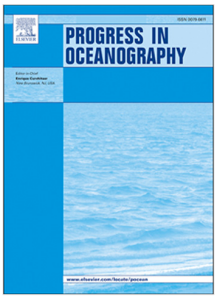
Journal Pre-proof



Swimming in an ocean of curves: A functional approach to understanding elephant seal habitat use in the Argentine Basin

Nadège Fonvieille, Christophe Guinet, Martin Saraceno,
Baptiste Picard, Martin Tournier, Pauline Goulet, Claudio Campagna,
Julieta Campagna, David Nerini

PII: S0079-6611(23)00163-5
DOI: https://doi.org/10.1016/j.pocean.2023.103120
Reference: PROOCE 103120

To appear in: *Progress in Oceanography*

Received date : 6 March 2023
Revised date : 30 June 2023
Accepted date : 5 September 2023

Please cite this article as: N. Fonvieille, C. Guinet, M. Saraceno et al., Swimming in an ocean of curves: A functional approach to understanding elephant seal habitat use in the Argentine Basin. *Progress in Oceanography* (2023), doi: https://doi.org/10.1016/j.pocean.2023.103120.

Swimming in an ocean of curves: A functional approach to understanding elephant seal habitat use in the Argentine Basin

Nadège Fonvieille[a,b,⋆], Christophe Guinet[b], Martin Saraceno[c,d], Baptiste Picard[b], Martin Tournier[b], Pauline Goulet[b], Claudio Campagna[e], Julieta Campagna[f] and David Nerini[a]

[a]Mediterranean Institute of Oceanography, Aix Marseille Université, CNRS/INSU, Université de Toulon, IRD, 13288, Marseille cedex 09, France

[b]Centre of Biological Studies Chizé (CEBC), Villiers-en-Bois, France

[c]CONICET - Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Ciencias de la Atmósfera y los Océanos, Mepukada, Buenos Aires, Argentine

[d]CNRS – IRD – CONICET – UBA. Instituto Franco- Argentino para el Estudio del Clima y sus Impactos (UMI 3351 IFAECI), Buenos Aires, Argentina

[e]Wildlife Conservation Society, Marine and Argentina Programs, Buenos Aires, Argentina

[f]Centro para el Estudio de Sistemas Marinos (CESIMAR), CENPAT-CONICET, Puerto Madryn, Argentina

Keywords: Southern elephant seals, Habitat use, Functional analysis, Model-based clustering, Brazil-Malvinas Confluence

Abstract

In recent decades, southern elephant seals (SES) have become a species of particular importance in ocean data acquisition. The scientific community has taken advantage of technological advances coupled with suitable SES biological traits to record numerous variables in challenging environments and to study interactions between SES and oceanographic features. In the context of big dataset acquisition, there is a growing need for methodological tools to analyze and extract key data features while integrating their complexity. Although much attention has been paid to study elephant seal foraging strategies, the continuity of their surrounding three-dimensional environments is seldom integrated. Knowledge gaps persist in understanding habitat use by SES, while the representativeness of a predator-based approach to understanding ecosys-

tem structuring is still questioned. In this study, we explore SES habitat use by using a functional data analysis approach (FDA) to describe the foraging environment of five female elephant seals feeding in the Southwestern Atlantic Ocean. Functional principal component analysis followed by model-based clustering were applied to temperature and salinity (TS) profiles from Mercator model outputs to discriminate waters sharing similar thermohaline structures. Secondly, in situ TS profiles recorded by the SES were employed to determine the habitat visited within the range of potential environments identified from the model data. Four Functional Oceanographic Domains (FOD) were identified in the Brazil-Malvinas Confluence, all visited, in varying proportion, by four of the five females studied. We found that the females favored areas where all the FODs converge and mix, generating thermal fronts and eddies. Prey-capture attempts increased in such areas. Our results are in accordance with previous findings, suggesting that (sub-)mesoscale features act as biological hotspots. This study highlights the potential of coupling FDA with model-based clustering for describing complex environments with minimal loss of information. As well as contributing to better understanding of elephant seal habitat use and foraging strategies, this approach opens up a wide range of applications in oceanography and ecology.

1 Introduction

Bio-logging is defined as the use of miniaturized animal-borne devices that provide data on animal movement, behavior, physiology or environment. Bio-loggers were first developed to investigate the at-sea behavior and distribution of enigmatic marine megafauna (Block et al., 2011; Boehme et al., 2012; Hays et al., 2016; Hussey et al., 2015; Jonsen et al., 2007; Payne et al., 2018). Nowadays, a wide range of small-size instruments is available to sample, at high frequency, fine-scale foraging behavior along encountered oceanographic conditions (Block et al., 2016; Evans et al., 2013; Fedak et al., 2004; Guinet et al., 2014; Harcourt et al., 2019). Large marine mammals, especially pinnipeds, have become species of particular importance in marine-environment data acquisition through bio-logging approaches (Bailleul et al., 2015; Fedak, 2013; Hindell et al., 2020; March et al., 2020; Roquet et al., 2013). In this context, southern elephant seals (*Mirounga leonina*, Linnaeus 1758, SES hereafter)

may currently represent the most significant marine animal contribution towards the collection of in situ oceanographic data. For instance, two-thirds of the temperature and salinity profiles available for the Southern Ocean are provided by SES equipped with Satellite Relay Data Loggers-CTD (SRDL-CTD) as part of the AniBOS network within the Global Ocean Observing System (McMahon et al., 2021; Roquet et al., 2014).

The most commonly-monitored environmental variables include physical (temperature, salinity, light), bio-geochemical (chlorophyll-*a*, oxygen) as well as biological variables such as bioluminescence (Bailleul et al., 2015; Guinet et al., 2014; Jaud et al., 2012; Vacquie-Garcia et al., 2012). More recently, the assessment of mid-trophic level organisms including SES preys, ranging in size from a few millimeters to a few centimeters, was made possible by implementing a miniature echo-sounding device (Goulet et al., 2019; Tournier et al., 2021). SES behavior information includes at-sea locations, fine-scale three-dimensional (3D) diving behavior, swimming effort, as well as prey-capture attempts (PCA, Le Bras et al., 2017; Le Bras et al., 2016). These technological advances have led to substantial multivariate environmental and behavioral datasets used by physicists to investigate fine-scale processes (Aubone et al., 2021; Carse et al., 2015; Roquet et al., 2013), and biologists to study the at-sea ecology of these marine predators, thus revealing critical foraging habitats (Labrousse et al., 2018) and providing key information on the distribution and accessibility of food resources (Goulet et al., 2019; Tournier et al., 2021). Finally, combining SES foraging behavior with oceanographic conditions encountered along an animal's trajectory is useful for understanding the trophic web structure and for studying organisms in relation to key physical processes (Bailleul et al., 2010; Cotté et al., 2015; Della Penna et al., 2015; Dragon et al., 2010; Rivière et al., 2019).

However, such a predator-based approach also has its limitations. The specific behavior of individual animals may induce selective bias of the sampled in situ environmental conditions (Conn et al., 2017; Dinsdale and Salibian-Barrera, 2019). Therefore, SES bio-sampling may not be suited for monitoring the full range of oceanographic conditions available to these animals. Consequently, it can be critical to consider the visited habitats as only part of a realm of possibilities when assessing the foraging habitat targeted by these predators. Although some species can reasonably be assumed to be mainly constrained by surface environmental conditions, like surface-

feeding seabirds (Schreer and Kovacs, 1997), diving species evolve in different layers of the water column. SES are deep divers (Hindell et al., 2016), and as such, one major difficulty when studying their foraging behavior is positioning it within a broader 3D oceanographic context. Satellite oceanographic data, while helpful for comparing marine predator tracks with sea surface temperature, salinity, sea surface height and/or ocean colors, cannot inform on the vertical variations of oceanographic conditions. Yet, for an equal distance, vertical gradients of physicochemical variables are much more accentuated than oceanographic gradients at the surface (Wunsch and Ferrari, 2004). Therefore, when investigating the foraging behavior of (deep) diving predators, it is crucial to contextualize this behavior in a 3D oceanographic environment. In this study, we aim to understand the habitat used by elephant seals by analyzing the 3D habitat visited in relation to oceanographic conditions at a regional scale.

Previous studies using classic oceanographic features tend to summarize the structure of the whole water column by a single discrete value like mean temperature or temperature at a given depth (Guinet et al., 2014; Hindell et al., 2016). However, temperature and salinity (hereafter TS) profiles are sampled by SES with SRDL-CTD (0.5 Hz frequency) which provide a considerable dataset over an animal's trip. Reducing these high-resolution profiles to single discrete descriptors leads to a substantial loss of information. Functional data analysis (FDA), first introduced by Ramsay (1982), allows manipulating curves or functions rather than scalars or vectors (discrete measures) used in classic data-analysis techniques (Ramsay and Silverman, 2005). Functional approaches encompass numerous methods and offer a wide range of applications in various fields of sciences (Cuevas, 2014; Ullah and Finch, 2013; Wang et al., 2016). This statistical framework has recently been applied in marine science and showed strong potential for describing vertical oceanographic features at a global scale (Pauthenet et al., 2017; Pauthenet et al., 2019). It is well suited to modeling TS profiles that are continuous variable functions of depth and can be applied to either in situ or modeled data without being affected by the sampling grid. Alongside FDA, various studies have proposed using profile classification models (PCM, Maze, Mercier, and Cabanes, 2017) to identify oceanographic regimes (Boehme and Rosso, 2021; Jones et al., 2019; Maze, Mercier, Fablet, et al., 2017; Rosso et al., 2020). PCMs take advantage of Gaussian mixture modeling (GMM) to find meaningful objective groups in data (Bouveyron et al., 2019). Yet, while these studies use temperature and/or

salinity profiles, they do not employ functional data methodology.

In this study, we propose to combine FDA with model-based clustering within a collection of temperature and salinity profiles. The proposed method shows promising results, allowing the identification of spatio-temporal-coherent oceanographic regimes sharing similar vertical thermohaline structures, named hereafter Functional Oceanographic Domains (FOD). Another originality of this paper resides in the coupling of TS profiles from Mercator model outputs and from head-mounted SES loggers. Profiles from the model are used to describe the vertical thermohaline oceanographic environment at a regional scale and identify FODs, while in situ profiles are used to determine the actual conditions visited by the SES within the range of possibilities provided by the model. This research offers new insights into the bias induced by SES choices in the representativeness of sampled environmental conditions. Moreover, prey-capture rates are analyzed to understand the impacts of oceanographic conditions on SES foraging strategies. This study is conducted on the Patagonian (Argentina) SES population known to forage within the Brazil-Malvinas Confluence (BMC), a region of major importance for that SES population (Campagna et al., 2021).

2 Method

2.1 Study area

The study is carried out on recently acquired data collected by SES breeding in southern Argentina (Península Valdés, 42°57' S, 63°59' W) and mainly feeding in the Southwestern Atlantic Ocean (Campagna et al., 2006; Campagna et al., 2021). This region includes the meeting zone between the Brazilian Current (BC) and the Malvinas Current (MC). The MC finds its origin in the Antarctic Circumpolar Current (ACC) that flows eastward around Antarctica (Piola and Gordon, 1989). After passing the Drake Passage, south of the Patagonian shelf, a portion of the northern branch of the ACC is deflected northward, crossing the North Scotia Ridge mainly through the eastern and western flanks of the Burdwood Bank and the Shag Rocks Passage (Artana et al., 2016) and giving origin to the MC. Rounding east of the Falklands (Malvinas) Islands, the MC finds a way northward along the continental

slope, following isopotential vorticity contours (Saraceno et al., 2004) and carrying cold and nutrient-rich subantarctic waters. Contoured by the Subantarctic Front (SAF), the MC collides with the BC, flowing southward along the Brazil coast at ∼ 38° S (Gordon, 1981; Saraceno et al., 2004). The collision generates a thermohaline front called the Brazil-Malvinas Confluence zone (Deacon, 1937). The BMC creates instabilities generating prominent (sub-)mesoscale structures, including eddies, making the Southwestern Atlantic one of the most energetic regions of the world ocean (Chelton et al., 1990).

The focus area is defined by the black rectangle displayed in Figure 1. This area covers the SES post-breeding trajectories in 2018 and 2019 and includes the BMC region. Land covers 9 % of the area. Within the ocean zone, 30 % is shallower than 200 meters and corresponds to the Patagonian Continental Shelf (Piola and Falabella, 2009). The continental slope ends at 6 000 meters depth, marking the beginning of the Argentine Abyssal Plain also named the Argentine Basin. The GEBCO14 grid is used to define the bathymetry in the area (30 arc second resolution; distributed by the British Oceanographic Data Centre; Weatherall et al., 2015).

2.2 Functional analysis

A bivariate principal component analysis for functional data (FPCA, functional principal component analysis) is performed on temperature and salinity (TS) profiles from Mercator model outputs to describe the vertical structure of the BMC area and extract the principal modes of variability at the regional scale. Model-based clustering is applied on the first principal components (PC) of the FPCA to define regimes sharing similar water properties, the Functional Oceanographic Domains (FOD). In situ TS profiles recorded by five female southern elephant seals (SES) are then projected onto the factorial map of the model-derived FPCA in order to assign each SES dive to one FOD. This step gained information on the oceanographic regimes visited by each female along her trajectory. Figure 2 summarizes the different steps of the method. All analyses are conducted in R (Team, 2021) and with Matlab R2021b.

2.2.1 Datasets

Profiles from elephant seal dives

Nine post-breeding female SES were tagged in october 2018 with different bio-loggers at Península Valdés, following the procedure described in McMahon et al. (2008). The females were captured and anesthetized with an intravenous injection of Zoletil®100 (1:1 combination of tiletamine and zolazepam) and tags were glued to the pelage using Araldite adhesive. The animal manipulations were in accordance with the "Use of Animals for Scientific Reasons (APAFIS)" ethics committee guidelines. Among phocids, the elephant seal is a particularly well-suited species for monitoring the oceans. It forages over extended distances (thousands of kilometers), diving continuously (i.e. 60 times a day) at depths generally ranging between 400 and 1 000 meters, and exceptionally up to 2 000 meters (Hindell et al., 2016). SES return to land to breed and molt at predictable cycles and places, facilitating device recovery and high-frequency data acquisition. Their large size allows scientists to equip them with several multi-sensor loggers with minimal disturbance (McMahon et al., 2008).

All individuals were equipped with a SRDL-CTD tag (Sea Mammal Research Unit, St Andrews, UK). The depth was logged with a 0.5 Hz frequency and 0.5 m resolution. A dive corresponds to a period during which the animal is deeper than 15 m continuously for at least 300 s (Le Bras et al., 2016). Each dive is associated with one temperature and one salinity profile, sampled and corrected from the ascending phase of the dive (Roquet et al., 2011; Siegelman et al., 2019). The CTD tag provides discrete TS measurements with a 0.5 Hz frequency and a 0.01 °C and 0.03 psu accuracy. Locations of the dives were recorded using a Global Positioning System (GPS) for four individuals equipped with a micro-sonar or a DTAG (Goulet et al., 2020; Goulet et al., 2019). The (approximately) 20 m resolution provided by the GPS offers accurate coordinates but its short battery lifetime led to non-localized TS profiles as the CTD continued to record data after the GPS stopped. Argos locations (2-10 km resolution; Lopez et al., 2015) provided by head-mounted SRDL-CTD or SPOT tags (Wildlife Computers, USA) were used to obtain dive coordinates for the other five SES. Head-mounted accelerometer sensors included in the SRDL-CTD, sonar and DTAGs are used to assess prey-capture attempts (PCA, Gallon et al., 2013; Viviant et al., 2010). The prey-capture rate (i.e. PCAs per unit time, referred to as PCR) is calculated from the dive duration (Guinet et al., 2014; Jouma'a et al., 2016) and used as a proxy of foraging performances. A day-night period is assigned to all localized dives based on the sun angle (Guinet et al., 2014). Day periods are defined when the

sun angle is > 12° above the horizon, while night periods are determined as < -12° below the horizon. The threshold of 12° is selected to ensure a clear distinction from dives occurring during the twilight period.

Out of the nine SES equipped, only five foraging within the BMC are kept for the present study (Figure 1). Among the discarded individuals, one of the nine females traveled southward, following the coast of Argentina to Chile, and two others spent the whole trip foraging over the Patagonian slope. A fourth female presenting out-of-range salinity data was also excluded. Those incorrect data were detected as outliers via the FCPA described in Section 2.2.3. To determine oceanographic conditions visited by the five selected SES, all localized and non-localized TS profiles deeper than 450 m are considered, accounting for a total of 9 537 dives. These recorded profiles constitute the in situ dataset used in this work (step 1 Figure 2).

Profiles from Mercator model outputs

The general three-dimensional oceanographic temperature and salinity context of the study area is obtained by the GLORYS12V1 product provided by the Copernicus Marine Environment Monitoring Service (CMEMS, https://doi.org/10.48670/moi-00021). The dataset used to define the environmental conditions, referred to as model-derived profiles, corresponds to daily averaged TS profiles covering the defined area (Figure 1) from October 21, 2018, to January 23, 2019, spanning a total of 95 days. The horizontal resolution of the dataset is 1/12°, which corresponds to approximately 8 km. The output products are displayed on a 193 × 133 horizontal grid (longitude × latitude) and 50 vertical irregular depth levels extended from 0.5 m (surface) to 5 728 m (bottom). Considering that the depth reached by the in situ profiles depends on the SES behavior, it was decided that statistical analysis would be applied on profiles (model and in situ) between 20 m and 450 m. This choice excludes profiles sampled on the Patagonian continental shelf (< 200 m) while capturing most of the vertical variability in the open ocean and retaining a large proportion of in situ profiles. The discrete model-derived profiles are linearly interpolated in order to obtain a uniform vertical grid with 44 x 10 m levels from 20 to 450 m. This step allows avoidance of spurious bumpiness in curve fitting (Pauthenet et al., 2019). Those profiles constitute the model-derived dataset used in this work (step 1 Figure 2).

2.2.2 TS profiles as continuous curves

Temperature and salinity profiles are continuous functions along depth. These data are always available as discrete values, induced by the sampling method. Yet, values within a profile $x(t)$ are ordered based on the parameter t. This link between two consecutive values implies particular considerations. First developed by Ramsay and Silverman (2005), statistical functional methods are explicitly designed to process and study data that are functions, considering one curve as a single entity. Each SES dive or each pixel (i.e. model-derived data) is described by two profiles (temperature and salinity), making the data bivariate functions. Those two profiles are intrinsically linked and regarded as a unique observation.

The first stage of the functional data analysis (step 1 to 2 Figure 2) is to transform the discrete TS profiles into functions $x(z)$, with $z \in [20; 450]$ being the depth in meters (library `fda`, Ramsay et al., 2009). This step is achieved by decomposing the profiles on a cubic B-spline basis (Wahba, 1990). Each profile n is expressed as a linear combination of basis functions $\phi_k(z)$:

$$x_n^i(z) = \sum_{k=1}^{K} \alpha_{n,k}^i \phi_k(z),\ i \in \{T, S\},\ n = 1, ..., N, \tag{1}$$

where the $\alpha_{n,k}^i$ coefficients are estimated by penalized least squares regression (see Supplementary Material) and N is the number of profiles. The number K of basis functions controls the vertical smoothness. Both model-derived and in situ profiles are expressed according to the same basis to facilitate their comparison. As the model-derived dataset used in this study is composed of more than 2×1.5 million profiles (T and S profiles) with 44 values per profile, one goal of the decomposition procedure is to reduce the dimension of the data. A basis of $K = 15$ functions is chosen, allowing sufficient vertical complexity to be kept for both kinds of profiles while summarizing the data with a small number of coefficients (see Supplementary Material). Henceforth, each observation n, consisting of two profiles $(x_n^T(z), x_n^S(z))$, can be expressed as a vector $\boldsymbol{\alpha}_n$ of dimension $2 \times K$, merging temperature and salinity coefficients:

$$\boldsymbol{\alpha}_n = (\alpha_{n,1}^T, ..., \alpha_{n,K}^T\ ;\ \alpha_{n,1}^S, ..., \alpha_{n,K}^S)'. \tag{2}$$

Model-derived coefficients are stored in a matrix $\mathbf{X}_N$ of size $N \times 2K$, with N the

number of model-derived observations (step 2 Figure 2). The mean model-derived observation $\overline{\boldsymbol{\alpha}} = (\overline{\alpha}_1^T, ..., \overline{\alpha}_K^T; \overline{\alpha}_1^S, .., \overline{\alpha}_K^S)'$ is evaluated with $\overline{\alpha}_k^i, i \in \{T, S\}, k = 1, ..., K$, the empirical mean of the N coefficients k.

2.2.3 Functional Oceanographic Domains (FOD)

Functional Principal Component Analysis

Functional principal component analysis (FPCA) is a powerful method to reduce the dimension of the data (Wang et al., 2016). It decomposes the thermohaline structure into modes of variability, allowing simultaneous analysis of the shape variation of temperature and salinity profiles. The search for the main modes of variability is achieved by solving the following eigenvalue problem:

$$\mathbf{VWM}\boldsymbol{b}_l = \lambda_l \boldsymbol{b}_l, \tag{3}$$

associating $\boldsymbol{b}_l$, the l^{th} eigenvector with the λ_l eigenvalue. This step allows us to find the unique decomposition of the matrix **VWM**. The block-structured covariance matrix **V** is computed from the matrix **C** of centered coefficients such as $\mathbf{V} = 1/N \times \mathbf{C'C}$. The centered matrix **C** is obtained by subtracting the coefficients of the average profile $\overline{\boldsymbol{\alpha}}$ from each row of $\mathbf{X}_N$. Matrix **W**, of size $2K \times 2K$, guarantees the metric equivalence between the functional problem (working on functions) and its discrete version (working on coefficients). **M** is the weighting matrix used to normalize the coefficients of temperature and salinity (see Supplementary Material). Each eigenvector generates two eigenfunctions (ξ_l^T, ξ_l^S), also called vertical modes (see more details in Supplementary Material and Pauthenet et al., 2017). A total of $2K$ eigenvectors are obtained, that can be sorted in ascending order according to their associated eigenvalues. The eigenvector associated with the largest eigenvalue corresponds to the first vertical mode. The main factors of variability can be seen as a perturbation of the mean function $(\overline{x}^T, \overline{x}^S)$ by adding or subtracting the eigenfunctions (ξ_l^T, ξ_l^S):

$$\overline{x}^i \pm \sqrt{\lambda_l}\xi_l^i, i \in \{T, S\}. \tag{4}$$

The FPCA is realized with the model-derived TS profiles (step 3 Figure 2). The observations can be projected into a 2D map using the first two principal components

(PC) computed as: $\boldsymbol{c}_l = \mathbf{C}\mathbf{M}^{-1/2}\mathbf{W}^{-1/2}\boldsymbol{b}_l, l = 1, 2$. PCs are the uncorrelated linear combinations of the original variables. They capture the variance of the system. For more details on the bivariate FPCA procedure, the reader is referred to Nerini et al. (2022) and Pauthenet et al. (2017).

Model-based clustering

Oceanographic regimes are identified using model-based clustering (library `mclust`, Scrucca et al., 2016). First introduced by Wolfe in 1963, model-based clustering is a clustering method based on a probability model defined by a finite mixture of multivariate Gaussian distributions, called components. The probability distribution of the P-variate observation $\boldsymbol{y}_n = (c_{n,1}, \ldots, c_{n,P})'$, can be seen as a weighted average of G conditional probability functions f_g, G being the number of components:

$$p(\boldsymbol{y}_n) = \sum_{g=1}^{G} \tau_g f_g(\boldsymbol{y}_n | \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g). \tag{5}$$

The parameter τ_g is the probability that an observation was generated by the g^{th} component while $\boldsymbol{\mu}_g$ and $\boldsymbol{\Sigma}_g$ are parameters that control the shape of f_g. The method allows us to assess uncertainty about the clustering. Each cluster found is modeled by its own probability distribution. As commonly used, the function f_g is chosen as a multivariate normal density ϕ_g:

$$\phi_g(\boldsymbol{y}_n | \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g) = \frac{\exp(-\frac{1}{2}(\boldsymbol{y}_n - \boldsymbol{\mu}_g)'\boldsymbol{\Sigma}_g^{-1}(\boldsymbol{y}_n - \boldsymbol{\mu}_g))}{(2\pi)^{P/2}|\boldsymbol{\Sigma}_g|^{1/2}}, \tag{6}$$

where P is the size of the vector $\boldsymbol{y}_n$, and the parameters $\boldsymbol{\mu}_g$ and $\boldsymbol{\Sigma}_g$ correspond respectively to the mean vector and the covariance matrix of the Gaussian density ϕ_g. The model potentially owns a large number of parameters, depending on the dimension and the number of groups, which may lead to computational issues in the estimation process. A common way use to address this problem is to control the geometric properties of the mixture components, known as Volume-Shape-Orientation decomposition. Details on the principle are found in Bouveyron et al. (2019). In order to give as much freedom as possible to the model computation, neither volume, shape nor orientation were constrained to be equal across the clusters. The choice of this model, identified by the letters VVV (for varying geometry), is in agreement with the Bayesian Information Criterion (BIC, Schwarz, 1978) and the Integrated

Completed Likelihood (ICL, Biernacki et al., 2000), showing the VVV model as the best combination.

The model-based clustering is applied in the space of the principal components obtained from the model-derived FPCA developed above (step 3 Figure 2). The expectation maximization algorithm (Fraley and Raftery, 2002) estimates the parameters of the model (Equation 6). The most appropriate number of group G is found using the ICL criterion which is recommended when the focus of the mixture analysis is clustering instead of density estimation (Biernacki et al., 2000). Initial values of the algorithm are selected by running the hierarchical model-based clustering with the VVV model (`mclust` package, Bouveyron et al., 2019). However, since model-based clustering can be sensitive to the initialization, we executed the model 100 times to choose the optimal pattern and assess the clustering variability.

2.2.4 FODs visited by the elephant seals

In order to study FODs visited by the SES, in situ TS profiles are projected as new observations onto the factorial plan of the FPCA realized with model-derived data (step 4 Figure 2). In situ TS profiles are first decomposed into T and S coefficients according to the procedure explained in Section 2.2.2. The Δ matrix, of size $M \times 2K$ stores the new coefficients centered with the vector $\overline{\boldsymbol{\alpha}}$ computed from the model-derived coefficients (see Section 2.2.2). Estimated scores $\widehat{\boldsymbol{c}}_l$ of the l^{th} principal component for the M new observations are computed as follows:

$$\widehat{\boldsymbol{c}}_l = \Delta \mathbf{M}^{-1/2} \mathbf{W}^{-1/2} \boldsymbol{b}_l, \tag{7}$$

with $\boldsymbol{b}_l$ as the lth eigenvector. The pair $(\widehat{c}_{m,1}, \widehat{c}_{m,2})$ gives the coordinates of the m^{th} in situ TS profile on the 2D map obtained from the model-derived FPCA. This step relates the environmental conditions visited by the seals to the actual range of possibilities given by the model-derived analysis.

3 Results

3.1 Vertical modes in the Brazil-Malvinas Confluence

The functional PCA is performed on 1 524 750 model-derived TS profiles covering the defined area (Figure 1) over 95 days (21 October 2018 to 23 January 2019 inclusive). The time period corresponds to the foraging trips of the studied SES. Only pixels where the bathymetry is deeper than 450 m are kept for the analysis, excluding the continental shelf area.

The first three vertical modes of variability resulting from the FPCA represent 98.77 % of the variance (Table 1). This value is of the same order as those obtained on the Kerguelen Plateau (Pauthenet et al., 2018) and in the Southern Ocean (Pauthenet et al., 2017). The higher modes capture less than 0.5 % of the variability and are not considered in the following analysis. Spatial distributions of the first and second principal components are displayed in Figure 3a and 3b for 21 October 2018. The maps show that while the FPCA is performed independently of location and day information, spatial patterns arise, accounting for temperature and salinity joint features. Ninety-five maps corresponding to each day of the study can be constructed similarly. Figure 3c and 3d represent the deformation of the mean profile associated with the two first modes.

The first mode, associated with the highest eigenvalue, alone summarizes as much as 94.5 % of the variance with an equivalent contribution of temperature and salinity variability. It involves a modification of the whole water column, going from cold and fresh waters unstratified under 100 m with a noticeable thermocline in the subsurface (positive PC1 in pink, Figure 3), to hot and salty waters (negative PC1 in orange, Figure 3). This mode marks a contrast between Subtropical waters transported by the Brazil Current and Subantarctic waters advected by the Malvinas Current, highlighting sharp thermohaline vertical gradients at the confluence (Figure 3a). The second mode (PC2) describes 3.55 % of the variance. High positive PC2 values (blue, Figure 3) present colder and fresher waters between 20 m and 200 m than high negative values (green, Figure 3). The trend reverses under 200 m. The spatial distribution of PC2 reveals mesoscale structures and a high dynamism produced by the meeting of the two currents (Figure 3b). This mode also opposes waters in the heart and at the

edges of the BC (dark blue vs green, Figure 3b). From 20 m to 200 m, temperature and salinity are greater at the edges than in the center part of the current. The third mode (see Supplementary Material) captures the seasonal thermocline development in response to surface heating. Figure 4 depicts the daily mean (continuous lines) and associated standard deviation (dashed lines) trends of the three first modes. A high daily standard deviation time series distinguishes the first mode (see y-axis) marked by a decrease between December and January. This decline can be related to the increase of the daily mean of the first component (pink line, Figure 4), indicating an average drop in temperature and salinity (effect of the first eigenfunction on the mean profiles). The daily mean of the second mode shows a decrease in stages over time, with a small daily standard deviation sharing a similar trend with the daily standard deviation of PC1. Even though the third mode represents less than 1 % of the variance (Table 1) its temporal evolution clearly shows a seasonal (spring to summer) increase of the subsurface temperature in the whole area.

In addition, a functional PCA is applied to the in situ TS profiles exclusively. The resulting shape decomposition is similar to the modes obtained with profiles from the model (see Supplementary Material) with comparable variance associated to the first modes (Table 1). This in situ FPCA was conducted independently of the model-derived FPCA and serves solely to compare the obtained modes of variability.

3.2 Functional Oceanographic Domains (FOD)

In order to cluster profiles sharing similar vertical structures in FODs, a model-based clustering is performed using PC1 and PC2 as input variables, outcome of the FPCA conducted on the model-derived profiles. PC3 is not included to avoid implying a seasonal water warming in the regime discrimination. An optimal number of four clusters is given by the ICL criterion. After running the model-based clustering 100 times, two distinct patterns emerged. We selected the main pattern, occurring in 68 % of cases and presenting low variability in the cluster composition. The FODs are then defined as groups containing profiles with a high probability of belonging to a given cluster. The threshold is arbitrarily fixed at 0.75. This step leads to the distinction of a fifth group composed of profiles with a probability lower than 0.75 and referred to as indeterminate profiles. Regimes 1 (purple), 2 (blue), 3 (green), 4 (yellow) and 5 (indeterminate) contain respectively 5.7 %, 18.4 %, 40.1 %, 24.2 %

and 11.6 % of the profiles considered over the 95 days. The proportions of the five regimes vary over time as shown in Figure 5, with a notable gradual disappearance of the first FOD (purple). The first FOD corresponds to cold and fresh waters with little vertical variability between profiles as seen by the shade around the mean profile (purple profiles, Figure 6a). The second FOD shows similar features to the first group, with waters warmer above 80 m, slightly colder deeper, and saltier above 100 m (blue profiles). The FOD numbered 4 corresponds to hot and salty waters with a high variability (yellow profiles), matching features of the BC. Finally, the third FOD corresponds to intermediate features with high variability decreasing with depth (green profiles, Figure 6a). This regime corresponds to waters that are formed by the mixture of Subantarctic and Subtropical waters. The spatial distribution of the clustering for the first day of the considered period is displayed in Figure 6b. Regimes obtained show very consistent spatial and temporal patterns. Hot and salty waters (yellow) form a coherent mass in the north part of the map, while the blue and purple FODs follow the Patagonian slope. The green regime occupies a vast area in the southeast part but with an indefinite shape. For both the yellow and green FODs, the daily probability density function of PC1 shifts toward positive values (see Supplementary Material), meaning that inside those regimes, the whole water column cools down and gets fresher over time (corroborating the observed drop in the daily trends, Figure 4). Gray filaments appear in Figure 6b and contain indeterminate profiles. They mainly correspond to spatial transitions between the different regimes and are considered as frontal zones.

3.3 Habitats used and foraging strategies

In situ profiles reaching 450 m depth account for 45.07 % of the total SES dives. Out of this dataset, 4 463 are attributed to the day-time period, while 767 are attributed to the night-time period. The remaining profiles occur either during the twilight period or are not geolocalized.

Figure 7 shows the projection of the in situ profiles superimposed onto the 2D mapping of the FPCA obtained from model-derived profiles. The marginal distributions reveal high-density modes and the high concentration of model-derived observations in the top right corner. The distribution of in situ TS profiles is pointed out by the 2D density estimation (red to yellow lines). Each in situ profile is assigned to one of

the five regimes (considering indeterminate profiles) based on their coordinates on the first factorial map (Figure 7). Figure 8 shows the five SES trajectories colored according to the FOD and the proportion of regimes visited by the five females. The spatial overlap of dives belonging to different FODs is explained by the displacement of regimes in space over time. Indeterminate profiles occupy a notable part of the trajectories (between 14 % and 25 %). All females except one visited the five regimes but at different percentages. The yellow FOD is not always seen on the trajectories meaning that this regime was visited by the females after the GPS stopped (i.e. the dives are not geolocalized). Since the analyses do not rely on the location and time of the profiles, it is indeed possible to assign a oceanographic regime to a non-localized SES dive. The female named 2018-46 quickly crosses the purple and blue FODs and spends a large part of her journey in the third regime identified (green). The other females linger more in the cold waters associated with regimes one and two (purple and blue), but are also found in the yellow regime, crossing highly contrasting waters along their trips (Figure 6a). The passage between two FODs with opposite features is sometimes performed in less than 24 hours (personal observations).

With the aim of understanding the SES habitat-visited proportions observed in Figure 8, the foraging strategy is studied within the scope of the FODs. The female identified as 2018-47 can not be included in the foraging analysis as data from the accelerometer is unavailable. Surprisingly, the foraging performances assessed by the prey-capture rate (PCR) of the seals do not show much variation according to the oceanographic regime (see Supplementary Material). At night, only the yellow regime shows a slight decrease in PCR. During the day, the PCR is lower in the green regime and slightly higher in the blue. The indeterminate group, corresponding to fronts between FODs, do not imply significant changes in the PCR in comparison with the other regimes. While the SES feed continuously along their pathways, the females studied seem to favor the area where the different FODs meet and mix, at around 41° S 56° W (Figure 8). Dives are gathered in this zone in particular as the speed decreases and the curvature increases (personal observations). To validate this assumption, a criterion is created to reveal areas of high spatial and temporal oceanographic variability (Figure 9). This criterion is determined by calculating, for a given pixel, the proportion of days (over the considered time-period) where the oceanographic domain changes from a given day to the previous one. The spatial distribution of this criterion

is presented in Figure 9. The map highlights one particular area with high probabilities of having oceanographic conditions changing at day+1. This area corresponds to the meeting of the FODs. Yellow filaments are seen along the Patagonian slope and correspond to the displacement of the blue FOD over time, gradually substituting the purple regime. Other areas of large probabilities are also observed in the northeastern and southern parts of the map, and can be associated with frontal regions induced by mesoscale features like the displacement of eddies carrying Subtropical waters in the area. The probability value is extracted for each SES dive location and related to the associated PCR. Figure 10 depicts the relationship between the PCR and the probability of changing regimes. The non-parametric regressions on the quantiles (Oh et al., 2011; Simonoff, 2012) reveal a positive relation between the two variables even if the data variability is high. The capture rate increases sharply to 0.15 with a slight acceleration for the high percentiles. A small decrease arises around 0.2, then the PCR gently increases with high probabilities. The median capture rate (red line) is multiplied by 3.78, going from 3.5 capture per hour to 13.24. This result supports the assumption that SES favor areas at the interface between water masses, where they increase their foraging success.

4 Discussion

In this study, we aim to understand the habitat use of five elephant seals equiped in 2018 in Argentina, and the influence of oceanographic conditions on their foraging strategies using descriptive statistics.

A functional principal component analysis followed by model-based clustering is used to describe the Malvinas-Brazil Confluence at a regional scale. The multivariate FPCA is an objective method developed in the oceanographic domain by Pauthenet et al. (2017) to identify the thermohaline modes of the Southern Ocean. This approach presents strong potential as it not only allows the description of oceanographic patterns at a global scale (Pauthenet et al., 2019) but also the definition of water-mass boundaries and variability (Pauthenet et al., 2018). The analysis extracts the principal modes of variability that describe and summarize the 3D thermohaline environment through joint variations in temperature and salinity. Modes of variability are orthogonal, considered as independent from each other. This multivariate technique

is extremely efficient in data reduction. The dataset used in this study is composed of more than 1.5 million combinations of temperature and salinity profiles, with each profile decomposed into 44 discrete values from 20 m to 450 m. We first projected the model-derived profiles onto a B-spline basis with a small number K of functions. Then, we worked on the two first principal components of the FPCA, which summarize 98 % of the curve's shape variability. The B-spline basis used can not capture the fine-scale variations of in situ profiles. Yet, since these profiles are projected into the first factorial map of the clean and smooth model-derived profiles (Section 2.2.4), the variations missed by the curve reconstruction do not impact the results in the oceanographic regime attribution. Furthermore, smoothing the profiles with a small number of basis functions is useful when data is potentially corrupted by noise, such as in the case of in situ profiles (Nerini et al., 2010). The functional approach is interesting because it integrates the vertical continuity of the data (considering the inherent link between two consecutive values), working on curve shape decomposition. Moreover, curve reconstruction allows reducing the dimension of the data while preserving the complexity of the curve shape. However, one constraint of the analysis is the requirement of equal depth range for each profile. Choice of the 450 m threshold was motivated by the desire to take into account a large part of the water column used by the SES. This limitation implies the removal of all profiles that do not reach that depth, resulting in the elimination of the continental shelf area. As females preferentially use the deep ocean basin (Campagna et al., 2021), this loss of area has limited consequences. The 450 m threshold also implies the loss of many night dives as these mostly do not reach that depth. Results on the foraging behavior observed in this study are thus mainly driven by daylight dives and deep night dives, necessitating cautious interpretation. One way to resolve this issue could be to reconstruct the missing part of shallow profiles by functional methods. This solution would allow consideration of a larger portion of night dives and could highlight potential nighttime foraging variabilities between FODs that were missed out in this study.

Results of the FPCA applied to model-derived TS profiles demonstrate the method's high capacity to decompose the BMC's main modes of variability. The three principal modes of vertical variability express the three main drivers of the oceanographic conditions in this area at that period, namely: (1) a positional mode (PC1) depicting the main vertical variability driven by the opposite features of the Brazil (north)

and Malvinas (south) Currents; (2) a dynamical mode (PC2) highlighting mesoscale structures inherent to the confluence of the two currents; and (3) a seasonal mode (PC3) revealing the spring to summer surface heating independent of the water's origin. Moreover, as the method is applied over 95 days, it allows us to follow each mode of vertical variability in space and time, giving key information about physical processes occurring in the area (Figure 4). The trends of the daily mean and associated standard deviation of PC1 and PC2 can be associated with seasonal oscillations and variability in Subantarctic and Subtropical influxes in the region but may also indicate sporadic events that change the average thermohaline features in the region.

The model-based clustering applied on the first two principal components of the FPCA led to the distinction of four FODs, whose thermohaline features share a common vertical structure. A FOD can be considered as an assemblage of water masses, integrating the vertical structure of the water column. The T-S diagram (see Supplementary Material) indicates the water-mass composition of each FOD, following definitions given by Gordon (1981), Maamaatuaiahutapu et al. (1994) and Piola and Gordon (1989). The spatial coherence of the FODs obtained and their consistency over time confirm the robustness of the analysis. The vertical features of the fourth group (yellow) correspond to Subtropical waters transported by the BC (Piola et al., 2001; Valla et al., 2018). In the upper layer of the yellow FOD, we find the warmest and saltiest Tropical Waters (TW) above the South Atlantic Central Waters (SACW, $\sigma_\theta < 27$ kg m^{-3}) corresponding with the nearly straight line in the diagram. The BC has been described as a high variability current (Valla et al., 2018), corroborated in this study by a large interquartile vertical domain (Figure 6a). This regime varies over time. The entire water column gets colder and fresher during the considered period, suggesting a greater influx of Subantarctic waters within the study area, while surface heating occurs (shown by the PC3). The third FOD (green) corresponds to mixed waters resulting from the confluence of the BC and the MC, as seen by the vertical features on the T-S diagram (see Supplementary Material). These mixed waters occupy almost half of the considered area. The first and second regimes (purple and blue) correspond to the Malvinas Current signatures (Piola et al., 2001). The two groups have very similar vertical characteristics. They gather cold and fresh profiles with low intra-cluster variability. The upper 100 m are composed of Subantarctic Surface Waters (temperature > 5 °C, salinity ~ 34 psu) while two water

masses are found deeper: the Subantarctic Mode Waters (temperature ~ 5 °C, salinity ~ 34.2 psu) and Antarctic Intermediate Waters (temperature < 4 °C, salinity < 34.2 psu) (see Supplementary Material). The MC is the only current to carry Subantarctic waters to Subtropical latitudes (Paniagua et al., 2021). It is generally accepted that the MC starts at 55° S, follows the SAF and meets with the BC at approximately 38° S (Artana et al., 2016), which corresponds to our observations (Figure 6b). The first FOD (purple) corresponds to pure Subantarctic waters transported by the MC while the second FOD (blue) may correspond to Subantarctic waters that have undergone some mixing with Subtropical waters and are carried southward, up to 49° S, by the Malvinas Return Current (MRC) (Saraceno et al., 2004). As the second regime substitutes the first FOD over time (Figure 5), it indicates that the waters along the Patagonian slope become colder under 100 meters depth, warmer above that threshold and saltier above 150 meters depth. This observation is consistent with the findings of Aubone et al. (2021) who studied the behavioral response of an elephant seal staying over a small portion of the Patagonian slope to vertical thermohaline variability of the MC. While the MC and the MRC are not distinguishable by an analysis performed on TS surface data (not shown), the functional analysis applied in this study succeeded in differentiating subtle patterns. In this way, the FODs obtained in this study are slightly different from those obtained by Tournier et al. (2021) who also performed a FDA to study the influence of oceanographic features on mid-trophic levels through active acoustics, and found three oceanographic regimes using SES in situ data. Features of those regimes are similar to the ones found in the present study but they did not differentiate the MRC and the MC, probably due to the clustering method and the use of selected in situ data to describe the thermohaline environment instead of model-derived data as in our case.

Different approaches have been explored to determine biogeographical regions in the BMC area but mostly using ocean surface variables from satellites (Gonzalez-Silvera et al., 2004; Saraceno et al., 2006; Saraceno et al., 2005). An extension of our work could comprise in integrating biogeochemical data like chlorophyll-a surface concentration with TS functional data in clustering, and comparing the results with previous findings. In other areas, different studies have also experimented with applying the model-based clustering framework to temperature and/or salinity profiles to identify spatial oceanographic patterns (Maze, Mercier, and Cabanes, 2017). However,

most of these studies did not combine a functional approach with a Gaussian mixture model (GMM), although this approach was suggested by Jones et al. (2019). The GMM was applied in the North Atlantic and the SO using temperature profiles from Argo floats (Jones et al., 2019; Maze, Mercier, Fablet, et al., 2017). More recently, other studies have used profile classification model (PCM) in the multivariate case to define regimes by their thermohaline structure (Boehme and Rosso, 2021; Rosso et al., 2020). Although they all reduce the dimensionality of the data using a classical principal component analysis, they do not preserve the original nature of the data (i.e. the functional properties). The eigenvalue decomposition of a matrix $\mathbf{T}$ in a classical PCA remains invariant under the permutation of columns or rows. However, a permutation of the values in a profile (i.e. the columns of matrix $\mathbf{T}$) changes the shape of the profile. Only a few papers leverage the functional properties of TS profiles to define oceanographic regions. For instance, Assunção et al. (2020) applied two distinct FPCAs to temperature and salinity profiles and used a functional hierarchical clustering to characterise the thermohaline structure of the Southwestern tropical Atlantic. Here, our results demonstrate the potential and effectiveness of combining multivariate FPCA and model-based clustering to extract key vertical thermohaline features and identify oceanographic domains. This functional data clustering approach falls under the category of *filtering methods* described in Jacques and Preda (2014a). This category encompasses clustering on either basis coefficients (James and Sugar, 2003) or on FPCA scores (Jacques and Preda, 2013, 2014b). Over the last decades, a vast amount of literature has been produced on methodological aspects and applications of the combination of FDA and clustering (Chamroukhi and Nguyen, 2019; Jacques and Preda, 2014a; Wang et al., 2016; Zhang and Parnell, 2023). Some studies have expanded hierarchical clustering and k-means partitioning methods to functional data (Abraham et al., 2003; Tokushige et al., 2007), while others have focused on functional clustering using mixture models (Korte-Stapff et al., 2022; Li et al., 2016; Schmutz et al., 2020). In the field of oceanography and limnology, while the application of functional data clustering may appear limited in the literature, functional data analysis has received more attention. Various studies are moving toward functional models (FM) to explain or predict biological variables with hydrological parameters (Ainsworth et al., 2011; Boudreault et al., 2021; Yen et al., 2015). It is possible within the frame of FDA to use functions as responses and/or

predictor variables in regression models (Ramsay and Silverman, 2005). Considering data as curves prevents information loss, improving models (Boudreault et al., 2021). Bayle et al. (2015) demonstrated the ability of FM to predict vertical chlorophyll-*a* from light profiles in the Antarctic region. Godard et al. (2020) have also shown that functional analysis helps to deepen the understanding of animal behavior as their study of SES dive shape allowed them to determine different dive patterns associated with foraging success. The FDA framework was recently extended to the acoustic domain (Ariza et al., 2022), showing promising results for improving our understanding of the links between oceanographic conditions and the structuring of sound-scattering layers.

Model-based clustering also offers the advantage of allowing for uncertainty regarding clustering (Bouveyron et al., 2019). In this present study, profiles with a probability of belonging to a cluster under 0.75 form frontal filaments (gray color, Figure 6b), which are in good agreement with intense fronts obtained with other criteria like thermal gradients or with the Finite Size Lyapunov Exponents criterion (d'Ovidio et al., 2010). The convergence of the MC and the BC with very different thermohaline characteristics leads to energetic mixing processes and intrusions of Subantarctic and Subtropical waters (Piola et al., 2001). These mesoscale structures are seen with the second mode of variability obtained by the FPCA (Figure 3b). The Atlantic portion of the Patagonian Shelf is marked by a large continental shelf extending from 34° S to 55° S and between the coastline and the 200 m isobath, followed by a slope plunging to 6 000 m depth (Weatherall et al., 2015). Several studies suggest that the interaction of the complex ocean topography with the MC flow generates upwellings along the shelf-break (Matano and Palma, 2008; Miller et al., 2011) and is also a source of mesoscale processes (Fu, 2006; Mason et al., 2017; Saraceno and Provost, 2012). The interaction of the shelfbreak, the nutrient-rich Subantarctic waters transported by the MC, and the complex physical processes induced by the BMC, contribute to the high phytoplankton biomass observed (Acha et al., 2004; Garcia et al., 2008). Indeed, this region is known for its high chlorophyll-*a* concentration during the austral spring and summer (Lutz et al., 2010; Romero et al., 2006; Saraceno et al., 2005). These conditions support the marine ecosystems in the region and sustain one of the most important fisheries in the southern hemisphere (Bogazzi et al., 2005; Martinetto et al., 2020; Rey and Huettmann, 2020). The confluence associated with shelf blooms

makes the Argentine Basin a particularly interesting area for large pinnipeds like the SES, whose females mainly feed there (Campagna et al., 2021).

The combination of FPCA with model-based clustering allows us to contextualize the environment in the foraging area of the elephant seals at the regional scale. The oceanographic domains identified are considered as distinct physical habitats and used to study conditions visited by the SES and their foraging strategies. By projecting in situ profiles on the factorial plan obtained from the FPCA performed with model-derived data, we are able to match the conditions visited by the females with the range of possibilities revealed by the analysis. We consider this approach to be more robust than directly performing clustering on in situ data. In fact, since model-based clustering can be sensitive to initialization, clustering the TS profiles from SES without taking precautions could result in highly variable groupings. Our results demonstrate that clustering over a large dataset ($>$ 1.5 million observations) yields to a stable model and provides more objective interpretations of the oceanographic regimes visited by the seals. The results show that SES foraging within the BMC tend to visit all FODs present in the study area but in unequal proportions (Figure 8). The results also show intervariability between the seals, which may indicate individual preferences. The percentages of the visited habitats observed (Figure 8) do not correspond to the spatial space occupied by the regimes (Figure 5). For example, the third group (green), which represents more than 40 % of the area, is proportionally less visited by the elephant seals. One major advantage of studying TS profiles through FDA is that it gives information about oceanographic conditions encountered with no need for spatial coordinates. The consideration of non-localized dives is necessary, first of all, to obtain an objective overview of habitat use, but is also useful for accessing information that could otherwise be missed. The results of the case presented here indicate that study of SES data allows us to find the four FODs observed at the regional scale. In this way, although our vision depends on their trajectory choice, the five seals offer a representative picture of the area. The decomposition of TS profiles through FPCA was also performed independently on in situ profiles from the SES. The results were similar whether using in situ or Mercator profiles (Table 1), confirming the representativeness of SES data. Furthermore, the similarity of results indicates that while modeled profiles did not capture the thermohaline complexity of the water column (e.g. the profiles are relatively smooth), they

were efficient in finding the main oceanographic patterns in the region. The chosen CMEMS product is a model derived from the Nucleus for European Modelling of the Ocean (NEMO), which assimilates in situ temperature and salinity profiles. Therefore, it is expected to be consistent with the recorded SES data. However, on any given day, the TS model-derived profiles can be very different from those recorded by the animal-borne device at the same geolocation. This phenomenon is particularly noticeable in the area where all the regimes meet and mix. This comes as no surprise as the model is not a perfect representation of reality and is likely to be less efficient for locating water-mass boundaries precisely—information more accurately extracted from the geolocalized in situ SES TS profiles.

The capture rate is used as a proxy of foraging success to investigate if the identified FODs have an influence on SES habitat choices. We find that foraging performances of the five SES studied are mostly unrelated to the visited regimes. The oligotrophic waters advected by the BC leads to a slight decrease in PCR during the night but not during the day. This can be explained by prey availability being affected by the warm waters of the epipelagic layer. The daily PCR is slightly higher in the blue FOD, which may indicate a nutrient input associated with the Malvinas Return Current and impacting higher trophic levels (Aubone et al., 2021). To better understand SES foraging strategies, we took advantage of the model-based clustering analysis to create a criterion of physical-condition variability by computing the probability of changing regimes. Interestingly, we found that foraging performance is associated with the probability of changing FODs (Figure 10), regardless of the specific FOD. The median prey-catch rate (PCR) is nearly quadrupled when there is a 50 % probability of changing FOD the following day. The background gray dots in Figure 10 illustrate a high variability in PCR, but all quantile regressions exhibit the same increasing trend. Our study suggests that instead of selecting a particular regime, SES may look for areas of high oceanographic variability found at the interface between FODs. This result is consistent with previous findings highlighting the importance of oceanographic fronts and eddies on movement behavior and foraging performances of SES (Bailleul et al., 2010; Cotté et al., 2015; Della Penna et al., 2015; Dragon et al., 2010; Rivière et al., 2019). However, the indeterminate group resulting from the model-based clustering and corresponding to frontal areas between FODs, do not imply a PCR increase although it represents a large part of the group proportion

(Figure 8 and Supplementary Material). This may indicate that only some particular frontal areas attract and are favored by SES. For more extensive studies, Lagrangian approaches may be complementary to comprehensively understand frontal-system impacts (Baudena et al., 2021; Bon et al., 2015). Finally, drawing general assumptions about the habitat use and foraging strategies of elephant seals based on only five individuals requires caution. To validate our findings regarding elephant seals' behavior in the Argentine Basin, it would be beneficial to incorporate a larger number of individuals in future studies.

5 Conclusion

Technological advances of the last decades have led to massive deployments of ocean-observing systems recording environmental variables at a very high frequency and accuracy. The huge datasets that result from such sensors need statistical tools to extract the main information and analyze the data integrating their complexity. In this paper, we explored functional data analysis combining model-derived data and in situ data collected from southern elephant seals (SES). Our results confirm the adequacy of using functional approaches for computing large and complex datasets. This framework followed by model-based clustering was appropriate for describing the foraging environment of five SES in 2018 and helpeing to understand their habitat use. New loggers, such as a head-mounted microsonar, should provide fine-scale insight into variations in the density of midtrophic level (MTL) organisms (i.e. macrozooplankton and micronekton) that could explain variability in elephant seal foraging behavior. On the basis of our current findings, we hypothesize that the greatest densities of MTL organisms are likely to be found at the interface between Functional Oceanographic Domains (FOD) rather than within specific FODs. Functional data analyses are promising methods for advancing study of the interaction between SES behavior, MTL organisms (including preys of SES) and environmental conditions.

Declaration of competing interest

None.

Acknowledgments

Through NF Ph.D, this work was funded by the French Ministry for Education and Research. We thank all people that have been involved in the fieldwork in Argentina (Península Valdés), part of the SABIO French-Argentinian collaborative program (PI Christophe Guinet and Martin Saraceno) supported by CNES-TOSCA (Centre National d'Etudes Spatiales). We also thank Mark Johnson for providing the tags, Fui Lee Luk for english correction, and Lloyd Izard for his advices and personal support.

Table and figures

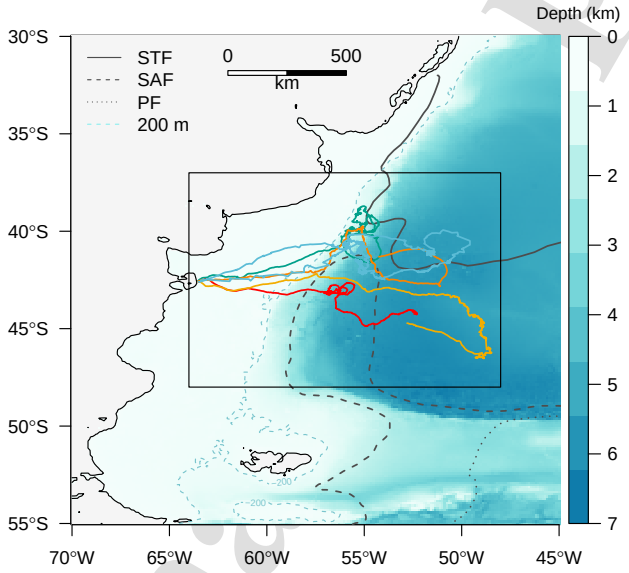


Figure 1: Colored trajectories represent the five SES females equipped in October 2018 at Península Valdés and used for the analysis. The black rectangle defines the focus area. Black lines delimit the Subtropical Front (STF), the Subantarctic Front (SAF) and the Polar Front (PF) from Orsi et al. (1995). The blue color gradient indicates the bathymetry, and the 200 m isobath (blue dashed line) defines the limit between the continental shelf and the slope.

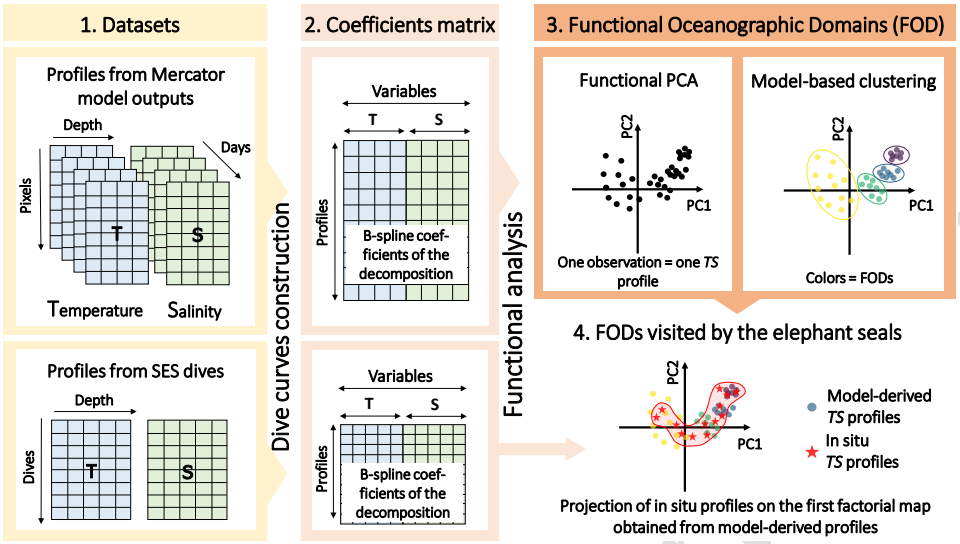


Figure 2: Workflow of the functional analysis. (1) Model-derived and in situ TS profiles are stored in different data frames. (2). TS profiles are approximated with continuous curves described by the coefficients of the B-spline decomposition. The coefficients are stored in new matrices. Each row corresponds to one observation. (3) The model-derived observations are decomposed into modes of variability with the FPCA and projected onto a space of reduce dimension. Model-based clustering is performed in the space of the first principal components to identify FODs (i.e. regimes sharing similar oceanographic structure). Steps (2) and (3) reduce the dimensionality of the initial data. (4) Finally, in situ observations from elephant seals are projected onto the factorial map as new statistical observations to associate each dive with one FOD and determine the habitat visited by the SES.

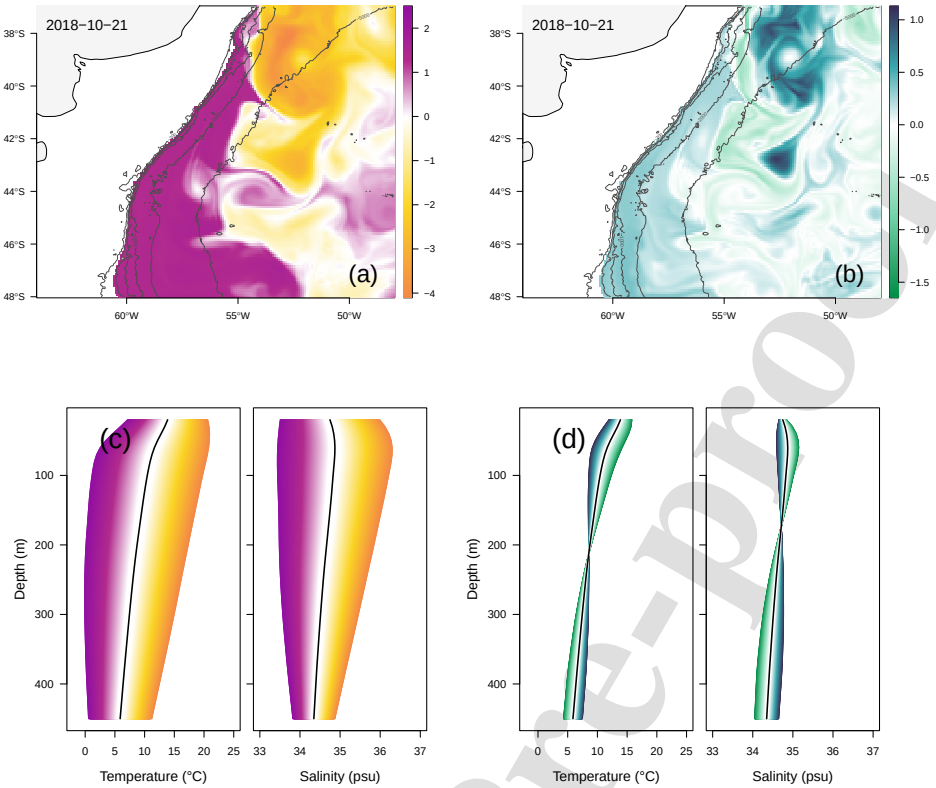


Figure 3: **(a,b)** Spatial distribution of the first and second PCs for 21 October 2018. Colors indicate the magnitude of the scores and match with color gradients in (c) and (d). The pink, orange, blue and green areas correspond with temperature and salinity profiles in pink, orange, blue and green (respectively) seen in (c). PC1 highlights the contrast between the BC (hot and salty) and the MC (cold and fresh), while PC2 reveals coherent mesoscale patterns. Bathymetry lines at 200, 500, 700, 1 000, 2 000 and 5 000 m define the Patagonian slope. **(c,d)** Vertical shapes induced by the first two eigenfunctions (respectively). The color gradients represent the effect of adding or subtracting (respectively pink or orange in (c), blue or green in (d)) each mode with a magnitude of two (see Pauthenet et al., 2017, for more details). The black line indicates the mean profile, and represents the center of gravity of the factorial map seen in Figure 7.

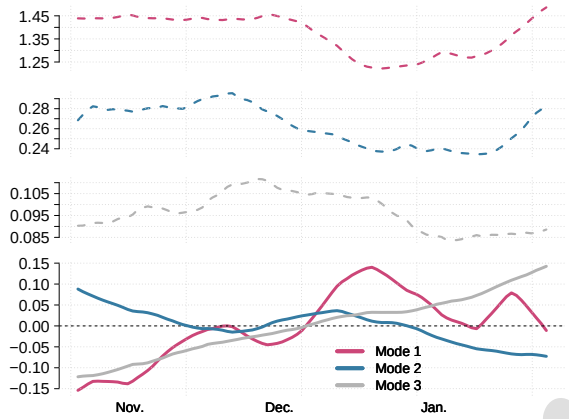


Figure 4: Trends of the daily mean (continuous lines) and associated daily standard deviation (dashed lines) for PC1 (pink), PC2 (blue) and PC3 (gray).

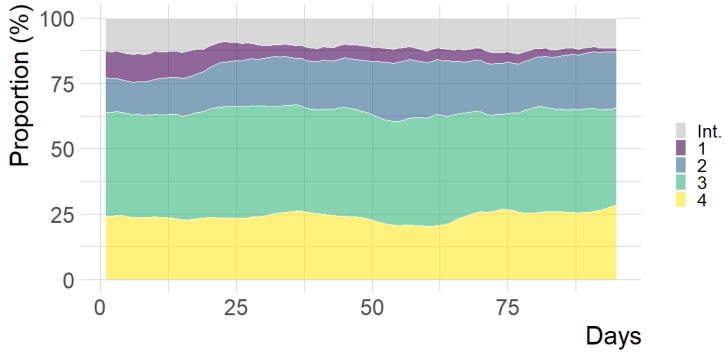


Figure 5: Proportion (%) occupied by the FODs in the considered area over the 95 days (21 October 2018 to 23 January 2019). The colors refer to the 4 FODs with the gray part corresponding to indeterminate profiles (Int.).

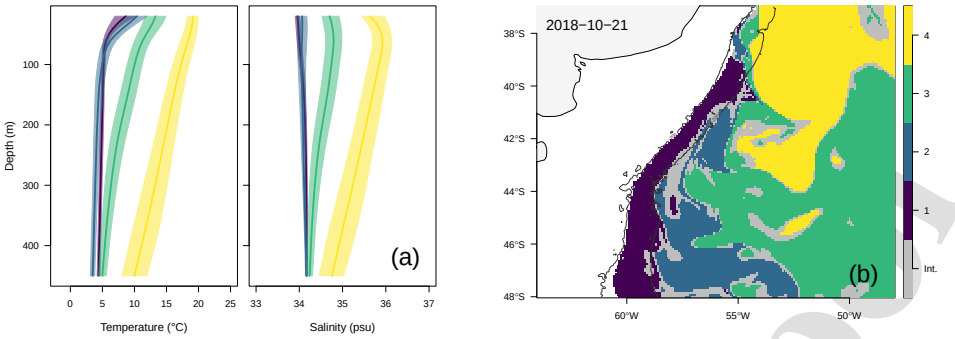


Figure 6: **(a)** Mean temperature and salinity profiles (solid lines) associated with the FODs obtained from the clustering (after removing profiles with a probability $<$ 0.75). Regimes are colored in purple (cluster 1), blue (cluster 2), green (cluster 3) and yellow (cluster 4). The envelopes contain 50 % of the profiles, delimiting the first and third quartiles calculated for each cluster. **(b)** Spatial distribution of the five regimes on 21 October 2018. The colors match with the profiles in (a). Profiles with a probability $<$ 0.75 of belonging to a group are colored in gray and form filaments between the regimes. Bathymetry lines at 200 and 2 000 m are drawn.

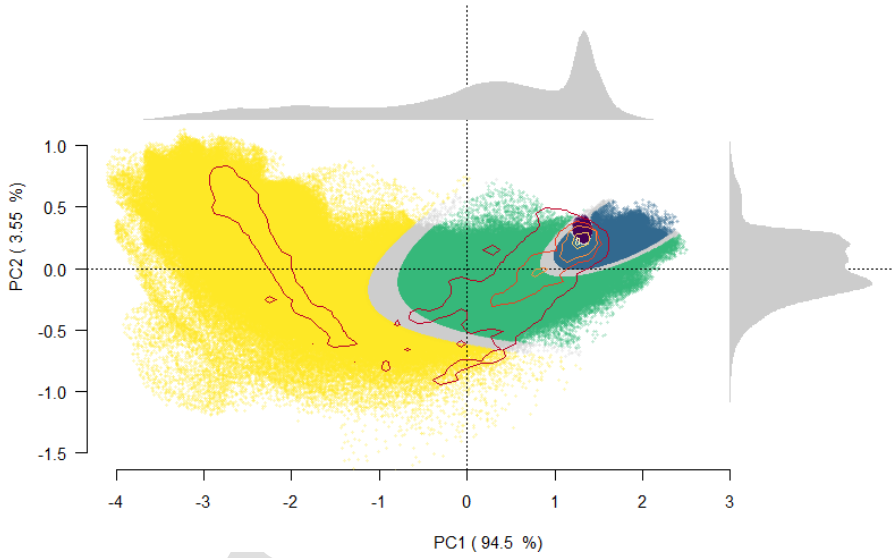


Figure 7: First factorial map of the functional PCA performed on model-derived data (colored dots) and clustering results. Gray dots correspond to observations with a probability of belonging to a cluster under 0.75. Marginal density probability of the first and second principal components are drawn (gray densities on the upper and right sides). They show modes of high density indicating areas of high concentrated model-derived observations. The 2D kernel density estimation (red to yellow lines) displays the projection of SES in situ TS profiles. The increasingly lighter successive lines contain respectively 2.5, 25, 50, 75 and 97.5 % of the dives centered around local density maximums. Each SES dive is associated to a FOD or to the indeterminate profiles group (gray).

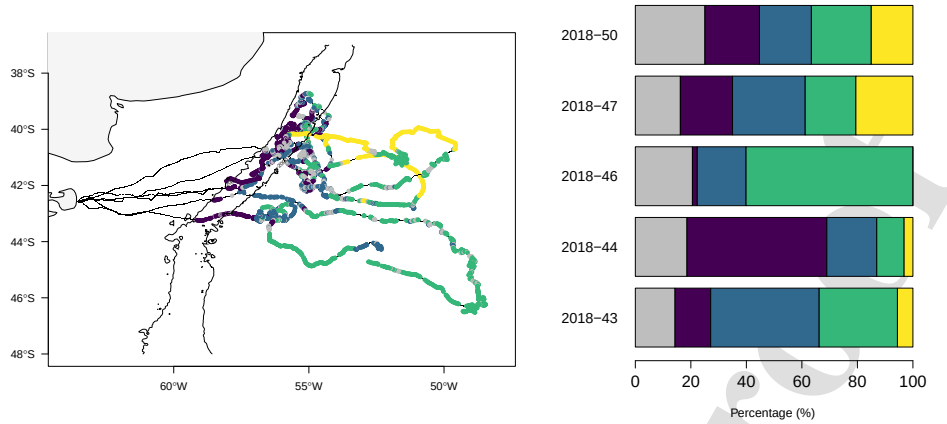


Figure 8: Trajectories of the five females colored according to the FODs identified and proportion of dives belonging to each regime including indeterminate profiles (gray group). Numbers 2018-43 to 50 correspond to SES identification. Refer to Figure 6 for FOD features and names. Bathymetry lines at 200 and 2 000 m are drawn.

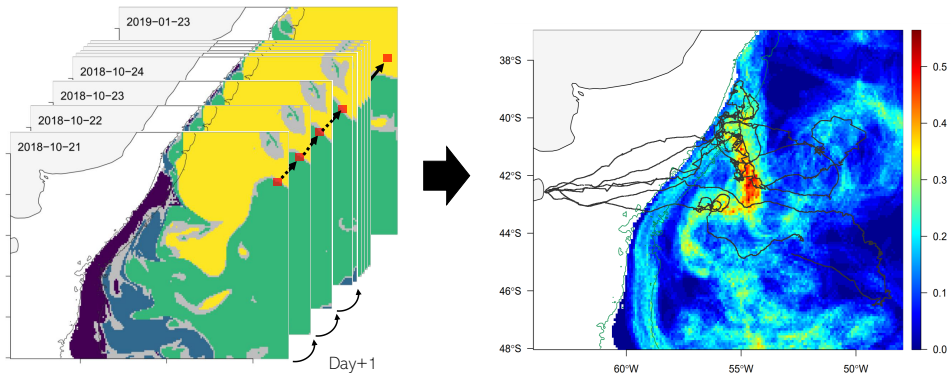


Figure 9: Probability of changing regimes at day+1. For a given pixel (red square), the proportion of days where the oceanographic regime change from a day to the next one is calculated (see the schematic view on the left). High probability values are colored in red and highlight a high variability area in the center of the map (right panel).

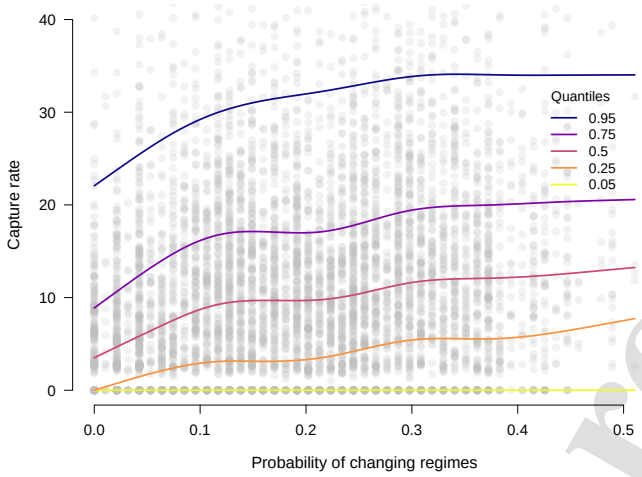


Figure 10: Scatter plot of prey-capture rate (captures per hour) against probability of changing regimes extracted for each SES dive. The lines indicate the non-parametric quantile regression at the quantiles of order 0.05, 0.25, 0.5, 0.75 and 0.95 respectively. For example, the red line corresponds to the median prey-capture rate and shows an increase from 3.5 to 13.24 captures per hour, which corresponds to a multiplication factor of 3.78, as the probability increases from 0 to 0.5.

Table 1: Variance (%) explained by the first principal components obtained from the functional PCA performed on model-derived and in situ profiles.

	Principal components		
	PC1	PC2	PC2
ACPF with model-derived TS profiles	94.50 %	3.55 %	0.72 %
ACPF with *in situ* TS profiles	88.14 %	5.93 %	1.98 %

References

Abraham, C., Cornillon, P.-A., Matzner-Løber, E., & Molinari, N. (2003). Unsupervised curve clustering using b-splines. *Scandinavian journal of statistics*, *30*(3), 581–595.

Acha, E. M., Mianzan, H. W., Guerrero, R. A., Favero, M., & Bava, J. (2004). Marine fronts at the continental shelves of austral south america: Physical and ecological processes. *Journal of Marine systems*, *44*(1-2), 83–105.

Ainsworth, L., Routledge, R., & Cao, J. (2011). Functional data analysis in ecosystem research: The decline of oweekeno lake sockeye salmon and wannock river flow. *Journal of agricultural, biological, and environmental statistics*, *16*(2), 282–300.

Ariza, A., Lengaigne, M., Menkes, C., Lebourges-Dhaussy, A., Receveur, A., Gorgues, T., Habasque, J., Gutiérrez, M., Maury, O., & Bertrand, A. (2022). Global decline of pelagic fauna in a warmer ocean. *Nature Climate Change*, *12*(10), 928–934.

Artana, C., Ferrari, R., Koenig, Z., Saraceno, M., Piola, A. R., & Provost, C. (2016). Malvinas current variability from argo floats and satellite altimetry. *Journal of Geophysical Research: Oceans*, *121*(7), 4854–4872.

Assunção, R. V., Silva, A. C., Roy, A., Bourlès, B., Silva, C. H. S., Ternon, J.-F., Araujo, M., & Bertrand, A. (2020). 3d characterisation of the thermohaline structure in the southwestern tropical atlantic derived from functional data analysis of in situ profiles. *Progress in Oceanography*, *187*, 102399.

Aubone, N., Saraceno, M., Alberto, M. T., Campagna, J., Le Ster, L., Picard, B., Hindell, M., Campagna, C., & Guinet, C. (2021). Physical changes recorded by a deep diving seal on the patagonian slope drive large ecological changes. *Journal of Marine Systems*, *223*, 103612.

Bailleul, F., Vacquie-Garcia, J., & Guinet, C. (2015). Dissolved oxygen sensor in animal-borne instruments: An innovation for monitoring the health of oceans and investigating the functioning of marine ecosystems. *PLoS One*, *10*(7), e0132681.

Bailleul, F., Cotté, C., & Guinet, C. (2010). Mesoscale eddies as foraging area of a deep-diving predator, the southern elephant seal. *Marine Ecology Progress Series*, *408*, 251–264.

Baudena, A., Ser-Giacomi, E., D'Onofrio, D., Capet, X., Cotté, C., Cherel, Y., & D'Ovidio, F. (2021). Fine-scale structures as spots of increased fish concentration in the open ocean. *Scientific Reports*, *11*(1), 1–13.

Bayle, S., Monestiez, P., Guinet, C., & Nerini, D. (2015). Moving toward finer scales in oceanography: Predictive linear functional model of chlorophyll a profile from light data. *Progress in Oceanography*, *134*, 221–231.

Biernacki, C., Celeux, G., & Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE transactions on pattern analysis and machine intelligence*, *22*(7), 719–725.

Block, B. A., Holbrook, C. M., Simmons, S. E., Holland, K. N., Ault, J. S., Costa, D. P., Mate, B. R., Seitz, A. C., Arendt, M. D., Payne, J. C., et al. (2016). Toward a national animal telemetry network for aquatic observations in the united states. *Animal Biotelemetry*, *4*(1), 1–8.

Block, B. A., Jonsen, I. D., Jorgensen, S. J., Winship, A. J., Shaffer, S. A., Bograd, S. J., Hazen, E. L., Foley, D. G., Breed, G., Harrison, A.-L., et al. (2011). Tracking apex marine predator movements in a dynamic ocean. *Nature*, *475*(7354), 86–90.

Boehme, L., & Rosso, I. (2021). Classifying oceanographic structures in the amundsen sea, antarctica. *Geophysical Research Letters*, *48*(5), e2020GL089412.

Boehme, L., Thompson, D., Fedak, M., Bowen, D., Hammill, M. O., & Stenson, G. B. (2012). How many seals were there? the global shelf loss during the last glacial maximum and its effect on the size and distribution of grey seal populations. *Plos one*, *7*(12), e53000.

Bogazzi, E., Baldoni, A., Rivas, A., Martos, P., Reta, R., Orensanz, J. M., Lasta, M., Dell'Arciprete, P., & Werner, F. (2005). Spatial correspondence between areas of concentration of patagonian scallop (zygochlamys patagonica) and frontal systems in the southwestern atlantic. *Fisheries Oceanography*, *14*(5), 359–376.

Bon, C., Della Penna, A., d'Ovidio, F., YP Arnould, J., Poupart, T., & Bost, C.-A. (2015). Influence of oceanographic structures on foraging strategies: Macaroni penguins at crozet islands. *Movement ecology*, *3*(1), 1–11.

Boudreault, J., St-Hilaire, A., Chebana, F., & Bergeron, N. E. (2021). Modelling fish physico-thermal habitat selection using functional regression. *Journal of Ecohydraulics*, *6*(2), 105–120.

Bouveyron, C., Celeux, G., Murphy, T. B., & Raftery, A. E. (2019). *Model-based clustering and classification for data science: With applications in r* (Vol. 50). Cambridge University Press.

Campagna, C., Piola, A. R., Marin, M. R., Lewis, M., & Fernández, T. (2006). Southern elephant seal trajectories, fronts and eddies in the brazil/malvinas conflu-

ence. *Deep Sea Research Part I: Oceanographic Research Papers*, *53*(12), 1907–1924.

Campagna, J., Lewis, M. N., González Carman, V., Campagna, C., Guinet, C., Johnson, M., Davis, R. W., Rodrıguez, D. H., & Hindell, M. A. (2021). Ontogenetic niche partitioning in southern elephant seals from argentine patagonia. *Marine Mammal Science*, *37*(2), 631–651.

Carse, F., Martin, M. J., Sellar, A., & Blockley, E. W. (2015). Impact of assimilating temperature and salinity measurements by animal-borne sensors on foam ocean model fields. *Quarterly Journal of the Royal Meteorological Society*, *141*(693), 2934–2943.

Chamroukhi, F., & Nguyen, H. D. (2019). Model-based clustering and classification of functional data. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(4), e1298.

Chelton, D. B., Schlax, M. G., Witter, D. L., & Richman, J. G. (1990). Geosat altimeter observations of the surface circulation of the southern ocean. *Journal of Geophysical Research: Oceans*, *95*(C10), 17877–17903.

Conn, P. B., Thorson, J. T., & Johnson, D. S. (2017). Confronting preferential sampling when analysing population distributions: Diagnosis and model-based triage. *Methods in Ecology and Evolution*, *8*(11), 1535–1546.

Cotté, C., d'Ovidio, F., Dragon, A.-C., Guinet, C., & Lévy, M. (2015). Flexible preference of southern elephant seals for distinct mesoscale features within the antarctic circumpolar current. *Progress in Oceanography*, *131*, 46–58.

Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *Journal of Statistical Planning and Inference*, *147*, 1–23.

Deacon, G. E. R. (1937). The hydrology of the southern ocean. *Discovery Rep.*, *15*, 3–122.

Della Penna, A., De Monte, S., Kestenare, E., Guinet, C., & d'Ovidio, F. (2015). Quasi-planktonic behavior of foraging top marine predators. *Scientific reports*, *5*(1), 1–10.

Dinsdale, D., & Salibian-Barrera, M. (2019). Modelling ocean temperatures from bioprobes under preferential sampling.

d'Ovidio, F., De Monte, S., Alvain, S., Dandonneau, Y., & Lévy, M. (2010). Fluid dynamical niches of phytoplankton types. *Proceedings of the National Academy of Sciences*, *107*(43), 18366–18370.

Dragon, A.-C., Monestiez, P., Bar-Hen, A., & Guinet, C. (2010). Linking foraging behaviour to physical oceanographic structures: Southern elephant seals and mesoscale eddies east of kerguelen islands. *Progress in Oceanography*, *87*(1-4), 61–71.

Evans, K., Lea, M.-A., & Patterson, T. (2013). Recent advances in bio-logging science: Technologies and methods for understanding animal behaviour and physiology and their environments.

Fedak, M., et al. (2004). Marine animals as platforms for oceanographic sampling: A" win/win" situation for biology and operational oceanography.

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

Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association*, *97*(458), 611–631.

Fu, L.-L. (2006). Pathways of eddies in the south atlantic ocean revealed from satellite altimeter observations. *Geophysical Research Letters*, *33*(14).

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*, 14–22.

Garcia, V. M., Garcia, C. A., Mata, M. M., Pollery, R. C., Piola, A. R., Signorini, S. R., McClain, C. R., & Iglesias-Rodriguez, M. D. (2008). Environmental factors controlling the phytoplankton blooms at the patagonia shelf-break in spring. *Deep Sea Research Part I: Oceanographic Research Papers*, *55*(9), 1150–1166.

Godard, M., Manté, C., Guinet, C., Picard, B., & Nerini, D. (2020). Diving behavior of mirounga leonina: A functional data analysis approach. *Frontiers in Marine Science*, *7*, 595.

Gonzalez-Silvera, A., Santamaria-del-Angel, E., Garcia, V. M., Garcia, C. A., Millán-Nuñez, R., & Muller-Karger, F. (2004). Biogeographical regions of the tropical

and subtropical atlantic ocean off south america: Classification based on pigment (czcs) and chlorophyll-a (seawifs) variability. *Continental Shelf Research*, *24*(9), 983–1000.

Gordon, A. L. (1981). South atlantic thermocline ventilation. *Deep Sea Research Part A. Oceanographic Research Papers*, *28*(11), 1239–1264.

Goulet, P., Guinet, C., Campagna, C., Campagna, J., Tyack, P. L., & Johnson, M. (2020). Flash and grab: Deep-diving southern elephant seals trigger anti-predator flashes in bioluminescent prey. *Journal of Experimental Biology*, *223*(10), jeb222810.

Goulet, P., Guinet, C., Swift, R., Madsen, P. T., & Johnson, M. (2019). A miniature biomimetic sonar and movement tag to study the biotic environment and predator-prey interactions in aquatic animals. *Deep Sea Research Part I: Oceanographic Research Papers*, *148*, 1–11.

Guinet, C., Vacquié-Garcia, J., Picard, B., Bessigneul, G., Lebras, Y., Dragon, A. C., Viviant, M., Arnould, J. P., & Bailleul, F. (2014). Southern elephant seal foraging success in relation to temperature and light conditions: Insight into prey distribution. *Marine Ecology Progress Series*, *499*, 285–301.

Harcourt, R., Sequeira, A. M., Zhang, X., Roquet, F., Komatsu, K., Heupel, M., McMahon, C., Whoriskey, F., Meekan, M., Carroll, G., et al. (2019). Animal-borne telemetry: An integral component of the ocean observing toolkit. *Frontiers in Marine Science*, *6*, 326.

Hays, G. C., Ferreira, L. C., Sequeira, A. M., Meekan, M. G., Duarte, C. M., Bailey, H., Bailleul, F., Bowen, W. D., Caley, M. J., Costa, D. P., et al. (2016). Key questions in marine megafauna movement ecology. *Trends in ecology & evolution*, *31*(6), 463–475.

Hindell, M. A., Mcmahon, C. R., Bester, M. N., Boehme, L., Costa, D., Fedak, M. A., Guinet, C., Herraiz-Borreguero, L., Harcourt, R. G., Huckstadt, L., et al. (2016). Circumpolar habitat use in the southern elephant seal: Implications for foraging success and population trajectories. *Ecosphere*, *7*(5), e01213.

Hindell, M. A., Reisinger, R. R., Ropert-Coudert, Y., Hückstädt, L. A., Trathan, P. N., Bornemann, H., Charrassin, J.-B., Chown, S. L., Costa, D. P., Danis, B., et al. (2020). Tracking of marine predators to protect southern ocean ecosystems. *Nature*, *580*(7801), 87–92.

Hussey, N. E., Kessel, S. T., Aarestrup, K., Cooke, S. J., Cowley, P. D., Fisk, A. T., Harcourt, R. G., Holland, K. N., Iverson, S. J., Kocik, J. F., et al. (2015). Aquatic animal telemetry: A panoramic window into the underwater world. *Science*, *348*(6240), 1255642.

Jacques, J., & Preda, C. (2013). Funclust: A curves clustering method using functional random variables density approximation. *Neurocomputing*, *112*, 164–171.

Jacques, J., & Preda, C. (2014a). Functional data clustering: A survey. *Advances in Data Analysis and Classification*, *8*, 231–255.

Jacques, J., & Preda, C. (2014b). Model-based clustering for multivariate functional data. *Computational Statistics & Data Analysis*, *71*, 92–106.

James, G. M., & Sugar, C. A. (2003). Clustering for sparsely sampled functional data. *Journal of the American Statistical Association*, *98*(462), 397–408.

Jaud, T., Dragon, A.-C., Garcia, J. V., & Guinet, C. (2012). Relationship between chlorophyll a concentration, light attenuation and diving depth of the southern elephant seal mirounga leonina.

Jones, D. C., Holt, H. J., Meijers, A. J., & Shuckburgh, E. (2019). Unsupervised clustering of southern ocean argo float temperature profiles. *Journal of Geophysical Research: Oceans*, *124*(1), 390–402.

Jonsen, I. D., Myers, R. A., & James, M. C. (2007). Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. *Marine Ecology Progress Series*, *337*, 255–264.

Jouma'a, J., Le Bras, Y., Richard, G., Vacquié-Garcia, J., Picard, B., El Ksabi, N., & Guinet, C. (2016). Adjustment of diving behaviour with prey encounters and body condition in a deep diving predator: The southern elephant seal. *Functional Ecology*, *30*(4), 636–648.

Korte-Stapff, M., Yarger, D., Stoev, S., & Hsing, T. (2022). A multivariate functional-data mixture model for spatio-temporal data: Inference and cokriging. *arXiv preprint arXiv:2211.04012*.

Labrousse, S., Williams, G., Tamura, T., Bestley, S., Sallée, J.-B., Fraser, A. D., Sumner, M., Roquet, F., Heerah, K., Picard, B., et al. (2018). Coastal polynyas: Winter oases for subadult southern elephant seals in east antarctica. *Scientific Reports*, *8*(1), 1–15.

Le Bras, Y., Jouma'a, J., & Guinet, C. (2017). Three-dimensional space use during the bottom phase of southern elephant seal dives. *Movement ecology*, *5*(1), 1–15.

Le Bras, Y., Jouma'a, J., Picard, B., & Guinet, C. (2016). How elephant seals (mirounga leonina) adjust their fine scale horizontal movement and diving behaviour in relation to prey encounter rate. *PLoS One*, *11*(12), e0167226.

Li, H., Deng, X., Dolloff, C. A., & Smith, E. P. (2016). Bivariate functional data clustering: Grouping streams based on a varying coefficient model of the stream water and air temperature relationship. *Environmetrics*, *27*(1), 15–26.

Lopez, R., Malardé, J.-P., Danès, P., & Gaspar, P. (2015). Improving argos doppler location using multiple-model smoothing. *Animal Biotelemetry*, *3*(1), 1–9.

Lutz, V. A., Segura, V., Dogliotti, A. I., Gagliardini, D. A., Bianchi, A. A., & Balestrini, C. F. (2010). Primary production in the argentine sea during spring estimated by field and satellite models. *Journal of Plankton Research*, *32*(2), 181–195.

Maamaatuaiahutapu, K., Garçon, V. C., Provost, C., Boulahdid, M., & Bianchi, A. A. (1994). Spring and winter water mass composition in the brazil-malvinas confluence. *Journal of Marine Research*, *52*(3), 397–426.

March, D., Boehme, L., Tintoré, J., Vélez-Belchi, P. J., & Godley, B. J. (2020). Towards the integration of animal-borne instruments into global ocean observing systems. *Global change biology*, *26*(2), 586–596.

Martinetto, P., Alemany, D., Botto, F., Mastrángelo, M., Falabella, V., Acha, E. M., Antón, G., Bianchi, A., Campagna, C., Cañete, G., et al. (2020). Linking the scientific knowledge on marine frontal systems with ecosystem services. *Ambio*, *49*(2), 541–556.

Mason, E., Pascual, A., Gaube, P., Ruiz, S., Pelegrı, J. L., & Delepoulle, A. (2017). Subregional characterization of mesoscale eddies across the b razil-m alvinas c onfluence. *Journal of Geophysical Research: Oceans*, *122*(4), 3329–3357.

Matano, R. P., & Palma, E. D. (2008). On the upwelling of downwelling currents. *Journal of Physical Oceanography*, *38*(11), 2482–2500.

Maze, G., Mercier, H., & Cabanes, C. (2017). Profile classification models. *Mercator Ocean Journal*, (55), 48–56.

Maze, G., Mercier, H., Fablet, R., Tandeo, P., Radcenco, M. L., Lenca, P., Feucher, C., & Le Goff, C. (2017). Coherent heat patterns revealed by unsupervised classification of argo temperature profiles in the north atlantic ocean. *Progress in Oceanography*, *151*, 275–292.

McMahon, C. R., Field, I. C., Bradshaw, C. J., White, G. C., & Hindell, M. A. (2008). Tracking and data–logging devices attached to elephant seals do not affect individual mass gain or survival. *Journal of Experimental Marine Biology and Ecology*, *360*(2), 71–77.

McMahon, C. R., Roquet, F., Baudel, S., Belbeoch, M., Bestley, S., Blight, C., Boehme, L., Carse, F., Costa, D. P., Fedak, M. A., et al. (2021). Animal borne ocean sensors–anibos–an essential component of the global ocean observing system. *Frontiers in Marine Science*, 1625.

Miller, R. N., Matano, R. P., & Palma, E. D. (2011). Shelfbreak upwelling induced by alongshore currents: Analytical and numerical results. *Journal of Fluid Mechanics*, *686*, 239–249.

Nerini, D., Manté, C., & Monestiez, P. (2022). Extending functional kriging when data are multivariate curves: Some technical considerations and operational solutions. *Geostatistical Functional Data Analysis*, 73–103.

Nerini, D., Monestiez, P., & Manté, C. (2010). Cokriging for spatial functional data. *Journal of Multivariate Analysis*, *101*(2), 409–418.

Oh, H.-S., Lee, T. C., & Nychka, D. W. (2011). Fast nonparametric quantile regression with arbitrary smoothing methods. *Journal of Computational and Graphical Statistics*, *20*(2), 510–526.

Orsi, A. H., Whitworth III, T., & Nowlin Jr, W. D. (1995). On the meridional extent and fronts of the antarctic circumpolar current. *Deep Sea Research Part I: Oceanographic Research Papers*, *42*(5), 641–673.

Paniagua, G. F., Saraceno, M., Piola, A. R., Charo, M., Ferrari, R., Artana, C., & Provost, C. (2021). Malvinas current at 44.7° s: First assessment of velocity temporal variability from in situ data. *Progress in Oceanography*, *195*, 102592.

Pauthenet, E., Roquet, F., Madec, G., Guinet, C., Hindell, M., Mcmahon, C. R., Harcourt, R., & Nerini, D. (2018). Seasonal meandering of the polar front upstream of the kerguelen plateau. *Geophysical Research Letters*, *45*(18), 9774–9781.

Pauthenet, E., Roquet, F., Madec, G., & Nerini, D. (2017). A linear decomposition of the southern ocean thermohaline structure. *Journal of Physical Oceanography*, *47*(1), 29–47.

Pauthenet, E., Roquet, F., Madec, G., Sallée, J.-B., & Nerini, D. (2019). The thermohaline modes of the global ocean. *Journal of Physical Oceanography*, *49*(10), 2535–2552.

Payne, N. L., Meyer, C. G., Smith, J. A., Houghton, J. D., Barnett, A., Holmes, B. J., Nakamura, I., Papastamatiou, Y. P., Royer, M. A., Coffey, D. M., et al. (2018). Combining abundance and performance data reveals how temperature regulates coastal occurrences and activity of a roaming apex predator. *Global change biology*, *24*(5), 1884–1893.

Piola, A. R., & Gordon, A. L. (1989). Intermediate waters in the southwest south atlantic. *Deep Sea Research Part A. Oceanographic Research Papers*, *36*(1), 1–16.

Piola, A., & Falabella, V. (2009). El mar patagónico (2009) atlas del mar patagónico: Especies y espacios. *Widlife Conservation Society and Birdlife International, Buenos Aires*, 56–75.

Piola, A., Matano, R., Steele, J., Thorpe, S., & Turekian, K. (2001). Brazil and falklands (malvinas) currents. *Ocean currents*, 35–43.

Ramsay, J. O. (1982). When the data are functions. *Psychometrika*, *47*(4), 379–396.

Ramsay, J. O., & Silverman, B. W. (2005). *Functional data analysis*. Springer 2nd Edn. New York, NY.

Ramsay, J., Hooker, G., & Graves, S. (2009). Introduction to functional data analysis. In *Functional data analysis with r and matlab* (pp. 1–19). Springer.

Rey, A. R., & Huettmann, F. (2020). Telecoupling analysis of the patagonian shelf: A new approach to study global seabird-fisheries interactions to achieve sustainability. *Journal for Nature Conservation*, *53*, 125748.

Rivière, P., Jaud, T., Siegelman, L., Klein, P., Cotté, C., Le Sommer, J., Dencausse, G., & Guinet, C. (2019). Sub-mesoscale fronts modify elephant seals foraging behavior. *Limnology and Oceanography Letters*, *4*(6), 193–204.

Romero, S. I., Piola, A. R., Charo, M., & Garcia, C. A. E. (2006). Chlorophyll-a variability off patagonia based on seawifs data. *Journal of Geophysical Research: Oceans*, *111*(C5).

Roquet, F., Charrassin, J.-B., Marchand, S., Boehme, L., Fedak, M., Reverdin, G., & Guinet, C. (2011). Delayed-mode calibration of hydrographic data obtained from animal-borne satellite relay data loggers. *Journal of Atmospheric and Oceanic Technology*, *28*(6), 787–801.

Roquet, F., Williams, G., Hindell, M. A., Harcourt, R., McMahon, C., Guinet, C., Charrassin, J.-B., Reverdin, G., Boehme, L., Lovell, P., et al. (2014). A southern indian ocean database of hydrographic profiles obtained with instrumented elephant seals. *Scientific data*, *1*(1), 1–10.

Roquet, F., Wunsch, C., Forget, G., Heimbach, P., Guinet, C., Reverdin, G., Charrassin, J.-B., Bailleul, F., Costa, D. P., Huckstadt, L. A., et al. (2013). Estimates of the southern ocean general circulation improved by animal-borne instruments. *Geophysical Research Letters*, *40*(23), 6176–6180.

Rosso, I., Mazloff, M. R., Talley, L. D., Purkey, S. G., Freeman, N. M., & Maze, G. (2020). Water mass and biogeochemical variability in the kerguelen sector of the southern ocean: A machine learning approach for a mixing hot spot. *Journal of Geophysical Research: Oceans*, *125*(3), e2019JC015877.

Saraceno, M., & Provost, C. (2012). On eddy polarity distribution in the southwestern atlantic. *Deep Sea Research Part I: Oceanographic Research Papers*, *69*, 62–69.

Saraceno, M., Provost, C., & Lebbah, M. (2006). Biophysical regions identification using an artificial neuronal network: A case study in the south western atlantic. *Advances in Space Research*, *37*(4), 793–805.

Saraceno, M., Provost, C., & Piola, A. R. (2005). On the relationship between satellite-retrieved surface temperature fronts and chlorophyll a in the western south atlantic. *Journal of Geophysical Research: Oceans*, *110*(C11).

Saraceno, M., Provost, C., Piola, A. R., Bava, J., & Gagliardini, A. (2004). Brazil malvinas frontal system as seen from 9 years of advanced very high resolution radiometer data. *Journal of Geophysical Research: Oceans*, *109*(C5).

Schmutz, A., Jacques, J., Bouveyron, C., Cheze, L., & Martin, P. (2020). Clustering multivariate functional data in group-specific functional subspaces. *Computational Statistics*, *35*(3), 1101–1131.

Schreer, J. F., & Kovacs, K. M. (1997). Allometry of diving capacity in air-breathing vertebrates. *Canadian Journal of Zoology*, *75*(3), 339–358.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461–464.

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). Mclust 5: Clustering, classification and density estimation using gaussian finite mixture models. *The R journal*, *8*(1), 289.

Siegelman, L., O'toole, M., Flexas, M., Rivière, P., & Klein, P. (2019). Submesoscale ocean fronts act as biological hotspot for southern elephant seal. *Scientific reports*, *9*(1), 1–13.

Simonoff, J. S. (2012). *Smoothing methods in statistics*. Springer Science & Business Media.

Team, R. C. (2021). R: A language and environment for statistical computing (r version 4.0. 3, r foundation for statistical computing, vienna, austria, 2020). *Google Scholar There is no corresponding record for this reference*.

Tokushige, S., Yadohisa, H., & Inada, K. (2007). Crisp and fuzzy k-means clustering algorithms for multivariate functional data. *Computational Statistics*, *22*, 1–16.

Tournier, M., Goulet, P., Fonvieille, N., Nerini, D., Johnson, M., & Guinet, C. (2021). A novel animal-borne miniature echosounder to observe the distribution and migration patterns of intermediate trophic levels in the southern ocean. *Journal of Marine Systems*, *223*, 103608.

Ullah, S., & Finch, C. F. (2013). Applications of functional data analysis: A systematic review. *BMC medical research methodology*, *13*, 1–12.

Vacquie-Garcia, J., Royer, F., Dragon, A.-C., Viviant, M., Bailleul, F., & Guinet, C. (2012). Foraging in the darkness of the southern ocean: Influence of bioluminescence on a deep diving predator.

Valla, D., Piola, A. R., Meinen, C. S., & Campos, E. (2018). Strong mixing and recirculation in the northwestern argentine basin. *Journal of Geophysical Research: Oceans*, *123*(7), 4624–4648.

Viviant, M., Trites, A. W., Rosen, D. A., Monestiez, P., & Guinet, C. (2010). Prey capture attempts can be detected in steller sea lions and other marine predators using accelerometers. *Polar biology*, *33*(5), 713–719.

Wahba, G. (1990). *Spline models for observational data*. SIAM.

Wang, J.-L., Chiou, J.-M., & Müller, H.-G. (2016). Functional data analysis. *Annual Review of Statistics and its application*, *3*, 257–295.

Weatherall, P., Marks, K. M., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J. E., Rovere, M., Chayes, D., Ferrini, V., & Wigley, R. (2015). A new digital bathymetric model of the world's oceans. *Earth and space Science*, *2*(8), 331–345.

Wolfe, J. H. (1963). *Object cluster analysis of social areas* (Doctoral dissertation). University of California.

Wunsch, C., & Ferrari, R. (2004). Vertical mixing, energy, and the general circulation of the oceans. *Annual Review of Fluid Mechanics*, *36*(1), 281–314.

Yen, J. D., Thomson, J. R., Paganin, D. M., Keith, J. M., & Mac Nally, R. (2015). Function regression in ecology and evolution: Free. *Methods in Ecology and Evolution*, *6*(1), 17–26.

Zhang, M., & Parnell, A. (2023). Review of clustering methods for functional data. *ACM Transactions on Knowledge Discovery from Data*, *17*(7), 1–34.

Highlights :

- Functional data analysis coupled with model-based clustering is a powerful method to identify oceanographic domains.
- Five functional oceanographic domains sharing similar thermohaline features are identified in the Brazil-Malvinas Confluence.
- Coupling data from model with in situ elephant seal profiles allows us to characterize habitats use by these predators.
- The studied elephant seals tend to visit all the oceanographic domains in varying proportion.
- The studied elephant seals increase their prey capture rate in transitions areas, i.e. where the probability of changing functional oceanographic domains is high.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: