Decomposing acoustic signal reveals the pelagic response to a frontal system between oceanic domains

Izard L. [1, *], Fonvieille N. [2], Merland C. [4, 5], Koubbi P. [4, 5], Nerini D. [2], Habasque Jeremie [3], Lebourges-Dhaussy Anne [3], Lo Monaco C. [1], Roudaut Gildas [3], D'ovidio F. [1], Charrassin J.-B. [1], Cotté Cedric [1]

[1] Sorbonne University, CNRS, IRD, MNHN, Oceanography and Climate Laboratory: Experiments and Numerical Approaches (LOCEAN-IPSL), 75005, Paris, France
[2] Aix-Marseille University, Toulon University, CNRS, IRD, Mediterranean Institute of Oceanography (MIO), Marseille, France
[3] LEMAR, UBO-CNRS-IRD-Ifremer IUEM, Plouzané, France
[4] UFR 918 Terre Environnement et Biodiversité, Sorbonne Université, Paris, France
[5] IFREMER, Centre Manche mer du Nord. Laboratoire Halieutique de Manche-Mer du Nord, Boulogne-sur-Mer, France

* Corresponding author : L. Izard, email address : lloyd.izard@locean.ipsl.fr

Abstract :

The pelagic zone is home to a large diversity of organisms such as macrozooplankton and micronekton (MM), connecting the surface productive waters to the mesopelagic layers (200-1000 m) through diel vertical migrations (DVM). Active acoustics complement net sampling observations by detecting sound-scattering layers (SL) of organisms, allowing to monitor the MM dynamics with a high spatio-temporal resolution. As the acoustic response of organisms is frequency-dependent, multi-frequency analyses are a pertinent approach to better integrate the rich diversity of organisms composing SLs and their respective dynamics. However, analysing simultaneously emitted acoustic signals with distinct depth ranges and separating spatial from temporal variability is challenging and needs adapted tools to be fully exploited. This study examines the pelagic realm in a transition zone between the Southern Ocean and the subtropical Indian Ocean, crossing the Saint-Paul and Amsterdam islands' natural reserve. We extended a Multivariate Functional Principal Component Analysis (mfPCA) to analyse the joint vertical variation of five frequencies from two oceanographic cruises (2016 and 2022), allowing the decomposition of the acoustic dataset into orthogonal vertical modes (VM) of variability. We found the first VM to be linked to the temporal variability due to DVM, while the following majorly depict patterns in spatial distribution. Overall, from the subantarctic to the subtropical zones, we observed (i) enrichment of densities in the surface layer (0–100 m), (ii) a decrease in densities in the intermediate layer during the daytime (100–300 m) and (iii) the apparition of an intensive deep scattering layer on the 38 kHz. We explored VMs' connection with in-situ environmental conditions by clustering our observations into three distinct environmental-acoustic regions. These regions were compared with vertically integrated nautical area scattering coefficient distribution, a proxy for marine organisms' biomass. Additionally, we analysed species assemblage changes from complementary cruises to further elucidate the observed acoustic distribution. We show that the mfPCA method is promising to better integrate the pelagic horizontal,

vertical and temporal dimensions which is a step toward further investigating the control of the environment on the distribution and structuring of pelagic communities.

Highlights

► Functional data analysis identifies variance in SL distribution via vertical modes. ► First mode is tied to temporal dynamics, next modes display spatial patterns. ► 3 regions link environmental surface data and SL from South Ocean to Southern Indian Ocean. ► 18 and 38 kHz differ in integrated acoustic density spatial patterns. ► Species assemblage help explain acoustic patterns in transition zones.

1. Introduction

The pelagic oceanic habitat harbours a rich assortment of life forms, including crustaceans, fish, and gelatinous plankton. Organisms ranging from ~2 to 20 cm and possessing the capability to swim against currents are referred to as micronekton (Kloser et al., 2009), while, zooplankton mostly drift horizontally. A large proportion of macro-zooplankton and micronekton (MM) performs diel vertical migrations (DVM; e.g. Brierley, 2014) which connects the epipelagic layer (0-200 m) with the mesopelagic zone (200-1000 m), leading to complex trophic interactions and ecosystemic services such as the active biological carbon pump (Giering et al., 2014). MM occupy a central place in the trophic web by consuming low trophic levels and being prey to top predators (Kozlov, 1995). The large volume

of mesopelagic fish (1-20 billion tons; Irigoien et al., 2014) supports the populations of top predators of patrimonial or economic value such as diving mammals (Daneri and Carlini, 2002) or tunas (Bertrand et al., 2002). Despite these crucial roles in ecological systems, information on the horizontal and vertical distribution of the mesopelagic fauna is globally scarce due to the difficulty of sampling and monitoring these organisms. Basic information is lacking on their vital rates, their stock and the links between oceanographic conditions and mesopelagic biomass and biodiversity (St. John et al., 2016). Consequently, the scientific community is urged to study the mesopelagic zone before human direct or indirect impacts bias our observations (Martin et al., 2020). Indeed, the global pelagic fauna is facing a loss of biomass at a global scale due to climate change (Ariza et al., 2022b) and increasing off-shore and deep industrial activities. Its large fish populations are also leading to a growing interest in commercial exploitation (e.g. Valinassab et al., 2007; Grimaldo et al., 2020).

Net sampling methods are extensively employed to investigate these ecosystems. They have shown that the distribution and abundance of MM are connected to productive conditions (Cotté et al., 2022), the presence of a thermocline (Youngbluth, 1975), larger-scale water mass distribution (major fronts; e.g. Hulley, 1981; Koubbi, 1993; Koubbi et al., 2011; Duhamel et al., 2014), and mesoscale oceanographic features (e.g. Pakhomov et al., 1994). However, the perpetual motion of the ocean, combined with intricate small-scale processes, results in a heterogeneous dispersion of MM, particularly within the upper ocean layers (Bertrand et al., 2014). This patch dynamic, with far-reaching impacts on predators, makes it challenging to fully grasp the underlying mechanisms governing marine ecosystems (Benoit-Bird and McManus, 2012).

To address this challenge, active acoustic serves as a complementary non-intrusive observation method that enables continuous monitoring of the pelagic fauna's distribution in both horizontal and vertical dimensions at high resolution (Trenkel et al., 2011; Benoit-Bird and Lawson, 2016). Continuous data collection allows for the detection of sound scattering layers (SL) of organisms, that can extend for tens to thousands of km (Proud et al., 2017). The high-frequency resolution enables the monitoring of both fine-scale (Baudena et al., 2021; Della Penna et al., 2021), and broader (Béhagle et al., 2016) changes in relation to oceanographic features. The vertical distribution of SL, and how it varies in time and space, are known to be affected by parameters such as oxygen minimum layers (Czudaj et al., 2021) and light (Aksnes et al., 2017). Their intensity and depth are correlated at a global scale to surface variables such as temperature, surface productivity and wind stress (Proud et al., 2017). It follows that the vertical distribution of SL has been found to be correlated to hydrological gradients and major fronts (Béhagle et al., 2016; Annasawmy et al., 2018; Kang et al., 2021; Ariza et al., 2022a; Chawarski et al., 2022), as well as to the seasonal cycle (Receveur et al., 2020b). However, the response of MM to acoustic stimulation is non-linear and some organisms are potentially less/not detected at certain frequencies (Lavery et al., 2007; Benoit-Bird and Lawson, 2016). Thus, the type of scatterers (the insonified organisms) greatly impacts the acoustic signal and the detection of SLs. Moreover, frequencies below or equal to 38 kHz are frequently employed, often using single-frequency analysis, to depict and characterise pelagic biota down to 1000 metres. Consequently, part of pelagic organisms remain unaccounted for due to their resonance at lower/higher frequencies. Considering simultaneous multifrequency acquisitions allows an effective description of pelagic community changes at distinct stations (Cotté et al., 2022) and along transects (Peña et al., 2014; Ariza et al., 2022a; García-Seoane et al., 2023; Assunção et al., 2023). In this context, our objective is to encompass the extensive variety of organisms comprising SLs within the epi- and mesopelagic zone, and their respective dynamics, using five frequencies (18, 38, 70, 120 and 200 kHz), even though their respective maximum depth ranges differ.

When dealing with abundant and complex data, particularly considering variables with distinct ranges, it is essential to develop statistical tools that objectively extract key components of variability and fully leverage the information in multi-frequency acoustic observations. For instance, the computation of integrated nautical area scattering coefficient

(NASC, m^2 nmi^{-2}; Maclennan, 2002), a proxy of organisms' relative biomass (e.g. Irigoien et al., 2014; Dornan et al., 2022), gives insight into the horizontal dispersion of biomass but information on the depth at which these changes occur are lacking from this metric. Investigating the vertical variability in acoustic profiles can be performed with Functional Principal Component Analysis (fPCA), part of Functional Data Analysis (FDA) methods (Ramsay and Silverman, 2005). To our knowledge only recent research has employed functional methods for analysing pelagic acoustic seascape variability through either single (Ariza et al., 2022b) or multiple (Ariza et al., 2022a) frequency analysis. In the same vein as these studies, our approach complements the latter by introducing an extended version of the Multivariate fPCA (mfPCA) outlined by Pauthenet et al. (2017). The methodology proposed here enhances statistical relevance to acoustic multiple frequencies by eliminating the need for vertical data segmentation prior to analysis when considering variables that share distinct ranges.

Implementing this unexplored method, this study investigates the structuring patterns of backscatter distribution over a broad area located at the mid-southern latitude in the Indian Ocean and known to be a transition area between the oligotrophic subtropical gyre and the more productive frontal zone that characterises the northern part of the Southern Ocean (Geisen et al., 2022). The study is conducted as part of the expansion of the French national reserve around Saint-Paul and Amsterdam islands (SPA), with a focus on identifying coherent ecological regions. Earlier investigations have distinguished distinct SL patterns at 38 kHz in the South-Western Indian Ocean (Béhagle et al., 2016; Annasawmy et al., 2018). Independent examination of 18 and 38 kHz datasets have similarly shown variations in SL patterns over time and space in the study region, depending on specific metrics and frequencies considered (Boersch-Supan et al., 2017).

The temporal facet of MM dynamics, driven by DVM, can obscure the identification of spatial trends, with repercussions up to large-scale biogeography (Sutton et al., 2017). Within the literature, certain studies focus on a specific period (e.g. Irigoien et al., 2014) or employ temporal segregation of their data, often using solar elevation as a discriminating factor (e.g. Béhagle et al., 2016, Ariza et al., 2022a). Another approach is to treat night and day as two complementary variables when the dataset permits the combination of day and night profiles into a single statistical observation (Ariza et al., 2022b). Finally, distinguishing between nighttime and daytime periods can be achieved using clustering methods (Boersch-Supan et al., 2017) or data reduction techniques (Receveur et al., 2020b). In this study, we employ the latter approach, hypothesising that by identifying temporal variability, it can be filtered to enhance our comprehension of spatial dynamics within the study area.

This study aims to (i) propose an extended application of mfPCA for analysing the backscatter across five frequencies with varying depth ranges, avoiding vertical segmentation; (ii) extract principal modes of variability that account for simultaneous vertical changes in all five profiles; and (iii) identify spatially coherent regions that link surface environmental conditions to the acoustic patterns by filtering out temporal variability. In accordance with previous studies, we expect an increase of the global acoustic densities in frontal areas and productive regions and contrasted structures of SL patterns in this transition zone. As complementary analyses, we compute the integrated NASC and analyse samples from additional cruises to enhance our understanding of observed acoustic biomass and communities. This research encourages discussions about adopting multifrequency and functional approaches to unravel vertical, temporal and spatial patterns in acoustic studies.

2. Methods

2.1. Study area

The study area ranges between 33-47° S and 72-80° E and englobes the French Economic Exclusive Zone (EEZ) of Saint-Paul and Amsterdam islands (SPA). This region covers contrasted oceanographic conditions and regimes from

the North of the Polar Front and subantarctic zones of the Southern Ocean to the subtropical zone of the South Indian
Ocean (Figure 1).
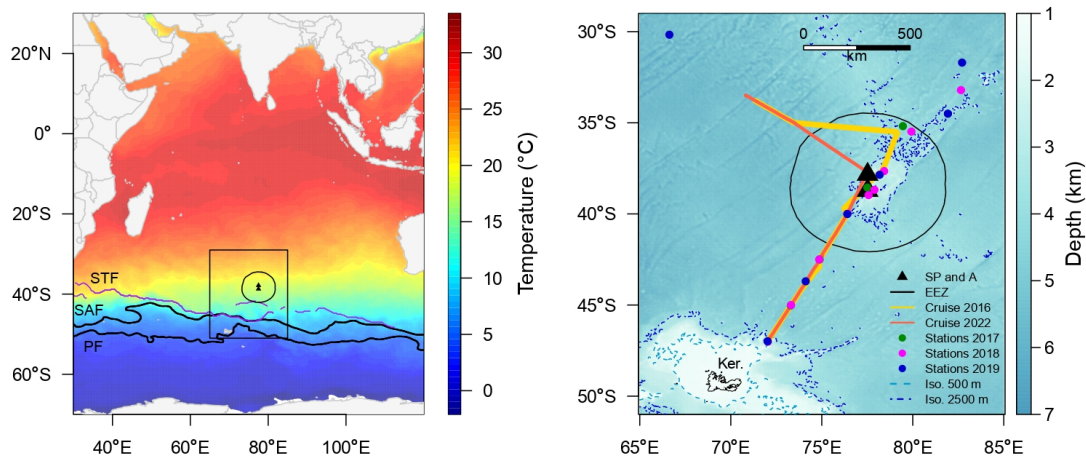
Figure 1: Location of the study area in the Southern Indian Ocean. The left panel shows the mean daily sea surface temperature from the 1st January 2016 to the 29th February 2016 (satellite product provided by the Copernicus Marine Environment Monitoring Service, https://doi.org/10.48670/moi-00165). The Polar Front (PF) and Subantarctic Front (SAF) are shown with black lines (Park et al., 2019). The dynamical Subtropical Front (STF) is shown with the purple lines (Graham and De Boer, 2013). The black rectangle defines the area of study. The right panel displays the bathymetry of the focus area. The 500 m and 2500 m isobath highlight the rise in bathymetry around the Kerguelen archipelago (Ker.) and Saint-Paul and Amsterdam islands (SP and A). The circular black line delimits the French Economic Exclusive Zone (EEZ). The yellow and orange lines correspond respectively to the transect of the THEMISTO cruises in 2016 and 2022. The dots correspond to the biological sampling stations coloured according to the year. Some stations are superimposed on others.

2.2. Data acquisition

All data acquisition took place aboard the R/V Marion Dufresne II and was carried out during austral summer.

2.2.1. Active acoustic data acquisition and processing

Acoustic data was continuously recorded during THEMISTO cruises in 2016 (https://doi.org/10.17600/16016100) and 2022 (https://doi.org/10.17600/18001847), selected among a time series of campaigns in the same area, starting in 2013, due to their high quality for the proposed analyses. Data was collected between February 19-24, 2016, and February 23-28, 2022, using an EK80 echosounder (Norway, SIMRAD), operating at 18, 38, 70, 120, and 200 kHz. The echosounder was calibrated annually before data acquisition (Demer et al., 2015). Calibrations were carried out off La Réunion using a 38.1 mm tungsten carbide sphere (6 % cobalt binder). During each procedure, the sphere was suspended 7-10 m below the transducers and measurements of this known target were recorded while moved throughout the beams. Acoustic parameters for each frequency are in Table 1. Note that the depth range differs between the frequencies. For both cruises, the average ping interval was 3 seconds and average ship speed was 6.2 m s^{-1}. Stationary data (speed < 1.5 m s^{-1}) and profiles collected at isobaths shallower than 1000 m were excluded. Data processing was performed using Matecho (Perrot et al., 2018) allowing data validation and manual removal of aliased seabeds by an expert, followed by the application of algorithms to filter attenuated pings, parasites, deep spikes, and correct background noise (Ryan et al., 2015; De Robertis and Higginbottom, 2007). Sound celerity was corrected using

Table 1

Surface offset, maximum acquired depth range and transmitted power of the 5 frequencies used by the EK80 echosounder during the two cruises. The echointegration ranges from the surface saturation zone down to the maximum acquisition range of frequencies. If two values are indicated, they correspond respectively to the 2016/2022 cruises.

	Frequency (kHz)				
	18	38	70	120	200
Surface offset (m)	25	25	25	25	25
Maximum depth range (m)	1000	820	500	230	110
Pulse power (10^3 W)	1/2	1	0.75	0.25	0.09/0.12

in situ temperature and salinity profiles from Conductivity-Temperature-Depth (CTD) casts deployed at depths > 1000 m.

Using the terminology from Maclennan (2002), we calculated the volume backscatter strength (S_v) in dB re 1 m^{-1}, representing marine organism density. Profiles of S_v were down-sampled to Elementary Sampling Units (ESU), with each echointegration cell being 2 m in height and 1 nautical mile in width (1 nautical mile = 1852 m), with a lower threshold of -100 dB. This spatial resolution balances finescale feature variation and coherent spatio-temporal structure in the dataset. We calculated the NASC (m^2 nmi^{-2}) from the smoothed S_v profiles (see paragraph (i) in Section 2.3.1) to ensure consistency between the acoustic profiles and the integrated NASC. The ESU periods were determined using solar elevation: daytime corresponds to solar elevation > 18° above the horizon and nighttime to solar elevation < 18° below the horizon. Twilight periods correspond to solar elevation ranging from -18° to 18°.

2.2.2. In situ environmental data

Subsurface (hull depth, 6 m) temperature (°C), salinity (psu), fluorescence (mg m^{-3}) and oxygen (μmol kg^{-1}) were continuously measured with an thermosalinograph, fluorometer and oxygen sensor. The data are provided by the OISO program (Océan Indien Service d'Observations; https://doi.org/10.18142/228). The four variables are used to investigate the link between the multi-frequency acoustic patterns and the oceanographic context. Bathymetry data used for analysis were obtained from the General Bathymetric Chart of the Oceans 2022 database (GEBCO, 2022; 15 arc-second grid resolution). Taking benefit from simultaneous hydrological *in situ* measurements, we associated multiple surface temperature and salinity discontinuities to the surface characteristics (Anilkumar et al., 2007) of Subantarctic Front (SAF) and two branches of the Subtropical Front (STF), separating distinct oceanic zones (e.g. Geisen et al., 2022).

2.2.3. Biological sampling and processing

Biological samples acquisition took place during the REPCCOAI (Réponse de l'écosystème pélagique aux changements climatiques dans l'océan Austral et Indien Sud - Response of the pelagic ecosystem to climate change in the Southern Ocean and South Indian) cruises in 2017, 2018 and 2019 (https://doi.org/10.18142/249). MM samples were collected using a pelagic Isaacs-Kidd Midwater Trawl (IKMT) net, which allows the retaining of large planktonic and micronektonic organisms. The net had a total length of 17 metres and a decreasing mesh size going from 3.5 cm at the entrance to 0.5 cm before the cod-end. The spinning speed was approximately 0.8 m s^{-1} and the vessel speed varied between 1 and 1.5 m s^{-1}. The number of IKMT samples by year in the area was n = 4 in 2017, n = 9 in 2018, and n = 7 in 2019. The net was obliquely towed at each station from 1000 m depth to the surface to obtain an integrated inventory of the macrozooplankton and micronekton community. This protocol was adopted for all nets in 2018 and 2019. In 2017, logistical and bad weather conditions meant that the same protocol could not be applied

to all nets. Out of a total of 20 nets, 3 were towed from a depth of 600 m during nighttime. The net was stabilised for three minutes at fishing depth before being raised to the surface at a turning speed of 0.6 m s^{-1}. Filtered volumes were calculated using a Hydrobios flowmeter.

Once the IKMT samples were collected, they were fixed in 5 % formalin buffered with seawater supplemented with sodium tetraborate for pH adjustment. Once species identifications were completed, the samples were reconditioned with Battaglia sauce for long-term preservation (Mastail and Battaglia, 1978). When the number of individuals in a sample was too large, the sample was split using a Motoda box (Motoda, 1959). Organisms were observed with a stereomicroscope and determined using identification keys (Baker et al., 1990; Boltovskoy, 1999; Kirkwood, 1982; O'Sullivan, 1983). Identification was performed down to species level and on rare occasions to the genus or the order when specimens were not in a good condition. Once counts were completed, the abundances (ind m^{-3}) were estimated using filtered volumes.

2.3. Statistics

All statistical methods were applied using R software version 4.2.0 (Team, 2022). The workflow displayed in Figure 2 summarises the statistical analyses and is used as a backbone to describe the methodology in the following sections. All acoustic data periods were analysed simultaneously (daytime, sunset, sunrise, nighttime), but 2016 and 2022 datasets were analysed separately.

2.3.1. Functional Data Analysis applied to multivariate acoustic data

The raw simultaneously acquired S_v profiles at the five frequencies (18, 38, 70, 120, and 200 kHz) were echointegrated onto a common grid. The resulting profiles are associated with the same ESU information, which includes identical GPS coordinates and time acquisition. To analyse the multiple frequencies joint variation, we term the five acoustic profiles as one acoustic observation (AO, Figure 2A). This consideration leads to $N = 720$ AOs in 2016 and $N = 962$ AOs in 2022.

(i) Discrete to functional acoustic data

An acoustic profile arrives as P pairwise values $(z_1, Sv_1), \ldots, (z_P, Sv_P)$, where z is depth (m) and Sv acoustic backscatter (dB re 1 m^{-1}). As depth is a continuum, it is assumed that values are sampled points of a single entity, here a continuous non-periodic function, whose shape characterises the vertical structure of MM distribution in the water column (Figure 2B). An acoustic value is possibly blurred by measurement errors (e.g. parasites, attenuation) and treatment choices (e.g. echointegration level). Hence, the backscatter Sv_p observed at depth z_p can be estimated with a continuous function f such that

$$Sv_p = f(z_p) + \varepsilon_p,$$

where the remainder ε_p is hoped to be as small as possible. It is assumed that the function f is expressed as a linear combination of K known basis functions $\phi_k(z)$:

$$f(z) = \sum_{k=1}^{K} \alpha_k \phi_k(z),$$

where the ϕ_k are continuous B-splines, piecewise polynomials of degree 3, joined end to end with continuity constraints at arguments values called knots (Ramsay and Silverman, 2005). Coefficients α_k are estimated by least squares regression when minimising the penalised sum of squared errors:

$$PENSSE = \sum_{p=1}^{P} \left(Sv_p - f(z_p) \right)^2 + \lambda \int_{\tau} \left[f^{(2)}(z) \right]^2 dz.$$

λ is a penalization parameter and $f^{(2)}$ is the second derivative of f whose square is integrated between surface to depth over the domain τ. The B-spline regression was applied with the R package `fda` (version 6.0.3).
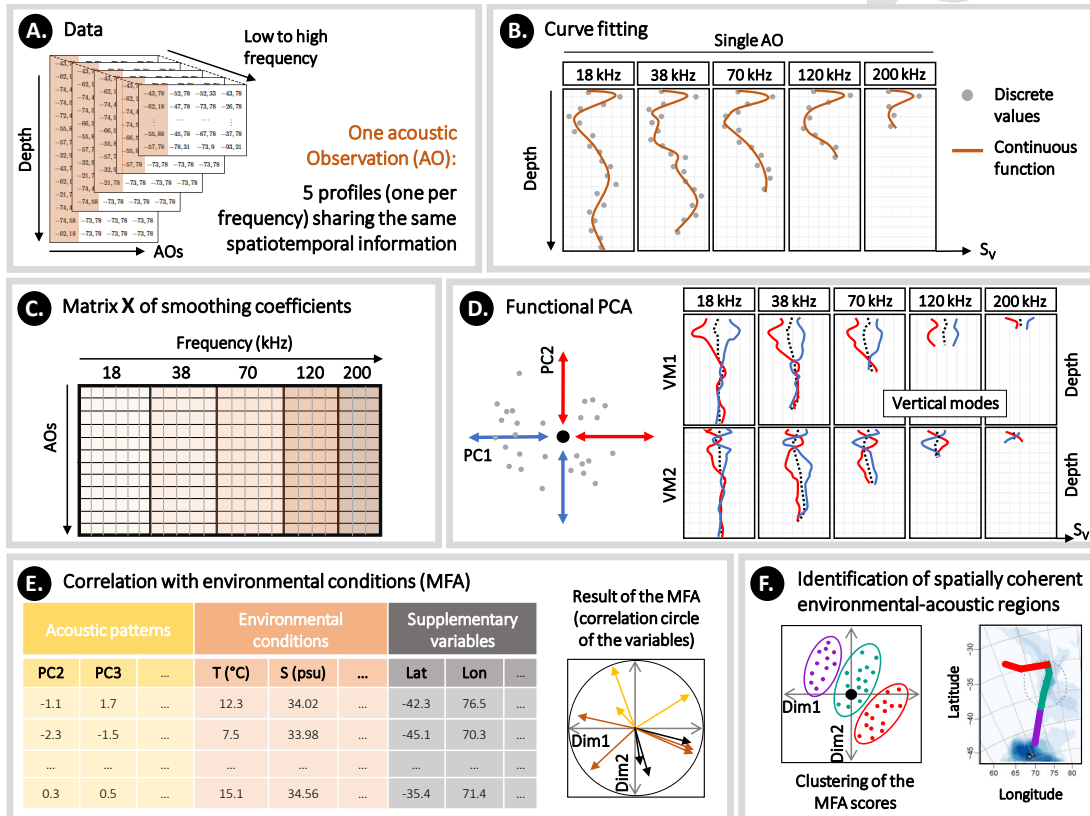
Figure 2: Workflow illustrating the methodology. (**A.**) An acoustic observation (AO) is defined by five echointegrated profiles, one per frequency (18, 38, 70, 120 and 200 kHz). Each of the five profiles differ in depth range. (**B.**) Each profile is smoothed in a B-spline basis system that is adapted to the frequency depth range. The grey dots are the discrete vertical values and the orange lines correspond to their associated functional curves. All profiles are expressed in backscatter strength Sv (dB re 1 m^{-1}). (**C.**) Each AO is summarised by a vector merging the set of coefficients that describe the continuous curves in B. The functional AOs are stored by row in a single matrix **X**, allowing us to consider the distinct frequencies covariance. (**D.**) A Multivariate Functional PCA (mfPCA) is applied on **X**. Each AO is decomposed into a combination of principal modes of variability, directly associated with the shape of the profiles. PC1 and PC2 correspond to the first and second principal components of the mfPCA while VM1 and VM2 are the associated first and second vertical modes of variability. (**E.**) A final dataset is created by merging the environmental conditions with the main modes of acoustic patterns defined by the scores of a selection of principal components. A Multivariate Factorial Analysis (MFA) is applied to the final dataset. The correlations between the different variables can be displayed and analysed. (**F.**) The MFA scores (lines of the matrix in E.) are clustered with a model-based clustering method. The groups can be projected into the geographical space and define environmental-acoustic regions. Dim1 and Dim2 refer to the dimensions of the MFA.

177 The number of basis functions controls the smoothness of the profiles. In the literature, it has been chosen either
178 to match data resolution (Ariza et al., 2022a) or minimise quadratic error (Godard et al., 2020). Accounting for the
179 varying depth ranges of each frequency (Table 1), we determine distinct values of K to achieve a knot spacing of 20
180 m, resulting in separate basis systems ($K^{18} = 50$; $K^{38} = 42$; $K^{70} = 26$; $K^{120} = 13$; $K^{200} = 7$). We set the penalty
181 parameter $\lambda = 0.05$ to balance solution smoothness and data fit. This smoothing penalty enables increasing K without
182 overfitting the curves.

183 Following this step, profile shapes are represented by coefficients. Each AO is summarised by a vector $\boldsymbol{\alpha}$ merging
184 the coefficients associated with the five frequencies:

$$\boldsymbol{\alpha} = (\alpha_1^{18}, .., \alpha_{K^{18}}^{18}; \alpha_1^{38}, .., \alpha_{K^{38}}^{38}; \alpha_1^{70}, .., \alpha_{K^{70}}^{70}; \alpha_1^{120}, .., \alpha_{K^{120}}^{120}; \alpha_1^{200}, .., \alpha_{K^{200}}^{200})'.$$

185 The successive AOs are then row combined into a single matrix $\mathbf{X}$ of dimension $N \times K^T$, where N is the
186 number of AOs and K^T the total number of coefficients associated with the five frequencies such that $K^T =$
187 $K^{18} + K^{38} + K^{70} + K^{120} + K^{200}$ (Figure 2C).

188 (ii) Principal modes of multivariate acoustic variability

189 The following section presents the mfPCA procedure by following the steps documented in Pauthenet et al. (2017).
190 Firstly, we compute the mean vector $\overline{\boldsymbol{\alpha}}$ of acoustic observations with

$$\overline{\boldsymbol{\alpha}} = (\overline{\alpha}_1^{18}, \ldots, \overline{\alpha}_{K^{18}}^{18}; \overline{\alpha}_1^{38}, \ldots, \overline{\alpha}_{K^{38}}^{38}; \overline{\alpha}_1^{70}, \ldots, \overline{\alpha}_{K^{70}}^{70}; \overline{\alpha}_1^{120}, \ldots, \overline{\alpha}_{K^{120}}^{120}; \overline{\alpha}_1^{200}, \ldots, \overline{\alpha}_{K^{200}}^{200})', \tag{1}$$

191 where $\overline{\alpha}_1^{18}$ is the mean of the N coefficients α_1^{18}. Subtracting this mean vector to each row of $\mathbf{X}$ forms the centred
192 matrix $\mathbf{C}$. The main modes of variability are obtained by solving the following eigenvalue problem:

$$\mathbf{VWb}_l = \lambda_l \mathbf{b}_l,$$

193 where $\mathbf{b}_l$ is the lth eigenvector associated with the eigenvalue λ_l. The crossed covariance matrix $\mathbf{V} = \frac{1}{N}\mathbf{C}'\mathbf{C}$ of
194 size $K^T \times K^T$ is a block matrix with entries

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_{18/18} & \mathbf{V}_{18/38} & \mathbf{V}_{18/70} & \mathbf{V}_{18/120} & \mathbf{V}_{18/200} \\ \mathbf{V}_{38/18} & \mathbf{V}_{38/38} & \mathbf{V}_{38/70} & \mathbf{V}_{38/120} & \mathbf{V}_{38/200} \\ \mathbf{V}_{70/18} & \mathbf{V}_{70/38} & \mathbf{V}_{70/70} & \mathbf{V}_{70/120} & \mathbf{V}_{70/200} \\ \mathbf{V}_{120/18} & \mathbf{V}_{120/38} & \mathbf{V}_{120/70} & \mathbf{V}_{120/120} & \mathbf{V}_{120/200} \\ \mathbf{V}_{200/18} & \mathbf{V}_{200/38} & \mathbf{V}_{200/70} & \mathbf{V}_{200/120} & \mathbf{V}_{200/200} \end{pmatrix},$$

195 where $\mathbf{V}_{18/38}$ is the covariance matrix of size $K^{18} \times K^{38}$ between coefficients of the 18 and 38 kHz frequencies.
196 Since the B-spline basis does not form an orthonormal basis, the $\mathbf{W}$ matrix of size $K^T \times K^T$ is defined to guarantee
197 the metric equivalence between the functional problem (working on functions) and its discrete version (working on
198 coefficients of the decomposition). The matrix is constructed by block as follows:

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_{18} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{W}_{38} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{W}_{70} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{W}_{120} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{W}_{200} \end{pmatrix}.$$

Each non zero entry is constituted with a matrix of scalar products of the basis functions. Note that unlike in Pauthenet et al. (2017), different basis systems have been used for the five profile expansion (the five frequencies), leading to $\mathbf{W}_{18} \neq \mathbf{W}_{38} \neq \mathbf{W}_{70} \neq \mathbf{W}_{120} \neq \mathbf{W}_{200}$. Finally, the matrix $\mathbf{M}$ ensure the normalisation step during the mfPCA, giving the same weight to each frequency.

The AO can now be projected in a space of reduced dimension when computing the principal component (PC) vectors $\mathbf{y}_l$ associated to each eigenvalue λ_l with

$$\mathbf{y}_l = \mathbf{C}\mathbf{M}^{-1/2}\mathbf{W}^{-1/2}\mathbf{b}_l.$$

The PCs capture the variance of the system. They correspond to the uncorrelated linear combinations of the original variables. In this mfPCA, a total of K^T eigenvectors are obtained and sorted in ascending order based on their associated eigenvalue. Each eigenvector generates five eigenfunctions ($\xi^{18}, \xi^{38}, \xi^{70}, \xi^{120}, \xi^{200}$), referred to as vertical modes (VM, Pauthenet et al., 2017). The first vertical mode (VM1) corresponds to the eigenvector with the largest eigenvalue.

We compute the mean functional profile for each frequency as

$$\overline{x}^j(z) = \sum_{k=1}^{K^j} \overline{\alpha}_k^j \phi_k(z), j \in \{18, 38, 70, 120, 200\}.$$

Here, the mean profile $\overline{x}^{18}(z)$ corresponds to the mean backscattering profile at 18 kHz (and so on for the other frequencies). The effect of the different VMs can be displayed by adding or subtracting the corresponding eigenfunction to the mean profile $\overline{x}^j(z)$ (Figure 2D):

$$\overline{x}^j(z) \pm \sqrt{\lambda_l}\xi_l^j(z), j \in \{18, 38, 70, 120, 200\}.$$

For example, the deformation of the mean profile $\overline{x}^{18}(z)$ associated with VM1 can be displayed computing $\overline{x}^{18}(z) \pm \sqrt{\lambda_1}\xi_1^{18}(z)$.

PC scores obtained for each AO were interpreted by examining the effect of the associated eigenfunctions on the mean profiles. The dynamic of PC scores was investigated by comparing them with spatial and temporal factors. The R library `circular` (version 0.4.95) was used to display and explore temporal patterns.

2.3.2. Coupling environmental variables and acoustic patterns

A Multiple Factor Analysis (MFA) was conducted using the R library `FactoMineR` (version 2.4) to explore the potential relationship between the identified acoustic patterns and oceanographic conditions. Each AO was time-associated with *in situ* temperature, salinity, oxygen and fluorescence sampled at 6 m below the surface. MFA is a statistical method that allows us to describe observations structured in groups of variables (Pagès, 2002). The analysis involved two groups: (i) the acoustic group, consisting of a selection of mfPCA principal components (see Section 2.3.1), and (ii) the environmental group, consisting of the surface environmental variables. A third group composed of latitude, longitude and bathymetry was included as supplementary variables to examine their relation with the MFA dimensions without affecting the results.

The whole set of variables constitute a final dataset displayed in Figure 2E. Standardisation was performed to ensure comparability among variables measured in different units within the environmental and supplementary groups. Since MFA is based on the core of the more classic PCA, the eigenvalues, variable contributions, and scores are investigated to identify the variables that contribute most to variations in the dataset.

2.3.3. Identifying environmental-acoustic regions

The scores of the MFA were clustered using a model-based method (R library `mclust` version 5.4.10, Scrucca et al., 2016) to define groups sharing similar acoustic vertical structures and environmental conditions, called hereafter environmental-acoustic regions (Figure 2F). Model-based clustering is based on a probability model defined by a finite mixture of multivariate Gaussian distributions (Bouveyron et al., 2019). The Bayesian Information Criterion (BIC, Schwarz, 1978) and the Integrated Completed Likelihood (ICL, Biernacki et al., 2000) were computed for defining the optimal number of groups in the dataset. Pairwise Wilcoxon tests and functional ANalysis Of VAriance (fANOVA, R library `fdANOVA` version 0.1.2, Górecki and Smaga, 2019) were used to confirm the significance of the clusters ($\alpha = 0.05$).

2.3.4. Biological data

The abundance and diversity of organisms were explored by geographically assigning each net station with its corresponding environmental-acoustic regions. Statistical analyses were performed on five major taxonomic groups: euphausiids, siphonophores and salps for the zooplankton, and Gonostomatidae and Myctophidae for the fish. The abundance of the different taxon between the regions were compared using pairwise Wilcoxon tests ($\alpha = 0.05$).

Species-level abundance data have been used for euphausiids, siphonophores, salps and Myctophidae, while no identification of Gonostomatidae was performed. To study the assemblages' structure, (i) the abundances were transformed with a log x+1 transformation, (ii) a Bray Curtis dissimilarity matrix was computed and (iii) a Non-Metric Multidimensional Scaling (NMDS) was performed in order to represent the pairwise dissimilarities between stations in multidimensional space. An ANalysis Of Similarities (ANOSIM) was also performed to test the significance of the differences in species composition based on abundances, among the groups classified by environmental-acoustic regions. As a non-parametric test, ANOSIM uses ranked dissimilarities instead of actual distances (Clarke, 1993).

3. Results

3.1. Multi-frequency acoustic analysis

The mfPCA allows the examination of the pairwise correlation (normalised version of the crossed covariance matrix $\mathbf{V}$) between the five frequencies. Results for acoustic data in 2016 are displayed in Figure 3. For each panel, we can read the correlation between two frequencies at different pairs of depth (z, s). The diagonal panels correspond to the correlation for a given frequency (for instance $\mathbf{V_{18/18}}$). They are symmetrical with a correlation of 1 on their diagonal. Regarding the diagonal panels of the three lowest frequencies (18, 38 and 70 kHz), a highly correlated layer (> 0.4) is observed between 100 m and 350 m, particularly visible in the centre part of the 70/70 kHz panel. For the 18/18 kHz and 38/38 kHz panels, a second correlated layer (> 0.4) is highlighted from 400 m to the maximum depth range, while no correlation (close to 0) is observed between the upper layer (25-400 m) and the deeper layer (400 m to the maximum depth range). For the highest frequencies (diagonal panels 120/120 and 200/200 kHz), the correlation is always above 0.2, indicating that backscatter between the sub-surface and 220 m are well correlated.

Regarding the cross-correlation panels (e.g. 18/70 kHz), the correlations are maximal for similar depths, unless between the 18 and 38 kHz. For instance, the 18 kHz frequency at 100 m is highly correlated (> 0.8) with the 120 kHz frequency at the same depth. For pairs of distant depths (e.g. 18 kHz at 600 m and 120 kHz at 150 m), the correlation between two frequencies is globally close to zero, meaning that backscatters measured near the surface at high frequency give little to no information on the low-frequency backscatters measured deeper. Negative correlation (< -0.4) only appears with the 38 kHz. Backscatters measured in a thin layer around 400 m at 38 kHz are negatively correlated with backscatters measured between 100 and 400 m with the other frequencies.

Particular attention should be given to the cross-correlation between the 18 and 38 kHz. The pairwise correlation (z, s) when $z = s$ is weakly or not correlated (around 0) when depth z is between 300 and 400 m or below 600 m. For a given depth, the backscatter information provided by these two frequencies seems complementary, which emphasises the need for multivariate acoustic analysis.

Results on the frequency correlations are similar for acoustic data recorded in 2022 (not shown).
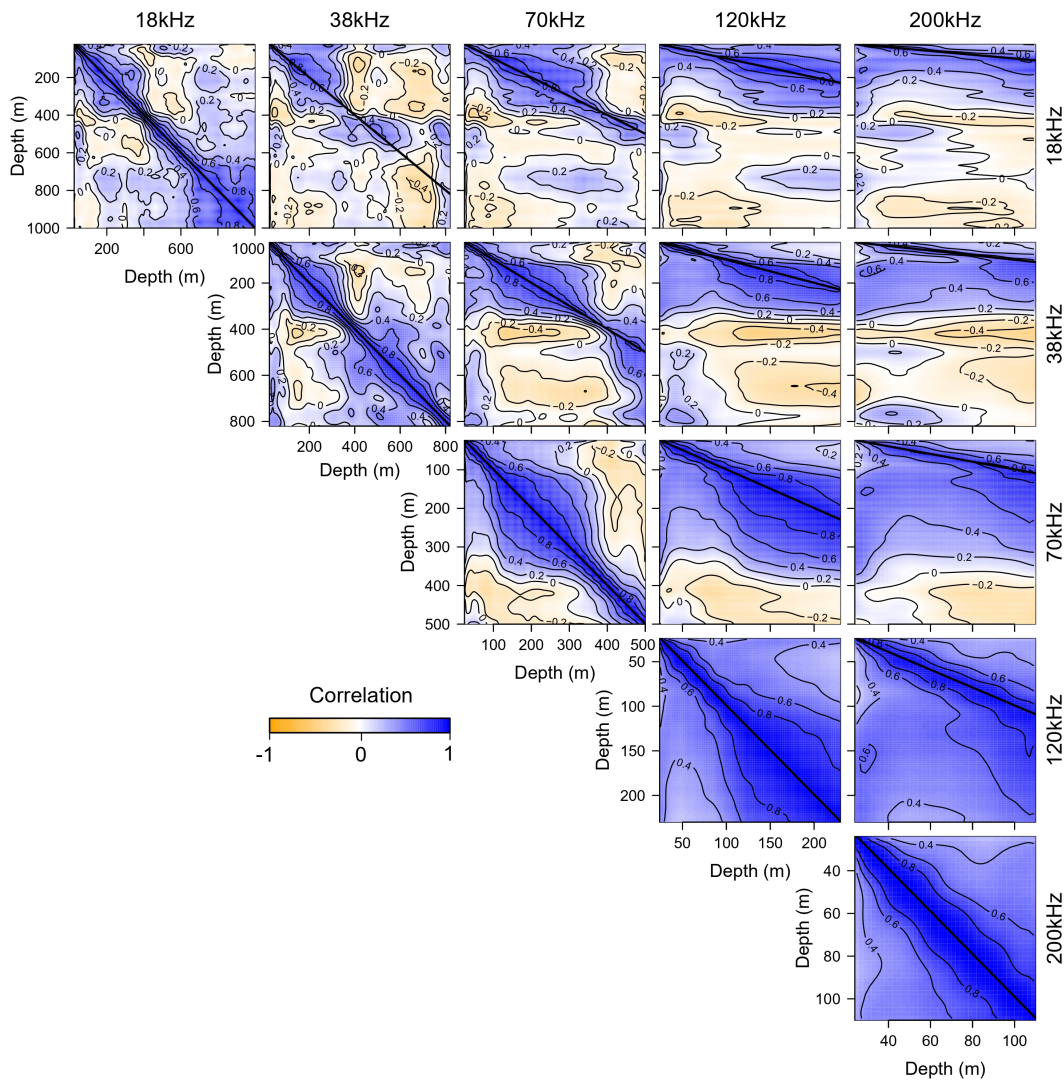
Figure 3: Contour plot of the normalized block-structured matrix **V** obtained for the 2016 acoustic data. **V** is computed on the coefficients of the B-spline expansion (projection of the acoustic observations in their functional space, Section 2.3.1). Each panel displays the correlation between pairwise frequencies (18, 38, 70, 120 and 200 kHz) at any depth. For example, the correlation between the 18 kHz at 700 m and the 200 kHz at 100 m is close to 0. Note that only the diagonal panels are symmetrical and square; the others are unsymmetrical and rectangle (as the range of the different frequencies changes). The square representation was chosen to facilitate the reading. The black line indicates the 1:1 depth.

3.2. Main modes of backscatter variability

The mfPCA identifies the main modes of variability associated with the changes in backscatter vertical distribution. Figure 4, Figure 5 and Figure 6 display the mfPCA results for 2016. Results for 2022 are shown in Supplementary Materials (Figure S1, Figure S2 and Figure S3). The first seven modes of variability were retained for subsequent analysis, accounting for 82.4 and 83.4 % of the variability in 2016 and 2022 (respectively) as done in Ariza et al. (2022a). In 2016, these modes represent 48.09, 12.54, 9.16, 3.93, 3.22, 2.92, and 2.58 % of the original variability (left panel Figure 4).
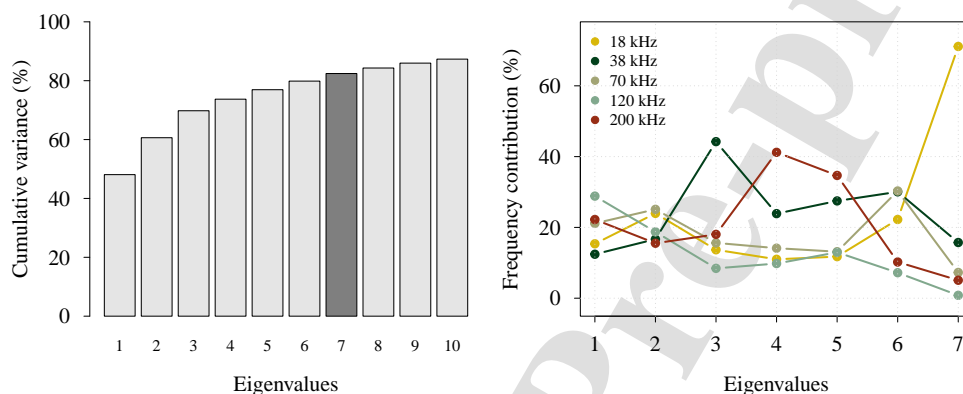
Figure 4: Results of the mfPCA performed on acoustic data in 2016. (Left) Cumulative percentage of variance explained by the 10 principal eigenvalues (the first seven add to 82.44 % of total inertia). (Right) Percentage of contributions of each frequency to the eigenvalues.

The contributions (%) of each frequency for a given mode are unequal and not constant across the frequencies (right panel Figure 4). The first eigenvalue contributions are distributed from low to high on the 38, 18, 70, 200 and 120 kHz. The first mode in 2022 is also primarily influenced by the 120 kHz frequency. The variation in contribution among frequencies decreases for the second eigenvalue, accompanied by a shift in their order of contribution (Figure 4). The contribution of the 38 kHz is particularly important for the third to the sixth eigenvalues (above 20 %). This contribution indicates modes of variability particularly driven by the vertical changes in MM distribution from organisms responsive to this frequency, also implying a large vertical range. The 200 kHz is also a large contributor to the fourth and fifth eigenvalues, while the 18 and 70 kHz are part of the 3 main contributors for the sixth eigenvalue. The seventh mode is finally mainly driven by the 18 kHz reaching 70 % of the variance explained.

For interpreting vertical modes (VM) of variability, we associate them with the mfPCA principal components (PC). The first three PCs' scores (left panels Figure 5) demonstrate the degree of deformation linked to their corresponding VM (right panels Figure 5). PC1's spatial pattern alternates between positive and negative values along the transect (Figure 5a). The associated VM1 reveals consistent deformation across the five frequencies, indicating periodic backscatter intensity changes from the surface down to 400 m (Figure 5d). PC2 scores transition from negative in the south to positive in the north (blue to red dots in Figure 5b). The corresponding VM2 opposes AOs with high or low backscatter in the upper 100 metres and an inversion between 100 and 400 m. Deformation is consistent across the five frequencies with varying intensity (e.g. more pronounced deformation from surface to 100 m for 38 kHz than 200 kHz; Figure 5e). PC3 scores are positive south of 42° S and north of 36° S (Figure 5c). This third mode is mainly influenced by the 38 kHz, opposing profiles with low or high backscatter between 200 and 800 m at this frequency
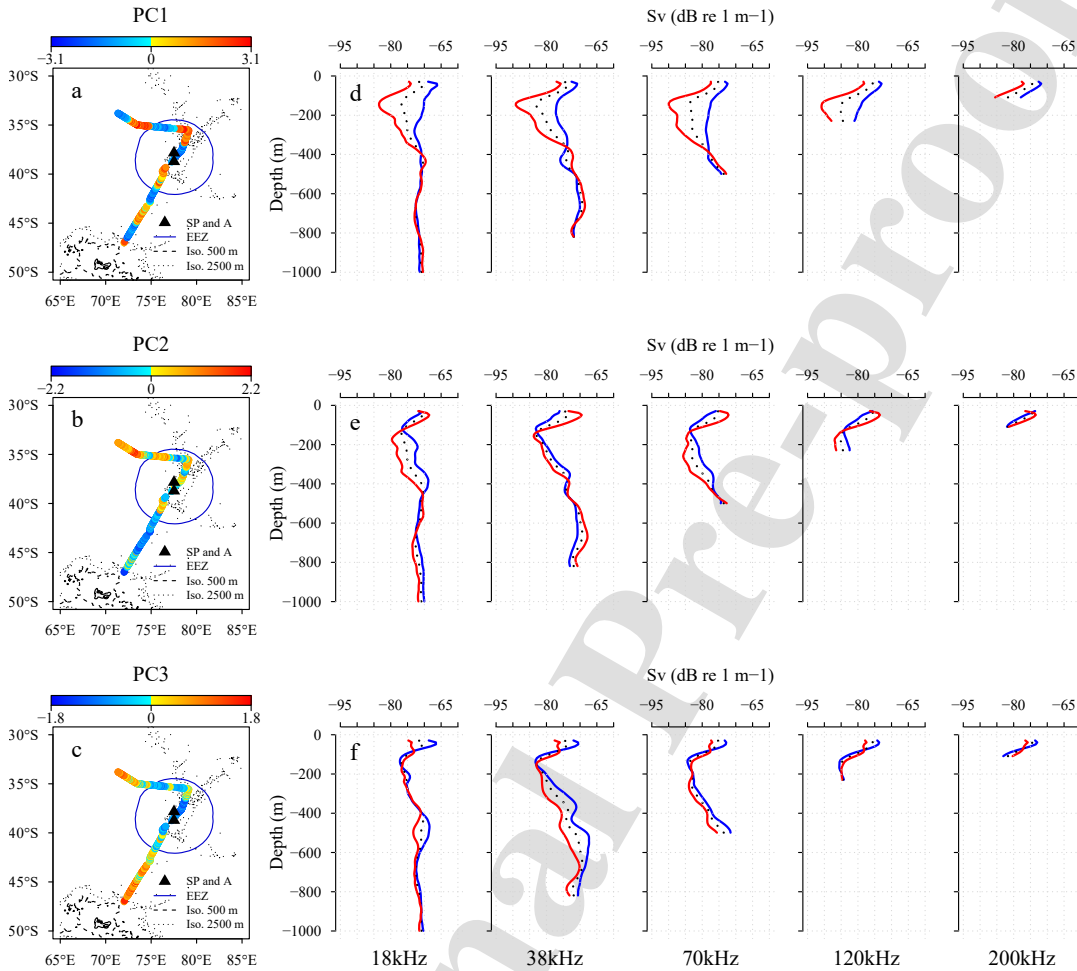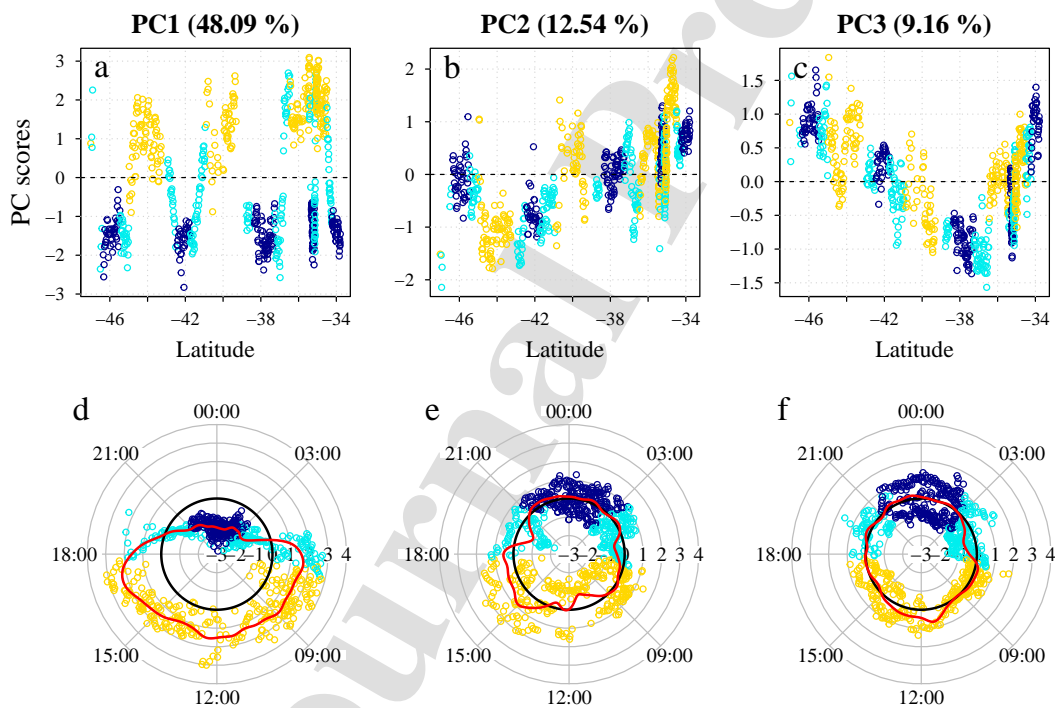
Figure 5: Results of the mfPCA performed on acoustic data in 2016. The three lines of panels correspond to the three first modes of variability. The left panels (**a**, **b**, **c**) display the spatial distribution of PC1, PC2 and PC3 along the ship trajectory while the right panels (**d**, **e**, **f**) show the deformation of the five mean profiles (black dotted line) associated with the corresponding vertical mode (VM). For a given line of panels, colors in the map match with the coloured profiles on the right side. For example, red dots along the trajectory in **a** correspond to acoustic observations with backscatter distribution close to the red profiles in **d** (i.e. low densities between the surface and 400 m for all frequencies). The color bar above each spatial panel is adjusted to the range of the corresponding PC. Bathymetry line at 500 and 2500 m surround the Kerguelen archipelago and Saint-Paul (SP) and Amsterdam (A) Islands. The blue circle define the Economic Exclusive Zone (EEZ).

(Figure 5f). Additionally, blue profiles (negative PC3) exhibit increased backscatter between the surface and 100 m across all frequencies.

To elucidate further the profile deformation patterns, scores of PC1, PC2, and PC3 were plotted against latitude and local time of the day (GMT+5; Figure 6, top and bottom panels, respectively).

The scores of PC1 present a latitudinal cyclic pattern that match the previous observation shown Figure 5a. The dots show a coherent alternance of periods along latitude due to the alternation of nighttime and daytime periods along the transect (Figure 6a). The same scores are displayed in a clockwise manner, with isolines circling the clock's zero level (black circle), depicting iso-shapes of AO deformation for the associated vertical mode. PC1 scores are positive during the daytime (yellow dots between 06:00 and 18:00) and negative at nighttime (dark blue dots between 21:00 and 03:00). Twilight periods (cyan dots between 03:00 and 06:00 and between 18:00 and 21:00) represent transitions between distinct stable temporal vertical structures. While PC1 primarily relates to the circadian cycle, the increasing amplitude between daytime and nighttime scores along latitude indicates that daytime period is more structured northwards, with lower densities recorded between the surface and 400 m (red curves in Figure 5d).

In contrast to PC1, PC2 and PC3 scores exhibit more distinct latitudinal trends (Figure 6). PC2 decreases from 46° to 44° S, then increases to 34° S. Notably, scores during the daytime show greater variability (PC2 $\in$ [-1.8;2.2]) than during the nighttime (PC2 $\in$ [-1.2;1.2]). For PC3, scores decrease until 37° S, then rise until the end of the transect, with an additional negative peak at 35°S. Additionally, PC2 and PC3 show a weak association with diel variability (Figure 6), as their scores cannot reliably distinguish between negative or positive values at a given time.

Figure 6: Results of the mfPCA performed on acoustic data in 2016. Projection of PC1 (**a, d**), PC2 (**b, e**) and PC3 (**c, f**) along latitudinal (top) and diel (bottom) variations. Yellow, cyan and dark blue dots correspond to day, twilight and night period (respectively) defined with the solar elevation. For both latitudinal and diel panels, the y-axis corresponds to the scores of the PC. For diel variability, the time is in UTC+5 (local time) and PC score scales are read horizontally ([-3;4]) from the center to the edge of the circles. The solid red line is the Local-Linear estimator for circular-linear data. The dashed black line (top panels) and the solid black line (bottom panels) highlight the 0-isoline. PC1 presents a clear link with local time, while PC2 and 3 present more of a latitudinal pattern.

Table 2
Results of the Multiple Factor Analysis (MFA) performed on the 2016 dataset. The lines correspond to the dimensions of the MFA with the contribution of each variable in percentage of variance explained (%). For example, PC2 is the major contributor of Dim1, reaching almost 42 %. PC2 to PC7 correspond to the principal components of the acoustic mfPCA (acoustic group). O, T, S and F correspond to oxygen, temperature, salinity and fluorescence (environmental group). Contributions above 10 % are highlighted in grey.

	Acoustic group						Environmental group			
	PC2	PC3	PC4	PC5	PC6	PC7	O	T	S	F
Dim1	41.96	5.45	0.74	0.22	0.07	0.02	15.69	15.19	13.93	6.74
Dim2	24.26	67.2	0.87	0.33	0.04	0.05	1.14	1.69	2.54	1.89
Dim3	7.22	17.05	43.66	9.51	0.86	0.93	3.23	4.86	7.33	5.36
Dim4	7.59	0.52	8.3	19.66	1.42	30.15	0.71	0.16	0.14	31.35
Dim5	4.50	1.41	21.39	44.66	8.5	11.08	1.88	1.22	0.01	5.35
Dim6	0.03	0.01	0.68	7.97	82.42	7.96	0	0.06	0.01	0.85

### 3.3. Acoustic and oceanographic context

Using Multiple Factor Analysis (MFA), we explored the relationship between two sets of variables. Our objective was to identify regions displaying shared environmental and acoustic patterns, while mitigating the influence of the circadian cycle. Consequently, the acoustic group of the MFA is composed of PC2 to PC7 scores, excluding PC1 (recognized as the main temporal mode) and consecutive modes that contributed less than 2% of the vertical variability. On the other hand, the environmental group is composed of surface temperature (°C), salinity (psu), oxygen (μmol kg^{-1}), and fluorescence (mg m^{-3}).

To prevent confusion with the mfPCA applied to acoustic data, the principal modes of MFA are termed *dimensions*. In 2016 and 2022, six and five primary dimensions were retained, respectively, to capture 94.69 and 91.32% of the total variability. The proportion of variance explained by the dimensions is 43, 20.6, 11, 7.4, 6.8, and 5.9 % for the 2016 cruise and 41.8 20.6, 12.1, 10.2 and 6.6 % for the 2022 cruise. In both cruises, the first dimension is evenly split between the environmental and acoustic groups, comprising 51.6 and 48.5 % of variance explained in 2016, and 52.8 and 47.2 % in 2022. Subsequent dimensions are primarily influenced by the acoustic group. Table 2 displays variable contributions (within each group) for each dimension. Among the first six MFA dimensions, all seven acoustic PCs contribute over 10 % to at least once. Temperature, salinity, and oxygen contribute to the first MFA dimension, while fluorescence contributes to the fourth dimension alongside PC5 and PC7 (Table 2).

The correlation circles in the left panels of Figure 7 offer a supplementary perspective on the connections between acoustic patterns and surface environmental conditions. Arrows closer to each other indicate positive correlation, while opposing directions signify negative correlation. The distance from the origin reflects variable representation quality for a specific axis. In both years, latitude, temperature, and salinity are highly correlated and show a negative correlation with oxygen. These four variables correlate with the first dimension of the MFA and PC2 (acoustic variable). This outcome suggests that the rise in backscatter within the upper 100 metres of the water column, linked to positive PC2 (Figure 5e, red curves), is positively linked to warmer, saltier but less oxygenated waters. In both years, PC3 is positively correlated with the second dimension.

Differences in correlations between the two years stem from two main reasons. First, the acoustic modes of the two mfPCA runs (for each year) may not correspond to the same profile deformations (see Figure 5 and Figure S2). For example, the deformation linked to PC2 in 2016 aligns with the deformation tied to PC3 in 2022. Second, deformations can be reversed (e.g. the red VM2 curve in 2016 corresponds to the blue VM3 curve in 2022). Connecting correlation circles with acoustic VMs reveals that fluorescence is correlated with lower backscatter density between the surface

349 and 100 m, and higher density between 100 and 400 m (blue curve in Figure 5 middle panels). Lastly, PC3 in 2016
350 exhibits a negative correlation with longitude and bathymetry (Figure 7), consistent with spatial variations observed
351 in Figure 5c.
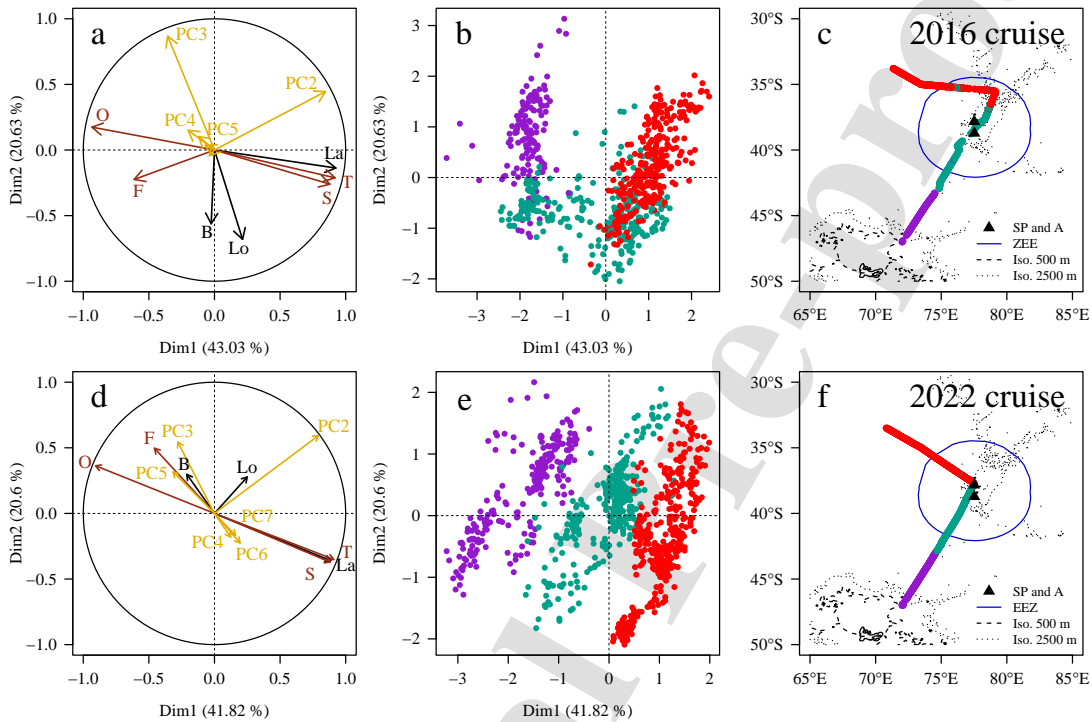
Figure 7: Results of the Multiple Factor Analysis (MFA) and model-based clustering (MBC) of the MFA scores. Top panels: 2016 cruise. Bottom panels: 2022 cruise. (**a**, **d**) Correlation circle between the variables and the first two dimensions of the MFA. The variables are separated in three groups : the environmental group (brown arrows), the acoustic group (yellow arrows) and the supplementary group (black arrows). The environmental group comprises 4 variables : temperature (T), salinity (S), oxygen (O) and fluorescence (F). The acoustic group comprises the second to the seventh principal components (PC2-7) of the acoustic mfPCA. In **a**, PC6 and 7 arrows are not indicated for more readability (arrows are close to (0,0)). The supplementary group gathers the latitude (La), longitude (Lo) and bathymetry (B). (**b**, **e**) First factorial map of the MFA with colors corresponding to the results of the MBC in three group. (**c**, **f**) Projection of the three clusters in the geographical space coloured with respect to the group. Bathymetry line at 500 and 2500 m surround the Kerguelen archipelago and Saint-Paul (SP) and Amsterdam (A) Islands. The blue circle define the Economic Exclusive Zone (EEZ).

352   The first factorial plan of the MFA (Figure 7b and e) projects closely similar observations. MFA scores were
353 clustered using a model-based clustering approach to identify regions with matching environmental conditions and
354 acoustic patterns. The optimal number of groups, denoted as G, was determined by examining the Integrated Completed
355 Likelihood (ICL) and Bayesian Information Criterion (BIC). With both criteria, the approximated integrated likelihood
356 gain when increasing the number of groups dropped after $G = 3$ in both years. To ensure a relevant and ecologically
357 interpretable model (without excessively dividing the data), we then decided to fix G to 3. The partition of MFA scores
358 in 3 groups was found to be highly stable (insensitive to initialisation) with a probability of belonging to a group always
359 above 0.5, supporting the choice of G.
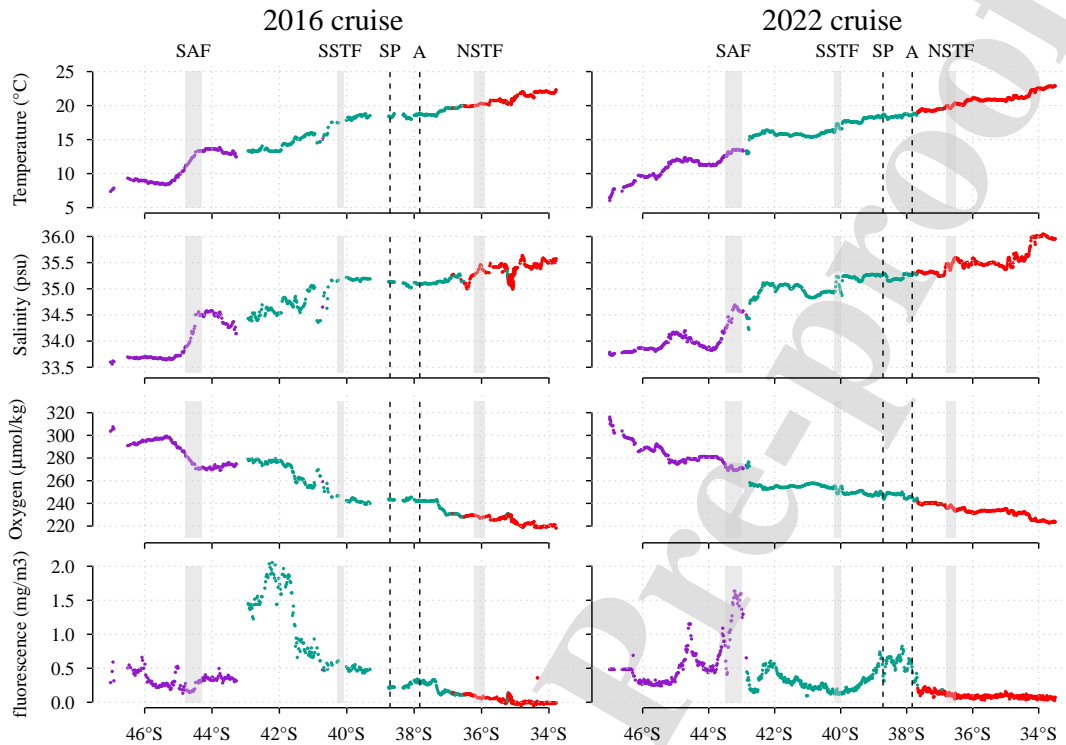
Figure 8: Environmental variables measured along the ship cruises as function of latitude. The colors correspond to the three regions defined with the model-based clustering. From top to bottom: sea surface temperature (°C), sea surface salinity (psu), surface oxygen (μmol kg^{-1}) and fluorescence (mg m^{-3}) for the cruises 2016 (left column) and 2022 (right column). From South to North, the grey shades approximately delimit the Subantarctic Front (SAF), the Southern Subtropical Front (SSTF) and the Northern Subtropical Front (NSTF). The fronts were determined based on hydrographic parameters compiled in the Southwest Indian Ocean (Anilkumar et al., 2007). The two black dashed vertical lines correspond to the latitude of Saint-Paul (SP) and Amsterdam (A) islands.

This clustering outcome exhibits geographical consistency across the two years, with the purple group located in the southern study area, the red group in the north, and the green group serving as a transitional zone (Figure 7c and f).

3.4. Environmental-acoustic regions

All four environmental parameters (temperature, salinity, oxygen and fluorescence) were highly variable along the transects and allowed the approximate location of fronts in the study area (Figure 8). In (Anilkumar et al., 2007), the northern branch of Subantarctic Front (SAF1) falls within 9-11 °C and 33.85-34 psu at surface, while the Southern Subtropical Front (SSTF) ranges from 11-17 °C and 34.05-35.35 psu, with its northern branch (NSTF) at 21-22 °C and a constant salinity of 35.5 psu. Here, we associated the SAF with a surface salinity shift located at ~44.55° S in 2016 and ~43.25° S in 2022. Salinity values increased from ~33.7-33.9 psu (south of the front) to ~34.5 psu (north of the front) in both years with increased temperature associated. The SSTF was consistently located around ~40.2° S in both cruises. SSTF exhibited mean temperatures of 17.7 °C and 16.4 °C in 2016 and 2022, coupled with mean salinities of 35.185 and 34.945 psu. The Northern SubTropical Front (NSTF) was positioned at 36.05° S in 2016 and
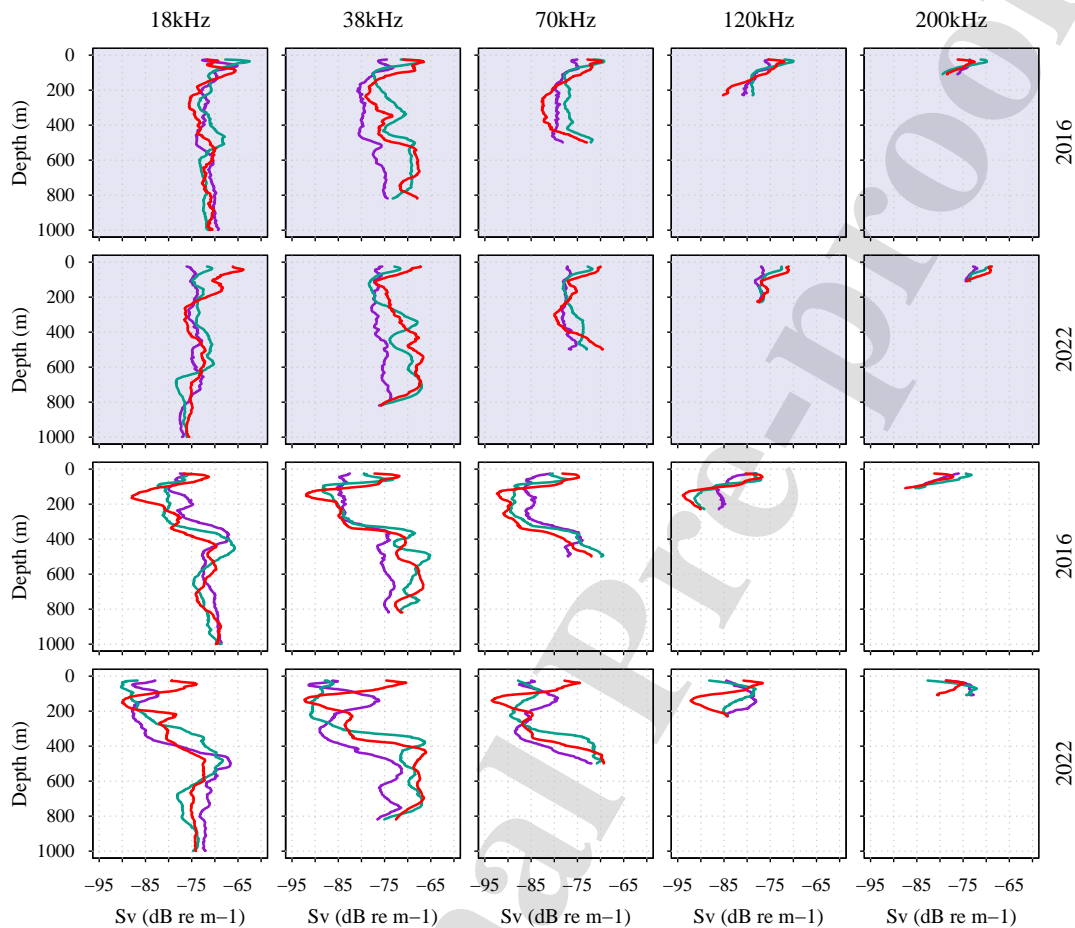
Figure 9: Median acoustic profiles computed in the linear domain for each region obtained with the model-based clustering. The median profile is calculated for each frequency (18, 38, 70, 120 and 200 kHz from left to right panels) depending on the night (blue shade) and day periods (white shade) and considering independently each year. The first (resp. second) and third (resp. fourth) lines of panels correspond to results for 2016 (resp. 2022). Purple profiles correspond to the southernmost region (PFZ), red profiles to the northernmost region (STZ) and green profiles to the region in between (SAZ) (Figure 7 right panels).

373 36.65° S in 2022 corresponding to a temperature > 20.15 °C and 20.05 °C and distinguishing areas with salinity >
374 35.3 psu in 2016 and 35.51 psu in 2022.

375   Following Geisen et al. (2022), we refer to the southern group (purple) as the Polar Front Zone (PFZ). The
376 green cluster located to the north of the SAF and to the south of the NSTF has been termed the SubAntarctic
377 Zone (SAZ), while the red cluster situated to the north of the NSTF is identified as the Subtropical Zone (STZ).
378 All four environmental variables (temperature, salinity, oxygen and fluorescence) were significantly different between
379 the different regions (Wilcoxon pairwise tests, p-value < 0.001). The PFZ (purple) shows colder, fresher and more
380 oxygenated surface waters. At the opposite, the STZ (red) corresponds to subtropical environmental conditions
381 characterised by hotter, saltier, less oxygenated surface waters with a weak fluorescence signal. The SAZ (green)

corresponds to a transition zone between two oceanographic systems. The fluorescence is significantly higher in the PFZ and SAZ than in the STZ, with a large peak around 42° S in 2016 belonging to the green region, and two peaks around 44.5 and 43° S in 2022 belonging to the purple region. The peak observed in 2022 matches the SAF location but not in 2016. A third fluorescence peak is observed in 2022 around Saint-Paul and Amsterdam islands.

We calculated the median acoustic profiles for both daytime and nighttime based on the environmental-acoustic region for both cruises (Figure 9). Pairwise fANOVA tests were used to confirm significant profile differences between regions according to the frequency and period. All pairwise tests showed significance (p-value < 0.001). Distinct differences between the median profiles are evident in the day-night comparison. Nighttime profiles present backscatter more uniformly distributed along the water column, without densities below -85 dB, whereas daytime profiles display lower average densities in the upper 400 m (matching PC1, Figure 5).

For nighttime profiles (blue background Figure 9), the PFZ (purple) presents less vertical variations of densities in the water column than the two other groups. The SAZ (green) is distinguished by a deep scattering layer (DSL) (400-600 m) on the 18 kHz and three scattering layers (SL) on the 38 kHz, one at the surface (above 100 m), also seen on the higher frequencies, one at ~350 m and a DSL between 500-800 m. The STZ (red) is characterised by an intense surface scattering layer (SSL) observed for all frequencies, low intensities in the intermediate layer (IL) and dense multi-layered DSL observed on the 38 kHz.

For daily profiles (white background Figure 9), the DSL observed on the 18 kHz around 400 m (2016) or 500 m (2022) is weaker in the STZ (red) while the large DSL observed on the 38 kHz (400-800 m) is weaker in the PFZ (purple). An intense SSL (0-100 m) on the 18, 38, 70 and 120 kHz characterises the STZ (red) in 2022 in comparison with the other regions. In 2016, this intense SSL is also observed in the SAZ (green) on the 38 and 70 kHz and in the SAZ and the PFZ on the 18 kHz. Finally, the PFZ (purple) is defined by higher densities in the IL (100-300 m) in 2016, observed as a peak around 100 m in 2022, while the STZ is particularly low at that layer.

Overall, we observe from the PFZ to the STZ (i) an increase of densities in the SSL (0-100 m), (ii) a decrease in densities in the IL during the daytime (100-300 m) and (iii) the apparition of an dense DSL on the 38 kHz. Even though variability can be observed, this main pattern is recurrent between years.

The median S_v profiles with interquartile ranges for each environmental-acoustic region are displayed Figure S4. Smoothed echograms for 2016 and 2022 are displayed in Figure S5 with associated environmental-acoustic regions.

3.5. Integrated NASC and taxons in the regions

We calculated the integrated NASC across all frequencies, from 25 m depth down to their maximum range. We focus on lower frequencies that penetrate greater into the mesopelagic zone, as the frequencies with a sampling depth ≤ 500 m (70, 120, and 200 kHz) were the first three contributors to the observed circadian cycle effect (Figure 4).

At nighttime in 2022, both 18 and 38 kHz show a significant increase in NASC values (Wilcoxon pariwise tests, p-value < 0.001) from the PFZ (purple) to the STZ (red) (Figure 10). In 2016, the NASC for the 18 kHz is significantly higher in the SAZ (green) but no difference is found between the PFZ and the STZ. At 38 kHz in 2016, the NASC is significantly lower in the PFZ but no difference is found between the two other regions. During the daytime, the NASC values at 38 kHz are approximately three times higher in the SAZ and STZ in comparison to the PFZ. The 18 kHz depicts an opposite pattern with the highest values observed in the southern region, gradually decreasing towards the north.

Concerning the biological composition in the area, net samples were spatially associated with the three environmental-acoustic regions. The abundance of Gonostomatidae and salps exhibit a distinct pattern, showcasing lower Gonostomatidae abundance and higher salps abundance within the Polar Front Zone (purple) in comparison to the northern regions (Figure 11). However none of the differences observed are significant (Gonostomatidae: ANOVA,

2016 Cruise 2022 Cruise
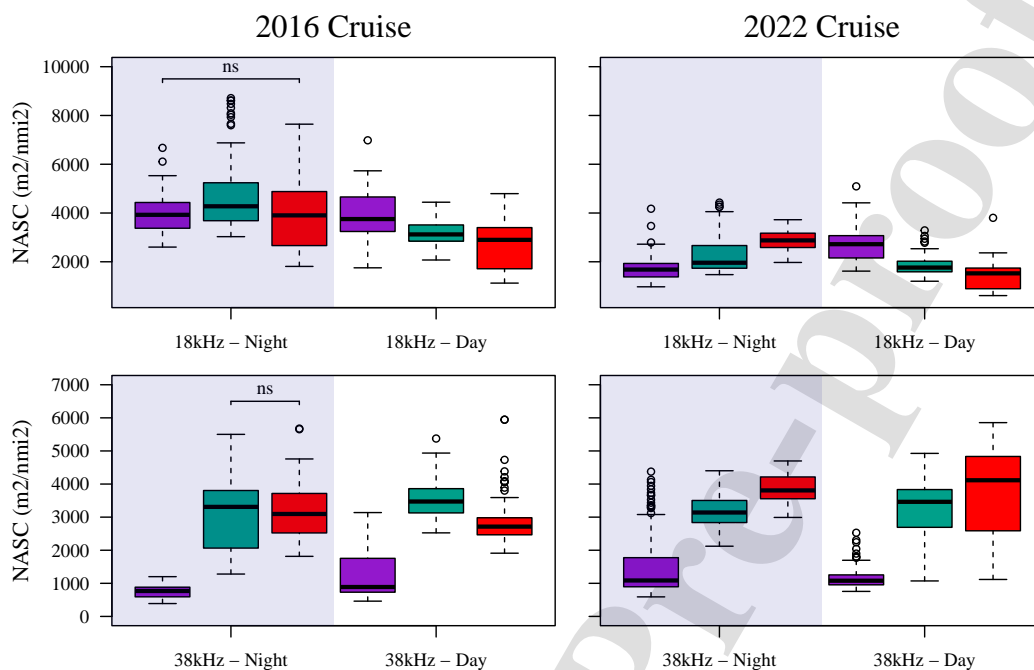
Figure 10: Boxplot of integrated NASC (m^2 nmi^{-2}) computed for each environmental-acoustic region defined with model-based clustering. The NASC is calculated for the 18 (top panels) and 38 kHz (bottom panels) over profiles sampled at night (blue shade) or at day (white shade). The y-axis is adjusted for each frequency. The left panels correspond to the 2016 cruise while the right panels correspond to the 2022 cruise. For a given year, frequency and period (night or day), all pairwise integrated NASC are significantly different (Wilcoxon test, p-value < 0.01) unless when indicated ns above two boxplots. The number of observations n per boxplot $\in$ [43,206].

$F^{2,13}$ = 3.487, p-value = 0.0613; Kruskal-Wallis, Chi-squared = 3.93, df = 2, p-value = 0.14). For euphausiids, siphonophores, and myctophids, no significant differences are observed.

We identified 51 species of myctophids, 32 species of euphausiids, 35 species of siphonophores and 11 species of salps from the net samples. The NMDS indicates a correct value of stress (0.11 for euphausiids and myctophids, 0.12 for siphonophores and 0.15 for salps) which means a good representation of the data (Figure 11). The NMDS projection shows a spatial distinction for the four groups (myctophids, salps, euphausiids and siphonophores) based on the environmental-acoustic regions. The ANOSIM carried out from stations separated by the environmental-acoustic regions shows significant differences between groups for myctophids, euphausiids and siphonophores (p-value = 0.001). However, only a significant difference between the Polar Front Zone (PFZ) and the SubAntarctic Zone (SAZ) is observed for salps.

The PFZ (purple) is dominated by Southern Ocean species, such as *Euphausia vallentini* and *Euphausia triacantha* for euphausiids, *Rosacea plicata* for siphonophores and *Salpa thompsoni* for salps. The dominant myctophid species were *Krefftichthys anderssoni*, *Protomyctophum bolini*, *Protomyctophum tenisoni*. Within the PFZ, a latitudinal shift was observed with clear endemic southern ocean species in the south, and more cosmopolitan in the northern stations. The SubAntarctic Zone is marked by the dominance of cosmopolitan species present throughout the transect, and the absence of Antarctica species that were identified in the PFZ: *Euphausia spinifera* and *Nematoscelis megalops* for

440 euphausiids, *Chelophyes appendiculata*, *Eudoxoides mitra* for siphonophores. The diversity of myctophids is higher in

441 this region compared to the PFZ with the presence of *Lampanyctus australis* and *Hygophum hanseni*. No dominance

442 of salp species was identified. Finally, several subtropical species such as the euphausiids *Thysanopoda aequalis*,

443 *Stylocheiron abbreviatum* or *Nematobrachion flexipes* characterise the SubTropical Zone. The siphonophore species

444 *Agalma okenii* and *Abyla trigona* are also dominant. *Ceratoscopelus warmingii* and *Lampanyctus pusillus* are the

445 dominant myctophid species. Similarly to the SAZ, no dominance of salp species was identified.
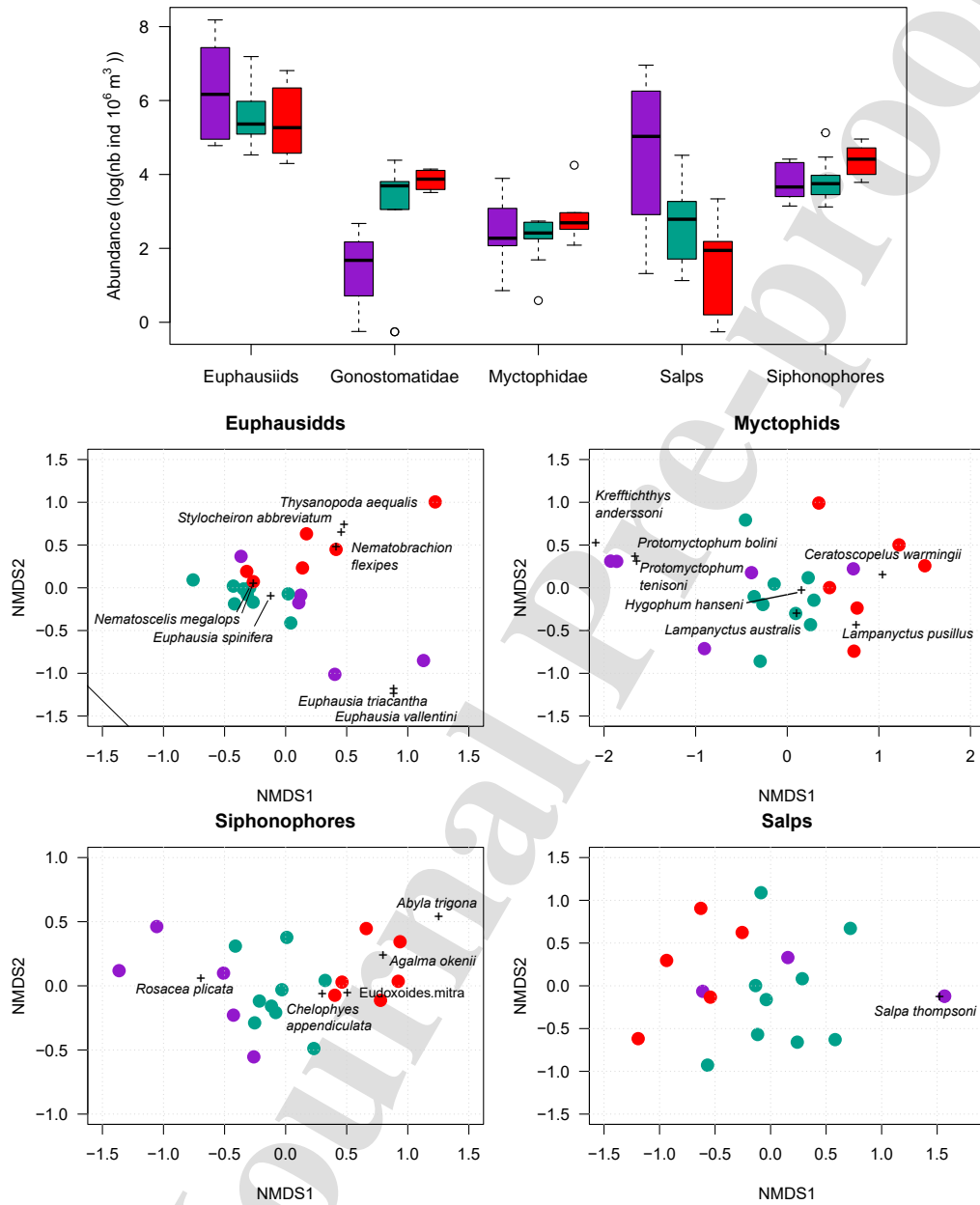
Figure 11: Biological communities and association with the environmental-acoustic regions (colors). Top panel : boxplots of organisms abundance ($\log(\text{nb ind } 10^6 \text{ m}^3)$) grouped in five taxa. For each taxon, no significant differences were found between regions (ANOVA or Kruskal-Wallis tests, p-value > 0.05). Bottom panels : ordination plot using Non-Metric Multidimensional Scaling (NMDS) of IKMT euphausiids, Myctophidae, siphonophores and salps captured in the area. Gonostomatidae could not be analysed unsing NMDS as no species-level identification could be performed. Each dot represent a IKMT station (n total = 20). Positive correlated objects are close together on the same side of the plot. Black crosses display the dominant species per region.

4. Discussion

We report here on the application of Functional Data Analysis methods applied to five acoustic frequencies sharing distinct depth ranges, acquired from a downward-facing echosounder. Our investigation delves into the inherent temporal variability of acoustic data, aiming to filter it to explore the spatial sound-scattering regimes evolution within a known transition zone. The analysis of the vertical distribution of sound-scattering layers can be achieved by vertically segregating the data prior to analysis (e.g. Béhagle et al., 2014). However, the mesopelagic vertical boundaries exhibit both spatial and temporal variability (Reygondeau et al., 2018), influenced by the specific oceanic province under consideration (Fuchs et al., 2022). This complexity in vertically delimiting oceanographic basins suggests that, instead of enforcing vertical separation prior to analysis, a more promising strategy is to adopt a comprehensive whole-water column approach (Sutton, 2013).

We believe the functional data analysis method employed here, which identifies the depth maximising vertical variability (Nerini et al., 2010), allows us to contribute to this approach. Furthermore, the originality of this work is the application of the fPCA to joint acoustic frequencies sharing distinct depths, complementing a recent similar approach (Ariza et al., 2022a). Such tools can be seen as valuable as it is becoming clearer that multi-frequency hydroacoustic analyses are needed to better understand the proportions of distinct groups composing sound scattering layers (e.g. siphonophores, crustaceans and fish), which can in turn help reduce the uncertainty in mesopelagic organisms biomass estimates from acoustic data (Proud et al., 2019). Indeed, part of hydroacoustics research on pelagic vertical structures has focused on the 38 kHz because it reaches deep depths and it corresponds to the frequency of air-filled organisms resonance (e.g. Béhagle et al., 2016; Klevjer et al., 2016; Receveur et al., 2020b). It follows that research has largely focused on the 38 kHz to model micronekton spatial dynamics (e.g. Lehodey et al., 2015; Ariza et al., 2022b) and estimate the global fish biomass from NASC integrated values (Irigoien et al., 2014; Dornan et al., 2022). Yet, a single-frequency approach might encounter constraints in capturing the wide spectrum of community composition variations (Dornan et al., 2019).

Decomposing and quantifying spatiotemporal variability in acoustic data

As emphasised in Boersch-Supan et al. (2017), "there is an imperative [...] in using analysis frameworks that adequately address the inherent space-time variability". We show here that the primary mode of variability is tied to temporal dynamics even across a pronounced hydrological transition zone where substantial spatial variability was anticipated. This diel dominance aligns with the well-recognized diel vertical migration (DVM) pattern, a major mode of variability in acoustic data (e.g. Receveur et al., 2020a), and corroborates the substantial impact of daylight on backscatter values in the southern Indian Ocean (Boersch-Supan et al., 2017). This outcome is consistent with the long-known response of marine organisms to light levels (e.g. Cotté and Simard, 2005; Brierley, 2014) and the observation of organisms adjusting their depth to remain within preferred irradiance ranges (Frank and Widder, 2002). Our results indicate a slight increase in the scores of the temporal mode during the day for both the 2016 and 2022 cruises, implying spatial implications within this diel pattern. This intricacy can become even more complex with longer timescales like the seasonal cycle (Urmy et al., 2012). Recent studies have introduced effective data-driven techniques to capture inherent temporal variability in acoustic datasets (Parra et al., 2019; Lee and Staneva, 2020). In this study, we enhance the interpretation of our multi-frequency database by considering how different frequencies respond to temporal changes in both the mesopelagic zone (18-38 kHz) and the epipelagic zone (across the 5 frequencies). Our findings indicate that the distribution of each acoustic observation can be understood as a linear combination of specific sound-scattering features (e.g. VM2, VM3), building on a main cyclic temporal mode of variability (VM1). This advancement could contribute to optimising parameters for marine ecosystem models (e.g. Lehodey et al., 2015).

487 The mfPCA approach also proves valuable in analysing the frequency contributions to individual vertical modes. As
488 illustrated in Figure 3, a high correlation exists within and between high frequencies (120-200 kHz). High frequencies
489 substantial influence on the temporal mode aligns with their limited sampling range, primarily capturing the epipelagic
490 dynamics, where organism migration significantly impacts abundance (e.g. Béhagle et al., 2017). The lower frequencies
491 (18 and 38 kHz) are also influenced by temporal variability, albeit less due to the possible persistent deep scattering
492 layers (DSL) between nighttime and daytime, associated with non-migratory organisms (particularly at the 38 kHz
493 frequency). Additionally, Figure 3 revealed no correlation at depth among the deep-ranging frequencies (18 and 38
494 kHz) probing the twilight zone. This absence of correlation results in an uneven contribution of the frequencies to each
495 VM. For instance, the change in the deep sound-scattering layers highlighted by VM3 in 2016 is majorly driven by
496 the 38 kHz, and could be associated with a change of community detected particularly at this frequency (e.g. a change
497 in mesopelagic fish community, Dornan et al., 2019). On the contrary, VM2 (2016) was driven by a combination of
498 frequencies, which suggest that the whole pelagic community structure varied with this pattern.

Numerous spatiotemporal patterns

500 Within the branch of Functional Data Analysis methods, Functional Principal Component Analysis (fPCA) has
501 largely been applied to oceanographic research as it allows the study of 3D structures (space and depth or time). In
502 analysing 3D hydrological patterns and front location, the initial factorial plan captured 92, 94, 87, and 79 % of inertia,
503 respectively in Pauthenet et al. (2017, 2018, 2019, 2021). Similarly, 99 % of variance was explained by the first three
504 vertical modes of variability in the Brazil-Malvinas confluence (Fonvieille et al., 2023). In Assunção et al. (2020) the
505 first axis alone of the fPCA explained 88 and 83 % of temperature and salinity variation. While studying Southern
506 Elephant seal dive patterns (Godard et al., 2020), three principal components were needed to encompass 78 % of total
507 variability. In our recent study, the first seven VMs were necessary to capture around 80 % of total variance in 2016
508 and 2022 cruises, which match recent fPCA applied to acoustic data (Ariza et al., 2022a). Comparing the first modes
509 of variability of a mfPCA applied to either biological or abiotic data (e.g. hydrological profiles) reveals that biological
510 patterns need more VMs than abiotic patterns to capture their complexity.

Sounds-scattering regimes are linked to surface environmental drivers

512 Grouping the result of the MFA with model based clustering provided evidence of 3 distinct environmental-acoustic
513 groups in the transition zone, with the same spatial distribution observed in both years. The approximate boundaries
514 of these regions fall near the Subantarctic Front (SAF) and the Northern branch of the Subtropical Front (NSTF),
515 delimiting from South to North the Polar Frontal Zone (PFZ), the SubAntarctic Zone (SAZ) and the SubTropical Zone
516 (STZ).

517 The 38 kHz dynamics from the PFZ to the STZ, matches previous large-scale studies conducted in the Southern
518 Indian Ocean (Béhagle et al., 2016; Boersch-Supan et al., 2017) and a recent shape-based global classification of
519 acoustic profiles (Ariza et al., 2022b). Notably, from South to North, the median profiles in each cluster showed an
520 increase of densities at the surface, denser DSL and a decrease of densities in the intermediate scattering layer. The
521 two-layer persistent DSL structure North of the SAF measured at the 38 kHz during both cruises matches with the
522 previous vertical distribution of SL found in the Indian Ocean (Béhagle et al., 2014; Klevjer et al., 2016). However,
523 a strong SL observed in daytime 2022 between ∼100-200 m depth could nuance the view of more evenly distributed
524 acoustic profiles South of the SAF (Boersch-Supan et al., 2017).

525 Although literature on the 18 kHz frequency in the region is limited, Boersch-Supan et al. (2017) also found an
526 increased 18 kHz backscatter between 35° and 45° S, spanning the surface and intermediate layers during the day and
527 throughout the water column at night. Although a decrease in densities within the DSL at 38 kHz was evident in the

southern group, there was no corresponding reduction in backscatter observed at 18 kHz. The difference in vertical structures between 18 and 38 kHz matches with recent observation at stations displaying contrasted biogeochemical conditions within the Southern Ocean (Cotté et al., 2022). These two distinct responses to the SAF could potentially indicate that the two different oceanographic domains are dominated acoustically by different frequencies. However, this assumption requires further investigation.

Detecting spatial patterns at higher frequencies poses a greater challenge due to the predominant sampling of the epipelagic layers, leading to the most notable contrast emerging between daytime and nighttime observations. However, during daytime in both years, between depths of 100 and 200 metres, a denser sound-scattering layer was detected on the 120 kHz in the PFZ in comparison to the STZ. This phenomenon might be attributed to the greater presence of dense crustacean communities in the Southern Ocean as indicated by previous studies (Demer and Conti, 2005), which were found to be particularly abundant in the 0–500 m water layer in waters surrounding Kerguelen (Cotté et al., 2022).

Sound-scattering layers and fluorescence

In comparison to the SAF, we observed that the Southern Subtropical Front (SSTF) and NSTF exhibited a weaker hydrological gradient which did not lead to an abrupt transition in the deep scattering layers between the two northern zones. However, the lowest fluorescence measured in the STZ delimits the entry into the oligotrophic regime of the Southern Indian Ocean. This frontal system is known to impact phytoplanktonic communities (Schlüter et al., 2011; Geisen et al., 2022).

Studies have found positive relationships between mesopelagic fish biomass and primary production (PP) (Irigoien et al., 2014) and between zooplankton biomass in the epi-, meso-, and bathypelagic layers and average net primary production (NPP) (Hernández-León et al., 2020). Here the appearance of the DSL to the north of the SAF could potentially correspond to an increased presence of deep micronekton biomass across the Subtropical Convergence Zone (defined as the region between 42° and 32° S) in the Southern Indian Ocean, as suggested by previous studies (Pakhomov et al., 1994; Boersch-Supan et al., 2017). However, we found that the DSL was not constrained to the SAZ but extended northward within the oligotrophic STZ. Additionally, denser surface scattering layers were found in the the same zone, where the lowest surface fluorescence was measured. A similar pattern was observed at 18 and 38 kHz in the Northeast Atlantic Ocean, with surface scattering layers being denser when surface fluorescence was the lowest, while no impact of the transition was observed on the DSL (García-Seoane et al., 2023).

Future research should focus on exploring the relationship between mid-trophic levels and low trophic levels, particularly investigating the connection between scattering layers and the intricate phytoplankton communities across this transition zone (Geisen et al., 2022).

Integrated acoustic biomass

The change in profile features can also reflect a change in integrated acoustic biomass. For higher frequencies, which sample the epipelagic layer only, the integrated NASC can be greatly impacted by the circadian cycle as the migrating organisms from the mesopelagic zone strongly increase the epipelagic densities. When the acoustic information fully integrates the mesopelagic scattering layers (∼200-1000 m), and for similar environmental conditions, the effect of the circadian cycle on the integrated NASC can be limited as the backscatter values can be rather reorganised in the water column; although organisms migrating from the depth > 1000 m was reported (Sutton, 2013).

Spatially, for the 38 kHz total NASC, values increased from subantarctic to subtropical conditions where the highest densities were previously observed, especially in surface waters, at a larger scale (Béhagle et al., 2016). Furthermore, the decline of acoustic backscatter DSL recorded at 38 kHz towards the southern ocean is in adequation with previous observations (e.g. Escobar-Flores et al., 2020; Dornan et al., 2022), and was recently observed across high latitude

569 fronts (with a drop of mean S_v at 38 kHz; Chawarski et al., 2022), showing evidence of a response of the mesopelagic
570 community to a temperature drop. Here, a significant inverse pattern in integrated biomass was seen between 18 and
571 38 kHz during the daytime, comforting that the drop of acoustic may be due to a change of community and not a drop
572 of biomass (Dornan et al., 2019).

573 **Abundance and diversity of marine organisms in environmental-acoustic regions**

574 Myctophids establish their dominance as the primary fish family in the mesopelagic and bathypelagic regions
575 of the Southern Ocean, leading in species diversity, abundance, and biomass. While we observed no variations in
576 myctophid abundance, their groupings effectively distinguish species within the Polar Front Zone from those in the
577 SubAntarctic and SubTropical Zones. This observation aligns with the distinct shift in fish communities documented in
578 the subantarctic zone (Koubbi et al., 2011) and the SAF' role as a significant biogeographic boundary for mesopelagic
579 fish larvae in the Indian sector of the Southern Ocean (Koubbi, 1993). Given the significant impact of organism
580 composition and physiology on backscatter resonance, we investigated if the observed shift in fish assemblages could
581 provide an explanation for the abrupt change in NASC values and alterations in vertical structure, particularly the
582 DSL drop at 38 kHz frequency towards the Southern Ocean. Indeed, swim bladders being one of the most resonant
583 organs of a fish, the absence of a gas-bearing swimbladder can lead to a near-no detection of certain species at low
584 frequencies (Foote, 1980). This proposition gains further support from the fact that the most abundant mesopelagic
585 fish species in the Southern Ocean lack gas-filled swimbladders or possess swimbladders filled with lipids (Dornan
586 et al., 2019). Notably, certain species found in the subtropical region, such as *Lobianchia dofleini*, do possess gas-filled
587 swimbladders (Kleckner and Gibbs, 1972), whereas species like *Gymnoscopelus fraseri*, identified in the SAZ are
588 known to lack swimbladders (Dornan et al., 2019). However, this hypothesis lacks support due to our investigation in the
589 PFZ which documented the presence of myctophid species with varying gas swimbladder status. For example, species
590 such as *Protomyctophum bolini*, *Electrona carlsbergi*, and *Krefftichthys anderssoni* possess gas swimbladders, whereas
591 *Electrona antarctica* and *Gymnoscopelus braueri* do not or exhibit an ontogenetic loss of gas swimbladder (Dornan
592 et al., 2019). The relation between acoustic backscatter levels and organisms' distribution may not be so evident as the
593 presence, type or size of swimbladders change among and within taxonomic groups, and potentially within genus and
594 species, depending on body length or developmental stage (Marshall, 1960; Dornan et al., 2019). Nonetheless, from a
595 biological perspective, the region identified as the PFZ might not be representative of the strict Southern Ocean fauna,
596 as species such as *Electrona antarctica* are commonly located within the vicinity of the Polar Front or to its south.
597 The PFZ region is likely associated with the northern margin of the Southern Ocean. The Gonostomatidae is the other
598 main fish family reported in this study. While Gonostomatidae can be found in epipelagic layers in the subtropical at
599 night, they were reported to be the most dominant species dwelling in the deep scattering layer (Marohn et al., 2021)
600 and to remain at these depths both during day and night time (McClain et al., 2001; Olivar et al., 2017). Furthermore,
601 they were found to be the main contributors to the 38 kHz DSL (between 400-600 m (Peña et al., 2014). Despite
602 the lack of a significant change in abundance, the scarce catches in the Southern Ocean align with a small number
603 of Gonostomatidae species extending as far south as the Southern Ocean (Broyer and Koubbi, 2014). Upon further
604 investigation, the persistent DSL we observed between night and day periods in the Southern Indian Ocean could be
605 imputed to these species.

606 Additionally, the southward dissolution of the 38 kHz does not imply a loss of biomass as acoustic-based estimation
607 of mesopelagic fish biomass greatly depends on the fish morphology (Proud et al., 2019) and acoustic signals are rather
608 sensitive to the scatterer's distribution than biomass distribution (Davison et al., 2015). As stated by : "the size structure
609 of mesopelagic fish is skewed with abundance driven by the smallest and biomass driven by the largest fishes" (Marohn
610 et al., 2021). In our study, the biomass is probably higher in the PFZ given the sizes of the individuals of each family

(myctophids are bigger than *Cyclothone* species which are the dominant Gonostomatidae in the area; Koubbi and Djian, pers. comm.).

The changes observed in the surface layer from the PFZ to the STZ could be explained by differences in assemblage for all taxa (except for salps). We encountered a large number of *Agalma okenii* in the STZ, a physonect siphonophore known to have relatively high acoustical target strengths (TS) at 24 and 120 kHz, caused by a gas inclusion in their pneumatophore (Warren, 2001). *Agalma okenii* vertical distribution extends from the surface down to 200 m depth (Boltovskoy, 1999) which would match the intense scattering layer observed at this depth ($\sim$100 m). The high resonance of this layer at 38 kHz would match previous observations (Klevjer et al., 2016).

Limitations <mark>and future perspectives</mark>

Globally, while the species assemblages have direct repercussions on acoustic backscatter, the primary factor of uncertainty in our study is the sound-scattering layers composition. The presence of highly resonant organisms, such as *Salpa thompsoni*, found in the Polar Frontal Zone (PFZ) and generally found in large aggregations, can be mistaken for krill or other zooplankton (Wiebe et al., 2010). Furthermore, the response of other fluid-like organisms such as euphausiids presents less difference in resonance between species. For instance, *E. vallentini* is found here to be a representative species of the PFZ and was found at $\sim$60 m in Kerguelen water (Cotté et al., 2022; Béhagle et al., 2017) but we did not identify a matching surface scattering layer in this zone. Moreover, the association between distinct years of net sampling and acoustic data introduce bias into the interpretation. Further simultaneous acoustic/sampling associations are required to clarify the response of MM to this transition zone.

The statistical method introduced in this study has a notable constraint: it demands a substantial number of principal components to effectively capture the intricate spatiotemporal interconnections and overall variability. To address this, the incorporation of additional spatiotemporal data could potentially reduce the number of required modes. Conducting additional acoustic surveys, especially zonal transects, could also help limit the overinterpretation of potential fine-scale structures (<10 km, e.g. fronts or filament) on the observed patterns as it was shown to have effect of increased fish concentration (Baudena et al., 2021), as well as mesoscale activity (Godø et al., 2012; Della Penna et al., 2022). Moreover, there is a need for further investigation into the combined impact of the transition from polar to subtropical waters and the presence of Saint-Paul and Amsterdam islands. The latter have been observed to produce internal waves, which could enhance the transfer of nutrients from deep waters to the surface (de Lavergne et al., 2020; Sergi, 2021). These processes have the potential to significantly influence the marine environment and habitats for marine organisms and predators (Bertrand et al., 2014). Finally, a more comprehensive understanding of depth-related dynamics could be attained by integrating environmental profiles into the analysis. The mfPCA approach has effectively untangled factors that contribute to vertical data variance. Further advancements employing machine learning techniques, particularly deep learning, could potentially uncover even more intricate dependencies (Bianco et al., 2019). This advancement holds the potential to provide deeper insights into the intricate spatio-temporal intricacies within acoustic datasets.

<mark>Finally, the work presented here was initiated under the scope of the extension of the protection status of Saint-Paul and Amsterdam islands to their Economic Exclusive Zone pelagic waters</mark> (Pagniez et al., 2021). <mark>Management plan includes a monitoring of ecosystems and the acoustic data has proven to be a valuable source of information in such remote regions for integrating knowledge on the distribution of the mid-trophic levels and contributing to the conservation and advisory objectives of policymakers. Moreover, understanding the dynamics of mid-trophic levels using echoregions from quasi-pristine areas can provide baseline information that can serve as a reference to detect the emerging impact of climate change on pelagic fauna.</mark>

Declaration of competing interest

None.

Acknowledgments

References

Aksnes, D. L., Røstad, A., Kaartvedt, S., Martinez, U., Duarte, C. M., and Irigoien, X. (2017). Light penetration structures the deep acoustic scattering layers in the global ocean. *Science Advances*, 3(5):e1602468.

Anilkumar, N., Pednekar, S. M., and Sudhakar, M. (2007). Influence of ridges on hydrographic parameters in the Southwest Indian Ocean. *Marine Geophysical Researches*, 28(3):191–199.

Annasawmy, P., Ternon, J., Marsac, F., Cherel, Y., Béhagle, N., Roudaut, G., Lebourges-Dhaussy, A., Demarcq, H., Moloney, C., Jaquemet, S., and Ménard, F. (2018). Micronekton diel migration, community composition and trophic position within two biogeochemical provinces of the South West Indian Ocean: Insight from acoustics and stable isotopes. *Deep Sea Research Part I: Oceanographic Research Papers*, 138:85–97.

Ariza, A., Lebourges-Dhaussy, A., Nerini, D., Pauthenet, E., Roudaut, G., Assunção, R., Tosetto, E., and Bertrand, A. (2022a). Acoustic seascape partitioning through functional data analysis. *Journal of Biogeography*, page jbi.14534.

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

Assunção, R., Lebourges-Dhaussy, A., Da Silva, A. C., Roudaut, G., Ariza, A., Eduardo, L. N., Queiroz, S., and Bertrand, A. (2023). Fine-scale vertical relationships between environmental conditions and sound scattering layers in the Southwestern Tropical Atlantic. *PLOS ONE*, 18(8):e0284953.

Assunção, R. V., Silva, A. C., Roy, A., Bourlès, B., Silva, C. H. S., Ternon, J.-F., Araujo, M., and 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.

Baker, A. C., Boden, B. P., and Brinton, E. (1990). A Practical Guide to the Euphausiids of the World. *Natural History Museum Publications*, page 96 pp.

Baudena, A., Ser-Giacomi, E., D'Onofrio, D., Capet, X., Cotté, C., Cherel, Y., and D'Ovidio, F. (2021). Fine-scale structures as spots of increased fish concentration in the open ocean. *Scientific Reports*, 11(1):15805.

Benoit-Bird, K. J. and Lawson, G. L. (2016). Ecological Insights from Pelagic Habitats Acquired Using Active Acoustic Techniques. *Annual Review of Marine Science*, 8(1):463–490.

Benoit-Bird, K. J. and McManus, M. A. (2012). Bottom-up regulation of a pelagic community through spatial aggregations. *Biology Letters*, 8(5):813–816.

Bertrand, A., Bard, F.-X., and Josse, E. (2002). Tuna food habits related to the micronekton distribution in French Polynesia. *Marine Biology*, 140(5):1023–1037.

Bertrand, A., Grados, D., Colas, F., Bertrand, S., Capet, X., Chaigneau, A., Vargas, G., Mousseigne, A., and Fablet, R. (2014). Broad impacts of fine-scale dynamics on seascape structure from zooplankton to seabirds. *Nature Communications*, 5(1):5239.

Bianco, M. J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M. A., Gannot, S., and Deledalle, C.-A. (2019). Machine learning in acoustics: Theory and applications. *The Journal of the Acoustical Society of America*, 146(5):3590–3628.

Biernacki, C., Celeux, G., and Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood.

695 Boersch-Supan, P. H., Rogers, A. D., and Brierley, A. S. (2017). The distribution of pelagic sound scattering layers across the southwest Indian
696   Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 136:108–121.

697 Boltovskoy, D. (1999). South atlantic zooplankton. page No. C/592 S6.

698 Bouveyron, C., Celeux, G., Murphy, T. B., and Raftery, A. E. (2019). *Model-based clustering and classification for data science: with applications*
699   *in R*, volume 50. Cambridge University Press.

700 Brierley, A. S. (2014). Diel vertical migration. *Current Biology*, 24(22):R1074–R1076.

701 Broyer, C. d. and Koubbi, P. (2014). Biogeographic atlas of the Southern Ocean. Num Pages: 1.

702 Béhagle, N., Cotté, C., Lebourges-Dhaussy, A., Roudaut, G., Duhamel, G., Brehmer, P., Josse, E., and Cherel, Y. (2017). Acoustic distribution
703   of discriminated micronektonic organisms from a bi-frequency processing: The case study of eastern Kerguelen oceanic waters. *Progress in*
704   *Oceanography*, 156:276–289.

705 Béhagle, N., Cotté, C., Ryan, T. E., Gauthier, O., Roudaut, G., Brehmer, P., Josse, E., and Cherel, Y. (2016). Acoustic micronektonic distribution
706   is structured by macroscale oceanographic processes across 20–50°S latitudes in the South-Western Indian Ocean. *Deep Sea Research Part I:*
707   *Oceanographic Research Papers*, 110:20–32.

708 Béhagle, N., du Buisson, L., Josse, E., Lebourges-Dhaussy, A., Roudaut, G., and Ménard, F. (2014). Mesoscale features and micronekton in the
709   Mozambique Channel: An acoustic approach. *Deep Sea Research Part II: Topical Studies in Oceanography*, 100:164–173.

710 Chawarski, J., Klevjer, T. A., Coté, D., and Geoffroy, M. (2022). Evidence of temperature control on mesopelagic fish and zooplankton communities
711   at high latitudes. *Frontiers in Marine Science*, 9:917985.

712 Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community. *Australian Journal of Ecology*, pages 117–143.

713 Cotté, C., Ariza, A., Berne, A., Habasque, J., Lebourges-Dhaussy, A., Roudaut, G., Espinasse, B., Hunt, B., Pakhomov, E., Henschke, N., Péron,
714   C., Conchon, A., Koedooder, C., Izard, L., and Cherel, Y. (2022). Macrozooplankton and micronekton diversity and associated carbon vertical
715   patterns and fluxes under distinct productive conditions around the Kerguelen Islands. *Journal of Marine Systems*, 226:103650.

716 Cotté, C. and Simard, Y. (2005). Formation of dense krill patches under tidal forcing at whale feeding hot spots in the St. Lawrence Estuary. *Marine*
717   *Ecology Progress Series*, 288:199–210.

718 Czudaj, S., Koppelmann, R., Möllmann, C., Schaber, M., and Fock, H. O. (2021). Community structure of mesopelagic fishes constituting sound
719   scattering layers in the eastern tropical North Atlantic. *Journal of Marine Systems*, 224:103635.

720 Daneri, G. and Carlini, A. (2002). Fish prey of southern elephant seals, Mirounga leonina, at King George Island. *Polar Biology*, 25(10):739–743.

721 Davison, P. C., Koslow, J. A., and Kloser, R. J. (2015). Acoustic biomass estimation of mesopelagic fish: backscattering from individuals,
722   populations, and communities. *ICES Journal of Marine Science*, 72(5):1413–1424.

723 de Lavergne, C., Vic, C., Madec, G., Roquet, F., Waterhouse, A. F., Whalen, C., Cuypers, Y., Bouruet-Aubertot, P., Ferron, B., and Hibiya, T. (2020).
724   A parameterization of local and remote tidal mixing. *Journal of Advances in Modeling Earth Systems*, 12(5):e2020MS002065.

725 De Robertis, A. and Higginbottom, I. (2007). A post-processing technique to estimate the signal-to-noise ratio and remove echosounder background
726   noise. *ICES Journal of Marine Science*, 64(6):1282–1291.

727 Della Penna, A., Llort, J., Moreau, S., Patel, R., Kloser, R., Gaube, P., Strutton, P., and Boyd, P. W. (2022). The Impact of a Southern Ocean Cyclonic
728   Eddy on Mesopelagic Micronekton. *Journal of Geophysical Research: Oceans*, 127(11).

729 Della Penna, A., Llort, J., Moreau, S., Patel, R. S., Kloser, R. J., Gaube, P., Strutton, P. G., and Boyd, P. W. (2021). The impact of a Southern Ocean
730   cyclonic eddy on mesopelagic micronekton. preprint, Oceanography.

731 Demer, D., BERGER, and WILLIAMSON (2015). Calibration of acoustic instruments.

732 Demer, D. A. and Conti, S. G. (2005). New target-strength model indicates more krill in the southern ocean. *ICES Journal of Marine Science*,
733   62(1):25–32.

734 Dornan, T., Fielding, S., Saunders, R. A., and Genner, M. J. (2019). Swimbladder morphology masks Southern Ocean mesopelagic fish biomass.
735   *Proceedings of the Royal Society B: Biological Sciences*, 286(1903):20190353.

736 Dornan, T., Fielding, S., Saunders, R. A., and Genner, M. J. (2022). Large mesopelagic fish biomass in the Southern Ocean resolved by acoustic
737   properties. *Proceedings of the Royal Society B: Biological Sciences*, 289(1967):20211781.

738 Duhamel, G., Hulley, P.-A., Causse, R., Koubbi, P., Vacchi, M., Pruvost, S., Vigetta, J.-O., Irisson, S., Mormède, M., Belchier, A., Detrich, H., Gutt,
739   J., Jones, C., Kock, K.-H., Lopez Abellan, L., and Van de Putte, A. (2014). Biogeographic atlas of the Southern Ocean.

740 Escobar-Flores, P. C., O'Driscoll, R. L., Montgomery, J. C., Ladroit, Y., and Jendersie, S. (2020). Estimates of density of mesopelagic fish in the
741   Southern Ocean derived from bulk acoustic data collected by ships of opportunity. *Polar Biology*, 43(1):43–61.

742 Fonvieille, N., Guinet, C., Saraceno, M., Picard, B., Tournier, M., Goulet, P., Campagna, C., Campagna, J., and Nerini, D. (2023). Swimming in an
743   ocean of curves: A functional approach to understanding elephant seal habitat use in the argentine basin. *Progress in Oceanography*, 218:103120.

744 Foote, K. G. (1980). Importance of the swimbladder in acoustic scattering by fish: A comparison of gadoid and mackerel target strengths. *The*
745   *Journal of the Acoustical Society of America*, page 2084.

Frank, T. and Widder, E. (2002). Effects of a decrease in downwelling irradiance on the daytime vertical distribution patterns of zooplankton and micronekton. *Marine Biology*, 140(6):1181–1193.

Fuchs, R., Baumas, C. M. J., Garel, M., Nerini, D., Le Moigne, F. A. C., and Tamburini, C. (2022). A RUpture-Based detection method for the Active mesopeLagIc Zone (RUBALIZ): A crucial step toward rigorous carbon budget assessments. *Limnology and Oceanography: Methods*, page lom3.10520.

García-Seoane, E., Klevjer, T., Mork, K. A., Agersted, M. D., Macaulay, G. J., and Melle, W. (2023). Acoustic micronektonic distribution and density is structured by macroscale oceanographic processes across 17–48° N latitudes in the North Atlantic Ocean. *Scientific Reports*, 13(1):4614.

Geisen, C., Ridame, C., Journet, E., Delmelle, P., Marie, D., Lo Monaco, C., Metzl, N., Ammar, R., Kombo, J., and Cardinal, D. (2022). Phytoplanktonic response to simulated volcanic and desert dust deposition events in the South Indian and Southern Oceans. *Limnology and Oceanography*, 67(7):1537–1553.

Giering, S. L. C., Sanders, R., Lampitt, R. S., Anderson, T. R., Tamburini, C., Boutrif, M., Zubkov, M. V., Marsay, C. M., Henson, S. A., Saw, K., Cook, K., and Mayor, D. J. (2014). Reconciliation of the carbon budget in the ocean's twilight zone. *Nature*, 507(7493):480–483.

Godard, M., Manté, C., Guinet, C., Picard, B., and Nerini, D. (2020). Diving Behavior of Mirounga leonina: A Functional Data Analysis Approach. *Frontiers in Marine Science*, 7:595.

Godø, O. R., Samuelsen, A., Macaulay, G. J., Patel, R., Hjøllo, S. S., Horne, J., Kaartvedt, S., and Johannessen, J. A. (2012). Mesoscale Eddies Are Oases for Higher Trophic Marine Life. *PLoS ONE*, 7(1):e30161.

Górecki, T. and Smaga, Ł. (2019). fdanova: an r software package for analysis of variance for univariate and multivariate functional data. *Computational Statistics*, 34:571–597.

Graham, R. M. and De Boer, A. M. (2013). The Dynamical Subtropical Front: The Dynamical Subtropical Front. *Journal of Geophysical Research: Oceans*, 118(10):5676–5685.

Grimaldo, E., Grimsmo, L., Alvarez, P., Herrmann, B., Møen Tveit, G., Tiller, R., Slizyte, R., Aldanondo, N., Guldberg, T., Toldnes, B., Carvajal, A., Schei, M., and Selnes, M. (2020). Investigating the potential for a commercial fishery in the Northeast Atlantic utilizing mesopelagic species. *ICES Journal of Marine Science*, 77(7-8):2541–2556.

Group, G. C. (2022). *GEBCO 2022 Grid*.

Hernández-León, S., Koppelmann, R., Fraile-Nuez, E., Bode, A., Mompeán, C., Irigoien, X., Olivar, M. P., Echevarría, F., Fernández de Puelles, M. L., González-Gordillo, J. I., Cózar, A., Acuña, J. L., Agustí, S., and Duarte, C. M. (2020). Large deep-sea zooplankton biomass mirrors primary production in the global ocean. *Nature Communications*, 11(1):6048.

Hulley, P. (1981). Results of the research cruises of FRV Herwig, Walther to South-America. 58. Family Myctophidae (Osteichthyes, Myctophiformes). pages 31:1–300.

Irigoien, X., Klevjer, T. A., Røstad, A., Martinez, U., Boyra, G., Acuña, J. L., Bode, A., Echevarria, F., Gonzalez-Gordillo, J. I., Hernandez-Leon, S., Agusti, S., Aksnes, D. L., Duarte, C. M., and Kaartvedt, S. (2014). Large mesopelagic fishes biomass and trophic efficiency in the open ocean. *Nature Communications*, 5(1):3271.

Kang, M., Kang, J.-H., Kim, M., Nam, S., Choi, Y., and Kang, D.-J. (2021). Sound Scattering Layers Within and Beyond the Seychelles-Chagos Thermocline Ridge in the Southwest Indian Ocean. *Frontiers in Marine Science*, 8:769414.

Kirkwood, J. M. (1982). *A guide to the Euphausiacea of the Southern Ocean*. Number 1 in ANARE research notes. Information Services Section, Antarctic Division, Dept. of Science and Technology, Kingston, Tas.

Kleckner, R. C. and Gibbs, R. H. (1972). Swimbladder structure of Mediterranean midwater fishes and a method of comparing swimbladder data with acoustic profiles. *Mediterranean Biological Studies*, pages 230–281.

Klevjer, T. A., Irigoien, X., Røstad, A., Fraile-Nuez, E., Benítez-Barrios, V. M., and Kaartvedt., S. (2016). Large scale patterns in vertical distribution and behaviour of mesopelagic scattering layers. *Scientific Reports*, 6(1):19873.

Kloser, R. J., Ryan, T. E., Young, J. W., and Lewis, M. E. (2009). Acoustic observations of micronekton fish on the scale of an ocean basin: potential and challenges. *ICES Journal of Marine Science*, 66(6):998–1006.

Koubbi, P. (1993). Influence of the frontal zones on ichthyoplankton and mesopelagic fish assemblages in the Crozet Basin (Indian sector of the Southern Ocean).

Koubbi, P, Moteki, M., Duhamel, G., Goarant, A., Hulley, P.-A., O'Driscoll, R., Ishimaru, T., Pruvost, P., Tavernier, E., and Hosie, G. (2011). Ecoregionalization of myctophid fish in the Indian sector of the Southern Ocean: Results from generalized dissimilarity models. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(1-2):170–180.

Kozlov, A. N. (1995). A REVIEW OF THE TROPHIC ROLE OF MESOPELAGIC FISH OF THE FAMILY MYCTOPHIDAE IN THE SOUTHERN OCEAN ECOSYSTEM. page 8.

Lavery, A. C., Wiebe, P. H., Stanton, T. K., Lawson, G. L., Benfield, M. C., and Copley, N. (2007). Determining dominant scatterers of sound in mixed zooplankton populations. *The Journal of the Acoustical Society of America*, 122(6):3304–3326.

Lee, W.-J. and Staneva, V. (2020). Compact representation of temporal processes in echosounder time series via matrix decomposition. *The Journal of the Acoustical Society of America*, 148(6):3429–3442.

Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., Hernandez, O., and Kloser, R. (2015). Optimization of a micronekton model with acoustic data. *ICES Journal of Marine Science*, 72(5):1399–1412.

Maclennan, D. (2002). A consistent approach to definitions and symbols in fisheries acoustics. *ICES Journal of Marine Science*, 59(2):365–369.

Marohn, L., Schaber, M., Freese, M., Pohlmann, J. D., Wysujack, K., Czudaj, S., Blancke, T., and Hanel, R. (2021). Distribution and diel vertical migration of mesopelagic fishes in the Southern Sargasso Sea — observations through hydroacoustics and stratified catches. *Marine Biodiversity*, 51(6):87.

Marshall, N. (1960). Swimbladder structure of deep-sea fishes in relation to their systematics and biology. *Discov. Rep.,*, pages 1–121.

Martin, A., Boyd, P., Buesseler, K., Cetinic, I., Claustre, H., Giering, S., Henson, S., Irigoien, X., Kriest, I., Memery, L., Robinson, C., Saba, G., Sanders, R., Siegel, D., Villa, M., and Guidi, L. (2020). Study the twilight zone before it is too late. page 3.

Mastail, M. and Battaglia, A. (1978). Amélioration de la conservation des pigments du zooplancton.

McClain, C. R., Fougerolle, M. F., Rex, M. A., and Welch, J. (2001). MOCNESS estimates of the size and abundance of a pelagic gonostomatid fish *Cyclothone pallida* off the Bahamas. *Journal of the Marine Biological Association of the United Kingdom*, 81(5):869–871.

Motoda, S. (1959). Devices of simple plankton apparatus.

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

Olivar, M. P., Hulley, P. A., Castellón, A., Emelianov, M., López, C., Tuset, V. M., Contreras, T., and Molí, B. (2017). Mesopelagic fishes across the tropical and equatorial Atlantic: Biogeographical and vertical patterns. *Progress in Oceanography*, 151:116–137.

O'Sullivan, D. (1983). *A guide to the Pelagic Tunicates of the Southern Ocean and Adjacent Waters*. Number 8 in Anare Research Notes. Department of Science and Technology, Antarctic Division, Kingston, Australia.

Pagniez, C., Fournier, S., Verdier, A.-G., Guéné, M., Hoarau, F., Allibert, S., and Lustenberger, F. (2021). Extension de la réserve naturelle nationale des terres australes françaises. Technical report.

Pagès, J. (2002). Analyse factorielle multiple appliquée aux variables qualitatives et aux données mixtes. page 34.

Pakhomov, E., Perissinotto, R., and McQuaid, C. (1994). Comparative structure of the macrozooplankton/ micronekton communities of the Subtropical and Antarctic Polar Fronts. *Marine Ecology Progress Series*, 111:155–169.

Park, Y., Park, T., Kim, T., Lee, S., Hong, C., Lee, J., Rio, M., Pujol, M., Ballarotta, M., Durand, I., and Provost, C. (2019). Observations of the Antarctic Circumpolar Current Over the Udintsev Fracture Zone, the Narrowest Choke Point in the Southern Ocean. *Journal of Geophysical Research: Oceans*, 124(7):4511–4528.

Parra, S. M., Greer, A. T., Book, J. W., Deary, A. L., Soto, I. M., Culpepper, C., Hernandez, F. J., and Miles, T. N. (2019). Acoustic detection of zooplankton diel vertical migration behaviors on the northern Gulf of Mexico shelf. *Limnology and Oceanography*, 64(5):2092–2113.

Pauthenet, E., Roquet, F., Madec, G., Guinet, C., Hindell, M., McMahon, C. R., Harcourt, R., and 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., and 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., and Nerini, D. (2019). The Thermohaline Modes of the Global Ocean. *Journal of Physical Oceanography*, 49(10):2535–2552.

Pauthenet, E., Sallée, J.-B., Schmidtko, S., and Nerini, D. (2021). Seasonal Variation of the Antarctic Slope Front Occurrence and Position Estimated from an Interpolated Hydrographic Climatology. *Journal of Physical Oceanography*, 51(5):1539–1557.

Perrot, Y., Brehmer, P., Habasque, J., Roudaut, G., Behagle, N., Sarré, A., and Lebourges-Dhaussy, A. (2018). Matecho: An Open-Source Tool for Processing Fisheries Acoustics Data. *Acoustics Australia*, 46(2):241–248.

Peña, M., Olivar, M. P., Balbín, R., López-Jurado, J. L., Iglesias, M., and Miquel, J. (2014). Acoustic detection of mesopelagic fishes in scattering layers of the Balearic Sea (western Mediterranean). 71.

Proud, R., Cox, M. J., and Brierley, A. S. (2017). Biogeography of the Global Ocean's Mesopelagic Zone. *Current Biology*, 27(1):113–119.

Proud, R., Handegard, N. O., Kloser, R. J., Cox, M. J., and Brierley, A. S. (2019). From siphonophores to deep scattering layers: uncertainty ranges for the estimation of global mesopelagic fish biomass. *ICES Journal of Marine Science*, 76(3):718–733.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis*. Springer series in statistics. Springer, New York, 2nd ed edition.

Receveur, A., Kestenare, E., Allain, V., Ménard, F., Cravatte, S., Lebourges-Dhaussy, A., Lehodey, P., Mangeas, M., Smith, N., Radenac, M.-H., and Menkes, C. (2020a). Micronekton distribution in the southwest Pacific (New Caledonia) inferred from shipboard-ADCP backscatter data. *Deep Sea Research Part I: Oceanographic Research Papers*, 159:103237.

Receveur, A., Menkes, C., Allain, V., Lebourges-Dhaussy, A., Nerini, D., Mangeas, M., and Ménard, F. (2020b). Seasonal and spatial variability in the vertical distribution of pelagic forage fauna in the Southwest Pacific. *Deep Sea Research Part II: Topical Studies in Oceanography*,

175:104655.

Reygondeau, G., Guidi, L., Beaugrand, G., Henson, S. A., Koubbi, P., MacKenzie, B. R., Sutton, T. T., Fioroni, M., and Maury, O. (2018). Global biogeochemical provinces of the mesopelagic zone. *Journal of Biogeography*, 45(2):500–514.

Ryan, T. E., Downie, R. A., Kloser, R. J., and Keith, G. (2015). Reducing bias due to noise and attenuation in open-ocean echo integration data. *ICES Journal of Marine Science*, 72(8):2482–2493.

Schlüter, L., Henriksen, P., Nielsen, T. G., and Jakobsen, H. H. (2011). Phytoplankton composition and biomass across the southern Indian Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 58(5):546–556.

Schwarz, G. (1978). Estimating the dimension of a model.

Scrucca, L., Fop, M., Murphy, Brendan, T., and Raftery, E., A. (2016). mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *The R Journal*, 8(1):289.

Sergi, S. (2021). *Apport en nutriments par les monts sous-marins et les vents hydrothermaux dans l'océan Austral: effets sur les écosystèmes pélagiques et implications pour la conservation/Nutrient input from seamounts and hydrothermal vents in the Southern Ocean: impacts on the pelagic ecosystems and implications for conservation*. PhD thesis, PhD Thesis, Université Sorbonne.

St. John, M. A., Borja, A., Chust, G., Heath, M., Grigorov, I., Mariani, P., Martin, A. P., and Santos, R. S. (2016). A Dark Hole in Our Understanding of Marine Ecosystems and Their Services: Perspectives from the Mesopelagic Community. *Frontiers in Marine Science*, 3.

Sutton, T. T. (2013). Vertical ecology of the pelagic ocean: classical patterns and new perspectives: vertical ecology of the pelagic ocean. *Journal of Fish Biology*, 83(6):1508–1527.

Sutton, T. T., Clark, M. R., Dunn, D. C., Halpin, P. N., Rogers, A. D., Guinotte, J., Bograd, S. J., Angel, M. V., Perez, J. A. A., Wishner, K., Haedrich, R. L., Lindsay, D. J., Drazen, J. C., Vereshchaka, A., Piatkowski, U., Morato, T., Błachowiak-Samołyk, K., Robison, B. H., Gjerde, K. M., Pierrot-Bults, A., Bernal, P., Reygondeau, G., and Heino, M. (2017). A global biogeographic classification of the mesopelagic zone. *Deep Sea Research Part I: Oceanographic Research Papers*, 126:85–102.

Team, R. C. (2022). R: A language and environment for statistical computing.

Trenkel, V., Ressler, P., Jech, M., Giannoulaki, M., and Taylor, C. (2011). Underwater acoustics for ecosystem-based management: state of the science and proposals for ecosystem indicators. *Marine Ecology Progress Series*, 442:285–301.

Urmy, S. S., Horne, J. K., and Barbee, D. H. (2012). Measuring the vertical distributional variability of pelagic fauna in Monterey Bay. *ICES Journal of Marine Science*, 69(2):184–196.

Valinassab, T., Pierce, G. J., and Johannesson, K. (2007). Lantern fish (Benthosema pterotum) resources as a target for commercial exploitation in the Oman Sea. *Journal of Applied Ichthyology*, 23(5):573–577.

Warren, J. (2001). In situ measurements of acoustic target strengths of gas-bearing siphonophores. *ICES Journal of Marine Science*, 58(4):740–749.

Wiebe, P. H., Chu, D., Kaartvedt, S., Hundt, A., Melle, W., Ona, E., and Batta-Lona, P. (2010). The acoustic properties of Salpa thompsoni. *ICES Journal of Marine Science*, 67(3):583–593.

Youngbluth, M. J. (1975). The vertical distribution and diel migration of euphausiids in the central waters of the eastern South Pacific. *Deep Sea Research and Oceanographic Abstracts*, 22(8):519–536.

<u>Revised manuscript with changes marked</u>

The manuscript has undergone substantial revisions; as a result of these extensive changes, we opted to provide only the new version, rather than a revised manuscript with changes marked.

Highlights:

o Functional data analysis identifies variance in SL distribution via vertical modes.
o First mode is tied to temporal dynamics, next modes display spatial patterns.
o 3 regions link environmental surface data and SL from South Ocean to Southern Indian Ocean.
o 18 and 38 kHz differ in integrated acoustic density spatial patterns.
o Species assemblage help explain acoustic patterns in transition zones.

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: