongside depth. In first approximation, it can be considered that the Chl *a* concentration is independent from depth and proportional to the coefficient q such that

$$C(z) \propto q.$$

However, this hypothesis does not hold, being in contradiction with the observations of the water column (Fig. 2). Light profiles do not decrease exponentially with depth and Chl *a* values are not linearly correlated with depth.

One way to release the hypothesis of homogeneous environment to broader assumptions is to consider that the coefficient of light attenuation is depending with depth. This implies that the Chl *a* concentration is proportional to $q(z)$ *i.e.*

$$C(z) \propto q(z).$$

This also implies that light attenuation is solution of the following differential equation

$$\frac{dL}{dz}(z) = -q(z)L(z), \quad L(z = Z_m) = L_0,$$

where $q(\cdot)$ is now a function of the depth z. The solution of the last equation is given with

$$L(z) = L_0 \exp\left(-\int_{Z_m}^{z} q(s)ds\right), \quad L(z = Z_m) = L_0, \tag{1}$$

which is a monotonic function: the light decreases as depth increases. Taking the derivative of that solution in logarithmic scale, we obtain

$$L'(z) = \frac{d\log(L(z))}{dz} = q(Z_m) - q(z),$$

the derivative of the log-light is then proportional to the attenuation coefficient $q(z)$. As $C(z) \propto q(z)$ and $L'(z) \propto q(z)$, one can properly assume that $C(z) \propto L'(z)$.

The most simple linear model that can be constructed to predict the Chl *a* functional variable with $L'(z)$ used as a covariate reads

$$C(z) = \alpha + \beta L'(z) + \varepsilon(z),$$

where α and β are real parameters that must be estimated with the sample, and $\varepsilon(z)$ is a remainder. A more general version of this simple model can be constructed by considering possible variations of parameters α and β alongside the depth:

$$C(z) = \alpha(z) + \beta(z)L'(z) + \varepsilon(z).$$

However, this model only considers interactions between variables at the same depth. From a predictive point-of-view, it can be relevant to construct a more flexible model which gives predictions of the Chl *a* at depth z by using the information of the whole curve

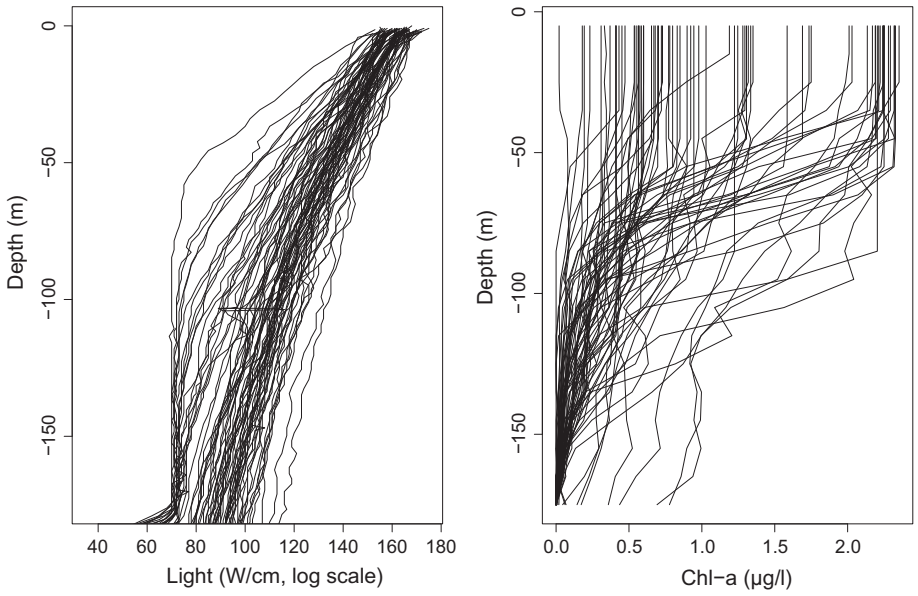

Fig. 2. Examples of light and Chl-*a* raw data sampled on the path of the elephant seal number 1. The objective is to construct smoothed pairs of observed profiles using pointwise vertical observations that include many sources of variability.

L'. For that purpose, we consider the following linear functional model

$$C(z) = \alpha(z) + \int_{Z_m}^{Z_M} \beta(s,z)L'(s)ds + \varepsilon(z). \quad (2)$$

The bivariate coefficient $\beta(s,z)$ can be interpreted as a loading function which gives the ability to take into account cross-dependencies between Chl *a* and light at different depths. The coefficient $\alpha(z)$ is a function that acts as an intercept. The last term $\varepsilon(z)$ is considered as a random error term. The properties of such a model and different ways to estimate the parameters have already been studied in Ramsay and Silverman (2005) and related references therein. In this study, both coefficients $\alpha(z)$ and $\beta(s,z)$ are estimated using the data samples at hands (see Appendix A). The advantage of such a model is that it includes the more simple versions of the linear model presented above but allows more flexibility if required.

Constructing Chl a and light observed profiles

Consider a sampled profile of variable C which arrives as p discrete observations $(z_j, c_j),\ j = 1,\ldots,p$. We wish to reconstruct this unknown observed profile $C(z)$ using these pointwise observations (Fig. 3). One way to proceed is to consider that an observed profile $C(z)$ is expressed as a linear combination of known basis functions $\phi_k,\ k = 1,\ldots,K$ such that

$$c_j = C(z_j) + \epsilon_j = \sum_{k=1}^{K} a_k\phi_k(z_j) + \epsilon_j.$$

The deterministic part $C(z) = \sum_{k=1}^{K} a_k\phi_k(z)$ of the data is entirely determined by the coefficients a_k which are estimated when minimizing the sum of squares of the errors ϵ_j. In our case, releasing all assumptions on a parametric form of the curves, we choose a B-spline basis to fit the raw data. A B-spline is defined as a piecewise polynomials of order 4 where coefficients are computed when minimizing the penalized cost function

$$1/p\sum_{j=1}^{p}(c_j - C(z_j))^2 + \theta\int (C''(u))^2 du.$$

The smoothing parameter θ gives trade-off between smoothness of the curve (norm of its second derivative) and closeness to raw data (Fig. 3). The smoothing parameter θ and the number K of basis functions can be chosen by cross-validation or by using the variability of the sampling devices as suggested in Craven and Wahba (1979), Nerini et al. (2010) and Hosseini-Nasab (2012). Once the polynomial regression is achieved for both observed variables L and C, we dispose of a sample of matched functions $\{(L_i, C_i),\ i = 1,\ldots,n\}$ for Chl *a* profiles and for light profiles. Each of these functions is entirely determined by the knowledge of its coefficients when expanded into the basis. The advantage of considering a smoothing spline basis expansion is that the derivative of any curve can be explicitly calculated, some required constraints (positivity, monotony, etc.) can be included as well. For instance, monotony constraints for light profiles fitting as suggested by Eq. (1) have been included when estimating the parameters of the spline functions using the data (see Meyer and splines (2012) for more technical details).

Prediction errors

Once the dataset of profiles is constructed, parameters α and β of model 2 can be estimated following the procedure in Appendix A. Considering a light profile $L(z)$ and the derivative of its logarithm $L'(z)$, the functional linear model gives a Chl *a* predicted curve denoted by $\widehat{C}$ such that

$$\widehat{C}_i(z) = \widehat{\alpha}(z) + \int_{Z_m}^{Z_M} \widehat{\beta}(s,z)L_i'(s)ds,$$

where $\widehat{\alpha}$ and $\widehat{\beta}$ are parameter estimations. The predictive capabilities of the model can be measured with the integrated mean squared error

$$\mathrm{SSE} = \frac{1}{n}\sum_{i=1}^{n}\int_{Z_m}^{Z_M}\left(C_i(z) - \widehat{C}_i(z)\right)^2 dz.$$

In the following, we will use expressions derived from this quantity in order to select the right number of basis functions and to test if differences appear between data sampled by different animals.

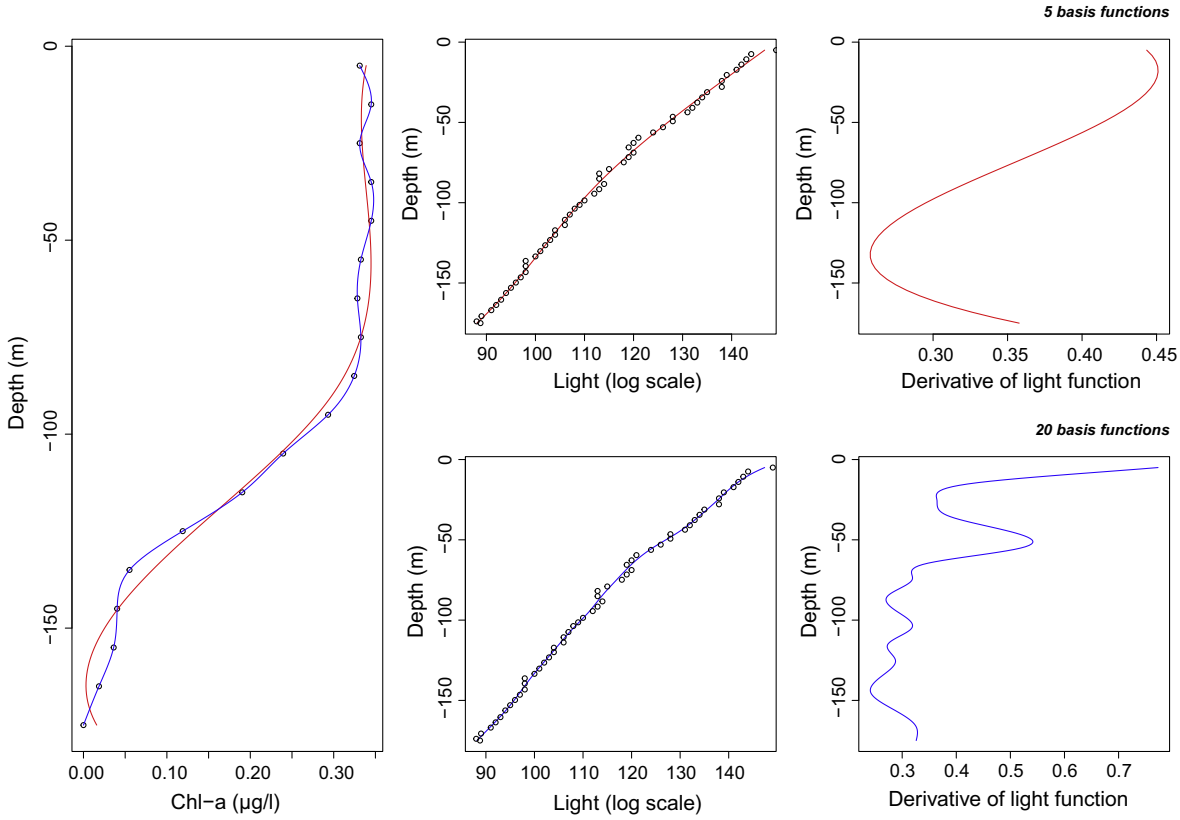

Fig. 3. Examples of Chl-*a*, light profiles and derivatives fitted with splines. Red curves show fits with 5 basis functions, and blue curves represent fits with 20 basis functions. The more basis functions number increases, the more accurate is the fit. Monotony constraints have been included for fitting light data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Choosing the right number of basis functions

Leave-one-out cross validation method is used to find the right number of basis functions for fitting Chl *a* profiles and light profiles from raw data (Efron and Tibshirani, 1993). Consider data from a unique elephant seal denoted as *e*. Consider the observed sample $\{(L_i, C_i),\ i = 1, \ldots, n_e\}$ of size n_e. These profiles have been constructed with K basis functions for the functions C_i and M basis functions for profiles L_i using the raw data. Let $\mathrm{SSE}_{CV}(K, M)$ be the mean prediction error integrated alongside depth and defined as

$$\mathrm{SSE}_{CV}(K, M) = \frac{1}{n_e} \sum_{i=1}^{n_e} \int_{Z_m}^{Z_M} \left(C_i(z) - \widehat{C}_{(-i)}(z) \right)^2 dz.$$

The curve $\widehat{C}_{(-i)}$ is the prediction of the observed profile C_i when the estimated parameters $\widehat{\beta}_{(-i)}(s, z)$ and $\widehat{\alpha}_{(-i)}(z)$ of the functional linear model have been computed using the whole set of observations except observation (L_i, C_i). The computation of $\mathrm{SSE}_{CV}(K, M)$ can be achieved for various values of pair (K, M) giving rise to a graph such as in Fig. 4. The right number of basis functions is provided by the pair $(K^\star, L^\star)$ which minimizes the cross-validated prediction error $\mathrm{SSE}_{CV}(K, M)$.

Cross validated prediction between animals

Once the number of basis functions has been selected, the whole set of n matched observed profiles is constructed. Using that curves, the parameters of the functional linear model (2) are estimated. This model can then be used to give prediction of a Chl *a* profile using the corresponding light profile. It can be interesting to assess error variance of prediction between animals. Consider again the elephant seal *e* and estimate the parameters of a linear model (2) using the n_e profiles $(L_i, C_i),\ i = 1, \ldots, n_e$. Prediction error can be computed using the $n - n_e$ remaining profiles from other elephant seals such that

$$\mathrm{SSE}_e = \frac{1}{(n - n_e)} \sum_{j=1}^{n-n_e} \int_{Z_m}^{Z_M} \left(C_j(z) - \widehat{C}_j^{(e)}(z) \right)^2 dz,$$

where $\widehat{C}_j^{(e)}$ is the predicted Chl *a* profile using the linear model whose parameters have been estimated with data from elephant seal *e*.

Results

Cross validation steps indicate that taking five basis functions for both Chl *a* and light profiles gives a reasonable minimum error between fitted and the predicted Chl *a* profiles (Fig. 4). With the set of 208 pairwise observed curves, parameters of the functional linear model are estimated as well as bootstrap prediction intervals (see Appendix A).

Several examples of these results are represented in Fig. 5. The first eight panels (Fig. 5.1–.8) display well predicted Chl *a* profiles. Black curves (observed profiles) and red curves (predicted profiles) have same shape and amounts of Chl *a* predicted by the functional linear model every depth match fairly well to raw observations.

However, two types of prediction errors can be distinguished. Firstly, there are prediction problems at surface or in depth (Fig. 5.9 and .10). These patterns are likely to be the consequence of a greater light attenuation. Particles in suspension in the water column other than phytoplankton can contribute to light attenuation, *e.g.* inorganic particles such as CDOM as well as zooplankton. Furthermore, fluorescence in itself is only a proxy of phytoplankton concentration and fluorescence response is known to vary

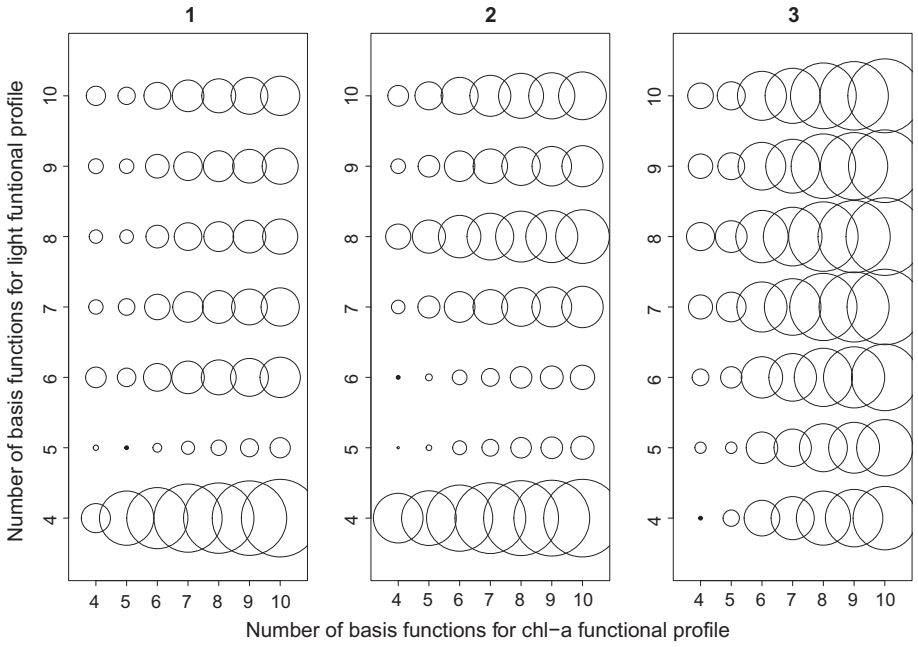

Fig. 4. Errors between observed and predicted Chl *a* profiles, according to number of basis function used to fit Chl *a* and light profiles for each elephant seal. The circle radius is proportional to the value of SSE_{CV}, equal to 0 for $(K^{\star}, L^{\star})$. Five basis functions are required to adjust Chl *a* and light profiles for mammal 1 (panel 1). For elephant 2 (panel 2), four basis functions for Chl *a* and six basis functions for light profiles. Taking four basis functions for elephant 3 gives minimum prediction error (panel 3).

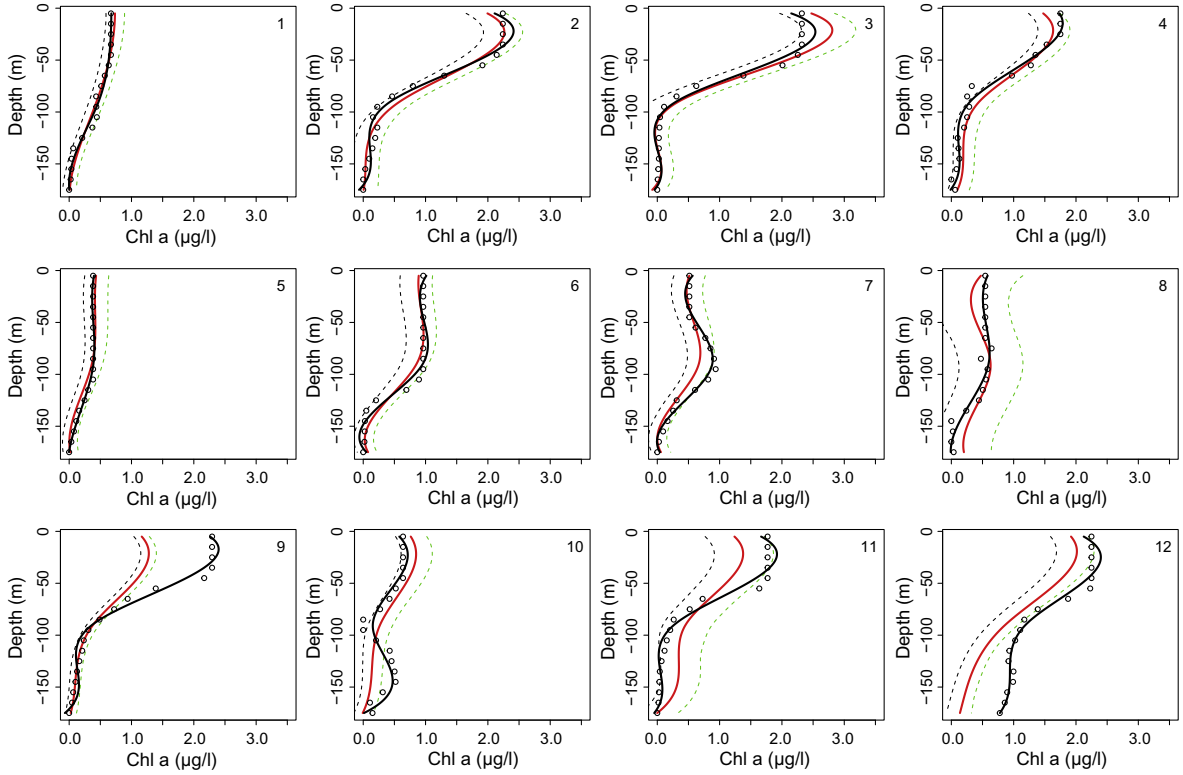

Fig. 5. Examples of prediction of Chl *a* profiles. Red curves are observed curves and black curves are predicted profiles. Profiles 1, 2, 3, 4, 9 and 12 were taken on the path of elephant seal 1; profiles 5, 6, 7, 10 and 11 recorded from elephant seal 2, and profile 8 from elephant seal 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

according to phytoplankton species and physiological state (Xing et al., 2012). Therefore, for a given Chl *a* concentration estimated from the fluorometer, we may expect a variation in light attenuation related to phytoplankton species. Secondly, profiles can be ill-predicted all along depth with a high prediction error. Predicted profiles can be either of very different shape (Fig. 5.11) or the whole predicted curve is shifted from the observed profile even if the shape has been well captured (Fig. 5.12). This last problem

refers to an offset problem: when the sensor does not measure Chl *a*, it returns to 0.

A global measure of goodness of fit can then be proposed by constructing a coefficient of determination. For each pair of profiles, define

$$R_i^2 = 1 - \frac{\frac{1}{n}\int \left(C_i(z) - \widehat{C}_i(z)\right)^2 dz}{\frac{1}{n}\int \left(C_i(z) - \overline{C}(z)\right)^2 dz},$$

where $C_i(z)$ is an observed Chl *a* profile, $\overline{C}(z)$ the mean function of the sample $\{C_1, \ldots, C_n\}$ defined as

$$\overline{C}(z) = \frac{1}{n}\sum_{i=1}^{n} C_i(z),$$

and $\widehat{C}_i(z)$ the predicted Chl *a* profile for observation $C_i(z)$. If the linear relationship between C and L' holds at every spatial position, this quantity ranges from 0 to 1 and is dimensionless.

Fig. 6 displays boxplots of R_i^2 between the predicted and observed Chl *a* profiles for each elephant seal path (1, 2 and 3). In the worst case (seal 2), 90% of the predicted profiles have got a R^2 above 0.41. Nevertheless, the above definition of the R_i^2 does not prevent negative values if the total variance is smaller than the sum of the predicted squared errors. This indicates that the assumptions of the linear model fail. The linear model predictions are worse than taking the mean function as prediction. Then, eight pairs have negative values (−0.008, −0.07, −0.13, −0.23, −0.43, −1.13, −2.15, −21.77) for path 1. There are twenty-six pairs of profiles with a negative R^2 with values ranging between −0.09 and −14.96 for path 2. Finally, path 3 gave thirteen pairs with a negative R^2 value (−0.0001, −0.11, −0.52, −0.71, −0.06, −0.11, −0.03, −0.08, −1.07, −0.31, −0.17, −0.12, −0.31). These ill-predicted profiles often appear in cases where observations of Chl *a* are close to zero, as mentioned earlier.

The values of the R_i^2 are also degraded when prediction errors are estimated in the worst case *i.e.* by cross-validation between animals. However, R_i^2 retains high values for path 1 and path 2 (not presented here).

As the prediction errors for predicted Chl *a* profiles are good, the functional linear model can be used to predict Chl *a* concentration in places where Chl *a* is not sampled but where light profiles are available. This is illustrated in Fig. 7.1 and .2 with two examples of predicted Chl *a* profiles along the daily course of an elephant seal. Successive predicted Chl *a* profiles reveal a high degree of temporal and therefore spatial variability. The mean travel distance between dives is 1.4 ± 0.8 km (range 0.08–5.4 km) under reasonable assumption that elephant seal swimming speed is constant. Wide variations in Chl *a* concentration are detected at the scale of 2–3 dives corresponding to a sub-mesoscale spatial variation (*i.e.* 3–5 km) of the phytoplankton.

Discussion

The main objective of this work was to predict Chl *a* profiles C from the derivative of log-light profiles L' and to highlight changes in Chl *a* concentration at fine spatial scale in the euphotic layer of the Antarctic ocean. For that purpose, we have proposed the construction of a linear model of the form:

$$C = \alpha + B(L') + \varepsilon, \tag{3}$$

where both C and L' are functions which is the originality of that work. The operator B acts on the predictive variable L' as a linear transformation such that:

$$B(L')(z) = \int_{Z_m}^{Z_M} \beta(s,z)L'(s)ds. \tag{4}$$

We have provided technical solutions for the estimation of parameters $\alpha(\cdot)$ and $\beta(\cdot,\cdot)$ which are functions (see Appendix A). We placed the problem in a statistical framework and more precisely in a forecasting framework. It is worthy to note that this approach is not intended to explain, from a mechanistic point-of-view, relationships between Chl *a* and light. Our main purpose is to propose a flexible statistical model constructed around physical assumptions, with good predictive capabilities while taking into account the functional nature of the data: both C and L are observed profiles that arrive as a discrete set of sampled points.

The proposed method is not equivalent to straightforwardly compare the observed pointwise data of fluorescence to those of light using multivariate statistical methods (canonical analysis, for instance). Purposely, the functional approach presents the great advantage to integrate the "vertical link" that connects pointwise

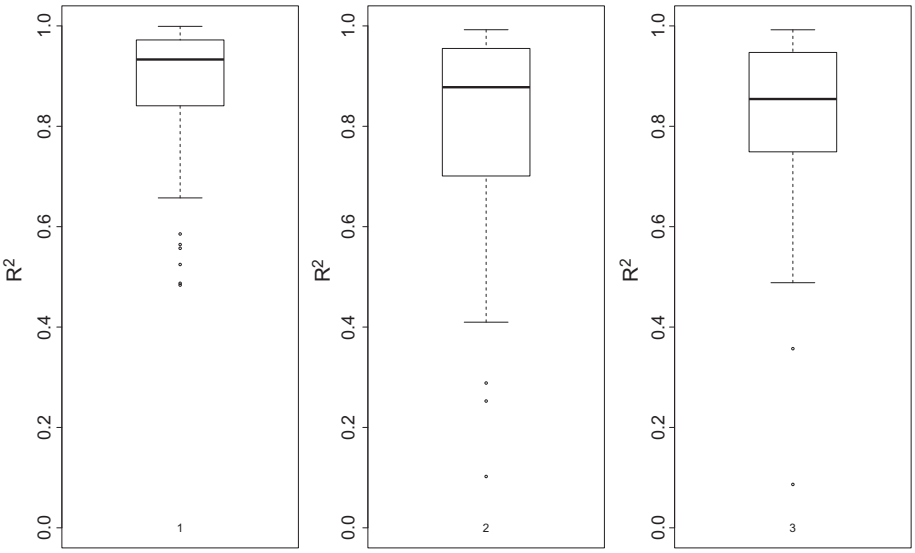

Fig. 6. Boxplots of R^2 between observed and predicted Chl *a* profiles for each path of elephant seal (1, 2 and 3). The best goodness of fit concerns the first elephant, with a median R^2 equals to 0.93 (1).

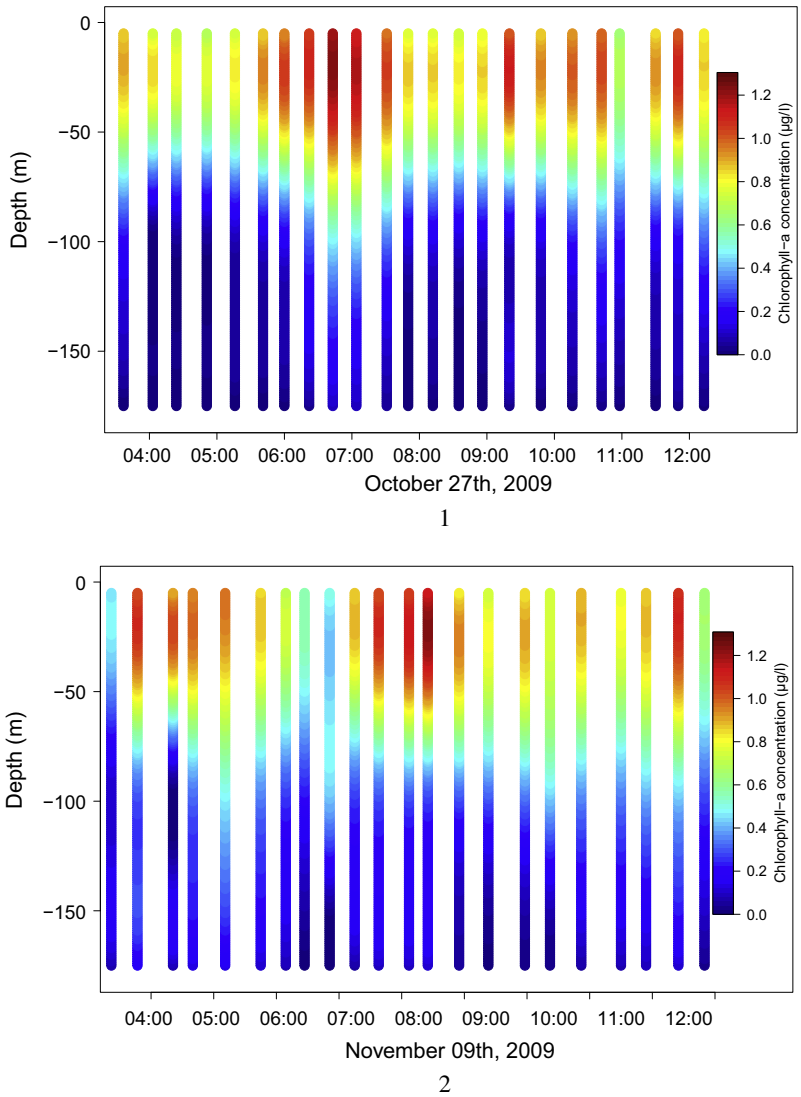

Fig. 7. Examples of prediction of Chl *a* concentration in a range between two measured Chl *a* profiles. Each predicted profile of Chl *a* is realized where light measure is available. Spatial structures occur, related to sub-mesoscale variations of Chl *a*. The more color is red, the more Chl *a* concentration is high. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

observations along depth, a way to include the shape of the profiles in the analysis. All classical multivariate methods used to compare block structured variables (fluorescence data *vs* light data along depth) provide invariant results under any permutation of observations.

Another important reason for the use of a functional approach relies on the fact that it is not an obvious task to match, at corresponding depth, light and Chl *a* observations when they have not been sampled simultaneously. This is not our case but many situations in oceanography face with the problem of comparing variables that are potentially never available at the same depth. Usually, the naive solution to construct a data table is to linearly interpolate each sampled profile between two successive vertical observations. A set of interpolated values is then resampled for every variables on a common mesh. This approach is not suitable for two reasons. Data are in most cases corrupted by noise: the form of the interpolated curve is entirely dependent of the noise level. And if profiles own complex structures, a great number of resampling depths (the mesh size) can potentially be needed. This often leads to construct data tables with a great number of artificially correlated variables, compared to the number of available observations.

The solution we proposed is a smoothing step with B-spline regression which allows to include several data-driven constraints when constructing a continuous profile. The B-spline regression gives the ability to construct profiles including both regularity constraints or monotonic constraints as required for light data. Each profile can then be expressed with a low number of basis coefficients. Moreover, using physical assumptions, we have shown that the statistical model uses the derivative of light data as predictive variable. This derivative makes sense because regularity conditions are met using a B-spline decomposition. Otherwise, computing the derivative of a profile using noisy pointwise raw data would lead to disastrous estimation. Finally, when fitting the data the regularization step eliminates some undesirable sources of variability such as those provided by the measurement devices (see Nerini et al. (2010) for more details).

This last remark leads us to discuss the remainder term ε of the linear model which potentially includes many types of random

errors as already suggested in Jaud et al. (2012). These errors are generated by many factors such as:

- the animal diving behavior that can change the orientation of the sensor and then, the amount of light detected by the instruments,
- the amount of Chl *a* which is a complex function of the fluorescence,
- the presence of non-algal particles such as CDOM of zooplankton that greatly deteriorates the measure of light,
- the sub-surface quenching correction,
- other unknown factors that potentially act as a perturbation.

Despite many sources of error listed above, results show pretty good predictions confirmed using two measures of accuracy: a cross-validated extension of the classical R^2 and the construction of bootstrapped prediction intervals. And the main objective of a predictive model is to provide accurate predictions, making a compromise between the complexity of the model (*i.e.* the number of parameters and state variables) and its forecasting capabilities. For most profiles, the proposed model works well: the shape of profiles and the total amount of Chl *a* in the euphotic layer are reasonably predicted.

Even if great care has been taken to select raw data sampled during daylight hours throughout the course of a seal, some particular profiles are problematic. The main errors came from profiles for which the amount of Chl *a* is close to zero. In an attempt to correct instrument bias and calibration issues, the data have been pre-processed before the statistical analysis. Data were corrected for the offset (shifting of the origin) and for the quenching (decrease of fluorescence intensity in the sub-surface area). This pretreatment has a relative influence on the Chl *a* profiles adjustment and future efforts should be made to improve these corrections. In the same way, temperature or salinity profiles could be added as covariates in the linear model. These variables are of critical importance in controlling the vertical distribution of phytoplankton. Therefore they may be of great help in assessing more precisely the depth of decline in Chl *a* concentration within the water column (Boyd, 2002; Chiswell, 2011; Taylor and Ferrari, 2011). It will also be interesting to assess the importance of physical variables when predicting the distribution of Chl *a* useful for quenching correction. For all that purposes, the linear functional model must be generalize to the multivariate case.

Regarding the global quality of prediction errors, we proposed to predict Chl *a* profiles from light measurements in areas where no fluorescence data were available. Because of the good predictive capabilities of the linear model, clear structures that appeared along the trajectory of seals are qualified of sub-mesoscale variability in Chl *a* concentrations. Nevertheless, a more complete analysis of these structures must be completed with additional data. And for profiles recorded at night, kriging methods when data are curves could be used as well (Nerini et al., 2010).

Finally, *in situ* high spatial resolution of vertical profiles of phytoplankton concentration provided by the Southern elephant seals from light attenuation profiles combined with remote sensing ocean color and sea surface temperature images represent a significant contribution in assessing the (sub-)mesoscale spatial structures of areas of ecological importance within the Kerguelen Region. The simultaneous collection of information on Southern elephant seals foraging success assessed from head-mounted accelerometers (Gallon et al., 2013; Guinet et al., 2013b) with concomitant fine scale oceanographic variables, including the spatial distribution of phytoplankton, should provide new insight into the spatial structuring of the prey field in relation to the oceanographic landscape at (sub-)mesoscale.

Acknowledgments

The authors wish to thank the anonymous reviewers for their valuable comments and for the improvement of this work. the Provence-Alpes-Côte-d'Azur Regional Council is also thanked for funding. This work was funded by the ANR project IPSOS-SEAL (French National Research Agency). The authors also thank IPEV and Total Foundation.

Appendix A. Parameters estimation for functional linear model with functional covariate

Suppose that we dispose of n pairs $\{(L_i, C_i),\ i = 1, \ldots, n\}$ which define a set of matched curves sampled from functional variables L and C. We suppose that these functions belong to $\mathcal{H} = \mathbb{L}^2[Z_m; Z_M]$ the space of square integrable functions defined on the bounded interval $[Z_m; Z_M]$. This functional space is equipped with the inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. We consider a functional regression model in which the function $L(s),\ s \in [Z_m, Z_M]$ is used as a covariate to explain the variation of the response curve $C(z),\ z \in [Z_m, Z_M]$.

The most general version of a functional linear model using a functional covariate is given by

$$C = \alpha + B(L) + \varepsilon,$$

where the intercept $\alpha(z)$ is a functional parameter and B is a linear operator such that

$$B(L)(z) = \int_{Z_m}^{Z_M} \beta(s, z) L(s) ds.$$

The kernel $\beta(s, z)$ of the operator B is a bivariate function that acts as a regression coefficient. It potentially gives the influence of $L(s)$ on $C(z)$ at any value of z. The functional remainder ε gives the error between the model and the function C. The coefficients α and β must be estimated using the sample at hands.

The search for estimates $\widehat{\alpha}$ and $\widehat{B}$ is achieved when minimizing the expectation of the quadratic error

$$\mathrm{SSE}(\alpha, B) = \mathbb{E}\left(\|\varepsilon\|^2\right).$$

The solution of that minimization problem leads to the normal equations for the functional linear model

$$\begin{cases} V_L B = V_{LC}, \\ \alpha = \mu_C - B(\mu_L), \end{cases} \tag{A.1}$$

where functions μ_L and μ_C are expectations of variables L and C respectively, V_L is the variance–covariance operator for variable L and V_{LC} is the cross-covariance operator between variables L and C. Empirical versions of estimators for both covariance operators and for mean functions are computed from the sample $\{(L_i, C_i),\ i = 1, \ldots, n\}$ as

$$\begin{cases} \widehat{V}_L = \frac{1}{n}\sum_{i=1}^{n} (L_i - \widehat{\mu}_L) \otimes (L_i - \widehat{\mu}_L), \\ \widehat{V}_{LC} = \frac{1}{n}\sum_{i=1}^{n} (L_i - \widehat{\mu}_L) \otimes (C_i - \widehat{\mu}_C), \\ \widehat{\mu}_L = \frac{1}{n}\sum_{i=1}^{n} L_i, \\ \widehat{\mu}_C = \frac{1}{n}\sum_{i=1}^{n} C_i, \end{cases}$$

where the tensor product of two elements X and Y of $\mathcal{H}$ is the rank one operator such that $[X \otimes Y](f) = \langle X, f \rangle Y$ for all $f \in \mathcal{H}$. The straightforward estimators $\widehat{B}$ of B and $\widehat{\alpha}$ of α are obtained by replacing empirical estimators in (A.1) such that

$$\begin{cases} \widehat{V}_L \widehat{B} = \widehat{V}_{LC}, \\ \widehat{\alpha} = \widehat{\mu}_C - \widehat{B}(\widehat{\mu}_L). \end{cases}$$

One way to achieve the above calculus of is to decompose both $L(z)$ and $C(z)$ as a linear combination of known basis functions. With this decomposition, the problem of estimation, including the calculus of the inverse of $\widehat{V}_L$, can then be handled as a known multivariate problem, working on the coefficients of the basis decomposition (He et al., 2010). New interesting solutions for solving normal equations, based on PCA decomposition, are also provided in Yao et al. (2005) and Crambes and Mas (2013).

Appendix B. Bootstrap pointwise prediction intervals

Once the parameters have been estimated, the functional linear model gives a predicted value of the response curve at depth z such that

$$\widehat{C}_i(z) = \widehat{\alpha}(z) + \int_{Z_m}^{Z_M} \widehat{\beta}(s,z) L_i(s) ds,$$

where $\widehat{\alpha}$ and $\widehat{\beta}$ are the functional parameters estimated using the initial sample $\{(L_i, C_i),\ i = 1, \ldots, n\}$ constructed with their right number of basis functions. Denote $\widehat{\varepsilon}_i(z) = C_i(z) - \widehat{C}_i(z)$ the estimated residuals at fixed depth z. The forecast of an independent observation $C_{n+1}(z)$ is then

$$\widehat{C}_{n+1}(z) = \widehat{\alpha}(z) + \int_{Z_m}^{Z_M} \widehat{\beta}(s,z) L_{n+1}(s) ds.$$

A bootstrap replication of the initial observations and a future value are given by the pairs $\left\{\left(L_i, C_i^\star\right),\ i = 1, \ldots, n+1\right\}$ such that

$$C_i^\star(z) = \widehat{\alpha}(z) + \int_{Z_m}^{Z_M} \widehat{\beta}(s,z) L_i(s) ds + \varepsilon_i^\star(z)$$

for $i = 1, \ldots, n$ and

$$C_{n+1}^\star(z) = \widehat{\alpha}(z) + \int_{Z_m}^{Z_M} \widehat{\beta}(s,z) L_{n+1}(s) ds + \varepsilon_{n+1}^\star(z),$$

where $\varepsilon_1^\star(z), \ldots, \varepsilon_n^\star(z)$ and $\varepsilon_{n+1}^\star(z)$ are obtained by sampling with replacement from the empirical distribution of the pointwise residuals

$$\widehat{F}_z(x) = \#\{\widehat{\varepsilon}_i(z) \leqslant x\}/n.$$

Let $\widehat{\alpha}^\star$ and $\widehat{\beta}^\star$ be the estimated parameters using the bootstrap sample $B = \left\{\left(L_i, C_i^\star\right),\ i = 1, \ldots, n\right\}$ and define the prediction error

$$e_{n+1}^\star(z) = C_{n+1}^\star(z) - \left[\widehat{\alpha}^\star(z) + \int_{Z_m}^{Z_M} \widehat{\beta}^\star(s,z) L_{n+1}(s) ds\right],$$

which has distribution $G_z^\star(\cdot; n+1)$. The pointwise bootstrap γ-prediction interval of observation $C_{n+1}(z)$ is given with

$$I_\gamma^{(B)}(z; n+1) = \left[\widehat{C}_{n+1}(z) + G_z^{\star -1}((1-\gamma)/2; n+1), \widehat{C}_{n+1}(z) + G_z^{\star -1}((1+\gamma)/2; n+1)\right],$$

where $G_z^{\star -1}(\cdot; n+1)$ is the quantile function that gives the $(1-\gamma)/2$ lower quantile and the $(1+\gamma)/2$ upper quantile $(0 \leqslant \gamma \leq 1)$.

The distribution of the prediction errors is then simulated using bootstrapped errors obtained by sampling the empirical distribution function $\widehat{F}_z$ in place of F_z and using $\widehat{\alpha}$ and $\widehat{\beta}$ as true coefficients of the regression instead of α and β. The estimation of a prediction interval can then be repeated for any value of depth z to form the functional prediction envelope (Fig. 3).

In our case, the covariate is the derivative of log-L.

References

Arrigo, K., Worthen, D., Lizotte, M., Dixon, P., Dieckmann, G., 1997. Primary production in Antarctic sea ice. Science 276 (5311), 394–397.

Behrenfeld, M., Falkowski, P., 1997. A customer's guide to phytoplankton primary productivity models. Limnology and Oceanography 42 (7), 1479–1491.

Behrenfeld, M., O'Malley, R., Siegel, D., McClain, C., Sarmiento, J., Feldman, G., Milligan, A., Falkowski, P., Letelier, R., Boss, E., 2006. Climate-driven trends in contemporary ocean productivity. Nature 444, 752–755.

Biuw, M., Boehme, L., Guinet, C., Hindell, M., Costa, D., Charrassin, J.-B., Roquet, F., Bailleul, F., Meredith, M., Thorpe, S., Tremblay, Y., McDonald, B., Park, Y.-H., Rintoul, S., Bindoff, N., Goebel, M., Crocker, D., Lovell, P., Nicholson, J., Monks, F., Fedak, M., 2007. Variations in behavior and condition of a Southern Ocean top predator in relation to in situ oceanographic conditions. Proceedings of the National Academy of Sciences 104 (34), 13705–13710.

Block, B., Costa, D., Boehlert, G., Kochevar, R., 2002. Revealing pelagic habitat use: the tagging of Pacific pelagics program. Oceanologica Acta 25 (5), 255–266.

Boehlert, G., Costa, D., Crocker, D., Green, P., O'Brien, T., Levitus, S., Le Boeuf, B., 2001. Autonomous pinniped environmental samplers: using instrumented animals as oceanographic data collectors. Journal of Atmospheric and Oceanic Technology 18 (11), 1882–1893.

Boehme, L., Lovell, P., Biuw, M., Roquet, F., Nicholson, J., Thorpe, S., Meredith, M., Fedak, M., 2009. Technical note: animal-borne CTD-satellite relay data loggers for real-time oceanographic data collection. Ocean Science 5 (4), 685–695.

Bouguer, P., 1729. Essai d'optique sur la gradation de la lumière. Claude Jombert, Paris.

Boyd, P., 2002. Environmental factors controlling phytoplankton processes in the Southern Ocean. Journal of Phycology 38 (5), 844–861.

Buesseler, K., Boyd, P., 2003. Will ocean fertilization work? Science 300 (5616), 67–68.

Charrassin, J., Hindell, M., Rintoul, S.R., Roquet, F., Sokolov, S., Biuw, M., Costa, D., Boehme, L., Lovell, P., Coleman, R., Timmermann, R., Meijers, A., Meredith, M., Park, Y.-H., Bailleul, F., Goebel, M., Tremblay, Y., Bost, C.-A., McMahon, C.R., Field, I.C., Fedak, M.A., Guinet, C., 2008. Southern Ocean frontal structure and sea-ice formation rates revealed by elephant seals. Proceedings of the National Academy of Sciences 105 (33), 11634–11639.

Chiswell, S., 2011. Annual cycles and spring blooms in phytoplankton: don't abandon Sverdrup completely. Marine Ecology Progress Series 443, 39–50.

Crambes, C., Mas, A., 2013. Asymptotics of prediction in functional linear regression with functional outputs. Bernoulli 19, 2627–2651.

Craven, P., Wahba, G., 1979. Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. Numerische Mathematik 31, 377–403.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York.

Fedak, M., 2013. The impact of animal platforms on polar ocean observation. Deep Sea Research Part II: Topical Studies in Oceanography 88, 7–13.

Fedak, M., Lovell, P., McConnell, B., Hunter, C., 2002. Overcoming the constraints of long range radio telemetry from animals: getting more useful data from smaller packages. Integrative and Comparative Biology 42 (1), 3–10.

Gallon, S., Bailleul, F., Charrassin, J.-B., Guinet, C., Bost, C.-A., Handrich, Y., Hindell, M., 2013. Identifying foraging events in deep diving southern elephant seals, Mirounga leonina, using acceleration data loggers. Deep Sea Research Part II: Topical Studies in Oceanography 88–89, 14–22.

Guinet, C., Xing, X., Walker, E., Monestiez, P., Marchand, S., Picard, B., Jaud, T., Authier, M., Cotté, C., Dragon, A.-C., Diamond, E., Antoine, D., Lovell, P., Blain, S., D'Ortenzio, F., Claustre, H., 2013a. Calibration procedures and first data set of Southern Ocean Chlorophyll-*a* profiles collected by elephant seal equipped with a newly developed CTD-fluorescence tags. Earth System Science Data 5, 15–29. http://dx.doi.org/10.5194/essd-5-15-2013.

Guinet, C., Vacquié-Garcia, J., Picard, B., Bessigneul, G., Lebras, Y., Dragon, A.-C., Viviant, M., Arnould, J., Bailleul, F., 2013b. Southern Elephant Seal foraging success in relation to temperature and light conditions: insight on their prey distribution, unpublished results.

He, G., Müller, H.-G., Wang, J.-L., Yang, W., 2010. Functional linear regression via canonical analysis. Bernoulli 16 (3), 705–729.

Holm-Hansen, O., Amos, A., Hewes, C., 2000. Reliability of estimating Chlorophyll *a* concentrations in Antarctic waters by measurement of in situ Chlorophyll *a* fluorescence. Marine Ecology Progress Series 1996, 103–110.

Hosseini-Nasab, M., 2012. Cross-validation approximation in functional linear regression. Journal of Statistical Computation and Simulation, 1–11. http://dx.doi.org/10.1080/00949655.2012.662502, <http://www.tandfonline.com/doi/abs/10.1080/00949655.2012.662502>.

Jaud, T., Dragon, A.-C., Garcia, J., Guinet, C., 2012. Relationship between Chlorophyll *a* concentration, light attenuation and diving depth of the southern elephant seal mirounga leonina. PLoS one 7 (10), e47444. http://dx.doi.org/10.1371/journal.pone.0047444.

Kiefer, D., 1973. Fluorescence properties of natural phytoplankton populations. Marine Biology 22, 263–269.

Marra, J., 1997. Analysis of diel variability in chlorophyll fluorescence. Journal of Marine Research 55, 767–784.

Maxwell, K., Johnson, G., 2000. Chlorophyll fluorescence – a practical guide. Journal of Experimental Botany 51 (345), 659–668.

McMahon, C.R., Autret, E., Houghton, J., Lovell, P., Myers, A., Hays, G., 2005. Animal-borne sensors successfully capture the real-time thermal properties of ocean basins. Limnology and Oceanography: Methods 3, 392–398.

Meyer, M.C., 2012. Constrained penalized splines. Canadian Journal of Statistics 40 (1), 190–206. http://dx.doi.org/10.1002/cjs.10137, ISSN 1708-945X.

Morel, A., 1988. Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). Journal of Geophysical Research: Oceans (1978–2012) 93 (C9), 10749–10768.

Nerini, D., Monestiez, P., Manté, C., 2010. Cokriging for spatial functional data. Journal of Multivariate Analysis 101, 409–418.

Queguiner, B., Brzezinski, M., 2002. Biogenic silica production rates and particulate organic matter distribution in the Atlantic sector of the Southern Ocean during austral spring 1992. Deep Sea Research Part II: Topical Studies in Oceanography 49, 1765–1786.

Ramsay, J., Silverman, B., 2005. Functional Data Analysis. Springer, New-York.

Taylor, J.R., Ferrari, R., 2011. Shutdown of turbulent convection as a new criterion for the onset of spring phytoplankton blooms. Limnology and Oceanography 56 (6), 2293–2307. http://dx.doi.org/10.4319/lo.2011.56.6.2293.

Teo, S., Kudela, R., Rais, A., Perle, C., Costa, D., Block, B., 2009. Estimating chlorophyll profiles from electronic tags deployed on pelagic animals. Aquatic Biology 5, 195–207.

Vacquié-Garcia, J., 2014. Variation spatio-temporelle de l'activité d'alimentation des éléphants de mer en relation avec les paramètres physiques et biologiques de l'environnement. Ph.D. thesis, Université Paul Sabatier.

Xing, X., Claustre, H., Blain, S., D'Ortenzio, F., Antoine, D., Ras, J., Guinet, C., 2012. Quenching correction for in vivo chlorophyll fluorescence measured by instrumented elephant seals in the Kerguelen region (Southern Ocean). Limnology and Oceanography: Method 10, 483–495. http://dx.doi.org/10.4319/lom.2012.10.483.

Yao, F., Müller, H.G., Wang, J.L., 2005. Functional linear regression analysis for longitudinal data. The Annals of Statistics 33, 2873–2903.