DOI: 10.1111/jbi.13149

ORIGINAL ARTICLE

WILEY Journal of Biogeography

Global biogeochemical provinces of the mesopelagic zone

Gabriel Reygondeau[1,2] | Lionel Guidi[3,4] | Gregory Beaugrand[5] | Stephanie A. Henson[6] | Philippe Koubbi[7] | Brian R. MacKenzie[2] | Tracey T. Sutton[8] | Martine Fioroni[3] | Olivier Maury[9,10]

[1]Aquatic Ecosystems Research Laboratory, Fisheries Centre, University of British Columbia, Vancouver, BC, Canada

[2]Center for Macroecology, Evolution and Climate, National Institute for Aquatic Resources (DTU Aqua), Technical University of Denmark, Kongens Lyngby, Denmark

[3]Sorbonne Universités, UPMC Université Paris 06, CNRS, Laboratoire d'oceanographie de Villefranche (LOV), Observatoire Océanologique, Villefranche-sur-Mer, France

[4]Department of Oceanography, University of Hawaii, Honolulu, HI, USA

[5]Centre National de la Recherche Scientifique, Laboratoire d'Océanologie et de Géosciences, UMR LOG CNRS 8187, Station Marine, Université des Sciences et Technologies de Lille – Lille 1, Wimereux, France

[6]National Oceanography Centre, Southampton,UK

[7]Sorbonne Universités, Université Pierre et Marie Curie, UMR BOREA (MNHN, UPMC, CNRS, IRD, UCBN), Paris, France

[8]Halmos College of Natural Sciences and Oceanography, Nova Southeastern University, Dania Beach, Fl, USA

[9]IRD (Institut de Recherche pour le Développement), UMR EME 212, CRH (Centre de Recherches Halieutiques Méditerranéennes et Tropicales), Sète cedex, France

[10]Department of Oceanography, University of Cape Town, Rondebosch, Cape Town, South Africa

Correspondence
Gabriel Reygondeau, Fisheries Centre, University of British Columbia, Aquatic Ecosystems Research Laboratory, Vancouver, BC, Canada.
Email: gabriel.reygondeau@gmail.com.

Funding information
Nippon fondation – UBC NEREUS program; Danish National Research Foundation

Editor: Michael Dawson

Abstract

Aim: Following the biogeographical approach implemented by Longhurst for the epipelagic layer, we propose here to identify a biogeochemical 3-D partition for the mesopelagic layer. The resulting partition characterizes the main deep environmental biotopes and their vertical boundaries on a global scale, which can be used as a geographical and ecological framework for conservation biology, ecosystem-based management and for the design of oceanographic investigations.

Location: The global ocean.

Methods: Based on the most comprehensive environmental climatology available to date, which is both spatially and vertically resolved (seven environmental parameters), we applied a combination of clustering algorithms (c-means, k-means, partition around medoids and agglomerative with Ward's linkage) associated with a nonparametric environmental model to identify the vertical and spatial delineation of the mesopelagic layer.

Results: First, we show via numerical interpretation that the vertical division of the pelagic zone varies and, hence, is not constant throughout the global ocean. Indeed, a latitudinal gradient is found between the epipelagic–mesopelagic and mesopelagic–bathypelagic vertical limits. Second, the mesopelagic layer is shown here to be composed of 13 distinguishable Biogeochemical Provinces. Each province shows a distinct range of environmental conditions and characteristic 3-D distributions.

Main conclusions: The historical definition of the mesopelagic zone is here revisited to define a 3-D geographical framework and characterize all the deep environmental biotopes of the deep global ocean. According to the numerical interpretation of mesopelagic boundaries, we reveal that the vertical division of the zone is not constant over the global ocean (200–1,000 m) but varies between ocean basin and with latitude. We also provide evidence of biogeochemical division of the mesopelagic zone that is spatially structured in a similar way than the epipelagic in the shallow waters but varies in the deep owing to a change of the environmental driving factors.

KEYWORDS
biogeochemical provinces, environmental division, macroecology, mesopelagic, Ocean, Twilight zone

1 | INTRODUCTION

Historically, the pelagic part of the global ocean has been vertically divided into three zones: epipelagic, mesopelagic and bathypelagic (Hardy, 1956; Hedgpeth & Ladd, 1957; Pérès, 1961; Pérès & Devèze, 1963). The epipelagic zone is usually distinguished from the mesopelagic layer at a depth of ~200 m, where solar radiation is too low for effective photosynthesis (Behrenfeld, 2010; Sverdrup, 1953). While this boundary is commonly accepted, the definition of the boundary between the mesopelagic and bathypelagic zone varies among marine science disciplines but is usually fixed at ~1,000 m. The definition mainly rests on observations of the vertical distribution of pelagic organisms (Hardy, 1956; Hedgpeth & Ladd, 1957; Pérès, 1961; Pérès & Devèze, 1963). The mesopelagic layer is characterized by the presence of marine species that perform diel migrations between shallow and deep waters for feeding and predator avoidance (Pérès & Devèze, 1963), while bathypelagic species remain in general in deep waters as adults (Proud, Cox, & Brierley, 2017).

In addition to this vertical dimension, the identification and quantification of spatial biogeographical delineations (i.e. large marine ecosystems, provinces and biomes) for the epipelagic zone has been established for some time thanks to the availability of satellite observations and some key global biogeochemical parameters. Three global partitions of the ocean are widely used: (1) the Longhurst Biogeochemical Provinces (BGCP; Longhurst, 2007), (2) the Marine Ecoregions of the World (MEOW; Spalding et al., 2007) and (3) the Large Marine Ecosystems (LME for coastal systems; Sherman, 2005). These classifications aim to delineate the main oceanographic, ecological and socio-economic patterns of the pelagic ocean, respectively.

However, due to the difficulty in gathering information in deep ocean environments, these horizontal oceanic divisions cannot be reliably extended to deeper waters and should be considered as purely epipelagic. Nonetheless, an increasing number of studies confirm that the temporal and spatial dynamics of deep environmental conditions (e.g. temperature, salinity) and biogeochemical processes (e.g. carbon and nutrient recycling, respiration rate) differ from the surface ocean (Kaiser et al., 2005; Martin et al., 2011). The most important biogeochemical changes described between shallow and deep layers are mainly linked to light and turbulence attenuation, and differences between water current dynamics (Robinson et al., 2010; Sarmiento & Gruber, 2006). This vertical change in environmental conditions also drives changes in marine community assemblage structure (Reygondeau et al., 2012; Robinson et al., 2010; Steinberg, Cope, Wilson, & Kobari, 2008; Stemmann et al., 2008). However, some recent studies have characterized local biotic and/or abiotic features to identify ecological deep-water provinces at a macroscale (Proud et al., 2017; Steinberg et al., 2008; Sutton et al., 2017).

In this study, we identify a three-dimensional biogeochemical division (*sensu* Longhurst, 2007) of the mesopelagic zone using a quantitative approach proposed by Reygondeau et al. (2013) for the epipelagic zone. First, we propose a statistical approach to capture the spatial change in the mesopelagic vertical boundaries at a global scale rather than using constant vertical divisions reported in the literature (200–1,000 m). Second, based on an objective clustering analysis (merged procedure between Oliver et al., [2004] and Reygondeau et al., [2013]), we propose a geographical delineation of the biogeochemical conditions of the mesopelagic zone. The methodology enables us to objectively identify and validate the main environmental conditions of the mesopelagic zone, here named Mesopelagic BioGeoChemical Province (MBGCP, *sensu* Longhurst, 2007), in three dimensions according to seven selected environmental parameters: temperature, salinity, dissolved oxygen concentration, nutrient concentrations (NO_3, SiO_2, PO_4) and particulate organic carbon (POC) flux. This study represents an important step towards a global deep marine ecosystem atlas, providing 3-D boundaries, as well as abiotic characteristics of each MBGCP (Costello et al., 2010).

2 | MATERIALS AND METHODS

2.1 | Data

Each environmental dataset gathered for the study is an annual climatology that covers the global ocean at a spatial resolution of 1° and at a vertical resolution standardized in the World Ocean Atlas

(102 depth layers from 0 to 5500 m, every 5 m from 0 to 100 m, every 25 m from 100 to 500 m, every 50 m from 500 m to 1,000 m to 2,000 m and every 100 m from 3000 m to 5500 m; see Locarnini et al., 2010a for more information). To accurately characterize the main biogeochemical conditions of the mesopelagic zone in a parsimonious way, seven major environmental parameters were assembled to approximate the major drivers influencing deep biogeochemical and biotic processes in the oceanic realm (Longhurst, 2007; Robinson et al., 2010). Only those environmental parameters (temperature, FPOC, salinity, nutrients and oxygen concentration) with full ocean coverage in the vertical and the horizontal dimensions, as well as a sufficient number of observations per season at a global scale, were chosen to implement a robust annual climatology.

Temperature (°C) was selected for its cardinal influence on each marine trophic level of the global ecosystem. Temperature is commonly used in macro-ecological studies (Reygondeau et al., 2013), and its spatial variations allow the characterization of many marine biomes. In addition, vertical temperature gradients approximate the influence of atmospheric conditions on the whole water column (De Boyer Montégut, Madec, Fischer, Lazar, & Ludicone, 2004). Salinity was selected because it represents a good index of water–mass interactions and hydrodynamics. Salinity enables a spatial discrimination between oceanic regions such as subtropical or subpolar gyres and regions influenced by river inflow. Dissolved oxygen concentration was chosen to delineate the different levels of biological activity in an ecosystem. In particular, this parameter allows the identification of oxygen minimum zones (Bertrand, Ballon, & Chaigneau, 2010), which are a common biotope in the mesopelagic layer (Sutton et al., 2017). Nutrient concentrations (NO_3, SiO_2, PO_4) are a limiting resource for epipelagic primary producers and hence ultimately determine the influx of carbon into the mesopelagic. Nutrient concentrations are also indicators of the various type of biogeochemical processes in the mesopelagic zone. Finally, POC flux approximates food availability for key components of pelagic food webs and remineralization rates in the mesopelagic zone (Guidi et al., 2008).

Annual climatologies (calculated using observations from 1955 to 2012) of temperature (°C) (Locarnini et al., 2010b), salinity (Antonov et al., 2010) and dissolved oxygen concentration (mL L^{-1}) (Garcia et al., 2010a), orthophosphate, nitrate and silicate concentration (Garcia et al., 2010b) were retrieved from the World Ocean Atlas 2013 (WOA13, https://www.nodc.noaa.gov/OC5/woa13) from 0 to 5500 m depth in discrete vertical levels (Locarnini et al., 2010a) at a spatial resolution of 1° × 1°. A second set of environmental parameters gathered during different oceanographic cruises (performed during the World Ocean Circulation Experiments and Joint Global Ocean Flux Study) was used as independent data sources for validation. All environmental datasets (temperature, salinity, dissolved oxygen concentration and FPOC) of each cruise have been gathered from the Particle Dynamic group website of the University of Texas: http://people.tamu.edu/~wgardner/~pdgroup/SMP_prj/DataDir/SMP-data.html.

To define vertical boundaries of the mesopelagic zone, three environmental parameters were gathered: mixed layer depth (MLD), euphotic depth (Z_{eu}) and the vertical gradient of the flux of particulate organic carbon, and MLD and Z_{eu} were used to identify the depth of the boundary between the epipelagic and mesopelagic layers. These parameters influence phytoplankton biochemical processes and thereby activity (Sverdrup, 1953). The annual climatology of MLD was gathered from de Boyer-Montegut (http://www.ifremer.fr/cerweb/deboyer/mld/Surface_Mixed_Layer_Depth.php; De Boyer Montégut et al., 2004) who used an algorithm based on temperature or density difference from the surface (~5 m) to identify MLD. The annual climatology of Z_{eu} was gathered from the GlobColour group website (http://hermes.acri.fr/index.php?class=archive). Values of Z_{eu} were computed using the downwelling irradiance at 490 nm as suggested by Morel et al. (2007). As no quantitative definition of the mesopelagic–bathypelagic boundary is available in the literature, we here propose to define the limit based on the vertical gradient of the flux of particulate organic carbon. We chose this criterion because it reflects change in biogeochemical conditions (Black & Shimmield, 2003) and food availability in the mesopelagic, and shapes the distribution and biodiversity of deep-dwelling organisms (Robinson et al., 2010; Woolley et al., 2016). The flux of POC (F_{POC}) and POC vertical profiles were gathered from Henson, Sanders, and Madsen (2012), on a 1° grid, and have been modelled following the classical Martin's curve (Martin, 1987). For more information, please read Appendix S1.

2.2 | Methodology

The approach used to identify spatially resolved (horizontally and vertically) provinces of the mesopelagic domain was based on merging procedures developed by Oliver et al. (2004) and Reygondeau et al. (2013). The different steps of the procedure are summarized in Fig. S1 in Appendix S2.

2.2.1 | Determination of the vertical boundaries

The upper boundary represents the depth where the epipelagic and the mesopelagic zone transition. This boundary is approximated as the maximum depth where primary production is possible. As no global climatology of vertical profiles of primary production is yet available, we have here quantified this depth using the vertical limitation of photosynthetic activity (Behrenfeld, 2010; Sverdrup, 1953). The upper mesopelagic boundary is set, for each geographical cell, at the shallowest depth between the Z_{eu} and MLD (see Fig. S2 in Appendix S2). The annual climatology of the upper mesopelagic boundary is mapped in Figure 1a.

The bottom boundary separates the mesopelagic and the bathypelagic layers. As no common definition has been found in the literature, we have here considered the vertical gradient of F_{POC}. We identified the bottom boundary as the depth where F_{POC} value is not significantly different than the last five consecutive depths. First, each profile is interpolated every 5 m between 0 and 5,000 m (cubic spline interpolation [Legendre & Legendre, 1998;]). Second, to identify the depth of the boundary, a test was performed on each F_{POC}

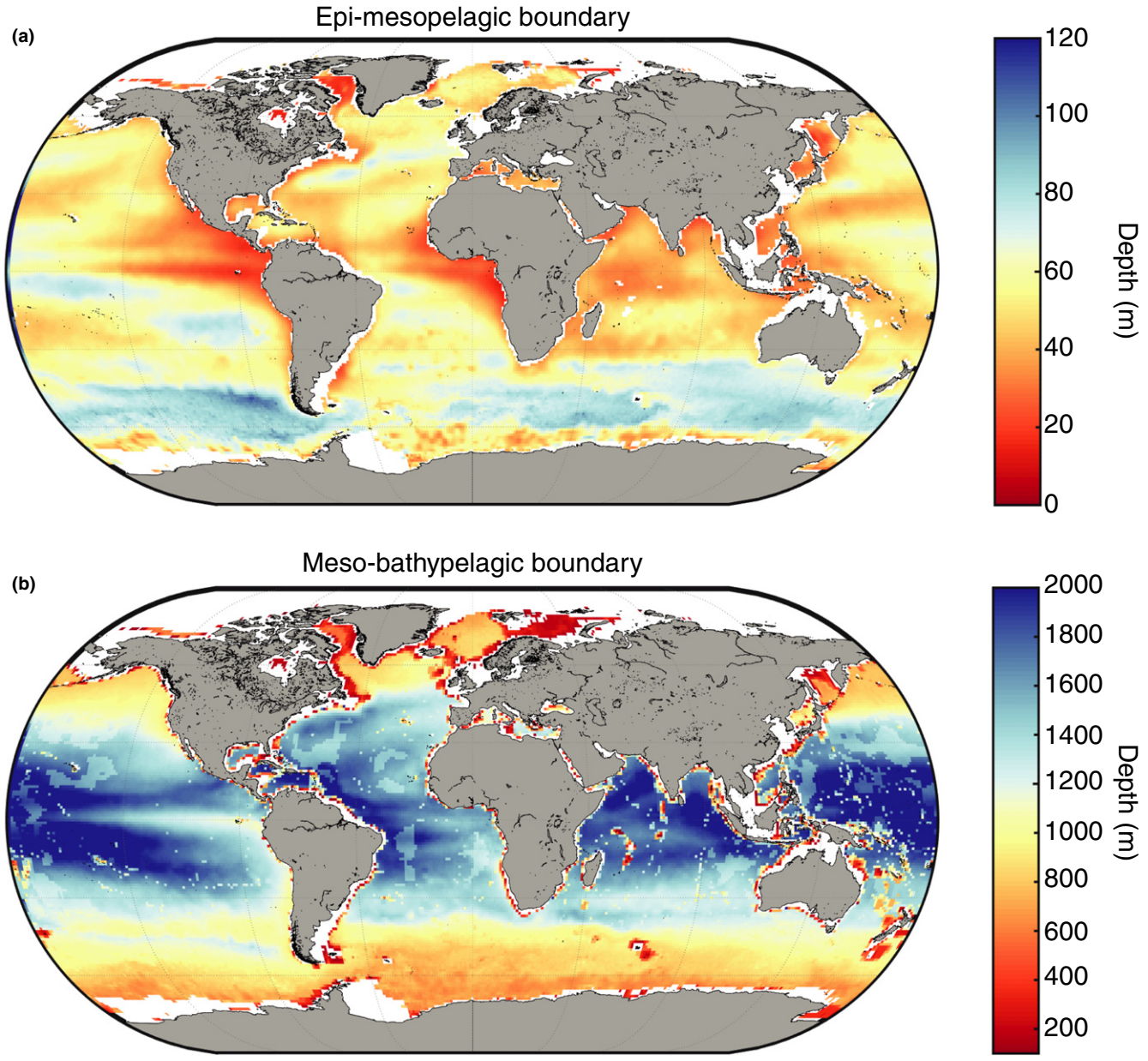


FIGURE 1 (a) Map of the annual climatology of the depth of the epi-mesopelagic boundary (m). (b) Map of the annual climatology of the depth of the meso-bathypelagic boundary (m)

profile. The depth derivative function of F_{POC} was computed, and the boundary was set at the depth where the decrease in F_{POC} between five consecutive points was <5% of the total variation of the parameter. This depth is mapped in Figure 1b. The interval of variation of the mesopelagic/bathypelagic depth was computed using the same methodology with thresholds of 2.5% and 7.5%, and results are mapped in Fig. S3 in Appendix S2.

2.2.2 | Identification of the spatial distribution of the MBGCP

In the global ocean, there is a strong anisotropy between vertical and horizontal environmental gradients, with gradients in the vertical exceeding those in the horizontal by orders of magnitude (if mesoscale variability is ignored, as it is in this study). Consequently, numerical procedures that partition space using environmental variance will emphasize a vertical division over a horizontal one. To identify an objective biogeochemical division of the mesopelagic layer with consideration of both geospatial and vertical environmental gradients, a new numerical procedure was implemented based on Oliver et al. (2004), Beaugrand, Lenoir, Ibañez, and Manté (2011) and Reygondeau et al. (2013). All specific information such as equations and mathematical explanations concerning these methodologies can be retrieved from these articles and is summarized below. All analyses were computed on Matlab ®.

First, two distinct environmental matrices were implemented; the first matrix, X_1, represented the depth-integrated environmental conditions (in temperature, salinity, dissolved oxygen concentration,

nutrients concentration and F_{POC}) between the upper and the bottom boundaries of the mesopelagic zone, while the second matrix, X_2, was vertically resolved for each geographical cell. Each environmental parameter of the matrix X_1 was then standardized (i.e. mean divided by standard deviation; Legendre & Legendre, 1998) (step 2.1; See Fig. S1 in Appendix S2). Second, to identify an objective environmental geospatial division, the methodology proposed by Oliver et al. (2004) was applied. This numerical procedure uses four types of clustering methodologies, here applied on the normalized matrix X_1: k-means (Hartigan & Wong, 1979), c-means (Quackenbush, 2001), k-medoids (or partitioning around the medoids; Kaufman & Rousseeuw, 1987) and agglomerative with a) Ward's linkage and b) complete linkage (Legendre & Legendre, 1998) (step 2.2; See Fig. S1 in Appendix S2). Each clustering algorithm was run to retrieve between two and 50 clusters using Euclidian distance. The c-means, k-means and k-medoids were repeated 999 times, and the division, which was retrieved most frequently, was selected. These four types of clustering algorithms were selected in the work of Oliver et al. (2004) for their ability to synoptically group similar environmental data and to handle low dissimilarity clusters.

The third step of the methodology consisted of the identification of the optimal number of clusters to consider. The figure of merit analysis (FOM; Yeung, Haynor, & Ruzzo, 2001) was used, as recommended in Oliver et al. (2004). This analysis allows quantification of the difference in the total environmental variance explained between two successive levels of clustering (i.e. here between two and 50) for each agglomerative methodology used. A normalized average slope function (ASF) was calculated using the FOM function computed for each clustering methodology. The ASF exhibits a negative exponential decrease with a rapid decrease for low numbers of clusters and slow decreases after a point k of inflection. The k points are considered to be the optimal number of clusters because the deviation between cluster means and the individual observations in each cluster are very small (Oliver et al., 2004). The k points named at the threshold of acceptable flatness (TAF) are defined as the first number of clusters where the decrease in ASF is <1% of the maximum ASF for three consecutive clusters. The FOM, ASF and TAF analysis are represented in Fig. S4 in Appendix S2. Here, the optimum number of clusters was found to be 13.

The fourth step of the numerical procedure aimed to identify an objective framework of the MBGCP distribution. The strength of the boundaries (denoted 'effectiveness') identified by each clustering methodology was first computed by taking into account all cluster delineations from two to k clusters. K-means and C-means procedures modify the spatial distribution of the clusters for each iteration, as these methodologies objectively divide the space in accordance with the centres of inertia detected. The map of the effectiveness (see Fig. S4 in Appendix S2) represents the percentage of occasions on which a boundary was retrieved in a given geographical cell for each methodology and from two to k clusters.

The last step of the procedure identified a composite distribution of the k clusters identified by the ASF. Contrary to the procedure of Oliver et al. (2004), which summed the cluster number retrieved with all clustering methodologies used, we attempted to find a trade-off between all the spatial partitions obtained. To do so, we used a methodology known as 'watershed' (Meyer, 1994), designed for imagery analysis. The watershed function tests the environmental variance similarity of each cluster computed with all methodologies within an area delineated by the boundary effectiveness computed previously (effectiveness greater than 25%; see Fig. S4 in Appendix S2). All clusters showing non-significant differences with adjacent clusters in their environmental variance were merged. The procedure stopped when k clusters (identified by the FOM) were retrieved. The resulting objective distribution of the 13 MBGCPs is shown in Figure 2.

2.2.3 | Characterization of the three-dimensional distribution of the MBGCP

The environmental envelopes (range of each parameter) of each MBGCP were quantified for all environmental parameters considered previously (temperature, salinity, nutrient concentrations, dissolved oxygen concentration and F_{POC}). For each province, the average environmental parameters in each province (Figure 2) and over the appropriate depth range (Figure 1) were extracted using matrix X_2. We subsequently computed all environmental values for each province using matrix X_2. We thus obtained 13 reference matrices (one for each province) named $Y^k_{n,p}$ with n = number of depths multiplied by the number of geographical cells, p = seven environmental variables and k ranging from 1 to 13.

To map the three-dimensional distribution of each MBGCP, we applied an ecological niche model called the nonparametric probabilistic ecological niche model (NPPEN; Beaugrand et al., 2011). NPPEN is based on the concept of the ecological niche of Hutchinson (Hutchinson, 1957), but was here applied as a test of environmental similarity as in Reygondeau et al. (2013). NPPEN can be summarized as a test that calculates the dissimilarity of a given sample set (here, any geographical cell and depth included in X_2) to a known set (here, one of the 13 MBGCP reference matrices, Y^k). In other words, we calculated the dissimilarity between the average environmental conditions within a cluster and all points belonging to that cluster. The method uses the generalized distance of Mahalanobis (Mahalanobis, 1936), which enables the correlation between variables to be taken into account. Distance values of the 13 MBGCP were subsequently calculated for each geographical cell and depth in matrix X_2. Then, for each geographical cell and depth, the lowest distance to the centroid of Y^k was identified and attributed to the MBGCP.

Based on the geospatial and vertical distribution of each MBGCP, the interval of environmental variation of each parameter is represented in Figure 3 and the geographical characteristics (mean latitude and volume) were assessed (Table 1). This procedure also allows an examination of the vertical distribution of each group along zonal and meridional transects (Figure 4). For each transect, a temperature-salinity diagram was plotted (see Fig. S5 in Appendix S2).

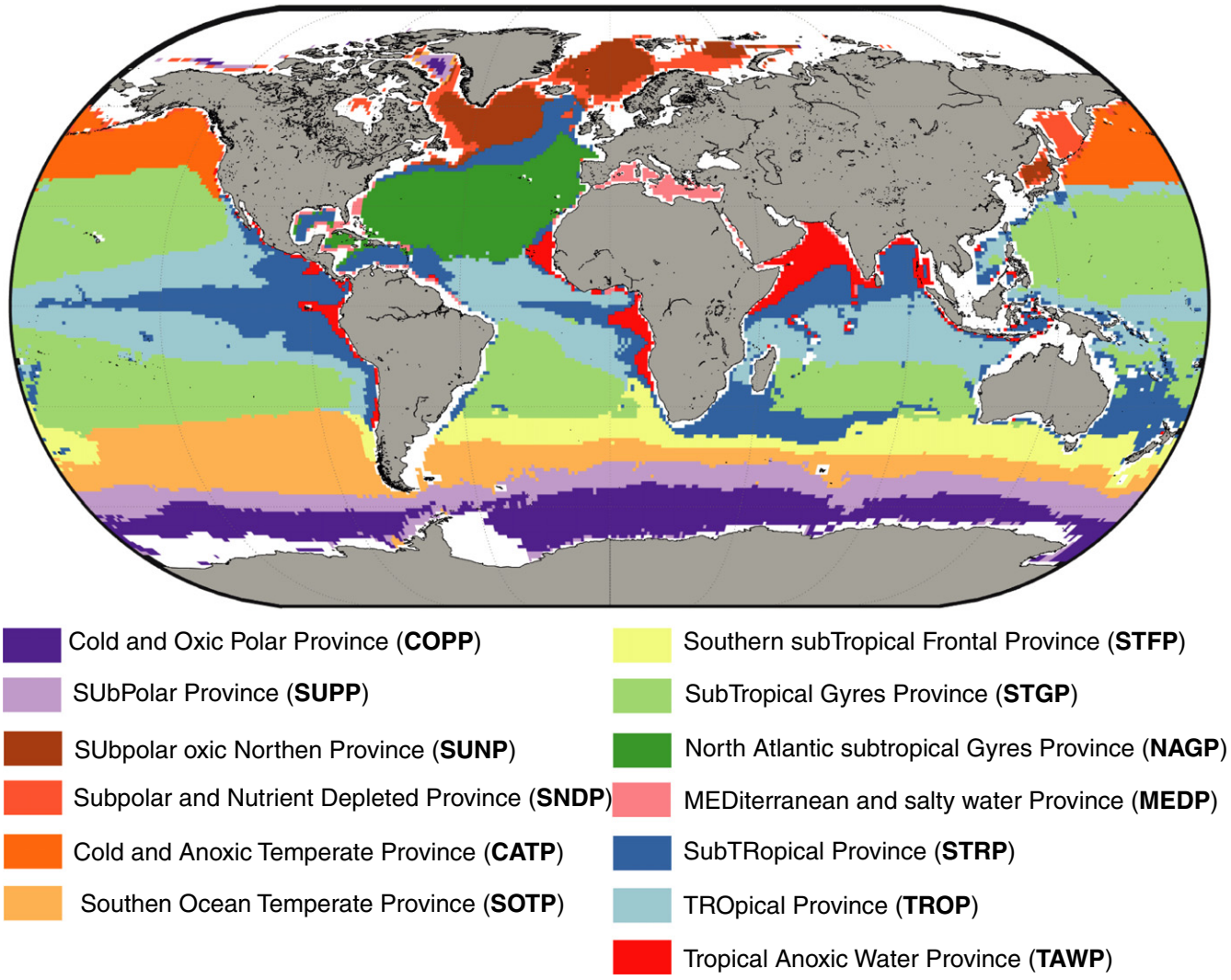


FIGURE 2 Map of the Mesopelagic BioGeoChemical Provinces (MBGCP). Names, annotations and corresponding colours of each MBGCP are provided. Geographical cells represented in white are areas where the methodology is not performed either because no data were available or the bathymetry was too shallow to retrieve the mesopelagic layer

2.2.4 | Validation of the vertical boundaries and distribution of the MBGCP

Based on all vertical environmental profiles (temperature, salinity, oxygen, FPOC) gathered for the World Ocean Circulation Experiment (WOCE) and the Joint Global Ocean Flux Study (JGOFS), a validation procedure of the vertical boundaries and spatial distribution of the MBGCPs was conducted (Figure 5a).

We have applied the methodology (see Determination of the vertical boundaries) on each in situ conductivity profile to quantify the meso-bathypelagic boundary in all sampling stations. We have then attributed to each sampling station the spatially closest value computed using the annual climatology (Figure 1b). A linear correlation was then calculated between in situ and climatological values of the depth of the boundary as well as a correlation index (r^2 and p-value). A plot of the relationship is shown in Figure 5b.

To validate the distribution of the MBGCPs, all environmental data (temperature, salinity, F_{POC} and oxygen concentrations) within the vertical mesopelagic bounds (Figure 2b) were first extracted. Average environmental conditions for each sampling site and parameter were then computed. Then, a principal component analysis (PCA) was performed on the mesopelagic average environmental values of all sampling sites (Figure 5c,d). According to Legendre and Legendre (1998), PCA is a statistical explorative methodology that reduces the variance of a large database (with more than 2 dimension/parameters) into a simplified dimension (principal components). Values of the coordinates of the first principal component (PC1, Figure 5c,d) are used as integrative index allowing to detect any multivariate environmental changes along the oceanographic cruises gathered. PC1 coordinates were mapped and confronted with the effectiveness of the MBGCP boundaries (Fig. S4 in Appendix S2) to evaluate whether environmental changes along a transect

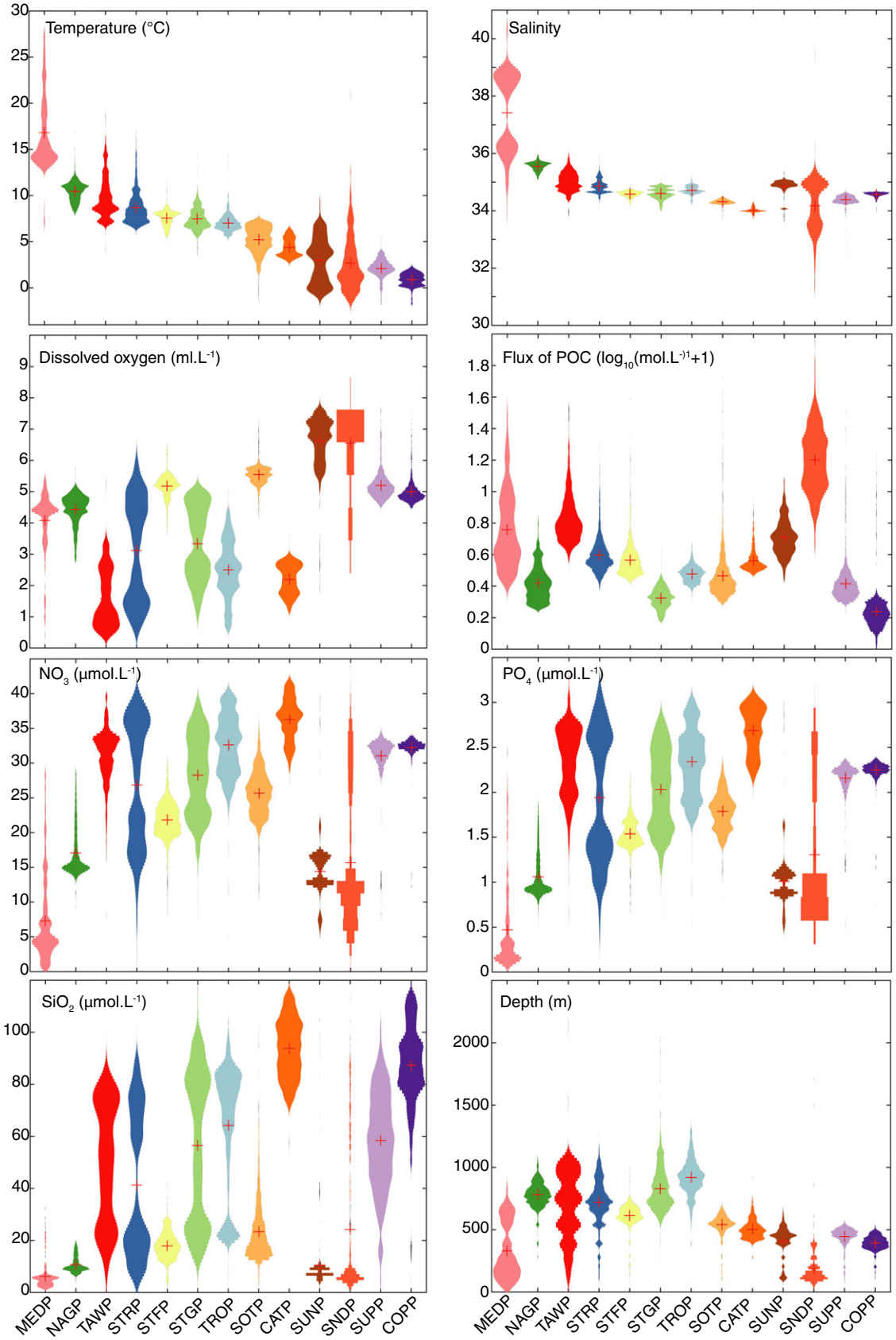


FIGURE 3 Violin plot of the environmental range of each Mesopelagic BioGeoChemical Provinces (MBGCP) for temperature (°C), salinity, dissolved oxygen concentration, flux of particulate organic carbon, nitrate, silicate and orthophosphate concentrations, and mean depth range

TABLE 1 Geographical characteristics of the identified 13 Mesopelagic Biogeochemical Provinces

	Latitude minimum	Latitude mean	Latitude maximum	Volume (km^3)	Percentage of the global ocean volume (%)
Cold and Anoxic Temperate Province (CATP)	18.5	61.3	81.5	1329049.097	0.10
MEDiterranean and salty water Province (MEDP)	0.5	27.8	42.5	1610639.34	0.12
SUbpolar oxic Northern Province (SUNP)	36.5	63.0	81.5	4392163.235	0.33
Subpolar and Nutrient Depleted Province (SNDP)	31.5	46.3	59.5	10773932.61	0.81
Cold and Oxic Polar Province (COPP)	−77 .5	−62.2	78.5	12035311.16	0.90
Tropical Anoxic Water Province (TAWP)	−38.5	5.6	40.5	12157284.49	0.91
SUbPolar Province (SUPP)	−76.5	−52.4	82.5	12918442.12	0.97
Southern subTropical Frontal Province (STFP)	−54.5	−40.2	−22.5	20709788.65	1.55
North Atlantic subtropical Gyres Province (NAGP)	11.5	29.7	51.5	25513605.18	1.91
Southern Ocean Temperate Province (SOTP)	−72.5	−46.2	75.5	29319580.88	2.20
SubTRopical Province (STRP)	−47.5	−4.2	62.5	60434448.04	4.53
TROpical Province (TROP)	−34.5	−2.5	36.5	91893036.24	6.88
SubTropical Gyres Province (STGP)	−39.5	−4.2	41.5	129302932	9.69
Mesopelagic (200–1,000 m)				254150000	19.03
Mesopelagic (in this study)				413350000	30.96
Global ocean				1335000000	100

corresponded with a change of MBGCP. A regional validation was also performed in the North Atlantic Ocean, and results are presented in Fig. S6 (Appendix S2).

However, there are caveats within this validation procedure that need to be taken into account: (1) the spatial division (Fig. 2) is implemented using six environmental parameters, while we only could gather an independent dataset with four parameters. Thus, the full environmental complexity summarized by the MBGCP division cannot be fully captured by the independent dataset; (2) MBGCP distribution is computed using annual average, while independent dataset is composed of a multitude of oceanographic cruises that have sampled each station at a specific time and location.

3 | RESULTS

3.1 | Vertical distribution of the mesopelagic layer

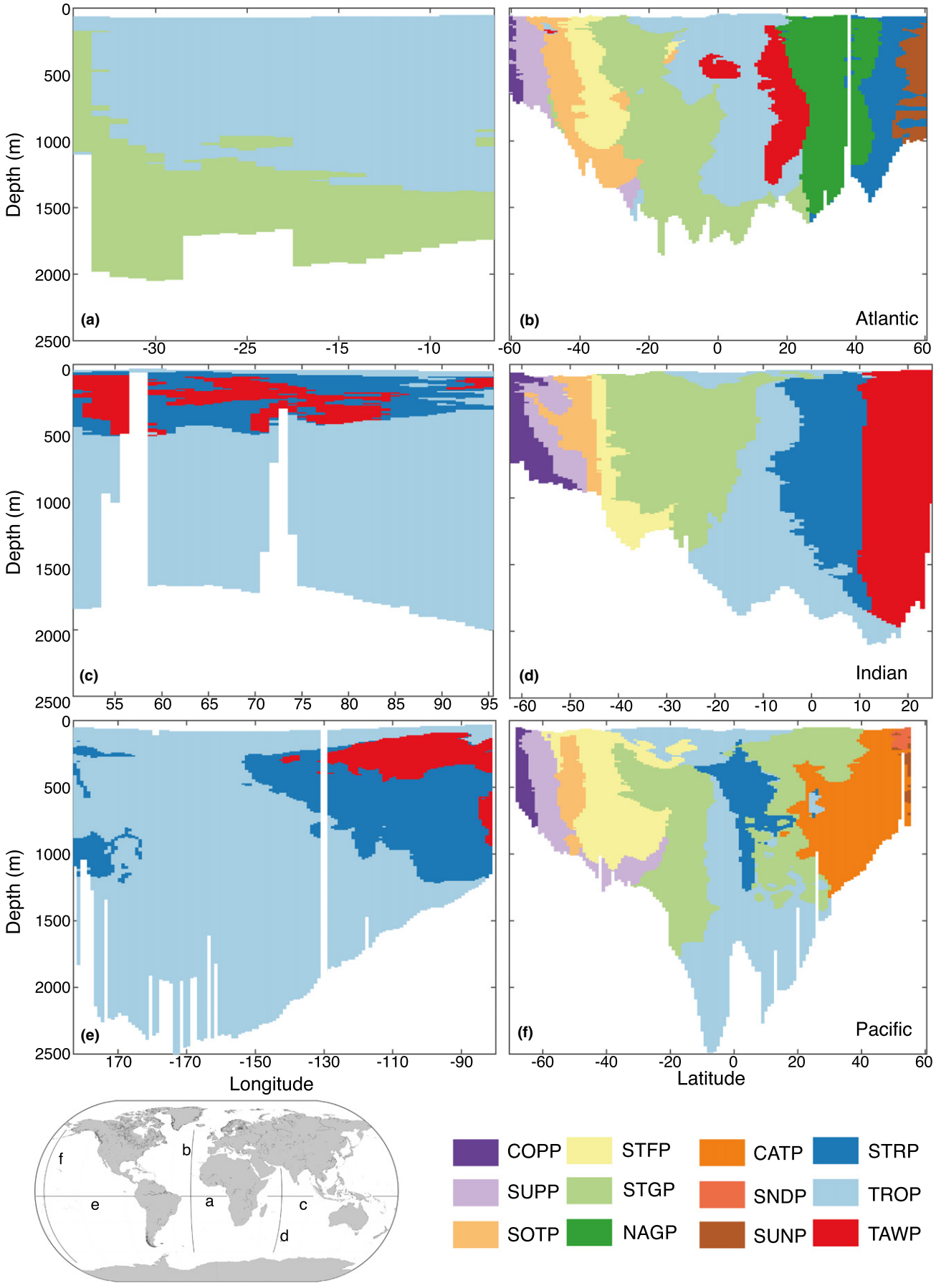
The upper boundary of the mesopelagic layer is approximated as the depth where the photosynthetic production of organic matter is limited by the low quantity of available light (Z_{eu}) or the low stability of the water column (MLD). We found that this boundary was equivalent to the euphotic depth in 71.4% of the global ocean, but only 28.6% of the mixed layer depth (see Fig. S2 in Appendix S2). The depth of the epipelagic/mesopelagic boundary varied between 10 m and 150 m (Figure 1a). The deepest upper mesopelagic boundaries were found in the south subtropical and subpolar regions and were driven by the MLD. In contrast, the shallowest depths of the boundary were found in the tropical and highly productive nearshore regions (Humboldt, equatorial upwelling) and was driven primarily by the Z_{eu} (Figure 1a).

The lower boundary separating the mesopelagic and the bathypelagic zone (Figure 1b) was set using the vertical gradient in F_{POC}. The annual climatology of the lower boundary of the mesopelagic layer (Figure 1b) showed a variation between 180 m and 2312 m. The depth of this deep boundary tended to decrease polewards (Figure 1b). The areas showing the deepest mesopelagic boundary were found in tropical regions between 30°N and 30°S, while the shallowest was located in temperate (40°–60°) and polar regions (>60°). Also, the vertical distribution of this boundary reflected local fluctuation in the bathymetry, especially near offshore islands.

3.2 | Spatial distribution of the MBGCPs and environmental characteristics

The mesopelagic zone was divided into 13 MBGCPs (Figure 2). The range of variation of parameters considered in the clustering analysis (temperature, salinity, dissolved oxygen concentration, nutrient concentrations and F_{POC}), as well as the average mean vertical

FIGURE 4 Spatial distribution of the Mesopelagic BioGeoChemical Provinces (MBGCP) as a function of depth, longitude and along (a) Atlantic Ocean longitudinal transect (from 50°W to 10°W at a latitude of 5°N) and (b) latitudinal transect (from 60°S to 60°N at a longitude of 25°W), (c) the Indian Ocean longitudinal transect (from 50°E to 95°E at a latitude of 5°N) and (d) latitudinal transect (from 65°S to 15°N at a longitude of 64° E) and (e) the Pacific Ocean longitudinal transect (from 130°W to 80°E at a latitude of 5°N) and (f) latitudinal transect (from 70°S to 60°N at a longitude of 170°W). White areas mark the depth of the lower boundary of the mesopelagic zone


Depth (m)
0
500
1000
1500
2000
2500
(a)
-30
-25
-15
-10
(b)
Atlantic
-60
-40
-20
0
20
40
60
(c)
55
60
65
70
75
80
85
90
95
(d)
Indian
-60
-50
-40
-30
-20
-10
0
10
20
(e)
170
-170
-150
-130
-110
-90
Longitude
(f)
Pacific
Latitude
a
b
c
d
e
f
COPP
SUPP
SOTP
STFP
STGP
NAGP
CATP
SNDP
SUNP
STRP
TROP
TAWP

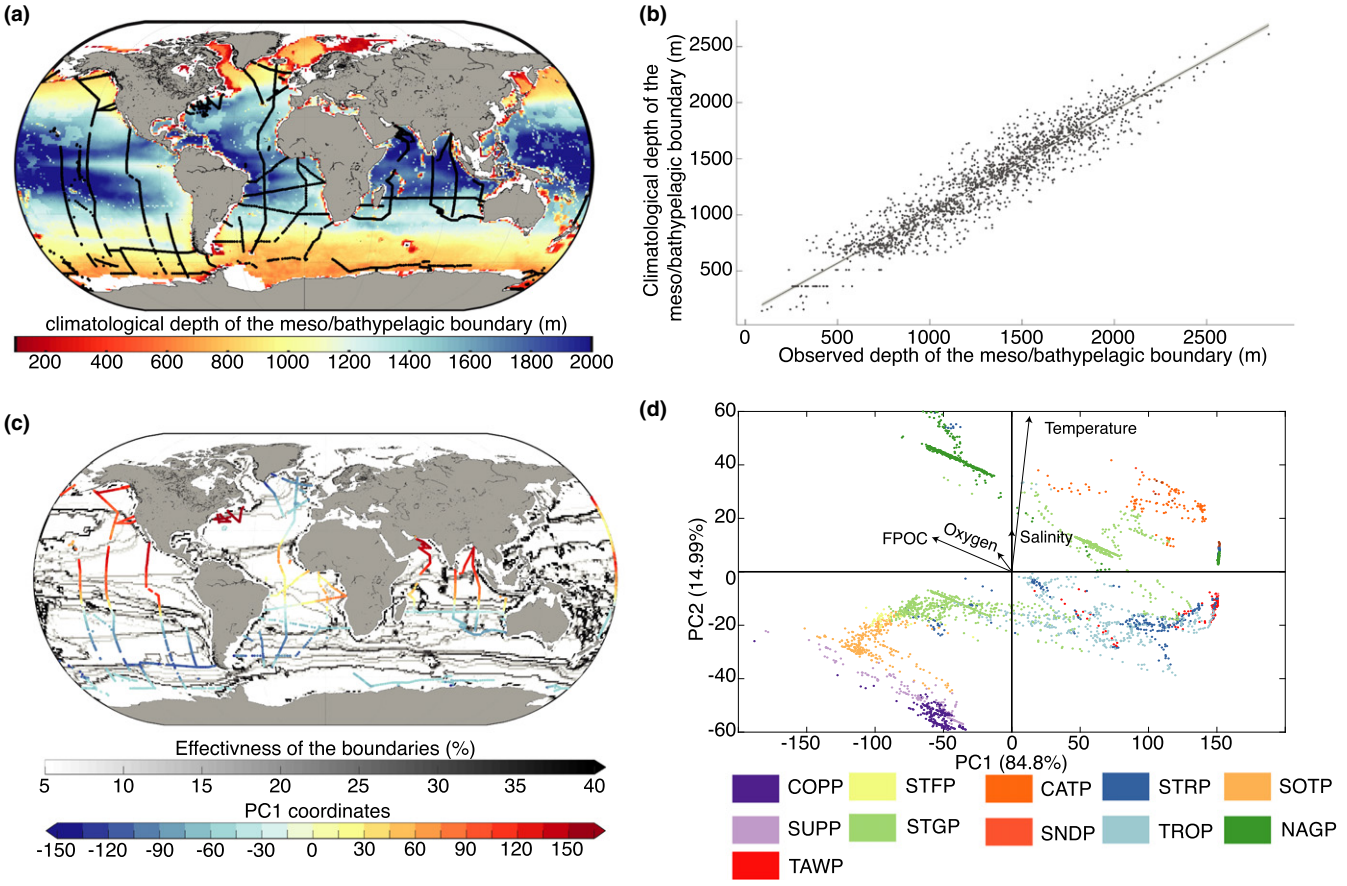


FIGURE 5 (a) Map of the location of the World Ocean Circulation Experiment (WOCE) and the Joint Global Ocean Flux Study (JGOFS) sampling stations. (b) Relationship between depth of the meso/bathypelagic boundary extracted from sampling stations and climatological values for the JGOFS-WOCE station. (c) Map of the PC1 coordinates value for PCA performed on mesopelagic environmental conditions of JGOFS/WOCE sampling stations, overlain with the map of boundary effectiveness between Mesopelagic BioGeoChemical Provinces. (d) Principal component analysis on mesopelagic environmental conditions of JGOFS/WOCE sampling stations. Each point represents a sampling station and marker colour represents the MBGCP where the station is located

distribution of each MBGCP, is shown in Figure 3. Geographical characteristics (latitude and volume) of each of the 13 MBGCPs are summarized in Table 1. Each MBGCP is here named according to their geospatial location and environmental characteristics.

Six provinces are located between 40°N and 40°S: TAWP (Tropical Anoxic Water Province), TROP (TROpical Province), STRP (SubTRopical Province), NAGP (North Atlantic subtropical Gyre Province), STGP (SubTropical Gyre Province) and MEDP (MEDiterranean [or 'salty water'] Province) (Figure 2). These provinces exhibit the highest temperatures and salinities of the mesopelagic layer. STRP and TROP are located around the equator. STRP is located in the Pacific, Atlantic and Indian Ocean seasonal upwellings and in the North Atlantic drift region, while TROP is distributed between equatorial upwellings and subtropical gyres, with an extension in the Kuroshio Current. STRP and TROP exhibit a similar environmental profile for temperature (5–12°C) and salinity (34 to 36 psu) as well as nutrient concentrations, but STRP diverges from TROP by delineating all deep anoxic regions of the global ocean, and with a higher F_{POC} (Figure 3). The provinces STGP and NAGP represent the mesopelagic part of the subtropical gyres. NAGP is located in the North Atlantic Ocean gyre, while STGP is located in the Pacific, Indian and South Atlantic basins. These provinces are characterized by high salinity (34–36 psu) and a moderate dissolved oxygen concentration, ranging from 1 to 5 mL L^{-1}. NAGP is distinguished by higher mean temperatures and lower nutrient concentrations than STGP.

Two provinces located within the tropical biome are characterized by highly specific environmental profiles: MEDP and TAWP. MEDP is situated in the Mediterranean and Caribbean Seas. While this province exhibits an environmental profile similar to subtropical gyre provinces (e.g. NAGP), the province shows the highest temperature, salinity and low nutrient concentrations of the global mesopelagic layer. The province TAWP is located in areas such as Arabian and next to New Zealand coasts, at the edge of the Amazon River run-off, and next to the main upwelling regions of the global ocean (Benguela, Canary, Humboldt and California). TAWP exhibits a high F_{POC} and a low dissolved oxygen concentration, while its temperature and salinity profiles are similar to other tropical provinces.

Regions located between 40° and 60° N/S were here divided into five MBGCPs: SUNP (SUbpolar oxic Northern Province), SOTP (Southern Ocean Temperate Province), STFP (Southern subTropical

Frontal Province), SNDP (Subpolar and Nutrient Depleted Province) and CATP (Cold and Anoxic Temperate Province). These provinces, representing the mesopelagic part of the temperate biome and extending to subplar regions, exhibit temperature values between −1°C and 9°C, salinity values between 31 and 35.8 psu, dissolved oxygen concentrations between 1 and 9 mL L^{-1}, high F_{POC} (between $\log_{10}$ (0.1 and 1.8 g C m^{-2})) and a high variation in nutrient concentrations. Also, this set of MBGCPs is characterized by a high variation within the cluster of each of the parameters considered. STFP and SOTP are located in the Southern Ocean. These provinces are the warmest of the 40°–60° latitude band and exhibit low nutrient concentrations. SUNP, SNDP and CATP are distributed in the Northern Hemisphere and represent the temperate and subpolar borders of the mesopelagic layer in these regions. These provinces are colder than STFP and SOTP and exhibit high F_{POC} and low nutrient concentrations. The environmental profile of CATP (mainly located in the Bering Sea) shows dissolved oxygen concentrations comparable to tropical regions such as TAWP and STRP, as well as high nutrient concentrations. SNDP and SUNP exhibit similar environmental conditions. However, SNDP is characterized by the highest F_{POC} of all MBGCPs.

Two MBGCPs are located in regions polewards to 60°: SUPP (SUbPolar Province) and COPP (Cold and Oxic Polar Province). The distribution of these provinces is bounded in the highest latitudes by the availability of environmental data considered in this study. Their environmental conditions are representative of a polar biome with cold temperature, low salinity and high oxygen concentration. Nonetheless, SUPP appears warmer, with higher F_{POC} than COPP, showing that this province marks the subpolar regions of the mesopelagic layer while COPP represents polar regions.

3.3 | Vertical distribution and volume of the MBGCPs

The zonal transects performed at 5°N (Figure 4a,c,e) in each of the Pacific, Atlantic and Indian Oceans reveal no common pattern in the mesopelagic thickness and MBGCP succession compared to meridional transects (Figure 4b,d,f). For instance, the zonal transect performed in the equatorial Pacific Ocean (Figure 4e) showed an increase of the upper and bottom boundaries, while upper boundaries in the Indian and Atlantic Oceans were constant (Figure 4a,c). Regarding the MBGCP distributions, TROP and STRP are generally found at depths ranging from 150 to 2,000 m, with TROP in the deepest region. However, no recurrent patterns were detected between basins and across different latitudes. These results suggest that regional environmental and oceanographic (i.e. circulation) characteristics of each basin drive the longitudinal MBGCP succession patterns.

The meridional transects showed that the mesopelagic layer becomes thicker from the pole to the equator in all oceanic basins (Figure 4). The upper boundaries are deeper above subtropical gyres and shallower in polar regions. Bottom boundaries decrease with latitude and exhibit high variation associated with a local change in bathymetry (e.g. islands). Examination of meridional transects (Figure 4c,d,f) reveals similar succession of the MBGCPs across oceanic basins, with relatively little variability in depth distribution. This result suggests that latitudinal environmental change driven by macro-ecological and global physical processes drive the 3-D poleward succession of MBGCPs rather than local and vertical oceanographic features.

The estimation of the mesopelagic volume differs greatly between our proposed division, which takes into account vertical environmental gradient (30.96% of the global ocean volume; Table 1), and the historical division with fixed boundaries at 200 and 1,000 m (19.03% of the global ocean volume; Table 1). The estimated volumes of water contained in the different MBGCPs range from 1 to 129 million km^3 (Table 1). The biggest mesopelagic provinces are generally located in the tropical regions (STGP, STRP and TROP) and collectively represent 65% of the ocean's mesopelagic habitat.

3.4 | Validation of the vertical and spatial boundary of the MBGCPs

Based on the environmental data collected over the global ocean during the World Ocean Circulation Experiments and Joint Global Ocean Flux Study, we have performed a vertical and spatial validation of the MBGCPs distribution (Figure 5).

The meso-bathypelagic boundary was tested here by comparing results from the methodology used to identify the depth where F_{POC} value does not significantly differ than the last five consecutive depths (Figure 2a) and in situ conductivity vertical profiles gathered at various location. A significant correlation of 0.88 (r^2, p-value <.01) was found showing that the observations and the climatological values estimated for the boundary were coherent despite the different sampling bias and potential seasonal fluctuation of this deep boundary (Figure 2b).

We have applied the procedure proposed by Oliver and Irwin (2008) to validate the spatial distribution of the MBGCPs by comparing in situ mesopelagic environmental values summarized by a PCA (Figure 5d) and the boundary effectiveness (see Methodology) resulting from the multiclustering methods (Figure 5c). This overlay revealed that along meridional transects (Pacific, Atlantic – Fig. S6 in Appendix S2 – and Indian Oceans) changes in the sampled environmental values (here summarized by PC1 coordinates) were located next to the boundary detected by the multiclustering methodology used in this study. The strongest changes were mostly located next to a boundary with a high effectiveness, as for instance next to subtropical gyres or subpolar–subtropical fronts. Along meridional transects, results revealed that environmental conditions had low variation within a province and changed when crossing a detected boundary. The fluctuation of the environmental values did not totally coincide with the detected boundary, suggesting an effect of natural seasonality on the MBGCPs that cannot be captured by our present methodology.

4 | DISCUSSION

In 1995, Longhurst established a theoretical approach to delineate environmentally homogenous regions with well-defined

oceanographic boundaries on the basis of available environmental data and expert knowledge. The approach was developed to identify all possible oceanic biotopes influencing the dynamics of the main biogeochemical processes and living marine organism distribution and interaction. The intended goal was to define an oceanic atlas of marine ecosystems with their spatial coverage providing a spatioecological framework for conservation management and for further ecological research (Longhurst, 2010). This conceptual framework was quantitatively refined by Reygondeau et al. (2013) for the epipelagic layer and by Watling, Guinotte, Clark, and Smith (2013) for the seabed using statistical methodologies. Here, we apply these numerical approaches to extend the biogeochemical division of Longhurst to the mesopelagic zone. This work should be considered as a first step towards the development of a biogeochemical atlas of the deep marine pelagic ecosystems in the same way as the work of Sathyendranth et al. (1995) for the global division of Longhurst (1998). Also, with the current availability of comprehensive 3-D environmental datasets and the use of exploratory statistical tools, we here proposed to consider variability in the vertical dimension in contrast to previous existing comparable studies (Hooker, Rees, & Aiken, 2000; Oliver & Irwin, 2008; Reygondeau et al., 2013).

The vertical partition of the water column was historically defined to account for the vertical distribution of communities (Hardy, 1956; Pérès, 1961; Pérès & Devèze, 1963; Sutton et al., 2017). Accordingly, the mesopelagic zone is usually defined by a uniform depth range (200–1,000 m) inhabited during the day by diel migratory organisms that feed at night on epipelagic organisms and by non-migrant organisms that stay in these deeper regions and feed on sinking material and other mesopelagic organisms (Angel & Pugh, 2000). This depth range does not account for spatial environmental or biological heterogeneity encountered over the global ocean and does not cover the full distribution of environmental or ecological features of deep-pelagic ecosystems. For instance, vertical ranges of deep oxygen minimum zones, oceanic equatorial upwelling zones, deep nutrient pumping, and subtropical and subpolar frontal systems are not fully covered by the 200–1,000 m range (Stramma, Johnson, Sprintall, & Mohrholz, 2008; Falkowski, Barber, & Smetacek, 1998; Proud et al., 2017; Sutton et al., 2017 [see Figure 2]). In this context, we have proposed to revisit the depth range used to partition the water column using an empirical methodology that considers the vertical environmental variability. For that purpose, we have defined the mesopelagic layer as the vertical region where light is highly attenuated but not null and where biological activity (e.g. grazing, remineralization of organic matter) is prevalent with respect to the bathypelagic zone. Accordingly, the upper and lower boundaries of the mesopelagic layer are defined as follows:

- Previous work on the vertical distribution of primary production and light attenuation (Behrenfeld, 2010; Morel et al., 2007; Sverdrup, 1953) defined the epipelagic layer as the vertical region dominated by the photosynthetically based production of organic matter. Consequently, the epi-mesopelagic boundary is established at the depth where light-driven primary production can no longer be supported by environmental conditions.
- We chose to account for both biogeochemistry and ecological studies to define the meso-bathypelagic boundary at the depth where a vertical change in F_{POC} is not significant over five consecutive 5-m bins. We assumed that this depth corresponds to the horizon where biogeochemical processes significantly reduce the organic carbon flux and where the influence of shallow waters on deep marine ecosystems is negligible (Woolley et al., 2016).

The resulting distribution of the boundaries (Figures 1 and 5) obtained by applying the present assumptions contrast with the currently accepted, constant vertical range for the mesopelagic zone (between 200 and 1,000 m). Overall, the results revealed that the vertical coverage of the mesopelagic zone may vary over the global ocean between 50 and 2,300 m with a greater thickness in low latitude regions, decreasing towards the poles, and with important variation across basins. These fluctuations in mesopelagic thickness can be attributed to the shallow-water and local influences on MLD or Z_{eu}, such as riverine run-off, detritus concentration and/or persistent mesoscale processes. In the deeper regions, the mesopelagic thickness is driven by the latitudinal increase of the meso-bathypelagic boundary that can be attributed to the change in the deep biological community and its efficiency in processing the sinking organic matter from the epipelagic zone (Guidi et al., 2009, 2015). The reinterpretation of mesopelagic vertical boundaries also influences the estimated size of the mesopelagic volume that, based on our calculations, comprises 30% of the total amount of ocean volume, representing a global increase of ca. 12% relative to volume estimates made by Longhurst using a constant depth range (Table 1). However, the volume estimates and boundary locations are somewhat uncertain due to the natural variability of the ocean-climate system, the lack of sampling in the mesopelagic part of the ocean and various assumptions which were used to derive these estimates; the precise volumes and boundaries should therefore be considered as preliminary estimates.

The present study is a first attempt to partition the mesopelagic zone using available 3-D observations of the environment at a global scale. For this purpose, comprehensive environmental data resolved in both horizontal and vertical dimensions have been gathered. Seven environmental variables were found to be suitable for a global partitioning of the mesopelagic layer because they met the following criteria: (1) they have a comprehensive 3-D global coverage; (2) they are computed from comparable field samples with a similar spatial heterogeneity of sampling sites; and (3) their annual climatologies are calculated using similar statistical interpolations. Also, we have analysed the covariability between each parameter (Table S1, Appendix S2) and evaluate their contribution in characterizing the MBGCPs by testing all possible combinations of environmental parameters in the multiclustering methodology. Results revealed that each variable had a significant contribution to the identification of

one or more MBGCPs. As a consequence, we subsequently retained the full set of environmental parameters in our analyses.

In addition, we are aware of the limitations of using one set of global F_{POC} to define the lower boundaries of the mesopelagic layer. Unfortunately, in situ F_{POC} data are limited to sediment trap flux measurements (Honjo, Manganini, Krishfield, & Francois, 2008), particle-reactive radionuclides proxies ($^{238}U/^{234}Th$ and $^{210}Pb/^{210}Po$; [Stewart, Moran, & Lomas, 2010; Owens, Pike, & Buesseler, 2015;]) and/or several optical/imaging proxies (e.g. Briggs, Slade, Boss, & Perry, 2013; Guidi et al., 2008; Iversen, Nowald, Ploug, Jackson, & Fischer, 2010; McDonnell & Buesseler, 2010). However, none of the above permit a global analysis of flux attenuation with depth with sufficient spatial/temporal resolution for the present study. Using a model approximation of F_{POC} was therefore the only alternative. While this dataset is not ideal, it is the best available to date to attempt a first estimation of the meso-bathypelagic boundary layer. Indeed, including F_{POC} is essential as it represents a critical component of the environment upon which mesopelagic communities rely (Woolley et al., 2016). We here provide a hypothetical definition of the mesopelagic zone boundary that was tested using several thresholds (see Fig. S3 in Appendix S2) and which may be improved when F_{POC} global estimates become more accurate. Indeed, in the past 10 years, new proxies of F_{POC} based on optical and imaging methods have emerged providing high vertical resolution profiles of F_{POC}, offering additional validation opportunities of the present MBGCP in the coming years.

The present study has identified, validated and mapped 13 mesopelagic biogeochemical environmental types over the global ocean (Figure 2). Each MBGCP represents a characteristic combination of environmental conditions (Figure 3) that covers all the known environmental specificities of the mesopelgic zone described in the literature (Robinson et al., 2010).We here project their potential 3-D coverage (Figure 4) and estimate their volumes (Table 1). In the same way as the biogeographical structure proposed by Longhurst, the mesopelagic provinces can be empirically regrouped by biomes based on their environmental characteristics (Figure 3): six belong to the tropical biome (STRP, TROP, STRP, NAGP, STGP, TAWP and MEDP), five to the temperate biome (SUNP, SOTP, STFP, SNDP and CATP) and two to the polar biome (COPP and SUPP). Also, each MBGCP shows variability in their environmental intervals (Figure 3) that can be attributed to oceanic basin specificities. Consequently, each MBGCP type can be subdivided by their basin location resulting in 28 regional mesopelagic provinces (two in the Southern Ocean, two in the Arctic Ocean, nine in the Pacific Ocean, 10 in the Atlantic Ocean and five in the Indian Ocean; Figure 2).

The average MBGCP distribution (Figure 2) reveals high spatial similarities with the partitions of the shallow environmental waters (Hardman-Mountford, Hirata, Richardson, & Aiken, 2008; Longhurst, 1995; Moore, Campbell, & Feng, 2002; Oliver & Irwin, 2008). Therefore, known macro-ecological concepts (*sensu* Longhurst) seem to be valid in the mesopelagic layer (Hardman-Mountford et al., 2008). Nonetheless, strengths and distribution of the boundaries between provinces are less pronounced than in the surface (Reygondeau et al., 2013) and fluctuate greatly with depth (Figure 4). This suggests that while the shallowest part of the mesopelagic zone is directly influenced by the same environmental parameters (temperature, salinity, primary production) and thus shows a similar spatial division as the epipelagic zone, the deepest part of the mesopelagic partition is driven by another set of variables (nutrients and oxygen concentrations) (Reygondeau et al., 2017). This is particularly marked when examining MBGCP distribution along zonal transects. Zonal transects reveal that two mesopelagic biotopes can co-exist in similar regions at different depths (Figure 4). This result shows clearly that meridional transects have strong vertical homogeneity, whereas zonal transects have strong horizontal homogeneity, emphasizing that the change in the environmental partition and influence of variables with depth need to be taken into account in further biogeographical studies.

Our new mesopelagic biogeochemical provinces likely influence the boundaries of living space for different mesopelagic biota and concomitantly mesopelagic biodiversity. Given the undersampling of the mesopelagic biota (Webb, Vanden Berghe, & O'Dor, 2010), these new biogeochemical provinces could potentially be used to estimate the geographical ranges of many mesopelagic species and functional groups (Dawson et al., 2013). New studies, which map the known distribution of mesopelagic biota relative to habitat requirements, would indicate the extent to which these new mesopelagic provinces also correspond to the distribution of mesopelagic communities and their biodiversity. One might hypothesize that the global species richness and biodiversity of the mesopelagic layer could be structured according to the new provincial biogeochemical boundaries, and the number and size of the different provinces. The proposed 3-D definitions, their abiotic properties and volume estimates open new possibilities for testing and understanding macro-ecological patterns of biodiversity distribution in the global ocean as well as defining a geographical framework for conservation management (Sutton et al., 2017).

Our approach has used available annual climatologies to partition the mesopelagic ocean. The procedure aggregates temporal variability in the mesopelagic boundaries and therefore proposes a baseline against which future changes may be measured. However, major changes in ocean circulation, such as variations in the global Meridional Overturning Circulation or seasonal oceanic currents, would probably affect the distribution of the boundaries between provinces in a similar way as in the epipelagic zone. New studies that will assess past interannual and seasonal temporal variability of the vertical boundaries of the zone and its associated MBGCPs, and how they might change under various future scenarios of global climate change and anthropogenic ocean use, are needed to improve our understanding of the mesopelagic layer in ocean ecology and biogeochemistry.

ACKNOWLEDGEMENTS

The authors are grateful to Daniel Dunn, William Cheung, Louis Legendre, Lars Stemman and Peter Søgaard Jørgensen, who improved our understanding of the methodology and definition used in this study and the interpretation of the results. We thank the Danish National Research Foundation for support to the Center for Macroecology,

Evolution and Climate. Gabriel Reygondeau is also thankful for the Nereus program (Nippon Foundation) for their financial support.

ORCID

Gabriel Reygondeau http://orcid.org/0000-0001-9074-625X

REFERENCES

Angel, M. V., & Pugh, P. R. (2000). Quantification of diel vertical migration by micronektonic taxa in the Northeast Atlantic. *Hydrobiologia*, *440*, 161–179.

Antonov, J. I., Seidov, D., Boyer, T. P., Locarnini, R. A., Mishonov, A. V, Garcia, H. E., … Johnson, D. R. (2010). *World Ocean Atlas 2009, Volume 2: Salinity*. Washington, D.C.: U.S. Government Printing Office.

Beaugrand, G., Lenoir, S., Ibañez, F., & Manté, C. (2011). A new model to assess the probability of occurrence of a species, based on presence-only data. *MEPS*, *424*, 175–190.

Behrenfeld, M. J. (2010). Abandoning Sverdrup's Critical Depth Hypothesis on phytoplankton blooms. *Ecology*, *91*, 977–989.

Bertrand, A., Ballon, M., & Chaigneau, A. (2010). Acoustic observation of living organisms reveals the upper limit of the oxygen minimum zone. *PLoS ONE*, *5*, e10330.

Black, K. D., & Shimmield, G. B. (2003). *Biogeochemistry of marine systems*. Oxford, UK: Blackwell.

Briggs, N. T., Slade, W. H., Boss, E., & Perry, M. J. (2013). Method for estimating mean particle size from high-frequency fluctuations in beam attenuation or scattering measurements. *Applied Optics*, *52*, 6710–6725.

Costello, M. J., Coll, M., Danovaro, R., Halpin, P., Ojaveer, H., & Miloslavich, P. (2010). A census of marine biodiversity knowledge, resources, and future challenges. *PLoS ONE*, *5*, e12110.

Dawson, M. N., Algar, A. C., Antonelli, A., Dávalos, L. M., Davis, E., Early, R., … Katharine, M. A. (2013). An horizon scan of biogeography. *Frontiers of Biogeography*, *5*, 130–157.

De Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., & Ludicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal Geophysical Research C*, *109*, C12003.

Falkowski, P. G., Barber, R. T., & Smetacek, V. (1998). Biogeochemical controls and feedbacks on ocean primary production. *Science*, *281*, 200–206.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O. K., Zweng, M. M., & Johnson, D. R. (2010a). *World Ocean Atlas 2009, Volume 3: Dissolved oxygen, apparent oxygen utilization, and oxygen saturation*. Washington, D.C.: U.S. Government Printing Office.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Zweng, M. M., Baranova, O. K., & Johnson, D. R. (2010b). *World Ocean Atlas 2009, Volume 4: Nutrients (phosphate, nitrate, silicate)*. Washington, D.C.: U.S. Government Printing Office.

Guidi, L., Jackson, G. A., Stemmann, L., Miquel, J. C., Picheral, M., & Gorsky, G. (2008). Relationship between particle size distribution and flux in the mesopelagic zone. *Deep Sea Research Part I: Oceanographic Research Papers*, *55*, 1364–1374.

Guidi, L., Legendre, L., Reygondeau, G., Uitz, J., Stemmann, L., & Henson, S. A. (2015). A new look at ocean carbon remineralization for estimating deepwater sequestration. *Global Biogeochemical Cycles*, *29*, 1044–1059.

Guidi, L., Stemmann, L., Jackson, G. A., Ibañez, F., Claustre, H., Legendre, L., … Gorskya, G. (2009). Effects of phytoplankton community on production, size, and export of large aggregates: A world-ocean analysis. *Limnology and Oceanography*, *54*, 1951–1963.

Hardman-Mountford, N. J., Hirata, T., Richardson, K. A., & Aiken, J. (2008). An objective methodology for the classification of ecological pattern into biomes and provinces for the pelagic ocean. *Remote Sensing of Environment*, *112*, 3341–3352.

Hardy, A. C. (1956). *The open sea. Its natural history: The world of plankton*. London: Collins.

Hartigan, J. A., & Wong, M. A. (1979). A k-means clustering algorithm. *Journal of the Royal Statistical Society, Series C*, *28*, 100–108.

Hedgpeth, J. W., & Ladd, H. S. (1957). Treatise on marine ecology and paleoecology. APA. National Research Council (U.S.).

Henson, S. A., Sanders, R., & Madsen, E. (2012). Global patterns in efficiency of particulate organic carbon export and transfer to the deep ocean. *Global Biogeochemical Cycles*, *26*, GB1028.

Honjo, S., Manganini, S. J., Krishfield, R. A., & Francois, R. (2008). Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: A synthesis of global sediment trap programs since 1983. *Progress in Oceanography*, *76*, 217–285.

Hooker, S. B., Rees, N. W., & Aiken, J. (2000). An objective methodology for identifying oceanic provinces. *Progress in Oceanography*, *45*, 313–338.

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposium Quantitative Biology*, *22*, 415–427.

Iversen, M. H., Nowald, N., Ploug, H., Jackson, G., & Fischer, G. (2010). High resolution profiles of vertical particulate organic matter export off Cape Blanc, Mauritania: Degradation processes and ballasting effects. *Deep Sea Research Part I: Oceanographic Research Papers*, *57*, 771–784.

Kaiser, K., Jennings, S., Thomas, D. N., Barnes, D. K. A., Brierley, A. S., Polunin, N. V. C., … Williams, P. J. B. (2005). *Marine ecology: Processes, systems, and impacts*. New York: Oxford University Press.

Kaufman, L., & Rousseeuw, P. J. (1987). Clustering by means of medoids. In Y. Dodge (Ed.), *Statistical data analysis based on the norm and related methods* (pp. 405–416). North-Holland, Amsterdam: Faculty of Mathematics and Informatics, https://books.google.com.mx/books/about/Clustering_by_Means_of_Medoids.html?id=HK-4GwAACAAJ&redir_esc=y

Legendre, P., & Legendre, L. (1998). *Numerical Ecology*. Amsterdam: Elsevier Science BV.

Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., … Johnson, D. R. (2010a). *World Ocean Atlas 2009*. Washington, D.C.: U.S. Government Printing Office.

Locarnini, R. A., Mishonov, A. V, Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., … Johnson, D. R. (2010b). *World Ocean Atlas 2009 Volume 1: Temperature*. Washington, D.C.: U.S. Government Printing Office.

Longhurst, A. (1995). Seasonal cycles of pelagic production and consumption. *Progress in Oceanography*, *36*, 77–167.

Longhurst, A. (1998). *Ecological geography of the sea* (1st edn). London: Academic Press.

Longhurst, A. (2007). *Ecological geography of the sea*. London: Academic Press.

Longhurst, A. R. (2010). *Ecological geography of the sea*. Academic Press.

Mahalanobis, P. C. (1936). On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, *2*, 49–55.

Martin, J. E. (1987). Slow aggregation of colloidal silica. *Physical Review A*, *36*, 3415–3426.

Martin, P., Lampitt, R. S., Perry, J. M., Sanders, R., Lee, C., & D'Asaro, E. (2011). Export and mesopelagic particle flux during a North Atlantic spring diatom bloom. *Deep Sea Research Part I: Oceanographic Research Papers*, *58*, 338–349.

McDonnell, A. M. P., & Buesseler, K. O. (2010). Variability in the average sinking velocity of marine particles. *Limnology and Oceanography*, *55*, 2085–2096.

Meyer, F. (1994). Topographic distance and watershed lines. *Signal Processing*, *38*, 113–125.

Moore, T. S., Campbell, J. W., & Feng, H. (2002). A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *39*, 1764–1776.

Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., & Franz, B. A. (2007). Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. *Remote Sensing of Environment*, *111*, 69–88.

Oliver, M. J., Glenn, S., Kohut, J. T., Irwin, A. J., Schofield, O. M., Moline, M. A., & Bissett, W. P. (2004). Bioinformatic approaches for objective detection of water masses on continental shelves. *Journal of Geophysical Research*, *109*, C07S04.

Oliver, M. J., & Irwin, A. J. (2008). Objective global ocean biogeographic provinces. *Geophysical Research Letters*, *35*. https://doi.org/10.1029/2008gl034238

Owens, S. A., Pike, S., & Buesseler, K. O. (2015). Thorium-234 as a tracer of particle dynamics and upper ocean export in the Atlantic Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, *116*, 42–59.

Pérès, J. M. (1961). *la vie benthique*. Paris: Presses Universitaires de France.

Pérès, J. M., & Devèze, L. (1963). *la vie pélagique*. Paris: Presses Universitaires de France.

Proud, R., Cox, M. J., & Brierley, A. S. (2017). Biogeography of the global ocean's mesopelagic zone. *Current Biology*, *27*(1), 113–119.

Quackenbush, J. (2001). Computational analysis of microarray data. *Nature Reviews Genetics*, *2*, 418–427.

Reygondeau, G., Guieu, C., Benedetti, F., Irisson, J. O., Ayata, S. D., Gasparini, S., & Koubbi, P. (2017). Biogeochemical regions of the Mediterranean Sea: An objective multidimensional and multivariate environmental approach. *Progress in Oceanography*, *151*, 138–148.

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., & Maury, O. (2013). Dynamic biogeochemical provinces in the global ocean. *Global Biogeochemical Cycles*, *27*, 1046–1058.

Reygondeau, G., Maury, O., Beaugrand, G., Fromentin, J. M., Fonteneau, A., & Cury, P. (2012). Biogeography of tuna and billfish communities. *Journal of Biogeography*, *39*, 114–129.

Robinson, C., Steinberg, D. K., Anderson, T. R., Arístegui, J., Carlson, C. A., Frost, J. R., … Koppelmann, R. (2010). Mesopelagic zone ecology and biogeochemistry—A synthesis. *Deep Sea Research Part II: Topical Studies in Oceanography*, *57*, 1504–1518.

Sarmiento, J., & Gruber, N. (2006). *Ocean biogeochemical dynamics*. Princeton, NJ: Princeton University Press.

Sathyendranath, S., Longhurst, A., Caverhill, C. M., & Platt, T. (1995). Regionally and seasonally differentiated primary production in the north Atlantic. *Deep-Sea Research part I*, *42*, 1773–1802.

Sherman, K. (2005). The large marine ecosystem approach for assessment and management of ocean coastal waters. In T. Hennessey, & J. Sutinen (Eds.), *Sustaining large marine ecosystems: the human dimension* (pp. 3–16). Amsterdam: Elsevier.

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M. A. X., … Lourie, S. A. (2007). Marine ecoregions of the World: A bioregionalization of coastal and shelf areas. *BioScience*, *57*, 573–583.

Steinberg, D. K., Cope, J. S., Wilson, S. E., & Kobari, T. (2008). A comparison of mesopelagic mesozooplankton community structure in the subtropical and subarctic North Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, *55*, 1615–1635.

Stemmann, L., Youngbluth, M., Robert, K., Hosia, A., Picheral, M., Paterson, H., … Gorsky, G. (2008). Global zoogeography of fragile macrozooplankton in the upper 100-1000 m inferred from the underwater video profiler. *ICES Journal of Marine Science*, *65*, 433–442.

Stewart, G. M., Moran, S. B., & Lomas, M. W. (2010). Seasonal POC fluxes at BATS estimated from 210 Po deficits. *Deep Sea Research Part I: Oceanographic Research Papers*, *57*, 113–124.

Stramma, L., Johnson, G. C., Sprintall, J., & Mohrholz, V. (2008). Expanding oxygen-minimum zones in the tropical oceans. *Science*, *320*, 655–658.

Sutton, T. T., Clark, M. R., Dunn, D. C., Halpin, P. N., Rogers, A. D., Guinotte, J., … Heino, M. (2017). A global biogeographic classification of the mesopelagic zone. *Deep Sea Research I*, *126*, 85–102.

Sverdrup, H. U. (1953). On conditions for the vernal blooming of phytoplankton. *ICES Journal of Marine Science*, *18*, 287–295.

Watling, L., Guinotte, J., Clark, M. R., & Smith, C. R. (2013). A proposed biogeography of the deep ocean floor. *Progress in Oceanography*, *111*, 91–112.

Webb, T. J., Vanden Berghe, E., & O'Dor, R. (2010). Biodiversity's big wet secret: The global distribution of marine biological records reveals chronic under-exploration of the deep pelagic ocean. *PLoS ONE*, *5*, e10223.

Woolley, S. N., Tittensor, D. P., Dunstan, P. K., Guillera-Arroita, G., Lahoz-Monfort, J. J., Wintle, B. A., … O'Hara, T. D. (2016). Deep-sea diversity patterns are shaped by energy availability. *Nature*, *533*, 393–396.

Yeung, K. Y., Haynor, D. R., & Ruzzo, W. L. (2001). Validating clustering for gene expression data. *Bioinformatics*, *17*, 309–318.

BIOSKETCH

Gabriel Reygondeau is a senior fellow at the University of British Columbia in the research team 'Changing Ocean Research Unit'. He mainly focuses on the effect of climate change and fisheries on ecosystem dynamic and biodiversity at a global scale.

Email: gabriel.reygondeau@gmail.com

Research team: Changing Ocean Research Unit

The objectives are to study the impact of global change on marine ecosystems, their governance and their exploitation and plan out evolution scenarios with the help of patterns and empirical analysis.

Website: www.nereusprogram.org

Author contributions: G.R. and O.M. conceived the idea; G.R., L.G. and S.H. collected the data; G.R. analysed the data; G.R. led the writing, in collaboration with all co-authors.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Reygondeau G, Guidi L, Beaugrand G, et al. Global biogeochemical provinces of the mesopelagic zone. *J Biogeogr*. 2017;00:1–15. https://doi.org/10.1111/jbi.13149