Estimation of seasonal and inter-annual variations in biomass and phytoplankton composition in the Southern Ocean over the past two decades and evaluation of their ecological consequences

Loïc Le Ster

Sorbonne Université

École doctorale des Sciences de l'Environnement d'Ile de France (SEIF) N°129

Centre d'Études Biologiques de Chizé (CEBC)
Équipe Prédateurs Marins

Laboratoire d'Océanographie de Villefranche (LOV)
Équipe Optique Marine et Télédétection : Applications à la Biogéochimie

Estimation des variations saisonnières et interannuelles de la biomasse et de la composition en phytoplancton du secteur indien de l'Océan Austral sur les deux dernières décennies et évaluation de leurs conséquences écologiques

Par Loïc Le Ster

Thèse de doctorat
Discipline: Océan Atmosphère, Climat et Observations Spatiales
Spécialité: Océanographie

Dirigée par Christophe Guinet et Hervé Claustre

Présentée et soutenue publiquement le 5 octobre 2023

Devant un jury composé de :
Shubha Sathyendranath - *Rapportrice*
Cédric Jamet - *Rapporteur*
Marina Lévy - *Examinatrice*
Anne Petrenko - *Examinatrice*
Stéphane Blain - *Président du jury*
Christophe Guinet - *Directeur de thèse*
Hervé Claustre - *Directeur de thèse*

Sorbonne Université

École doctorale des Sciences de l'Environnement d'Ile de France (SEIF) N°129

Centre d'Études Biologiques de Chizé (CEBC)
Marine Predators Team

Laboratoire d'Océanographie de Villefranche (LOV)
Marine Optics & Remote Sensing Group

Estimation of seasonal and inter-annual variations in biomass and phytoplankton composition in the Southern Ocean over the past two decades and evaluation of their ecological consequences

by Loïc Le Ster

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY
Discipline: Ocean Atmosphere, Climate and Spatial Observations
Speciality: Oceanography

Directed by Christophe Guinet and Hervé Claustre

Presented and publicly defended on 5th October 2023

Before a jury composed of:
Shubha Sathyendranath - *Rapporteur*
Cédric Jamet - *Rapporteur*
Marina Lévy - *Examiner*
Anne Petrenko - *Examiner*
Stéphane Blain - *Examiner*
Christophe Guinet - *Thesis supervisor*
Hervé Claustre - *Thesis supervisor*

Résumé

L'objectif principal de cette thèse est d'examiner les échelles de temps associées aux variations de la concentration en phytoplancton dans l'Océan Austral et de fournir des données clés pour comprendre comment les forçages externes conditionnent la croissance du phytoplancton à différentes échelles spatio-temporelles. En outre, cette thèse ouvre la voie à l'évaluation de l'influence des variations de production primaire sur l'écosystème. Le premier objectif de cette thèse visait à isoler les différentes échelles temporelles de variation présentes dans les séries temporelles à long terme de concentration en chlorophylle a estimées à partir de données satellite. Le premier objectif consistait spécifiquement à caractériser les différents types de phénologies phytoplanctoniques dans l'Océan Austral, à étudier leur distribution spatiale ainsi qu'à évaluer de leur évolution dans le temps. Deuxièmement, les travaux de cette thèse se sont concentrés sur le secteur Indien de l'Océan Austral afin d'étudier plus spécifiquement les liens entre la dynamique de la production primaire et le fonctionnement de l'écosystème. Dans le deuxième objectif, nous avons défini des biorégions basées sur la phénologie la chlorophylle a telle qu'estimée par satellite. La dymanique de ces biorégions a été dans un premier temps étudiée en termes de composition de la communauté phytoplanctonique à partir d'un modèle satellitaire. Puis des données caractérisant le comportement de recherche de nourriture d'un prédateur supérieur - l'éléphant de mer du Sud (Mirounga leonina) ont permis d'étudier la dynamique des biorégions aux niveaux trophiques supérieurs. Enfin, dans un troisième objectif, nous avons proposé une méthode basée sur une variable optique de base - l'atténuation de la lumière dans la colonne d'eau - pour améliorer la précision et la résolution spatiale des mesures de fluorescence de la chlorophylle a dérivées du bio-logging. Dans le contexte de missions satellitaires à venir focalisées sur l'observation à haute résolution de circulation océanique, la méthode que nous présentons est prometteuse pour l'étude des interactions bio-physiques à (sub)mésoéchelle dans l'océan Austral.

Abstract

The main objective of this thesis is to examine the timescales associated with variations in phytoplankton concentration in the Southern Ocean and provide key data to understand how external forcing conditions phytoplankton growth at various spatio-temporal scales. Furthermore, this thesis opens the way to the assessment of the influence of variations in primary production on the ecosystem. The first objective of this thesis was to disentangle the different timescales of variation present in long-term time series of chlorophyll-a concentration estimated from satellite data. The first objective was specifically to characterize the different types of phytoplankton phenologies in the Southern Ocean, to study their spatial distribution and to assess their evolution in time. Secondly, the work in this thesis focused on the Indian sector of the Southern Ocean to study more specifically the linkages between primary production dynamics and ecosystem functioning. In the second objective, we defined bioregions based on phenology of satellite chlorophyll-a concentration estimates. The dynamics of these bioregions were first studied in terms of phytoplankton community composition using a satellite model. Then, indicators derived from tracking data of a top predator - the southern elephant seal (*Mirounga leonina*) - were used to infer potential influence of variations in primary production on higher trophic levels. Finally, in a third objective, we proposed a method based on a cost-effective optical variable - namely, light attenuation in the water column - to improve accuracy and spatial resolution of bio-logging-derived chlorophyll a fluorescence measurements. In the context of upcoming high-resolution ocean-circulation satellite missions, the method shows promising for the studying of bio-physical interactions at (sub)mesoscale in the Southern Ocean.

Acknowledgments

Firstly, I would like to thank my thesis rapporteurs Shubha Sathyendranath and Cédric Jamet for accepting to review this manuscript. I thank my thesis jury members Marina Lévy, Anne Petrenko and Stéphane Blain for accepting to evaluate my work.

I would like to thank my thesis supervisors, Christophe Guinet and Hervé Claustre for offering me this great opportunity to dive into a research project, for guiding on my path during my thesis, stimulating my curiosity and sharing with me their knowledge and experience.

This thesis was partly funded by the Centre National d'Études Spatiales (CNES) and REFINE (European Research Council, Grant agreement 834177) project. This thesis was supported by a joint CNES-CNRS doctoral grant. The elephant seal work was supported as part of the SNO-MEMO and by the CNES-TOSCA project Elephant seals as Oceanographic Samplers of submesoscale features led by C. Guinet with support of the French Polar Institute (programmes 109 and 1201). The research presented in this manuscript was carried out at the Centre d'Études Biologiques de Chizé (CEBC) and at the Laboratoire d'Océanographie de Villefranche (LOV).

I thank the CEBC and the Institut de la Mer de Villefranche (IMEV) for hosting my thesis and providing me with an ideal environment in which to advance in my research, acquire new skills and meet extraordinary colleagues.

I thank my colleagues in the Prédateurs Marins team at the CEBC and in the OMTAB team at the LOV for all the fruitful professional interactions and friendly moments that made this experience unforgettable.

I would like to thank all my "Mission 71" and "Mission 72" colleagues met at Kerguelen Islands where I worked on the field from August to December 2021. The experience of working in such an extreme environment, discover such a remote and wild world is exceptional. Particularly, I would like to thank Christophe Guinet for giving me the opportunity to take part to the mission and share his knowledge, passion and skills about this fascinating region of the world.

Last but not least, I would like to thank my friends and family for their unconditional support throughout this PhD project, and for giving me the motivation and strength I needed to complete the amazing journey of this thesis.

Table of Contents

General Introduction

Contents of General Introduction

The Southern Ocean (SO): A key component of the Earth climate system

Connected to the Atlantic, Pacific and Indian Oceans, the Southern Ocean (SO) is a major component of the Earth system. The SO is an essential link in the closure of the global overturning circulation (Talley et al., 2011; Marshall & Speer 2012; Talley, 2013; Rintoul, 2018), as represented in the schematic of Figure 1, adapted from Talley et al., 2011. The circulation scheme of the SO influences the global heat balance and biogeochemical cycles in the world's oceans (Toggweiler & Samuels, 1995; Rintoul et al., 2011; Morrison et al., 2015). Indeed, SO ventilation processes (Sallée et al., 2010; Morrison et al., 2022) are key drivers of air-sea CO_2 fluxes in the SO (Bushinsky et al., 2019), which prove to be strongly supported by biological activity (Huang et al., 2023). Therefore, SO processes are central to the climate system and play a key role in the context of climate change (Williams et al., 2023).

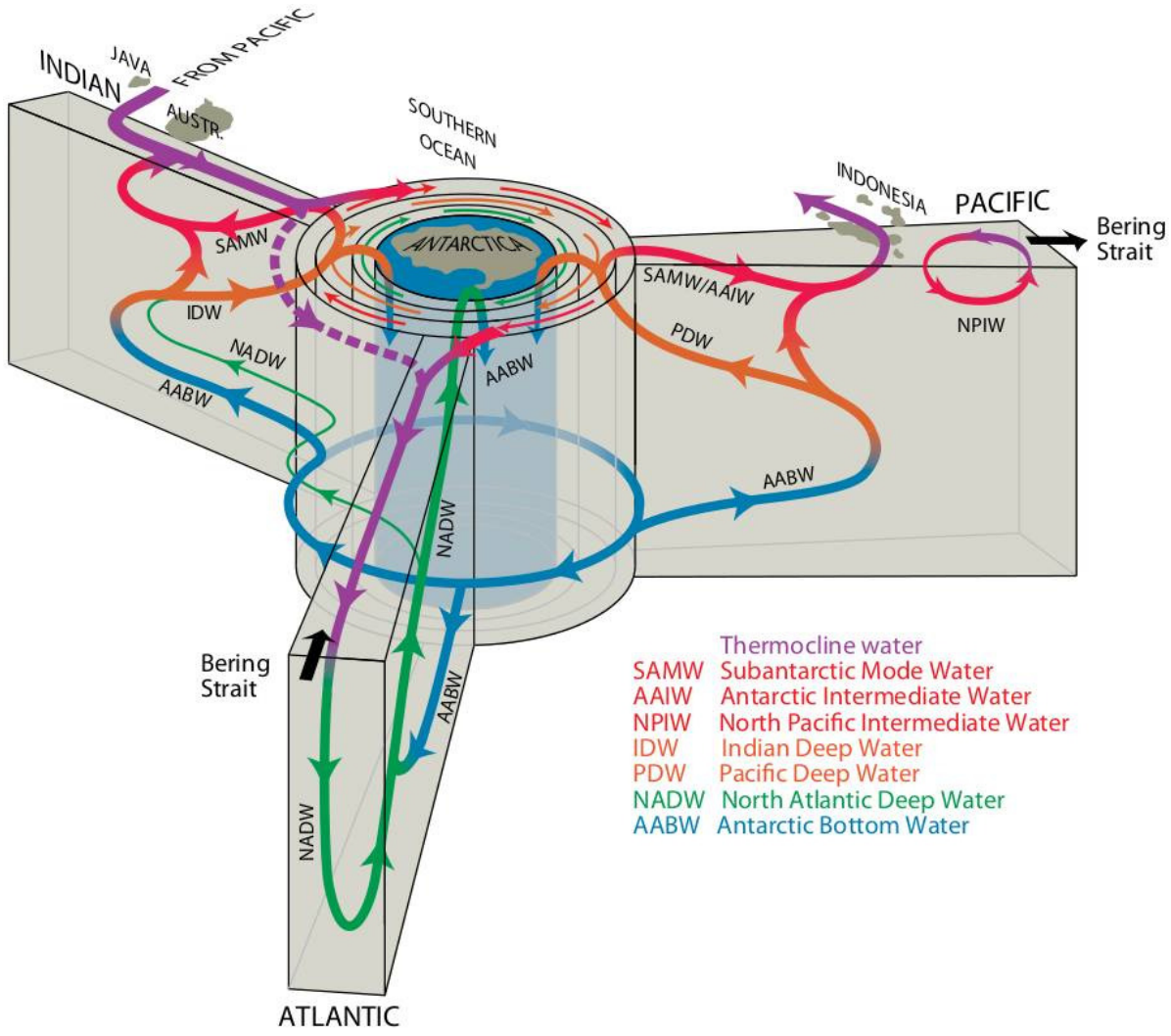


Figure 1: Schematic of the overturning circulation from a Southern Ocean perspective (from Talley et al., 2011).

SO processes affected by climate change

Revealed over the past few decades, significant changes attributed to human activities are currently critically affecting the SO (Fyfe, 2006; Sigmond et al., 2011; Shi et al., 2018; Swart et al., 2018). These changes include warming of the SO (Gille, 2008; Sallée et al., 2018; Shi et al., 2021), with direct consequences on SO heat uptake (Shi et al., 2018; Bourgeois et al., 2022). However, the large contribution of the SO to global ocean heat uptake (Frölicher et al., 2015; Shi et al., 2018) is crucial in the context of climate change. The SO is also the main oceanic sink of CO_2 (Mikaloff Fletcher et al., 2006; Gruber et al., 2019a). Yet, the SO carbon cycle is affected by climate change through modification of the ocean's physico-chemical properties and biological activity (Sarmiento et al., 1998; Le Quere et al., 2007; Hauck & Volker, 2015; Hauck et al., 2015; Bourgeois et al., 2022; Williams et al., 2023). Finally, Antarctic sea ice might as well be impacted by SO warming (Eayrs et al., 2021), with potential large-scale consequences on water masses properties (Sadai et al., 2020), and hence, on heat uptake and ocean circulation.

Focusing attention on the circulation scheme of the SO

The circulation scheme of the SO (Figure 2) is dominated by the Antarctic Circumpolar Current (ACC) system (Rintoul et al., 2001; Sokolov & Rintoul 2009a). Wind forcing by Southern Hemisphere westerlies is considered the main driver of the circumpolar eastward ACC flow (Wunsch, 1998; Lovenduski & Gruber, 2005; Allison et al., 2010). The Southern Annular Mode (SAM), the leading mode of climate variability in the Southern Hemisphere (Thompson & Wallace, 2000), is closely related to the intensity of winds in the Southern Hemisphere. Positive phases of the SAM (or Antarctic Oscillation, AAO) are associated with significant strengthening of the westerlies (Lovenduski & Gruber, 2005). Changes in SAM (and/or in the El Niño-Southern Oscillation, ENSO, see Sallée et al., 2008) therefore strongly influence the position of ACC fronts (Sallée et al., 2008; Gille, 2008; Sokolov & Rintoul 2009b). Influence of SAM on SO physical processes was even described at smaller scale through impact on eddy activity (Meredith & Hogg, 2006). Climate change-induced modifications of the ACC were recently pointed out, as a consequence of both an intensification of westerly winds (Böning et al., 2008; Beech et al., 2022; Behrens & Bostock, 2023) and ocean warming (Toggweiler & Russell, 2008; Shi et al., 2021). The key role of ACC fronts in physical processes and ocean biology in the SO (Lovenduski & Gruber, 2005; Chapman et al., 2020) emphasizes the critical link between atmospheric variability, ocean circulation and ecosystem functioning (Figure 3) in the current context of a changing climate system.

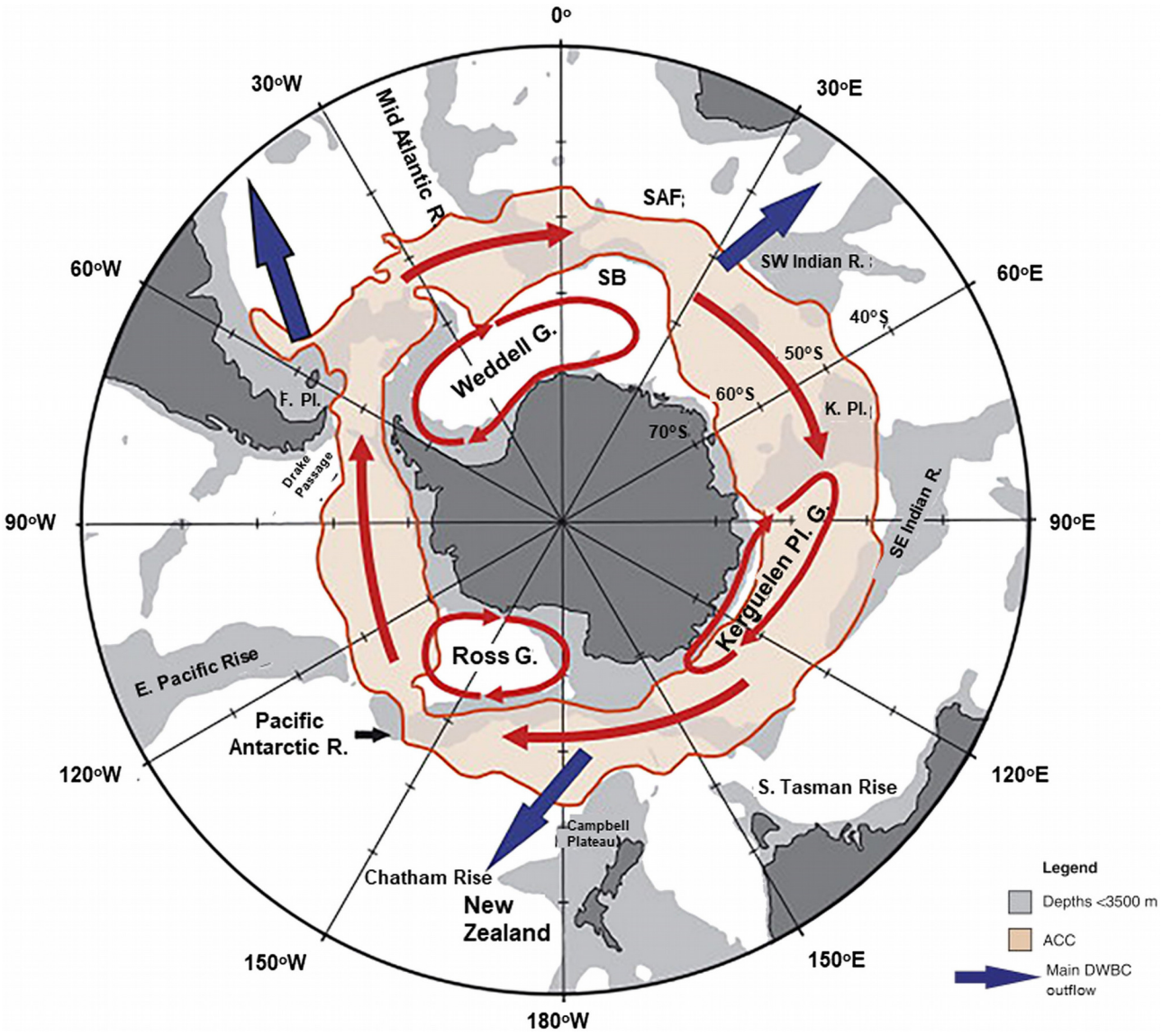


Figure 2: Main oceanographic features in the Southern Ocean (from Vernet et al., 2019). ACC: Antarctic Circumpolar Current; SAF: Subantarctic Front; SB: southern limit of Upper Circumpolar Deep Water or Southern Boundary; DWBC: Deep Western Boundary Currents; R.: ridge; K. Pl.: Kerguelen Plateau; F. Pl.: Falkland/Malvinas Plateau; and G.: gyre.

SO ecosystems: a compromised equilibrium

On global scales, SO processes largely support the functioning of marine ecosystems due to the worldwide connectivity of SO ecosystems (Murphy et al., 2021). Notably through the transport of nutrients to other regions of the world's oceans (Sarmiento et al., 2004; Vernet et al., 2019), the SO circulation scheme is key in sustaining a large range of ecosystems. In addition, SO ecosystems rely on a large (threatened) SO biodiversity (Ropert-Coudert et al., 2014; Chown et al., 2015;

Hughes et al., 2020). However, SO warming and changes in Antarctic sea ice cover compromise the equilibrium of SO ecosystems.

In the sea ice zone, breeding success of species like Adélie penguin (*Pygoscelis adeliae*) highly depend on seas ice cover (Barreau et al., 2019; Watanabe et al., 2020). Antarctic birds like the Adélie penguin or Antarctic seals like the Weddell seal (*Leptonychotes weddellii*) are considered sentinel species as changes in their breeding success, foraging effort or body condition can be used to study occurring environmental changes (Ropert-Coudert et al., 2015; Wege et al., 2021). The impact of climate warming on populations of antarctic birds has been clearly demonstrated (Forcada et al., 2006; Barbraud & Weimerskirch, 2006; Ropert-Coudert et al., 2015). Although poorly known, the sea ice zone also hosts a large phytoplankton biomass, that supports ecosystem functioning and has a significant role in sea ice biogeochemical cycles (Taylor et al., 2013; Vancoppenolle et al., 2013; Deppeler & Davidson, 2017; Moreau et al., 2020). As a consequence of warming in the sea ice zone, modified (enhanced) primary production has already been evidenced (Moreau et al., 2015).

In the open ocean, the effects of climate change also impact the ecosystem at various levels, from primary production (Deppeler & Davidson, 2017; Pinkerton et al., 2021; Henson et al., 2021), to zooplankton (Loeb & Santora 2015; Johnston et al., 2022) up to top predators (e.g. Volzke et al. 2021). Due to the importance of synchronization processes between functional groups within ecosystems (Edwards & Richardson, 2004), any impact on one compartment of the trophic network has ecosystem-wide consequences. Studying ecosystem changes through analysis of variability in the areal extent of biomes (Boyd & Doney, 2002; Sarmiento et al., 2004; Fay & McKinley, 2014) is an efficient way to identify both pivotal regions and key drivers of ecosystem changes.

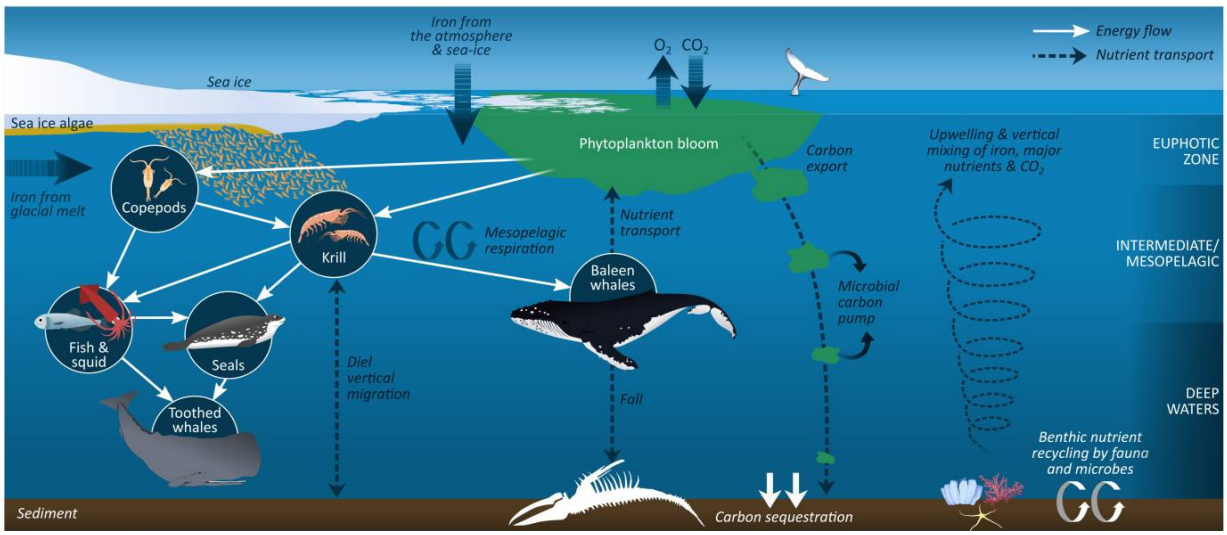


Figure 3: Key biogeochemical and ecosystem processes in the Southern Ocean (from Henley et al., 2020)

Drivers of ecosystem changes

The drivers of ecosystem changes in the SO are linked to changing physical environment as well as direct pressure from human activities (Morley et al., 2020). As previously mentioned, factors linked to changes in the physical environment mainly include increased heat content in the SO due to warming, alteration of SO circulation through intensification of westerly winds, and modification of water properties due to increased sea ice melt. Direct pressures by human activities are mainly linked to increased marine traffic and intensification of fishing effort enabled by the weakening of ocean barriers in the SO (Morley et al., 2020). The demand for activities related to tourism might also increase in the future (Bender et al., 2016; Rogers et al, 2020; Cavanagh et al., 2021).

As a consequence, a long list of ecosystem services, including food security, climate regulation and recreation (Murphy & Hofmann, 2012; Murphy et al., 2021; Cavanagh et al., 2021), with large socio-economic impact on human societies, are likely to be affected by climate change (Rogers et al., 2020; McCormack et al., 2021; Steiner et al., 2021; Cavanagh et al., 2021).

Phytoplankton: a pivotal role

As a consequence, through modification of habitats, concurrent pressure on resources, introduction of non-indigenous species and finally, transport of anthropogenic pollutants, SO marine biota might be critically affected in the future (Constable et al., 2014; Poloczanska et al., 2016; Morley et al, 2020). Yet, the response of SO ecosystems to climate change largely depends on how primary production will be affected by future changes (Deppeler & Davidson, 2017; Henley et al., 2020). Indeed, phytoplankton is at the base of antarctic food webs (Deppeler & Davidson, 2017) and plays a pivotal role in SO biogeochemical cycles (Arrigo et al., 2008; Henley et al., 2020). A key driver of the spatio-temporal variability of SO biogeochemistry, the biological carbon pump (BCP) is the oceanic process through which - by gravity, diurnal and seasonal vertical migration of living organisms or advection by water masses - the organic matter is transferred to depth (Volk & Hoffert, 1985; Boyd et al., 2019; Siegel et al., 2023). Primary producers, by converting inorganic carbon into organic carbon, hence have an essential function in the BCP. However, the little agreement in climate models projections regarding sensitivity of the SO BCP to climate change reveals our lack of understanding in the sequence of processes that drive the BCP (Thomalla et al., 2023).

Therefore, to address future challenges implied by the strong environmental changes that deeply affect the Earth system and human societies, monitoring phytoplankton is essential to (i) better describe the link between SO processes and primary production, and (ii) assess the future response of phytoplankton to climate change in the SO.

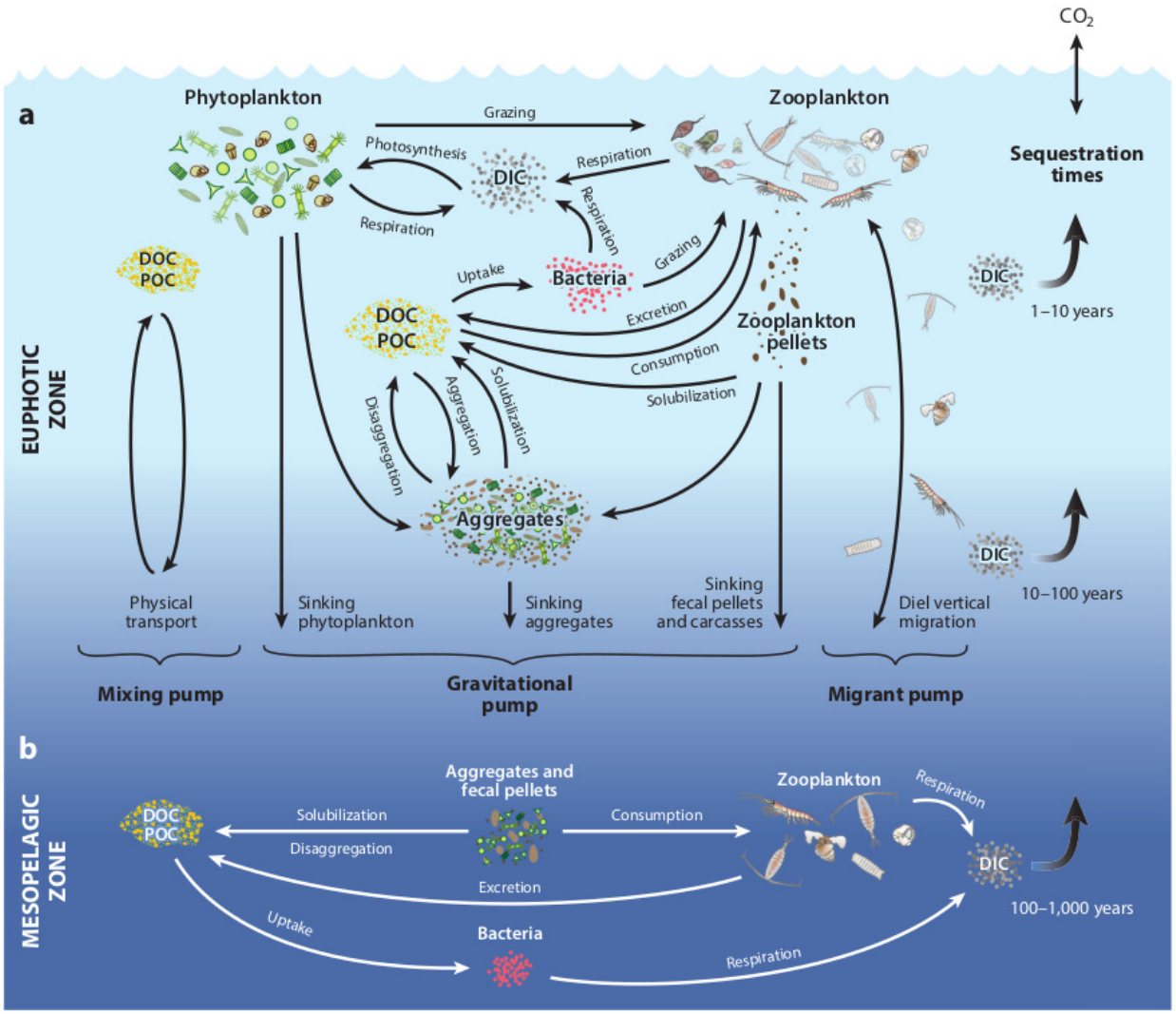


Figure 4: Food-web diagram illustrating the three export pathways of the biological pump, their regulating processes, and the timescales for carbon sequestration (a) in the euphotic zone, (b) beneath the euphotic zone (from Siegel et al., 2023).

Monitoring phytoplankton concentration

Phytoplankton concentration is commonly monitored through detection of chlorophyll-a, an essential pigment involved in photosynthesis (Butler, 1978). High Pressure Liquid Chromatography (HPLC) is a reliable method to assess phytoplankton pigment composition in a water sample, and hence, chlorophyll-a concentration.

In addition, detection of chlorophyll-a fluorescence - the fluorescence emitted by chlorophyll-a pigments following light excitation - is widely used as a non-destructive tool to estimate *in-situ* sea water chlorophyll-a concentration (Lorenzen, 1966; Huot & Babin, 2010; Roesler & Barnard, 2013). Therefore, although chlorophyll-a fluorescence is not directly proportional to phytoplankton biomass in all conditions (e.g. Serôdio and Lavaud, 2011; Xing et al.,

2012; Schallenberg, et al., 2022), it is commonly and widely used as a proxy for phytoplankton concentration to monitor ocean biogeochemistry at large scale (e.g. Chai et al., 2020).

The content in chlorophyll-a in the ocean can also be assessed from space with ocean-colour data. The remote sensing reflectance of sea water measured by ocean-colour sensors enable retrieval of chlorophyll-a concentration (Morel & Gordon, 1980; Morel & Gentili, 1996). However, satellite-based estimates of chlorophyll-a concentration (CHL) are only representative of ocean the surface layer (Gordon and McCluney, 1975; Morel, 1988), generally assumed to be ~15-20 m thick.

Chlorophyll-a fluorescence: a signal with multiple time scales of variation

Estimates of chlorophyll-a concentration from ship-based measurements combined with remote sensing data (e.g. Blain et al., 2008), augmented by the large development of autonomous platforms equipped with a fluorescence sensor (e.g. Blain et al., 2013; Guinet et al., 2013; Testor et al., 2019; Claustre et al., 2020; Chai et al., 2020), enable observation of the spatio-temporal variability in chlorophyll-a concentration in the SO at various time scales. Describing phytoplankton dynamics at suitable spatio-temporal resolution is crucial to dissociate spatial variability from temporal variability and examine the linkages between SO processes and primary production. In the SO, time scales of variation in phytoplankton concentration extend from paleoclimatic (Hillenbrand & Cortese, 2006) to sub-seasonal (e.g. Pellichero et al., 2020). In addition, the fluorescence signal emitted by phytoplankton cells naturally varies - spatially and temporally - according to species composition and physiological state (Bricaud et al., 2004). In particular, physiological processes involved in phytoplankton photoacclimation can occur on time scales of seconds to minutes (Li et al., 2009). Consequently, due to the complexity of the measured chlorophyll-a fluorescence signal, disentangling the time scales of variation in phytoplankton concentration in the SO constitutes a real challenge.

Objectives of the study

The objectives of the present study are:

(i) examine the time scales of variation in phytoplankton concentration in the SO,

(ii) assess the propagation of variations in primary production throughout the ecosystem,

(iii) provide key data to understand how external forcing conditions phytoplankton growth at various spatio-temporal scales.

For this purpose, we propose to analyze variations in chlorophyll-a fluorescence at specific spatio-temporal scales, suitable to examine the linkages between phytoplankton dynamics and oceanic processes, and their influence at ecosystem level. In Chapter I, we firstly investigate multi-

annual variations in satellite-based estimates of chlorophyll-a concentration (CHL) in the SO. We show that the confidence level associated with estimates of long-term CHL biomass trends in the SO is low. We then focus on smaller time scales of variation through examination of CHL phenology. We show that besides seasonal cycles, sub-seasonal variations are determinant in shaping annual cycles of CHL, giving high importance to (sub)mesoscale processes. In Chapter II, we propose a bioregionalization of the SO to investigate at regional scale the linkages between CHL biomass, CHL phenology and phytoplankton community composition. We then use indicators derived from tracking data of a top predator - the southern elephant seal (*Mirounga leonina*) - to infer potential influence of variations in primary production at ecosystem level. In Chapter III, we show the promising potential of data derived from bio-logging devices placed on marine mammals for applications linked to the study of the interaction between physical and biological processes at (sub)mesoscale.

Chapter I

Analysis of Phytoplankton Phenology in the Southern Ocean from Ocean-Colour Data

Contents of Chapter I

I.1. Introduction

The Southern Ocean (SO) is a major contributor to global carbon export (Sabine et al., 2004; Khatiwala et al., 2009; Gruber et al., 2019; Long et al., 2021), to which phytoplanktonic organisms critically contribute through their role in the biological carbon pump (Boyd et al., 2019; Siegel et al., 2021; Thomalla et al., 2023). Moreover, primary production has strong implication in SO ecosystem functioning (Deppeler & Davidson; Henley et al., 2020; McCormack et al., 2021). Long term monitoring of phytplankton in the SO is hence crucial to study climate change impacts and assess its role in future climate scenarios (Hauck & Volker, 2015; Hauck et al., 2015; Krumhardt et al., 2022).

Ocean-colour sensors use remote sensing reflectance to infer chlorophyll-a concentration (CHL) from space, a proxy for phytoplankton concentration. With first deployment launched in 1978, satellite ocean-colour (OC) sensors have proved essential for monitoring phytoplankton distribution at large scale (McClain et al., 2009). Nonetheless, achieving exploitable long term time series of remotely-sensed CHL is constrained by discontinuities due to successive generations of OC sensors (Pope et al., 2017; Henson et al., 2017). Disentangling discontinuities due to OC sensor change from true ecological changes is hard to resolve. Some studies overcome the issue by using the data of only one specific sensor (e.g. Vantrepotte & Mélin, 2011) or by developing a specific algorithm per OC sensor “era” (e.g. Xi et al., 2020). The way discontinuities are handled is decisive in the results provided by long-term analysis, and notably, estimates of long term trends (Beaulieu et al. 2013; Saulquin et al., 2013; Mélin 2016; Gregg et al., 2017; Hammond et al., 2020; Lim et al., 2022). Evidence of instrumental issues for assessment of long term changes in CHL biomass is illustrated in Figure I.1. The use of two analogous ocean-colour CHL products, namely the products distributed by Copernicus Marine Service GlobColour and the European Space Agency Ocean Color Climate Change Initiative (ESA OC-CCI) (see Section I.2.d) reveal similar shape, but a different mean value (~15%), and above all, opposite long-term trends.

Consequently, attempting to detect and quantify raw biomass trends in the SO is subject to strong limitations due to the above-mentioned continuity issues. In the present study, we propose to circumvent the issue by characterizing CHL variations at shorter time scales (from seasonal to sub-seasonal) through phenology.

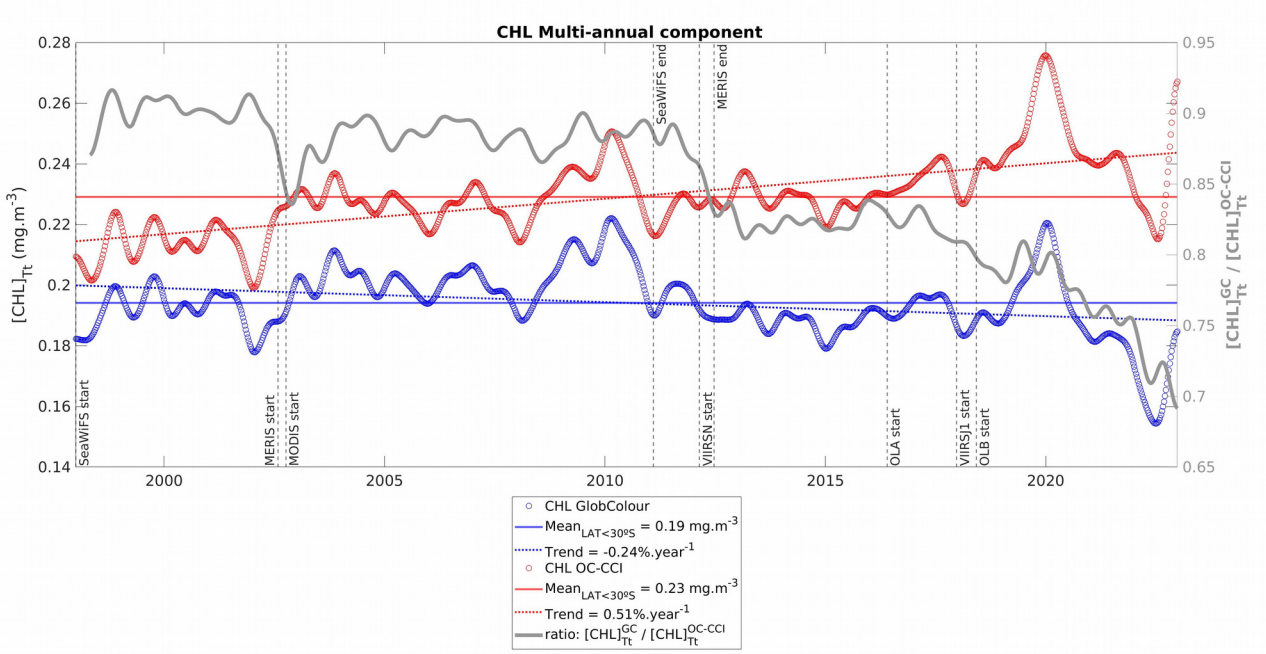


Figure I.1. Assessment of long term changes in CHL biomass. Comparison of the multi-annual component (see details in Section I.3.a) of two analogous ocean-colour CHL products, namely the products distributed by Copernicus Marine Service GlobColour (blue circles) and the European Space Agency Ocean Color Climate Change Initiative (OC-CCI, red circles) (see details in Section I.2.d). Continuous lines show climatological 1998-2022 value and dashed lines show long-term trends, with same color code as circles. Continuous gray line shows the ratio between CHL from GlobColour and CHL from OC-CCI. At each date, the data displayed in Figure I.1 corresponds to the mean value computed for the study zone (Southern Ocean, south of 30ºS). Trends in CHL biomass are estimated as differences in CHL per year and were only displayed for qualitative appreciation of diverging trends (no absolute values of trends were provided).

The study of the timing of periodic biological events in relation to their environmental conditions (i.e. phenology, see Schwarz et al., 2003; Ji et al, 2010) is a common-sense approach to detect repeatability as well as to track changes in the ecology of species (Rosenzweig et al., 2007). The presence of various phytoplankton phenologies across world's oceans is enabled at large scale by remote-sensing of CHL (e.g. Racault et al., 2011). Several studies characterized seasonal dynamics of phytoplankton from space using metrics derived from the annual cycle of CHL (e.g. Vargas et al., 2009) or relied on bloom phenology to define trophic regimes (e.g. D'Ortenzio & Ribera d'Alcalà, 2009, Ardyna et al., 2017). Phenology of phytoplankton is a key ecological indicator (Platt & Sathyendranath, 2008) to track climate change-driven trends (Henson et al., 2017). Phenology of phytoplankton is also a valuable indicator for marine resources management due to the trophic interactions (through synchronization) of the primary producers' compartment with higher trophic levels (Platt et al., 2003; Edwards & Richardson, 2004).

Many methods were proposed to “quantify” phytopklankton phenology with timing indices such as bloom initiation date, apex date, season duration (Ji et al., 2010). However important issues regarding the accuracy and precision of these timing indices were highlighted (Cole et al., 2012; Ferreira et al., 2014). Accuracy issues, notably strong in case of missing data (Brody et al., 2013), are central to the studying of phytoplankton phenology in the SO, where satellite coverage is highly constrained by the presence of clouds. As a response to the low accuracy of classical metrics at high latitudes (Cole et al., 2012), the use of functional data, by including (spatial or temporal) relations between variables, is particularly adapted in case of missing data (Ramsay & Silverman, 2005; Nerini et al., 2010). Here we propose a method based on functional data analysis (FDA) to characterize phytoplankton phenology in the SO.

In the present study, we demonstrate that studying CHL phenology appears as a robust way to investigate long-term time series in the presence of discrepancies at multi-annual scale. Firstly, we use a comparison between two different OC products (namely GlobColour and ESA OC-CCI), to emphasize discrepancies related to multi-annual variations (Section I.1). We then study seasonal to sub-seasonal time scales of variation through examination of CHL phenology and its evolution in time. The functional approach proposed in the present study reveals the main modes of variation of CHL phenology, from seasonal to sub-seasonal scales, their spatial distribution and their evolution in time (Section I.4). The main seasonal patterns identified in our study were validated through a series of case studies in the SO (Section I.5.a). The case studies scrutinized in our study also support the proposed underlying processes and the hypothesized role of (sub)mesoscale processes in sub-seasonal variations in CHL. We then investigate more closely the sub-seasonal component and discuss its influence on CHL phenology (Section I.5.b). Finally, we discuss long-term variations in CHL phenology and the potential induced implications for CHL biomass in the SO (Section I.5.c).

I.2. Data

I.2.a. Study period, area of interest and geographic sectors

The study period extends from 1998 to 2022. The area of interest of the present study is the Southern Ocean. The Southern Ocean was defined as the oceanic region located South of the 30ºS parallel. The region was divided into geographic sectors, named after the ocean they are connected to. The sectors of the Southern Ocean defined in the present study were the Pacific sector, the Atlantic sector and the Indian sector (see supplementary material, Figure I.7).

All data processed in the present study were mapped with a 0.25º x 0.25º resolution (~20 km in the region of the study), which we judged adapted for the purpose of the study.

I.2.b. Coastline and bathymetry data

We used the 50 m resolution coastline data product (version 5.1.0) distributed by Natural Earth (available at https://www.naturalearthdata.com/downloads/50m-physical-vectors/). The ocean bathymetry data was based on ETOPO1 1 Arc-Minute Global Relief Model data from National Centers for Environmental Information at the National Oceanic and Atmospheric Administration (NCEI NOAA; available at https://www.ngdc.noaa.gov/mgg/global/relief/ETOPO1/data/).

Grid points with bathymetry shallower than -1000 m were not considered in the study. In total, 210 893 grid points (81%) were retained (out of 260 821 grid points in the study zone).

I.2.c. Fronts in the Southern Ocean

Three main fronts are commonly used to delineate water masses boundaries within the Antarctic Circumpolar Current (ACC) (Orsi et al., 1995; Rintoul et al., 2001; Kim & Orsi, 2014; Park et al., 2019). These fronts are the Subantarctic Front (SAF), the Polar Front (PF) and the Southern ACC Front (SACCF). In the present study, the location of the three major fronts of the ACC was based on the fronts proposed in Park et al. (2019) (available at https://www.seanoe.org/data/00486/59800/). An additional front was added north of the SAF, namely the Subtropical Front (STF), based on the definition by Orsi, Whitworth & Nowlin (1995) applied to the sea temperature 1998-2022 climatology computed from Multi Observation Global Ocean 3D Temperature Salinity Height Geostrophic Current and MLD ARMOR3D product distributed by the Copernicus Marine Service (ARMOR3D; available at https://doi.org/10.48670/moi-00052). In reference to the eastward flow direction of the ACC, the terms “upstream” and “downstream” were further used as equivalents to west and east, respectively.

I.2.d. Chlorophyll-a concentration (CHL) satellite data

We used 8-day averaged Level 3 CHL products from January 1998 to December 2022 from two different data archives: the CHL product distributed by the Copernicus Marine Service GlobColour project (available at http://www.globcolour.info/) and the CHL product distributed by the European Space Agency Ocean Color Climate Change Initiative (ESA OC-CCI; available at http://www.oceancolour.org/). Both products merge the data from multiple sensors. The native resolution of both satellite-derived CHL products is 4 km. We mapped the CHL data with a 0.25º x 0.25º resolution. At each grid point, the Chla record constitutes a time series. Time series with more than 50% missing values over the full record were removed and the corresponding grid points were not considered in the study. In total, 124 627 grid points (48%) were retained (out of 260 821 grid points in the study zone). Therefore, most of the data are located North of 60ºS (see white mask around Antarctic continent in Figure I.3). Time series of CHL were linearly interpolated for missing values.

I.3. Methods

I.3.a. Temporal Decomposition

A time series decomposition was applied at each grid point. We used the temporal decomposition described in Keerthi et al. (2020). The method is based on the Census X11 algorithm developed by the US Bureau of the Census. The Census X11 algorithm was initially developed for the seasonal adjustment of economic time series (Shiskin et al., 1967). The decomposition method was then re-adapted for global time series of sea surface temperature (Pezzulli et al., 2005) and, more recently, for time series of CHL (see Vantrepotte & Mélin, 2009; Keerthi et al., 2020). Briefly, the method is based on the iterative application of bandpass filters (including moving average, Henderson filter) that aim at decomposing the time series into several components. Each component of the time series corresponds to a different time scale of variation. The time series of CHL was decomposed into a multi-annual (CHL_{MA}), a seasonal (CHL_{Seas}) and a sub-seasonal (CHL_{SS}) components, such that

$$CHL(t) = CHL_{MA}(t) + CHL_{Seas}(t) + CHL_{SS}(t) \quad (I.1)$$

where t is time. A detailed description of the decomposition method is provided in Keerthi et al. (2020).

The time series decomposition allows to describe and analyze separately the time scales of variations in the CHL signal, with the multi-annual CHL_{MA} describing time scales greater than one year, the seasonal CHL_{Seas} describing time scales between a year and 3 months, and the sub-seasonal CHL_{SS} describing time scales shorter than three months (see examples of corresponding power spectra in Keerthi et al., 2022, Extended Data Fig. 2). The components of the temporal decomposition fed specific analyses, with a purpose specific to each component. The multi-annual CHL_{MA} was used to analyze long-term variations in CHL (see Section I.1 and Figure I.1) but most of the focus was placed on shorter time scales. The analysis of CHL phenology was performed with the seasonal component CHL_{Seas}. Time scales shorter than seasonal were analyzed by assessing the contribution of the sub-seasonal component CHL_{SS} to the total variance of the CHL signal. The specific methods applied to each component of the time series are described below.

I.3.b. CHL Phenology

The phenology of CHL was characterized by analyzing the seasonal component of the CHL time series (CHL_{Seas}). Removing the multi-annual component prior to FPCA allows for an objective shape analysis that remains independent of long-term variations (e.g. trends in CHL biomass). The sub-seasonal component CHL_{SS} was also removed to analyze separately the part of the CHL signal related to time scales shorter than seasonal (i.e. shorter than 3 months, see previous section).

Normalization and segmentation into annual cycles

Seasonal time series CHL_{Seas} were normalized so that the signal of all grid points vary in a comparable range, independently of the biomass. At each grid point, normalization of the seasonal time series CHL_{Seas} was obtained by subtracting the mean and then dividing by the standard deviation of CHL_{Seas}. For each grid point with seasonal time series $CHL_{Seas}(t)$, the normalized seasonal time series $NCHL_{Seas}$ is expressed as

$$NCHL_{Seas} = \frac{\left(CHL_{Seas} - \overline{CHL_{Seas}}\right)}{STD\left(CHL_{Seas}\right)} \quad \text{(I.2)}$$

where STD is the standard deviation. $NCHL_{Seas}$ has mean zero and standard deviation equal to one.

Additionally, the time series $NCHL_{Seas}(t)$ were segmented into successive annual cycles. Annual cycles of CHL are defined from July to July, with a time resolution of 8 days. Consequently, each grid point is associated with a time series composed of 24 annual cycles, extending from 1st July 1998 to 30th June 2022.

Functional Principal Component Analysis (FPCA)

The time series were normalized and segmented into individual annual cycles to enable focusing only on CHL phenology (i.e. shape of the annual cycle) through Functional Principal Component Analysis (FPCA). FPCA is a tool used in Functional Data Analysis (FDA). FDA consists in handling the objects of interest as continuous curves (Ramsay & Silverman, 2005). The curves are generated from the raw data with specific constraints (e.g. smoothing). The parametrizable level of roughness of the curves (Meyer, 2012) introduces point-to-point relationship, which enables extraction of data properties related to shape (Pauthenet et al., 2017, 2019; Godard et al., 2021). In the present study, the handling of the objects of interest (annual cycles) as curves (functional data) enable the extraction of their shape, particularly adapted to the studying of CHL phenology. Each annual cycle was considered as an individual object in the functional analysis described in the present section. The annual cycles (N = 124 627 grid points x 24 years = 2 991 048 annual cycles) were all defined on the same time interval (i.e. a time vector with a duration of one year and a resolution of 8 days) and were expressed in a functional basis as a linear combination of continuous curves. FPCA was performed for the 2 991 048 annual cycles from the normalized seasonal time series SN(t).

FPCA outputs and terminology

The main results of the FPCA are the Functional Principal Components (FPCs). FPCs describe the main modes of variation in CHL annual cycles (FPCs are hereafter indifferently referred to as FPCs, modes of variation and FPCA axes). FPCs are numbered from 1 to the number of selected FPCA axes, in descending order of the associated explained variance. The number of

FPCA axes to be retained is determined according to the variance explained by each axis. Below a user-defined threshold of explained variance, FPCs may be discarded.

FPCA enables the creation of a new basis defined by the FPCs. Projection of an annual cycle on a FPC is hereafter referred to as its score. Absolute value of the score is not directly interpretable in terms of CHL.

FPCA-derived time series

Each grid point is assigned a score each year based on the phenology of the annual cycle recorded at the corresponding location. To each grid point hence corresponds a time series of 24 score values for each FPC (see example in supplementary material, Figure I.11). The evolution of CHL annual cycles phenology over time was examined at each grid point using the time series of its FPC scores and classical metrics derived from these, namely mean (hereafter referred to as "climatological score"), standard deviation, and trend (see Section I.3.c).

I.3.c. Trend analysis

The methods described in the present section produced time series of FPC scores characterizing the phenology of the annual cycle (see Section I.3.b). The time series were composed of one scalar for seasonal cycle of the study period (1998-2022, i.e. 24 values). Detection of significant trends in the time series was performed using Mann-Kendall trend test with significance level alpha = 0.05. Trends were quantified using an ordinary least squares (OLS) regression. Uncertainties in computed trends were defined based on the width of the 95% confidence interval.

I.4. Results

In this section we develop our interpretation of the modes of variation obtained after running the FPCA. We then examine the geographical distribution of each mode of variation on a climatological basis. Finally, we examine changes over time of the geographical distribution of the main modes of variation (see Section I.3.b for terminlogy of FPCA-derived variables).

I.4.a. Interpretation of the Functional Principal Components (FPCs)

The FPCs obtained after running the FPCA (Figure I.2) correspond to the main modes of variation that determine the phenology of CHL annual cycles in the SO. The mode of variation that explains most of the variance (FPC1, 54%) opposes annual cycles further referred to as “subtropical” and “subpolar” phenological types.

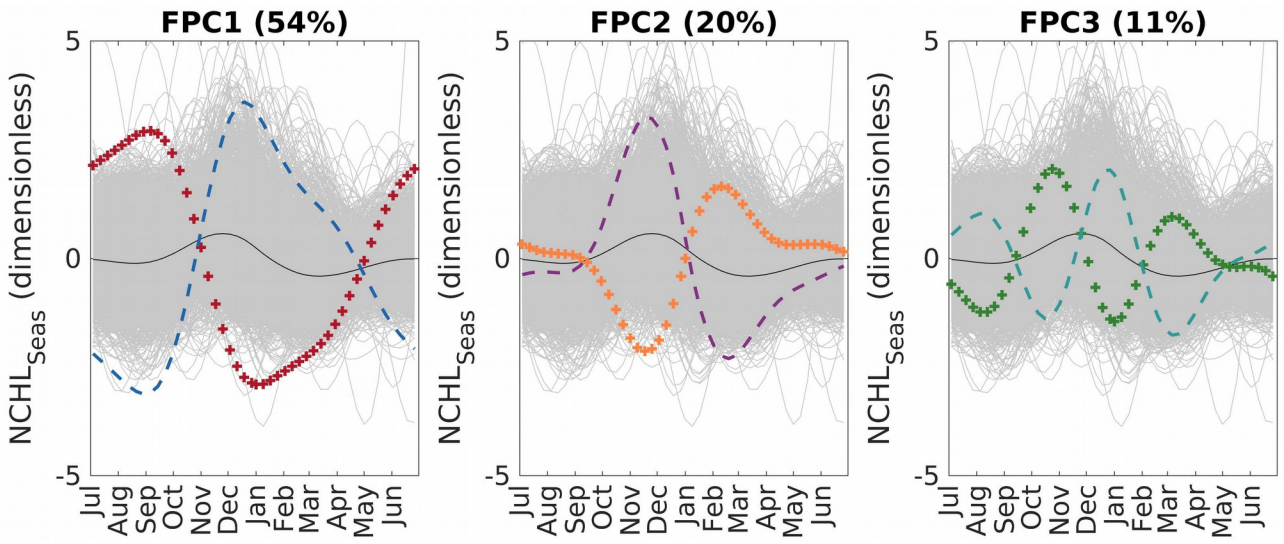


Figure I.2. Three first FPCs (modes of variation) obtained after running FPCA (see Section I.3.b). The value in brackets above each panel is the percentage of variance explained by the corresponding FPC. In each panel, the curve depicted by “-” (“+”) symbols shows a typical phenology associated with a negative (positive) FPC score. Negative (positive) scores are represented in blue, purple and cyan (red, orange and green) for FPC1, FPC2 and FPC3, respectively. The continuous black line displayed in every panels correspond to the average phenology in the study zone (Southern Ocean, south of 30ºS).

The subtropical phenological type (red curve in Figure I.2, left panel) is characterized by an annual maximum in early spring (around September) and annual minimum in summer (around January). Subtropical-like annual cycles show slow increase from summer minimum to early spring

maximum with relative high CHL abundance in winter, followed by a sharp decrease during late spring.

The subpolar phenological type (blue curve in Figure I.2, left panel) is characterized by an annual maximum around summer solstice (December-January) and annual minimum in late winter (around September). Subpolar-like annual cycles show a steep increase before the summer annual maximum and a slower decrease towards the late winter annual minimum, with even a break in the slope in late summer/early autumn (February-March). FPC1 hence not only differentiates annual cycles according to the timing of the annual maximum, but also highlights distinct annual cycle shapes between subpolar and subtropical phenologies.

Following the seasonal differentiation obtained with FPC1, the second mode of variation (FPC2, 20%) discriminates at finer temporal scale the timing of the summer annual maxima. Negative (positive) scores relatively to FPC2 indicate an earlier (later) annual maximum, occurring around November/December (February/March), represented by the blue (red) curve in Figure I.2 (middle panel).

From the third axis of the FPCA (FPC3, 12%) and beyond, interpretations of associated modes of variations becomes more delicate due to the less clear shape of the annual cycles described by the corresponding FPCs (Figure I.2, right panel, see also further axes of FPCA in supplementary material, Figure I.8). Although the phenology analysis was carried out with the seasonal component of the CHL time series only (Section I.3.b), influence of the sub-seasonal component on CHL phenology may be noticeable, as discussed in Section I.5.b. In the present study, only the first and the second modes of variation were interpreted and examined, totaling 74% of explained variance.

I.4.b. Geographical distribution of FPC scores

Following the interpretation of the FPCs, the geographical distribution of the climatological scores (1998-2022) of each grid point relatively to the two first modes of variation were examined. Spatial distribution of FPC1 and FPC2 climatological scores are shown in Figure I.3 and Figure I.4, respectively (see supplementary material, Figure I.12, right panel for FPC3 climatological scores). Grid points with highly positive average FPC1 score depict an outer ring that covers the subtropical region. Highly negative FPC1 scores are mainly found in the Atlantic and in the Indian sectors of the SO, in an inner ring delimited by the STF to the north, and the SAF to the south. South of the SAF, FPC1 scores are less negative. In the Pacific sector, the zone delineated by the STF and the SAF appears as a large transition area between positive and negative values of FPC1 scores. In the Atlantic and Indian sectors, the transition from subtropical to subpolar phenologies is sharper than in the Pacific sector and occurs north of the STF. The spatial structure of the phenologies described

by FPC1 scores were further examined through regional studies to support the spatial patterns described in the present section (see Section I.5.a).

Examination of grid points' contribution to the creation of the first axis of the FPCA (see supplementary material, Text S2) reveal that subtropical phenologies are the main contributors to the creation of FPC1 (see supplementary material, Figure I.9, left panel). Map of grid points' contribution to the creation of the first axis of the FPCA suggests that the FPC1 axis primarily distinguishes subtropical phenologies from the rest of the represented phenologies. Conversely, subtropical-like grid points generally have a low score and low contribution relatively to FPC2 (see supplementary material, Figure I.9, right panel). Therefore, no refined interpretation of FPC2 scores for subtropical-like grid points should be sought. Indeed, as highlighted in Section I.4.a, FPC2 mainly discriminates subpolar-like annual cycles according to the timing of their annual summer maximum. The geographical distribution of FPC2 scores (Figure I.4) clearly reveals large areas with an earlier summer annual maximum (e.g. south of the SAF in the Indian sector), compared with other areas associated with a later summer annual maximum (e.g. between the SAF and the STF in the in the Atlantic sector and in the eastern part of the Indian sector, south of the STF in the eastern part of the Pacific sector).

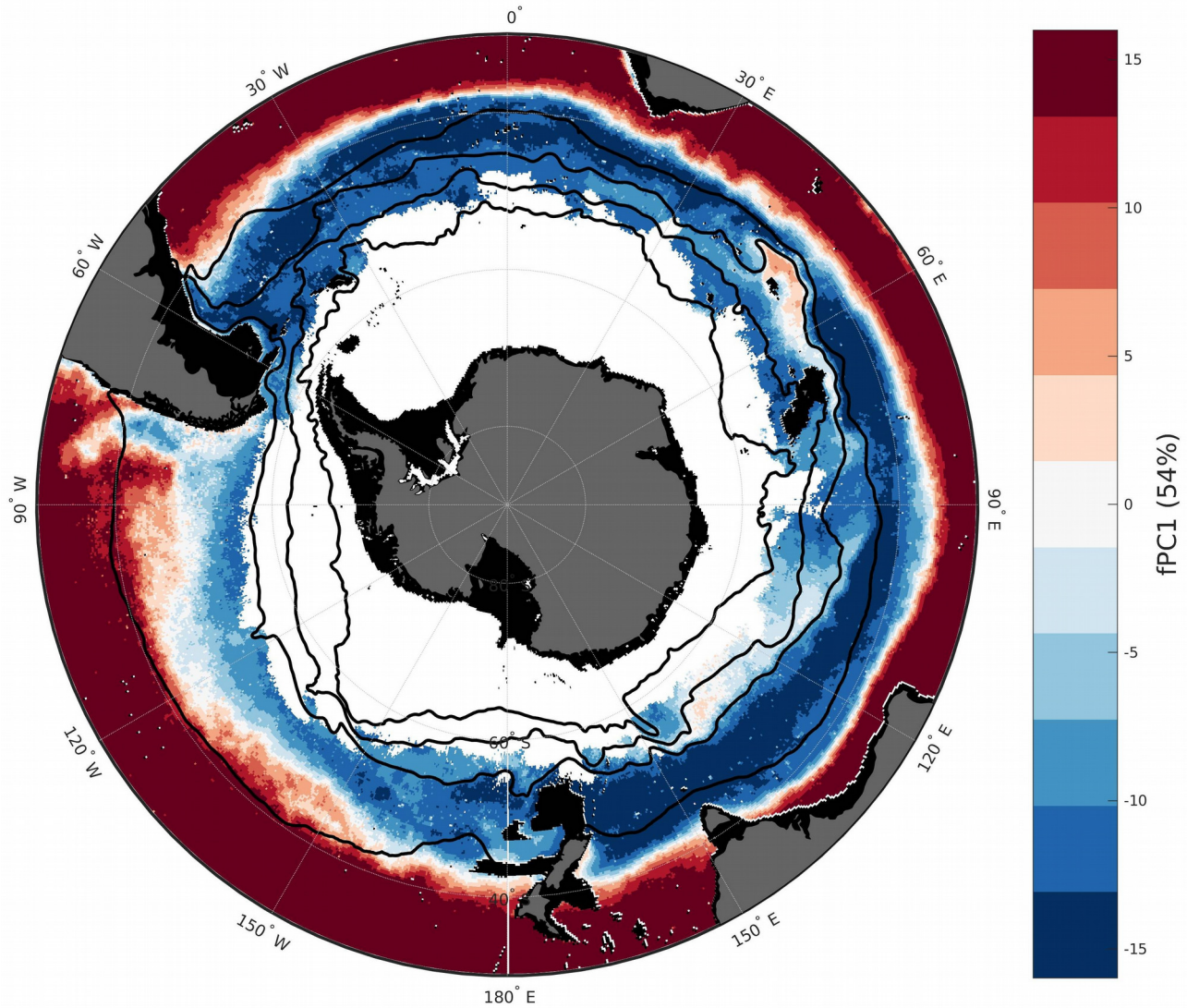


Figure I.3. Spatial distribution of FPC1 climatological scores (see Section I.3.b). Black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 time series (see Section I.2.d) are represented in white (white area around Antarctica, south of ~60ºS).

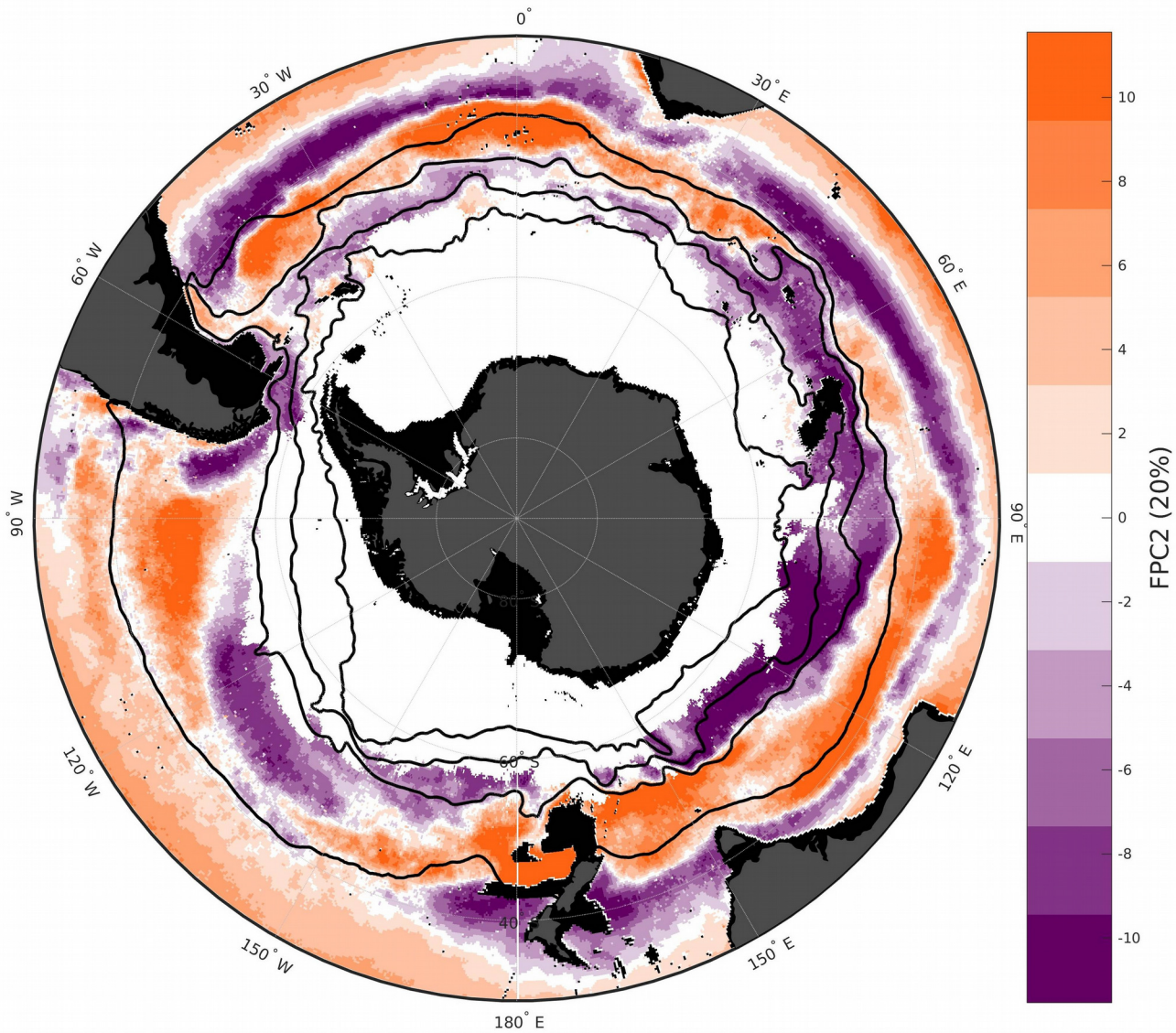


Figure I.4. Spatial distribution of FPC2 climatological scores (see Section I.3.b). Black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 time series (see Section I.2.d) are represented in white (white area around Antarctica, south of ~60ºS).

I.4.c. Time series of FPC scores

Time series of FPC1 scores were used to examine the behavior of CHL annual cycles phenology over time. Firstly, the steadiness of the annual cycle phenology over time was characterized at each grid point through the computation of the standard deviation of its FPC1 score time series ($FPC1_{STD}$). Secondly changes in CHL phenology were further examined through analysis of trends associated with FPC1 scores.

Detecting seasonal instability over time

The standard deviation of FPC score time series was computed at each grid point. High standard deviation in the time series of FPC scores is associated with unsteady seasonal pattern. The areas identified by the sutropical (subpolar) phenological type were generally associated with lower (higher) seasonal instability, revealed by examination of FPC1 score time series (Figure I.5). Geographical regions at the interface between the zones defined by the subtropical and subpolar phenological types (see Section I.4.b) were generally associated with high standard deviations relatively to FPC1 score time series. Thus, a large area of strong seasonal instability, located between the STF and the SAF, is clearly identified in the Pacific sector, while in the Atlantic and in the Indian sectors the zone of seasonal instability is located north of the STF. Geographical patterns of unsteadiness relatively to the modes of variation described by FPC1 and FPC2 were generally very similar (Figure I.5 and supplementary material, Figure I.13). However, detailed examination of the geographical region located between the SAF to the north and the PF to the south (only visible in the Atlantic and the Indian sector, due to lack of data in the Pacific sector, see Section I.2.d) reveal an area marked by a relative stability with respect to FPC1 scores, whereas FPC2 scores show high instability. Such contrasted behavior highlights the finer temporal resolution of the modes captured by FPC2. Thus, the continuous distribution of scores along the FPC1 and FPC2 axes allows assessment of seasonal to sub-seasonal changes in phenology over time.

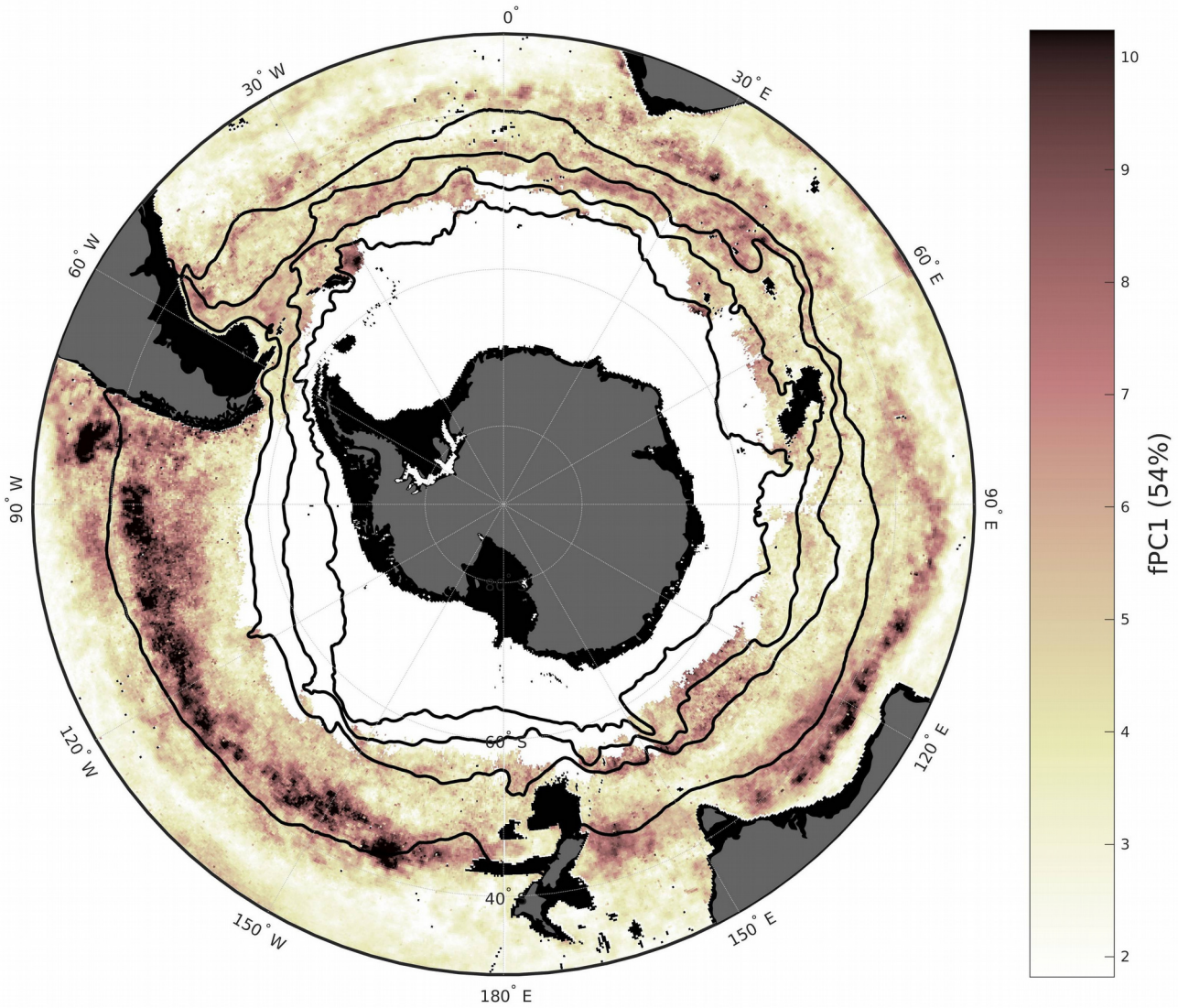


Figure I.5. Spatial distribution of $FPC1_{STD}$, the standard deviation of the time series of FPC1 scores (see Section I.3.b). Black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 time series (see Section I.2.d) are represented in white (white area around Antarctica, south of ~60ºS).

Assessing long-term changes in CHL phenology

The characterization of CHL phenology enabled by projection on FPC1 axis was further investigated through analysis of trends associated with FPC1 scores. Locations associated with significant positive (negative) trend along FPC1 axis tend to tend to have a changing phenology towards more subtropical-like (subpolar-like) annual cycles. Trends in FPC1 scores show large

spatial coherence (Figure I.6). In the Indian sector of the SO, the geographical distribution of FPC1 trends is complex but it broadly depicts positive trends in the east and negative trends in the west. In the Atlantic sector, positive FPC1 trends were mostly detected north of the SAF, while the area located between the SAF and the PF was generally associated with negative FPC1 trends, notably in the eastern part. FPC1 trends in the Pacific sector are marked by the presence of a geographical band of highly negative trends south of the STF, which extends north of the STF in the east. North and south of this geographical band, FPC1 trends in the Pacific sector of the SO are mainly positive. At the scale of the SO, it should be noted that some areas locally discord with the broad spatial patterns of FPC1 trends described in the present section (e.g. the Great Australian Bight in the Indian sector). In total, 16.3% of grid of points were associated with negative FPC1 trend, 16.2% with positive FPC1 trend, and 67.5% with no significant trend, which globally for the SO results in no FPC1 trend (see Figure I.6). Finally, it is also worth noting that the geographical areas associated with the highest climatological mean values of CHL (see 1998-2022 CHL climatology in supplementary material, Figure I.7), namely the Patagonian Plateau, South Georgia, Crozet Islands, the Kerguelen Plateau, up- and downstream New Zealand, and offshore Chile, were all associated with positive FPC1 trends.

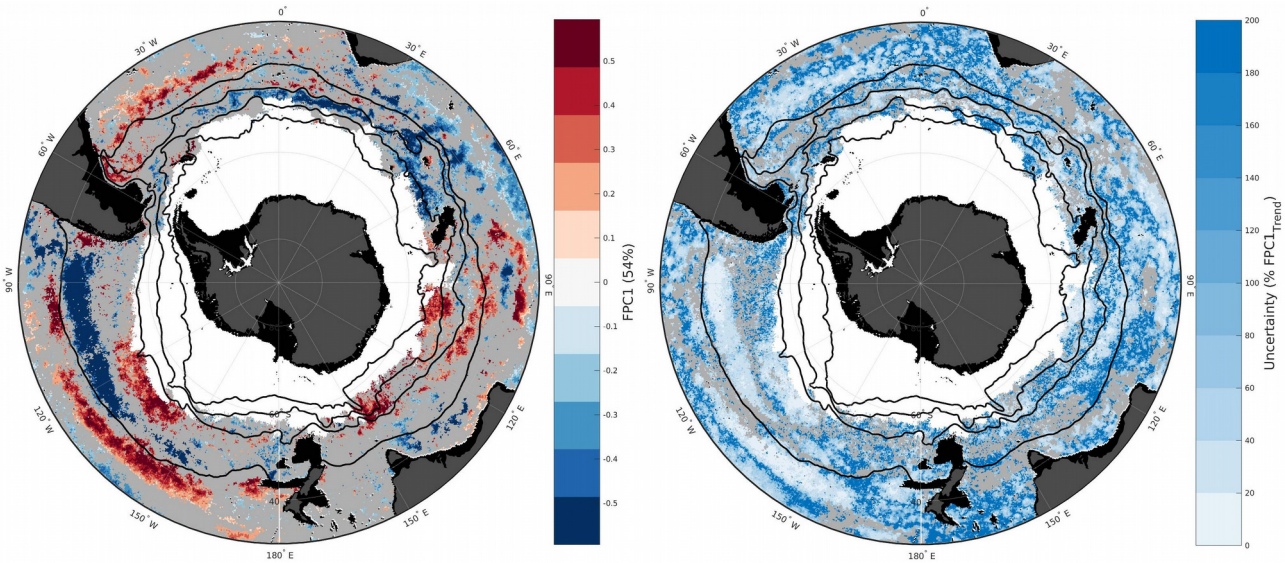


Figure I.6. Spatial distribution of (left panel) FPC1 trend estimates and (right panel) their uncertainties (width of 95% confidence interval) for the 1998-2022 GlobColour CHL record. Light gray regions indicate that the trend is not statistically different from zero. Black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 time series (see Section I.2.d) are represented in white (white area around Antarctica, south of ~60ºS).

I.5. Discussion

I.5.a. Processes supporting CHL production in the SO: a descriptive approach based on case studies

The analysis based on the shape of the annual cycle of CHL proposed in the present study highlights the main modes of variation that drive the phenology of annual CHL production in the SO. A clear contrast between two major phenological types, defined as subtropical and subpolar phenologies (Section I.4.a), is highlighted by projection on the first axis of the FPCA (see Figure I.3 and supplementary material, Figure I.14). At first order, the two phenological types can be differentiated by the timing of their annual CHL maximum (Figure I.2, left panel). However, there is a continuum between the two extremes of the FPC1 axis, spanning different regimes of phytoplankton dynamics in the SO. These regimes are the result of specific regional processes. In this section, we use case studies focused on different regions of the SO to illustrate the phenological types described in the present study (Section I.4.a) and their spatial distribution (Section I.4.b). The case studies support our results and provide hypotheses for understanding the processes driving CHL dynamics in the SO.

Region of New Zealand: A transition from subtropical to subpolar phenologies

The geographical distribution of FPC1 scores can partially be described using the major fronts of the SO (see Section I.4.b). However, some obvious geographical discrepancies between a transition based on a purely biological criterion (phenology) and water masses boundaries (climatological position of the fronts) were highlighted in Section I.4.b. These discrepancies could possibly be due to the year-to-year meandering of the SO fronts (Sallée et al., 2008; Sokolov & Rintoul, 2009b; De Boer et al., 2013; Kim & Orsi, 2014; Pauthenet et al., 2017). Focusing on the STF, one way to describe the subtropical/subpolar phenological transition, while removing variability in STF position, is to examine specific zones where isolated high bathymetry areas constrain the transport of water (Rintoul et al., 2018) and act as a real lock for major fronts (e.g. Roquet et al., 2009; Sokolov & Rintoul, 2009a; Forcén-Vázquez et al., 2017; Park et al., 2019). Such a zone is encountered around New Zealand, where an extension of the continental shelf eastward from South Island, materialized by the Chatham Rise, constitutes a real bathymetry lock for the STF (Hadfield et al., 2007). Extensively described in Murphy et al. (2001) using SeaWiFS ocean color images, CHL dynamics around New Zealand reveal three main patterns, referred to as subtropical, subtropical frontal, and subantarctic waters.

Firstly, the region associated with highly positive FPC1 scores located north of New Zealand (Figure I.3), characterized by elevated CHL abundance during winter reaching a maximum during spring, and falling to a minimum during summer (see Section I.4.a and red curve in Figure I.2, left

panel), is in line with the subtropical oligorophic regime identified in Murphy et al. (2001) in the same area. The processes that support the CHL dynamics associated with the subtropical phenology are typically a nutrient enrichment of the surface layer by winter mixing, followed by rapid consumption of nutrients in spring, while winter production is limited by light availability (Longhurst, 2007; Fay & McKinley, 2014).

Secondly, subtropical frontal waters are found at the interface between subantarctic and subtropical waters, where the mixing of cold, nutrient-rich, iron-depleted subantarctic waters with warm, nutrient-poor, iron-rich subtropical waters is hypothesized to support CHL production (Pinkerton et al., 2005; Chriswell et al., 2013; Behrens & Bostock, 2023). As a consequence of their frontal position, these waters are characterized with possible presence of CHL almost throughout the year (Boyd et al., 1999), which translates into intermediate phenologies, and therefore FPC1 scores close to zero. Combination of low absolute values of FPC1 scores (white, light red and light blue regions in Figure I.3) and low standard deviation of FPC1 scores (Figure I.5) on either sides of the Chatham Rise confirm the quasi-permanent regime of relatively high CHL abundance detected in subtropical frontal waters at the STF in Murphy et al. (2001). It is worth to note that the evidence of a bathymetry lock at the Chatham Rise is reinforced by the notably low standard deviation of FPC1 scores north and south of the STF around new Zealand, in contrast to the rest of the Pacific sector (Figure I.5).

Thirdly, the region south of New Zealand is characterized by the presence of subantarctic waters (Boyd et al., 1999). CHL dynamics in subantarctic waters present low annual amplitude, with a CHL minimum in the late winter and a maximum in summer (Banse & English, 1997; Boyd et al., 1999; Murphy et al., 2001). Phytoplankton growth in subantarctic waters seems to be clearly limited by light in winter, while during other seasons, light, iron and silicate can limit (or co-limit) CHL production (Boyd et al., 1999; Boyd et al., 2001; Hutchins et al., 2001), leading to a variety of phenologies, from highly marked solstice annual maximum, to phenologies with multiple peaks (Murphy et al., 2001). Annual variations in CHL in subantarctic waters correspond to the phenologies described by negative FPC1 scores (see Section I.4.a and blue curve in Figure I.2, left panel). In the southern part of the region located between the STF and the SAF, FPC1 scores ranging from negative to highly negative (see Figure I.3) are consistent with the presence of subantarctic waters. At the scale of the SO, subantarctic waters are further investigated and differentiated into two main regimes of annual CHL production (see next section).

South of the Subtropical Front: a HNLC system with local exceptions

As described in the previous section, it was demonstrated that CHL production in the area downstream New Zealand is highly dependent on iron delivery (Banse & English, 1997). More generally, the open Southern Ocean south of the STF is mainly a High Nitrate - Low Chlorophyll (HNLC) system. By definition, HNLC waters are characterized by low CHL biomass despite high

macro-nutrient (nitrate) concentrations (Chisholm & Morel, 1991). The limitation of phytoplankton growth by iron availability in HNLC regions has been largely demonstrated (Martin et al., 1990; Sedwick et al.,1999; Boyd et al., 2000; Boyd & Law 2001; Blain et al., 2007; Venables & Moore, 2010). Nevertheless, the control of micro-nutrients (iron) availability on CHL production is released downstream of some shallow bathymetry areas and subantarctic islands of the SO, due to natural iron enrichment. The main exceptions to the general HNLC regime of the subantarctic SO are the regions downstream of South Georgia (Korb & Whitehouse, 2004; Prend et al., 2019), Crozet (Planquette et al., 2007), and the Kerguelen Plateau (Blain et al., 2001). Based on FPC1 score, we highlight in the present section two main regimes of CHL production dynamics south of the STF, namely open-ocean HNLC regions and naturally iron-fertilized regions.

Firstly, regions with very negative FPC1 scores (dark blue, FPC1 score ~-15 in Figure I.3) were identified as typical open-ocean HNLC regions. As described in Section I.4.a, very negative FPC1 scores are associated with a steep increase in CHL biomass in spring and an annual maximum at the summer solstice. In the absence of land inputs, iron delivery in open-ocean HNLC regions is highly dependent on vertical mixing (Llort et al., 2015; du Plessis et al., 2019). The iron supply to surface waters ensured by the seasonal deep winter mixing (Thomalla et al., 2011; du Plessis et al., 2017, 2019) is an essential driver for phytoplankton growth in HNLC regions (Tagliabue et al., 2014; Mtshali et al., 2019). Following the winter mixing, the seasonal stratification is synonym of increased light availability, which results in enhanced CHL production in spring (Swart et al., 2015; Thomalla et al., 2015; du Plessis et al., 2019). The duration of the productive season is highly dependent on the iron stocks brought to the surface layer by winter mixing (Boyd et al., 2001; Thomalla et al., 2011; Mtshali et al., 2019), which can be assessed by comparing the relative depths of the ferricline and of the mixed layer in winter (i.e. the intensity of the winter mixing). In addition, examples of spring bloom being delayed (du Plessis et al., 2019), or summer CHL production being maintained (Swart et al., 2015 ; Carranza & Gille, 2015) by (sub)mesoscale activity reveal the importance of (sub)mesoscale physical dynamics in supporting CHL production in HNLC waters. The progressive decline of CHL production after the summer maximum described for the subpolar phenological type (see Section I.4.a) or the presence of subsidiary peaks (Murphy et al., 2001) could be a marker of disruption at sub-seasonal temporal scales of the subpolar-like annual cycle by (sub)mesoscale activity. Finally, in a given year and locally, less negative FPC1 scores observed in typical HNLC regions could result from stronger modifications of the annual cycle induced by (sub)mesoscale activity (see supplementary material, Figure I.10). Such year-to-year variations were highlighted in du Plessis et al. (2019) in the SOSCEx experiment where gliders were deployed for several years on a fixed site located between the STF and the SAF in the eastern part of the Atlantic sector of the SO to monitor the seasonal evolution of the vertical stratification.

Secondly, regions with slightly less negative FPC1 scores than the HNLC regions described above depict an inner ring south of the SAF (blue, FPC1 score ~-10 in Figure I.3). The inner ring south of the SAF characterized by less negative FPC1 scores (Section I.4.b) was identified as a

region with similar limiting factors to phytoplankton growth than the HNLC regions in winter, but with significantly increased iron inputs due to interactions of the ACC with shallow bathymetry areas and islands (Blain et al., 2001; Korb & Whitehouse, 2004; Lucas et al., 2007; Venables et al., 2007). Similarly to HNLC waters, the (sub)mesoscale activity-induced modifications at sub-seasonal temporal scale of the subpolar-like annual cycle might result in a less marked phenology, and hence, less extreme (negative) FPC1 scores. Downstream of the Kerguelen Plateau, natural iron fertilization supports an annual phytoplankton bloom roughly starting in November, reaching CHL maximum in December and January, and collapsing in February, as captured during the KEOPS study (Blain et al., 2007). KEOPS2 experiment focused on the same region and revealed the structuring role of (sub)mesoscale horizontal stirring in the phytoplankton bloom observed downstream of the Kerguelen Plateau (d'Ovidio et al., 2015). Evidence of (sub)mesoscale physical dynamics supports the spatial patterns described by the first axis of the FPCA, with less negative values of FPC1 scores observed east of the Kerguelen Plateau (blue, FPC1 score ~-10 in Figure I.3). Furthermore, investigations into the Crozet bloom carried out during the CROZEX (Pollard et al., 2007) and KEOPS2 (Sanial et al., 2014) experiments revealed the occurrence of an earlier bloom compared with other regions like the Kerguelen Plateau. North of the Crozet Plateau, bloom onset was detected in early-mid September, maximum CHL levels were encountered in October, followed by a rapid decrease in November (Venables et al., 2007; Pollard et al., 2009). Although the factors that control CHL annual production north of the Crozet Plateau are very similar to those described in the region of the Kerguelen Plateau (Lucas et al., 2007; Sanial et al., 2014), the typical CHL phenological type encountered around Crozet, with an annual maximum in October, appears to be closer to the subtropical phenology (with annual maximum in September, see Section I.4.a), which results in slightly positive FPC1 scores observed north of the Crozet Plateau (Figure I.3).

As discussed in Abott et al. (2000), other biotic or abiotic processes influence the dynamics of CHL production in the HNLC SO. The succession of the different factors driving CHL production in the SO leads to a diversity of phenologies, directly induced by local and temporary characteristics, which manifest at sub-seasonal scale in the CHL signal. Although these factors cannot be directly identified by inspecting the spatial distribution of FPC scores, we propose a list of possible underlying processes and discuss them in the next section.

Processes with sub-seasonal influence on CHL annual cycle

Interaction between water masses and landmass, vertical mixing, and horizontal advection of water masses were the main drivers of surface-water iron enrichment described in the previous section. However, open-ocean iron supplies can be induced by other mechanisms such as atmospheric dust input (Duce & Tindale, 1991; Jickells et al. 2005; Boyd & Ellwood, 2010), hydrothermal vents (Tagliabue et al., 2010; Guieu et al., 2018; Ardyna et al., 2019) and internal metal transformation (Boyd et al., 2017). These localized or intermittent mechanisms of natural iron enrichment in the HNLC SO manifest at sub-seasonal level in the phenology of the annual cycle of CHL.

In addition, other co-limitations interact with iron limitation in the dynamics of phytoplankton growth in HNLC waters. These limitations are mainly driven by light (Boyd et al., 2001; Edwards et al. 2004; Venables & Moore, 2010; Blain et al., 2013; Arteaga et al., 2014) and silicic acid (Boyd et al., 1999; Hutchins et al., 2001; Sedwick et al., 2002; Lannuzel et al., 2011) availability.

Furthermore, other ecological processes such as grazing (Banse, 1996; Abott et al., 2000; Gutiérrez-Rodríguez et al., 2020; Arteaga et al., 2020) or control by microbial activity (Boyd et al., 2012) might be responsible for sub-seasonal variations of CHL production.

Finally, the sensitivity of CHL dynamics to variations at sub-seasonal scales in the SO (discussed in the previous section for the HNLC SO) is reinforced by the important role of (sub)mesoscale physical processes in enhancing CHL production (Lévy et al., 2012, 2018; McGillicuddy et al., 2016; Uchida et al., 2020), with is expressed in the CHL signal from sub-seasonal (Keerthi et al., 2020) to seasonal (Uchida et al., 2020) and inter-annual (Lévy et al., 2014; Keerthi et al., 2022) time scales.

The phenological types identified in the present study (Section I.4.a) and their sub-seasonal variations (discussed in the present section) were obtained on the unique basis of the seasonal component of the decomposed time series (see Sections I.3.a and I.3.b). The sub-seasonal component of the time series of CHL (see Section I.3.a), with potential role in driving both CHL phenology and inter-annual CHL variability (Prend et al., 2022; Keerthi et al., 2022) is further investigated in Section I.5.b.

I.5.b. Investigating the sub-seasonal component

The weight of high-frequency variability

Time series decomposition (Section I.3.a) was used to disentangle the different time scales of variation contributing to the CHL signal. Long term variations were analyzed using the multi-annual component (Figure I.1), while the phenology of CHL annual cycle was studied through the seasonal component (Section I.3.b). As the main objective of the present study was to identify the predominant seasonal patterns that drive annual cycles of CHL in the SO, the sub-seasonal component was not included in the phenology analysis.

A sensitivity test to the inclusion of the sub-seasonal component in the FPCA analysis was performed to identify the weight of the sub-seasonal component in the phenology analysis. A similar FPCA analysis to that applied to the seasonal component (Section I.3.b) was performed with annual cycles of CHL defined by the sum of the seasonal plus the sub-seasonal components. The

first two FPCs obtained in both FPCA analyses were very similar (see supplementary material, Figure I.16), which demonstrates the efficacy of the functional approach (see Section I.3.b) in extracting the main modes of variation in annual cycles of CHL (Section I.4.a). However, the variance associated with the FPCs was differently shared. In the FPCA analysis including the sub-seasonal component, FPC1 and FPC2 were associated with 37% and 16% of explained variance, respectively (see supplementary material, Figure I.16, first two panels), while in the FPCA analysis including only the seasonal component, FPC1 and FPC2 were associated with 54% and 20% of explained variance (see Figure I.2, first two panels). The lower explained variance of the first two modes of variation indicates the presence of the sub-seasonal component in the analyzed signal. Similarly, the variance explained by the other axes (FPC3 and beyond) was higher in the FPCA that included the sub-seasonal. The sensitivity test described in the present section demonstrates the considerable proportion of variability attributable to the sub-seasonal component in the CHL signal. However the sensitivity test also highlights that the phenology analysis developed in the present study is inadequate for detecting accurate patterns of high-frequency variability.

Nonetheless, a possible expression of the sub-seasonal variability is detectable in the seasonal patterns of CHL identified in the FPCA, as presented in the next section.

The role of the sub-seasonal component in defining CHL phenology

The time series decomposition method applied in the present study is not obtained by application of strict frequency bandpass filters, but instead, by iterative filtering of the time series (Section I.3.a). This allows for a slight overlap in the frequency domains covered by the different components (Keerthi et al., 2022), which leads to potential interaction between seasonal and sub-seasonal scales and the role of the latter in defining the phenology of the annual cycle.

In this section, interactions between phenology of annual cycle of CHL, and relative contribution of the seasonal component to total CHL variability were examined. For this purpose, the percentage of variance explained by the components of the time series were computed for each grid point in the study zone (see supplementary material, Text S3. The relative contribution of the seasonal component (versus the sub-seasonal and the multi-annual), expressed as a percentage of the total CHL variance, was hereafter denoted VAR_{Seas}.

Interestingly, the standard deviation associated with the time series of FPC1 scores ($FPC1_{STD}$) was well correlated with the percentage of variance explained by the seasonal component (R^2=0.43, N=124 627, see supplementary material, Figure I.15). Correlation was negative, meaning that grid points associated with high standard deviation relatively to time series of FPC1 scores (high $FPC1_{STD}$ values), previously identified as locations with high seasonal instability (see Section I.4.c), generally have low values of VAR_{Seas} (i.e. low contribution of the seasonal component).

Conversely, high relative contribution of the seasonal component to total CHL variance corresponds to steady seasonal patterns (low $FPC1_{STD}$ values).

Geographical distribution of $FPC1_{STD}$ (Figure I.5) suggests higher seasonal repeatability at subtropical latitudes, while high latitudes are associated with more unsteady seasonal patterns. Our results are in line with previous studies (Demarcq et al., 2012; Thomalla et al., 2011; Jönsson et al., 2023). Moreover, Prend et al. (2022) and Keerthi et al. (2022) demonstrated that in the SO, non-seasonal variability at high latitudes was mainly imputable to sub-seasonal time scales. The proportion of variance carried by the sub-seasonal component can hence indirectly be assessed by examining $FPC1_{STD}$.

The relative high contribution of the sub-seasonal component to total CHL variance at high latitudes was revealed in our study by the spatial distribution of $FPC1_{STD}$ values (Figure I.5). Sub-seasonal variations in CHL have been linked to (sub)mesoscale processes (Swart et al., 2015; du Plessis et al., 2017), with wind stress from synoptic storm events (< 10 days) considered the main forcing factor (Carranza & Gille, 2015; Thomalla et al., 2015; du Plessis et al., 2019; Keerthi et al., 2021). However, the chaotic nature of small-scale physical processes should not be ignored (Lévy et al., 2014; Keerthi et al., 2021; Cravatte et al., 2021; Hogg et al., 2022) to explain part of the high-frequency variability of CHL.

In this section, we demonstrated that contribution of sub-seasonal variations to the phenology of the annual cycle of CHL is effective, with implications on inter-annual CHL biomass dynamics. Evolution of seasonal patterns of phytoplankton growth should hence be considered to study long term trends of primary production in the SO, discussed in the next section.

I.5.c. Trends in phenology and implications for phytoplankton biomass

The optical signal interpreted as CHL (chlorophyll-a concentration) can be subject to variations that may not be linked to an effective change in phytoplankton biomass. In this section we discuss the difficulties involved in assessing long-term trends in phytoplankton biomass from CHL measurements and how the phenology approach developed in the present study partially circumvents these difficulties.

Factors of biological origin

Uncertainties related to the measurement of chlorophyll-a fluorescence can affect long term estimates of phytoplankton biomass due to a series of factors that are inherent to phytoplankton biology. Firstly, long-term variations of measured chlorophyll-a fluorescence might include variations related to phytoplankton community composition (Henson et al., 2016; Sathyendranath et

al., 2017; Sun et al., 2023) with direct implications on biomass estimates due to different carbon : chlorophyll ratios (Bradford-Grieve et al., 1997; Henson et al., 2016). Moreover, long-term monitoring of phytoplankton biomass through chlorophyll-a fluorescence might be affected by varying phytoplankton physiological response to light due to environmental conditions (Suggett et al., 2009; Brewin et al., 2019; Schallenberg et al., 2022; Ryan-Keogh et al., 2023). Finally, the estimates of CHL concentration provided by ocean-colour sensors are limited to the surface. However, variations in surface CHL do not necessarily reflect integrated biomass change due to lag between surface and depth-integrated biomass (Sallée et al., 2015) or non homogeneity of phytoplankton concentration in the mixed layer (Carranza et al., 2018).

Instrumental biases

The overall CHL record used in the present study is the result of different ocean-color sensors eras (Section I.1). Despite the overlap between the different periods and all the efforts made by the scientific community to homogenize the data, the measurements were carried out using different instruments with their own biases. Whether the differences (or the absence of differences) regarding decadal variations (Figure I.1) are to be attributed to different ocean-color sensors eras or to true biological changes, the question remains in some way open. Additionally, beyond the instrumental issue, Sathyendranath et al. (2017) point out the crucial role of implemented ocean-colour algorithms for detection of ecosystem response to climate change. In view of the diverging (product-dependent) trends highlighted in Section I.1, we emphasize the relevance of using a method that removes the quantitative aspect of the data (i.e. a “dimensionless” CHL) for more robust results with regard to sensor interoperability issues.

Robustness of the phenology analysis

Given the difficulty of disentangling the above-mentioned factors, we judged relevant to work with a normalized CHL with no unit and focus on changes related to CHL phenology rather than biomass. To test the robustness of the phenology analysis proposed in the present study, we performed a similar analysis to that described in Section I.3.b, however using the ESA OC-CCI ocean-colour data (see Section I.2.d). Phenology analyses based on GlobColour data (Section I.4) and OC-CCI data (this section) produced similar results. The modes of variation obtained with both FPCA (see Figure I.2 and supplementary material, Figure I.17), the spatial distribution of FPC scores (see Figures I.3-I.4 and supplementary material, Figure I.18, left panel), as well as trends in FPC1 scores (see Figure I.6 and supplementary material, Figure I.18, right panel) were very analogous. As a consequence, while trends in CHL biomass proved to be product-dependent, both in terms of amplitude (our study, see Section I.1) and geographical distribution (Lim et al., 2022), analyzing phenology independently of biomass removed the dependence on the product used.

The retrieval of quasi-identical FPCA results with two different ocean-colour products emphasizes the strength of the phenology analysis proposed in the present study in the absence of

reliable long-term conclusions regarding CHL biomass. However, monitoring the evolution of phytoplankton biomass in the SO from a quantitative perpective - whether for questions related to ecosystem functioning, carbon export or biogeochemical cycles (Deppeler & Davidson, 2017; Henley et al., 2020) - remains essential. In the following section we discuss the fitness for purpose of the phenology analysis proposed in the present study for long-term CHL biomass assessments.

Can FPCA results be retrospectively linked to biomass?

To assess the link between phenology and biomass, FPC1 scores were compared to the CHL biomass accumulated at each location during the year, computed as the integral over time of each annual cycle (Σ_{CHL}, see supplementary material, Text S6.). Although the two variables were not strongly correlated, comparison of FPC1 scores and Σ_{CHL} clearly revealed two clusters (see supplementary material, Figure I.19, top panel). Primarily induced by the bi-modal distribution of FPC1 scores (see Sections I.4.a, I.5.a and supplementary material, Figure I.14), these clusters distinguish a first group with higher values of Σ_{CHL} and low FPC1 scores (the so-called subpolar phenology, see Section I.4.a), from a second one with lower values of Σ_{CHL} and higher FPC1 scores (the so-called subtropical phenology).

Combining the trends observed for FPC1 scores (Section I.4.c), with the results presented in the present section, we can hypothesize that migration from one cluster to another could be synonym of a change in annual CHL biomass. For example, grid points with decreasing FPC1 scores over time (i.e. negative trend in FPC1, blue color in Figure I.6) could on average experience an increase in CHL biomass. Conversely, grid points with increasing FPC1 scores over time (i.e. positive trend in FPC1, red color in Figure I.6) could on average experience a decline in CHL biomass.

To test the hypothesis that migration from one cluster (e.g. subtropical, less productive) to another (e.g. subpolar, productive) could be seen as proxy for a presumed change in CHL biomass, we recommend to concomitantly analyze trend in phenology and trend in CHL biomass on a “stable” period with respect to ocean-colour sensor configuration (e.g. 2002-2012, see gray line in Figure I.1). Nonetheless, it is worth to note that an important fraction of the grid points may remain associated with non significant trend, either for FPC1 (67.5% in the case of FPC1 score on the 1998-2022 period, gray color in Figure I.6) or for CHL.

No direct causality was sought through the matching between change in phenology and change in biomass. We merely considered that a trend observed not on the absolute value of CHL but on its phenology is a stronger result when issues related to sensor type are encountered (see previous section). However, drivers of such changes are specific to each region (Racault et al., 2011). For this reason, we recommend to adopt a more local approach to improve observations of ongoing changes and enhance the understanding of the underlying processes.

I.6. Conclusion

Monitoring primary production in the SO is essential but obtaining reliable long-term quantitative assessments of CHL biomass evolution is challenging (Henson et al., 2016). In the present study, analyzing phenology independently of biomass was proposed as a robust way to study the dynamics of CHL in the SO and changes inferred from long-term ocean color time series. Time series of CHL were decomposed and segmented into annual cycles to separately observe the various time scales expressed in the remotely-sensed CHL signal. A shape-based functional analysis was performed on the segmented, normalized CHL annual cycles to disentangle apparent changes in CHL biomass (Figure I.1) from changes in phytoplankton phenology (Figure I.6). Firstly, our results highlighted the potential of the applied functional methodology (FPCA) to extract the main seasonal patterns driving CHL dynamics in the study area. In the SO, two main phenological types were identified, characterized by subtropical- and subpolar-like annual cycles of CHL (Section I.4.a). Although pointing out these two seasonal patterns is apparently trivial, the strength the results lies in the unsupervised nature of the method (i.e. identification of the main phenologies was not imposed). Moreover, robustness of the results was reinforced by their insensitivity to a change in ocean-colour algorithm (Section I.5.c). As a result, the phenology analysis proposed in the present study enables efficient examination of the spatial distribution of seasonal patterns of CHL (Figure I.3), as well as their variability over time (Figure I.5).

The continuum of shapes used to characterize CHL phenology revealed the presence of various different seasonal patterns of CHL dynamics in the SO. The identified seasonal patterns were supported locally by case studies (Section I.5.a). The case studies used to support the results of the phenology analysis also provided key elements for the understanding of possible underlying processes and the time scales at which they affect CHL production. In particular, although the proposed methodology focused on CHL dynamics through analysis of the seasonal component (Section I.3.b), sub-seasonal variations were clearly detected and described in the present study (Section I.5.b). Highlighted in recent studies (Prend et al., 2022; Keerthi et al., 2022), the substantial contribution of small-scale, sub-seasonal forcing in driving inter-annual variations in CHL in the SO, emphasizes the crucial need to study more closely the interactions at meso- to submeso-scales between physical processes and primary production (Lévy et al., 2012; McGillicuddy et al., 2016). Attribution of specific (sub)mesoscale processes to variations in CHL (e.g. d'Ovidio et al., 2015; Uchida et al., 2020) and detection biological hotspots associated with fine-scale structures (e.g. Siegelman et al., 2019; Rivière et al., 2019; Baudena et al., 2021) emphasize the structuring role of fine-scale processes in sustaining marine ecosystems (Lévy et al., 2018).

Our study suggests that regional approaches scrutinizing phytoplankton dynamics at high spatial and temporal resolution, using a relevant combination of high-resolution *in-situ* and remote-sensing data (d'Ovidio et al., 2019) are necessary to understand long-term variations in primary production and underlying mechanisms.

Supplementary material for Chapter I

Text S1. CHL climatology

CHL climatology was computed for each pixel as the mean value of CHL during the overall CHL record of the study period (1998-2022). CHL climatology was derived from GlobColour 8-day averaged Level 3 product and computed with a 0.25º x 0.25º resolution.

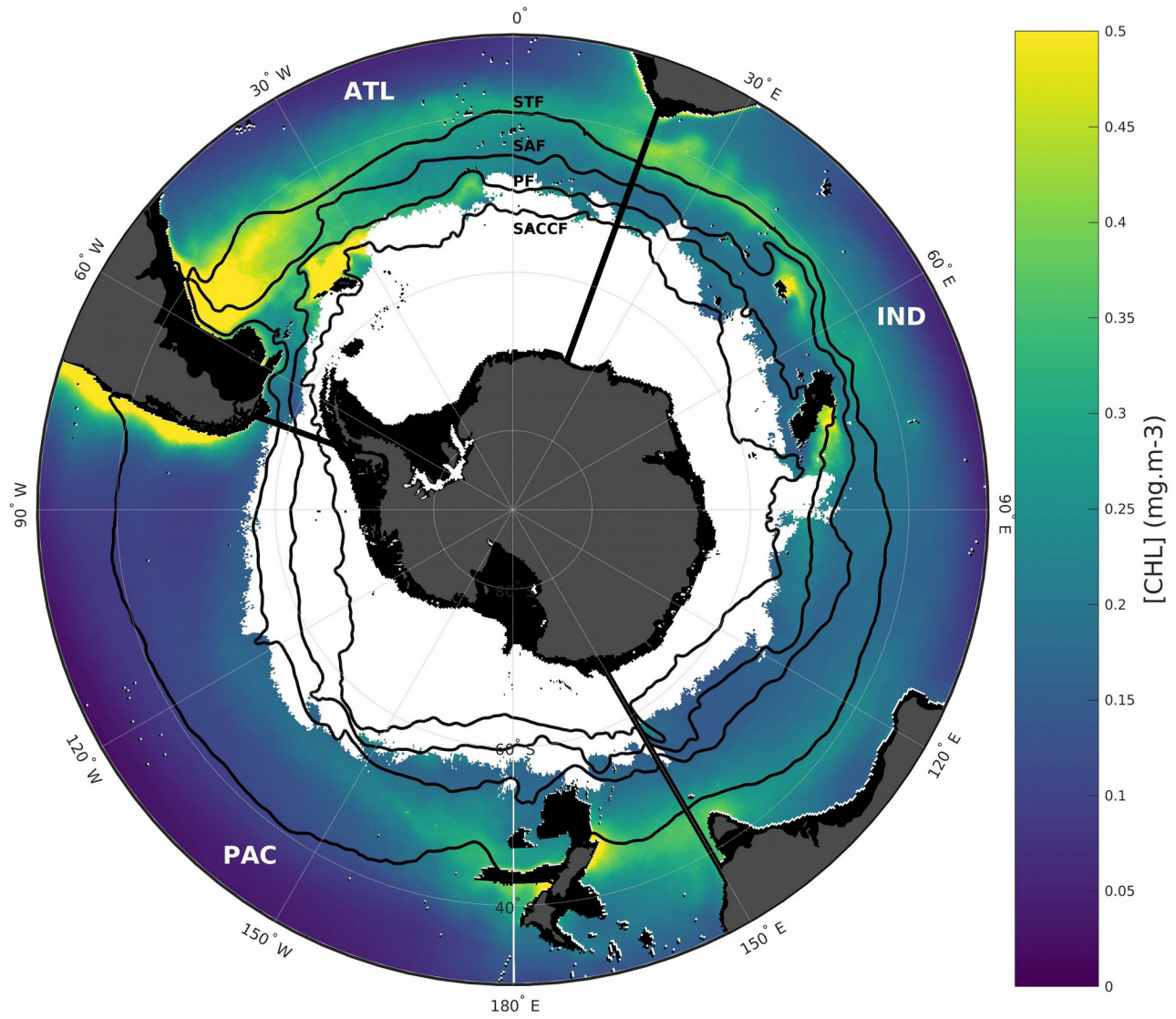


Figure I.7: CHL climatology (1998-2022) in the Southern Ocean derived from GlobColour 8-day averaged Level 3 product with 0.25º x 0.25º resolution. Sectors of the SO are separated by thick, straight black lines. Sectors of the SO are PAC: Pacific, ATL: Atlantic, IND: Indian. Continuous black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 time series (see Section I.2.d) are represented in white (white area around Antarctica, south of ~60ºS).

Text S2. FPCA outputs and terminology (2)

FPCA enables the creation of a new basis defined by the FPCs (see main FPCs in Figure I.2 and extended representation of the ten first FPCs in Figure I.8). Projection of an annual cycle on a FPC is referred to as its score. The contribution of a grid point relatively to the creation of a FPC can be assessed through the ratio between the sum of the absolute value of its scores and the sum of the absolute values of the scores of all other grid points (see Figure I.9). Each grid point is assigned a score each year based on the phenology of the annual cycle recorded at the corresponding location. To each grid point hence corresponds a time series of 24 score values for each FPC (see example in supplementary data, Figure I.11). The evolution of CHL annual cycles phenology over time was examined at each grid point using the time series of its FPC scores and classical metrics derived from these, namely mean (referred to as "climatological score", see FPC1, FPC2 and FPC3 climatological scores in Figures I.3-I.4, and supplementary material Figure I.15, respectively), standard deviation (Figures I.5 and I.13), and trend (see Section I.3.c).

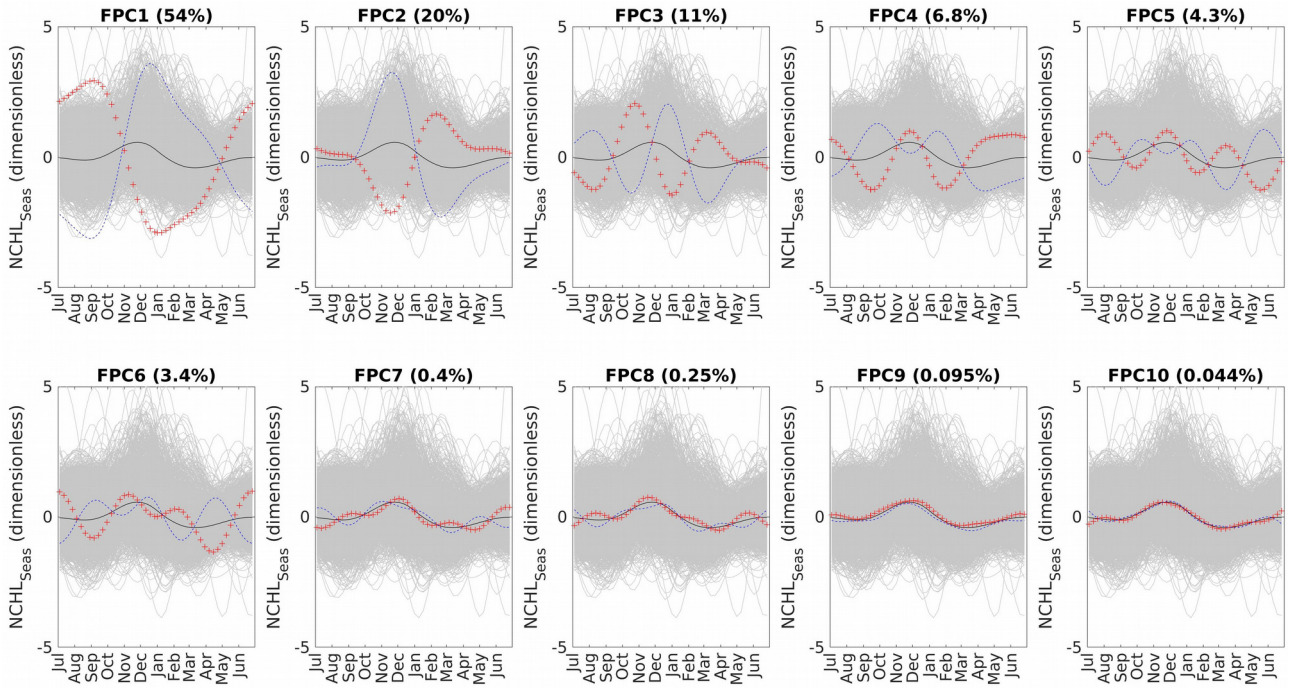


Figure I.8: Ten first FPCs (modes of variation) obtained after running FPCA (see Section I.3.b). The value in brackets above each panel is the percentage of variance explained by the corresponding FPC. In each panel, the blue (red) curve depicted by "-" ("+") symbols shows a typical phenology associated with a negative (positive) FPC score. The continuous black line displayed in every panels correspond to the average phenology in the study zone (Southern Ocean, south of 30ªS).

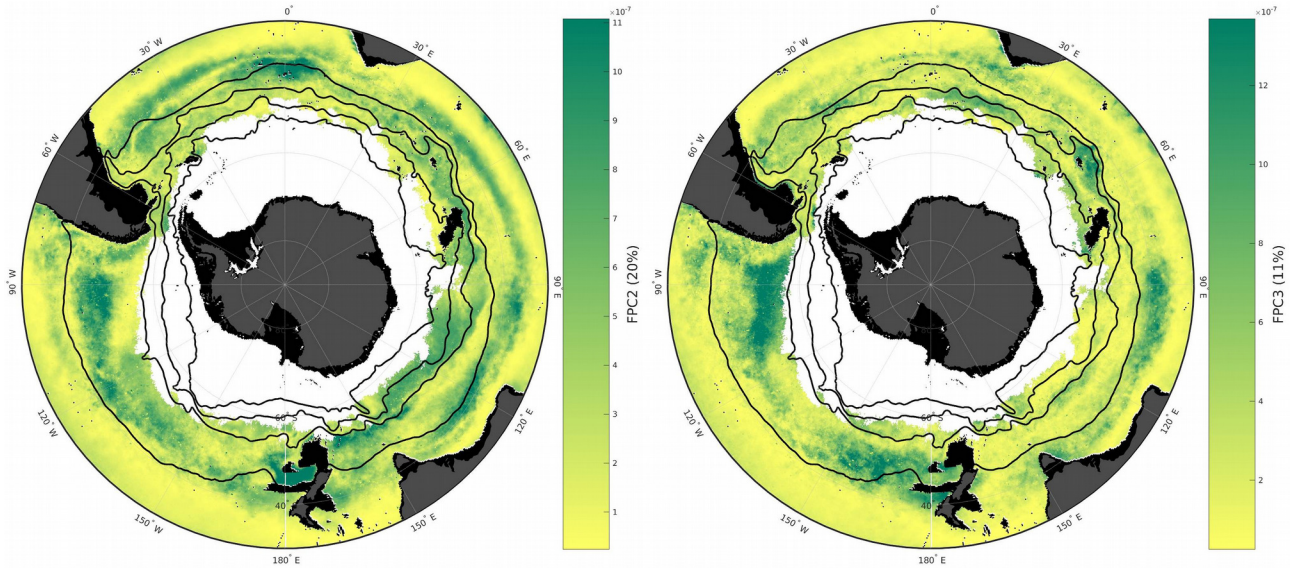


Figure I.9: Contribution of grid points to the creation of first (left panel) and second (right panel) axes of FPCA (see computation details in Text S2). Black lines represent the major fronts of the SO. Gray areas represent land areas. Black areas represent marine areas with bathymetry shallower than -1000 meters. White areas represent regions with less than 50% valid CHL values in the 1998-2022 time series (see details in caption of Figure I.7).

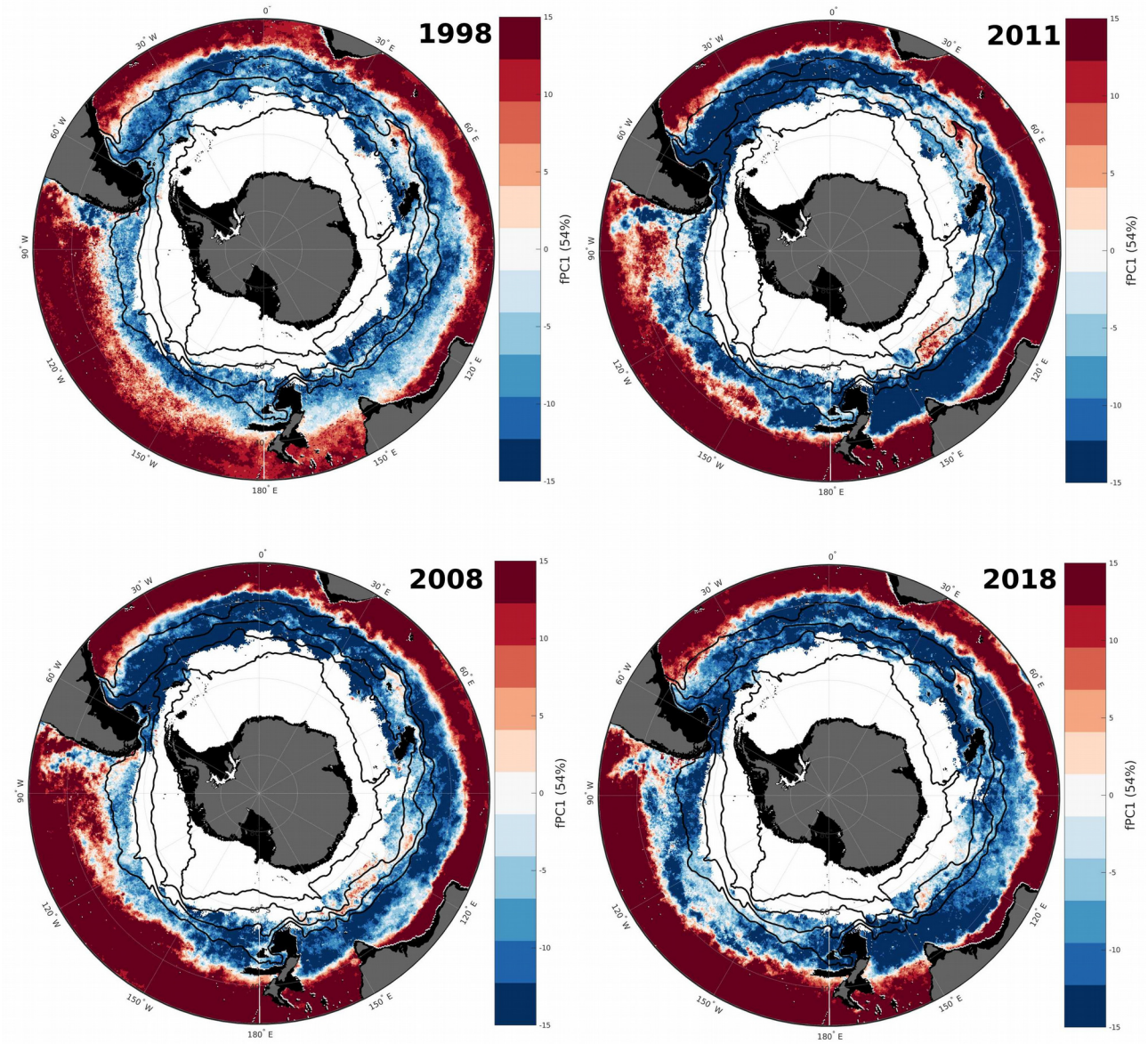


Figure I.10: Example of inter-annual variability of spatial distribution of FPC1 scores. Spatial distribution of FPC1 scores for years 1998 (top left), 2008 (bottom left), 2011 (top right), 2018 (bottom right). Black lines represent the major fronts of the SO. Gray areas represent land areas. Black areas represent marine areas with bathymetry shallower than -1000 meters. White areas represent regions with less than 50% valid CHL values in the 1998-2022 time series (see details in caption of Figure I.7).

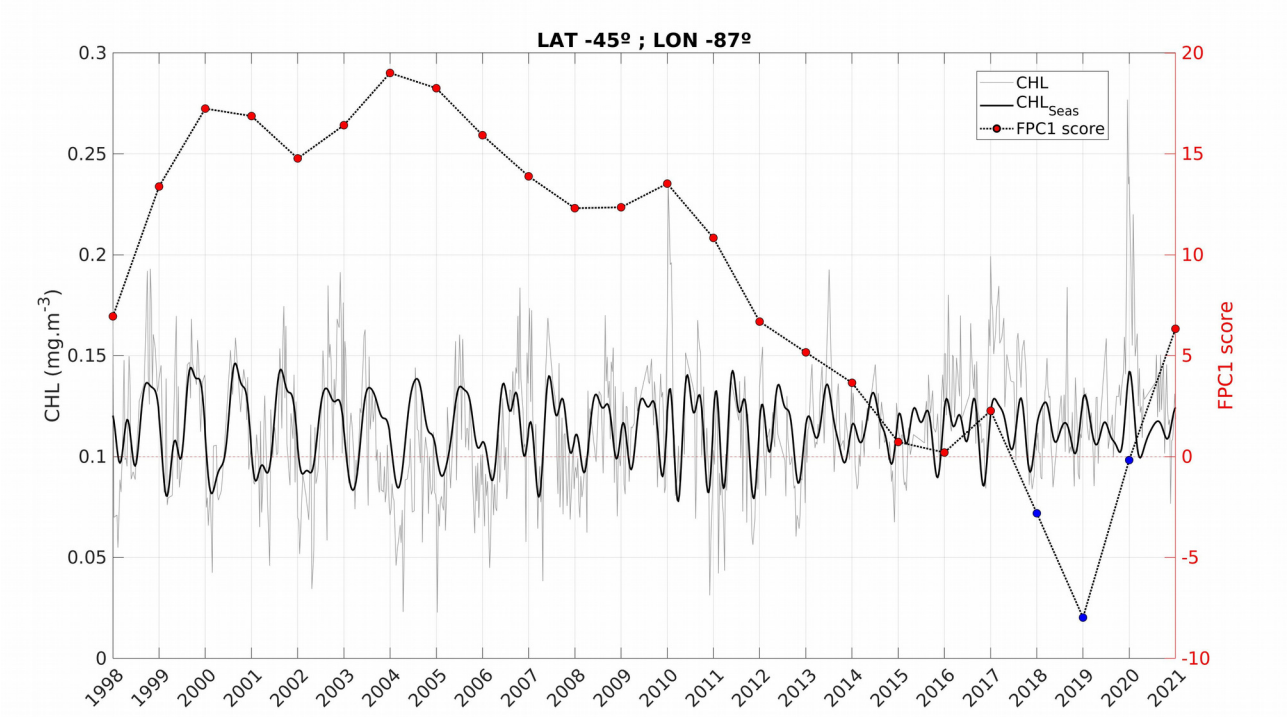


Figure I.11: Example of time series of CHL with FPC1 score. Interpolated CHL time series (continuous gray line) are represented together with their seasonal component (CHL_{Seas}, continuous black line) and FPC1 score (dotted black line with colored circles) for a grid point located in the Pacific sector of the SO (LAT 45ºS, LON 87ºW). The color of the circles code the sign of FPC1 score each year, with blue (red) circles representing negative (positive) FPC1 scores. The change in CHL phenology is highlighted by decreasing values of FPC1 score over time.

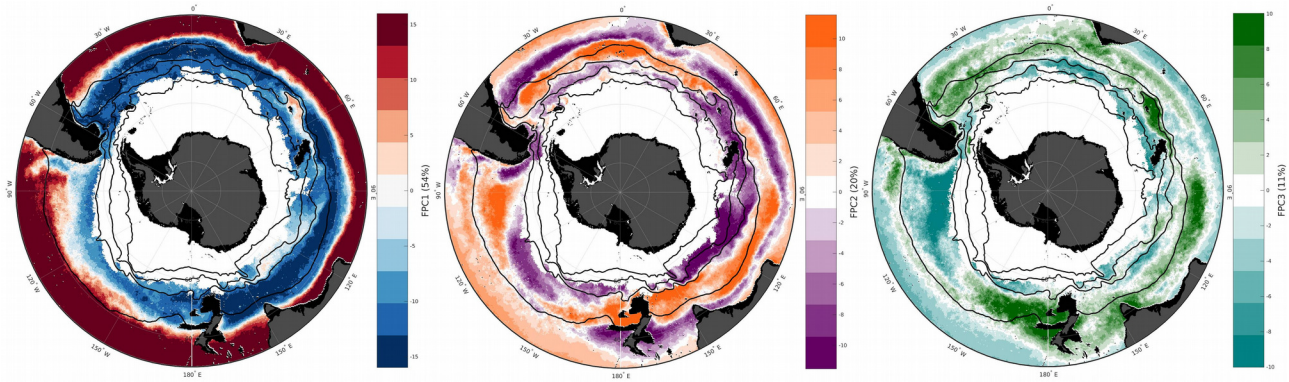


Figure I.12: Spatial distribution of FPC1 (left), FPC2 (middle) and FPC3 (right) climatological score. The colors used to represent FPC scores are the same as in Figure I.2. Black lines represent the major fronts of the SO. Gray areas represent land areas. Black areas represent marine areas with bathymetry shallower than -1000 meters. White areas represent regions with less than 50% valid CHL values in the 1998-2022 time series (see details in caption of Figure I.7).

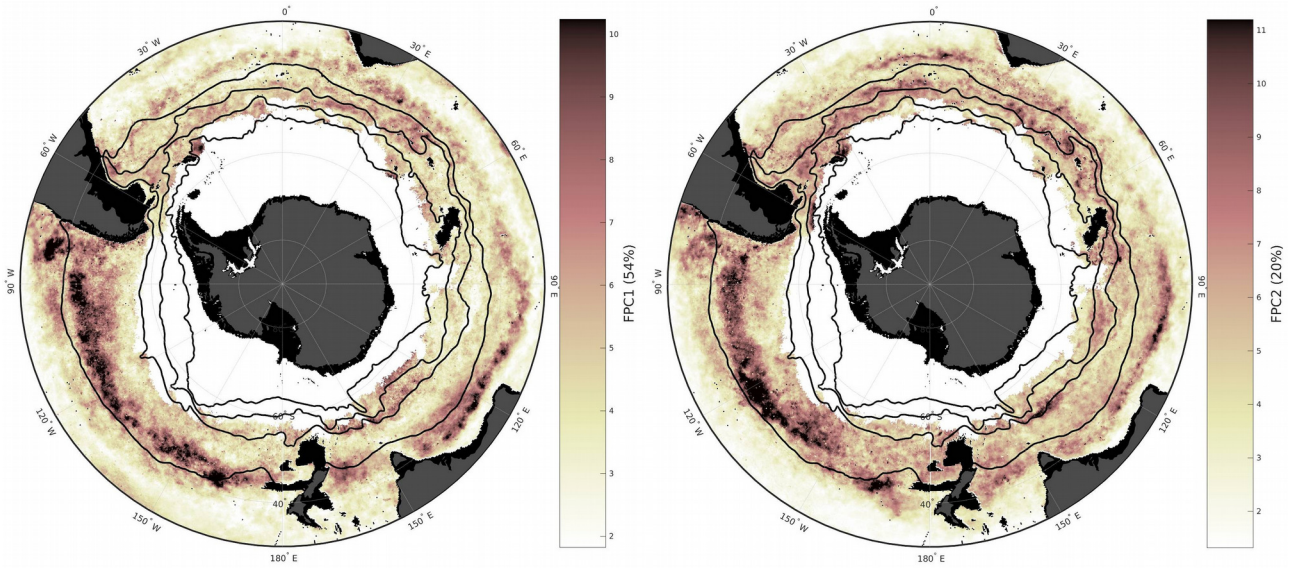


Figure I.13: Standard deviation of FPC1 (left panel) and FPC2 (right panel) scores. Black lines represent the major fronts of the SO. Gray areas represent land areas. Black areas represent marine areas with bathymetry shallower than -1000 meters. White areas represent regions with less than 50% valid CHL values in the 1998-2022 time series (see details in caption of Figure I.7).

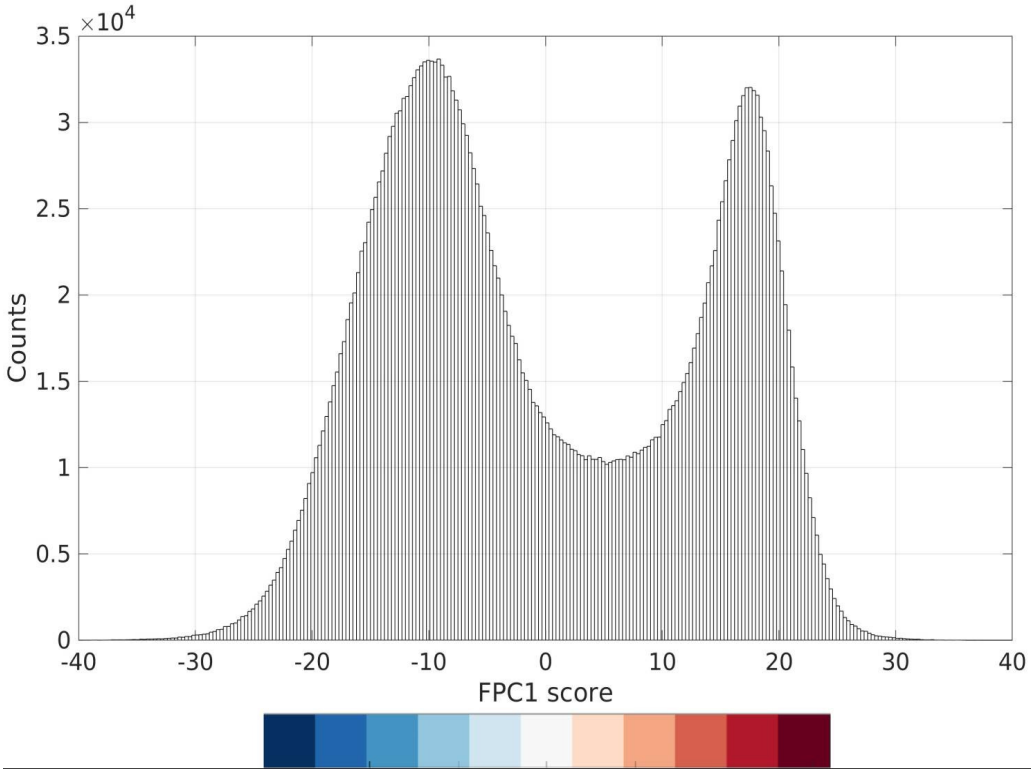


Figure I.14: Histogram of FPC1 scores. The horizontal colorbar repeats the colors used in Figure I.3.

Text S3. Contribution of seasonal component to the total CHL variance

The sub-seasonal time series $CHL_{SS}(t)$ of each grid point were segmented into 24 time series on the same basis as the individual annual cycles defined in Section I.3.b. The total variance of each CHL annual cycle can be expressed as the sum of the variance in the multi-annual, the seasonal and the sub-seasonal components of the time series (see Section I.3.a), and the covariances between these components. The covariance terms between the components of the time series are - by construction - small because they describe variations in the signal corresponding to different frequency domains. Consequently, most of the variance of an annual cycle can be examined through the individual variances of the time series components (Prend et al. 2022). The contribution of the seasonal component relatively to the total CHL annual cycle variance was therefore defined as:

$$VAR_{Seas} = var(CHL_{Seas}) / (var(CHL_{MA}) + var(CHL_{Seas}) + var(CHL_{SS})) \quad \text{(I.3)}$$

where VAR_{Seas} reflects the contribution of the seasonal component, expressed as percentage of the total CHL annual cycle variance. One value of VAR_{Seas} is computed for each annual cycle, so that each grid point is associated with a time series of 24 VAR_{Seas} values. As for FPC1 scores, the climatological mean value of VAR_{Seas} was computed for each grid point.

We then compare for each grid point the mean percentage of variance explained by the seasonal component to the standard deviation of the time series of FPC1 scores (supplementary material, Figure I.15).

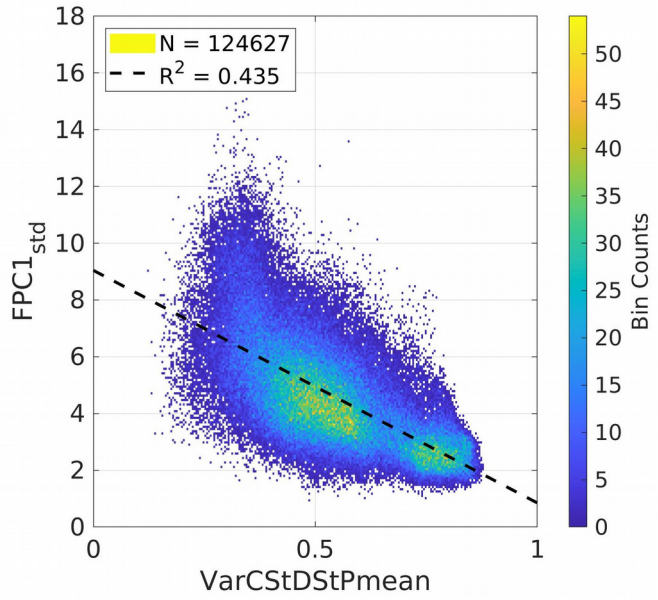


Figure I.15: Scatter plot of $FPC1_{STD}$ versus climatological mean VAR_{Seas}

Text S4. Sensitivity test to the inclusion of the sub-seasonal component in the FPCA

A sensitivity test to the inclusion of the sub-seasonal component in the FPCA analysis was performed to identify the weight of the sub-seasonal component in the phenology analysis (discussed in Section I.5.b). A similar FPCA analysis to that applied to the seasonal component (Section I.3.b) was performed with annual cycles of CHL defined by the sum of the seasonal (CHL_{Seas}) plus the sub-seasonal (CHL_{SS}) components. For each grid point, the normalized sum of the seasonal plus sub-seasonal time series $NCHL_{Seas+SS}$ is expressed as

$$NCHL_{Seas+SS} = \frac{\left(CHL_{Seas}+CHL_{SS}-\overline{CHL_{Seas}+CHL_{SS}}\right)}{STD\left(CHL_{Seas}+CHL_{SS}\right)} \quad \text{(I.4)}$$

where STD is the standard deviation. $NCHL_{Seas+SS}$ has mean zero and standard deviation equal to one. The results of the sensitivity test are examined through the modes of variation resulting from the FPCA and their associated explained variance (Figure I.16).

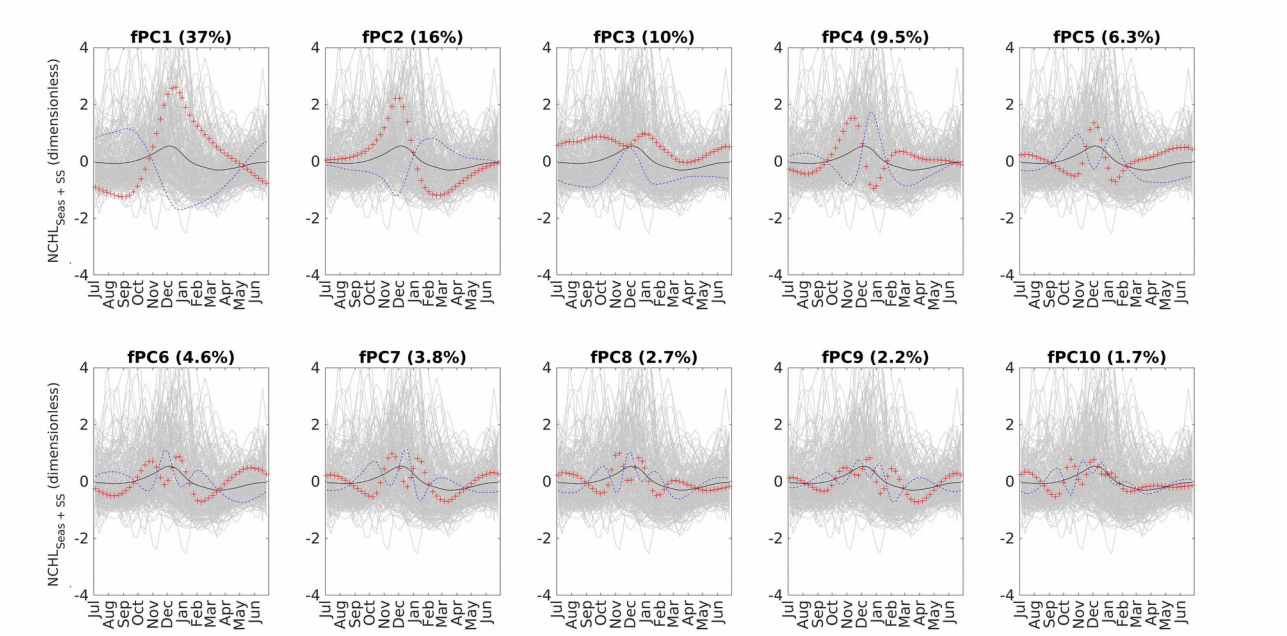


Figure I.16: Ten first FPCs (modes of variation) obtained after running FPCA with sub-seasonal component included in the time series (see Section I.5.b). The value in brackets above each panel is the percentage of variance explained by the corresponding FPC. In each panel, the blue (red) curve depicted by "-" ("+") symbols shows a typical phenology associated with a negative (positive) FPC score. The continuous black line displayed in every panels correspond to the average phenology in the study zone (Southern Ocean, south of 30ºS).

Text S5. Robustness of the phenology analysis

To test the robustness of the phenology analysis proposed in the present study, we performed a similar analysis to that described in Section I.3.b, however using the ESA OC-CCI ocean-colour data (see Section I.2.d). Results of phenology analyses based on GlobColour data (Section I.4) and OC-CCI data (Section I.5.c) were compared through the modes of variation resulting from both FPCAs and their associated explained variance (Figure I.17), the spatial distribution of the scores on the first axis (Figure I.18, left panel) and the long-term trend of the scores on the first axis (Figure I.18, right panel).

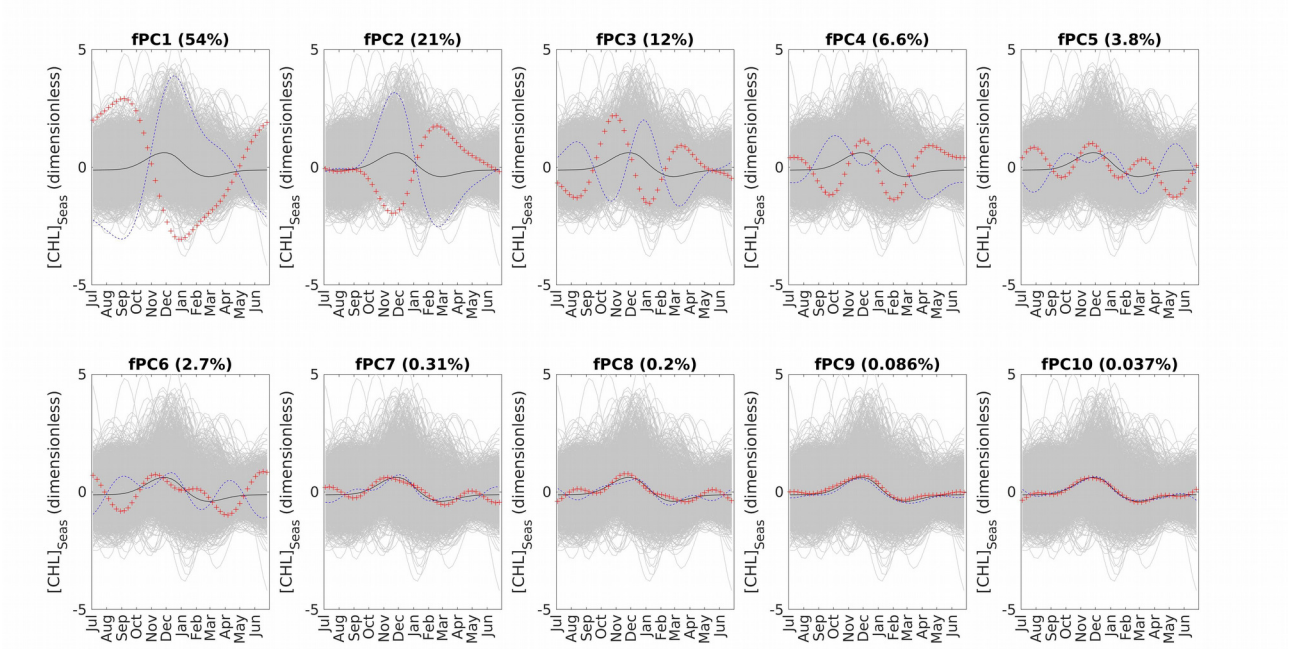


Figure I.17: Ten first FPCs (modes of variation) obtained after running FPCA with OC-CCI ocean-colour data (see Section I.2.d). The value in brackets above each panel is the percentage of variance explained by the corresponding FPC. In each panel, the blue (red) curve depicted by "-" ("+") symbols shows a typical phenology associated with a negative (positive) FPC score. The continuous black line displayed in every panels correspond to the average phenology in the study zone (Southern Ocean, south of 30ºS).

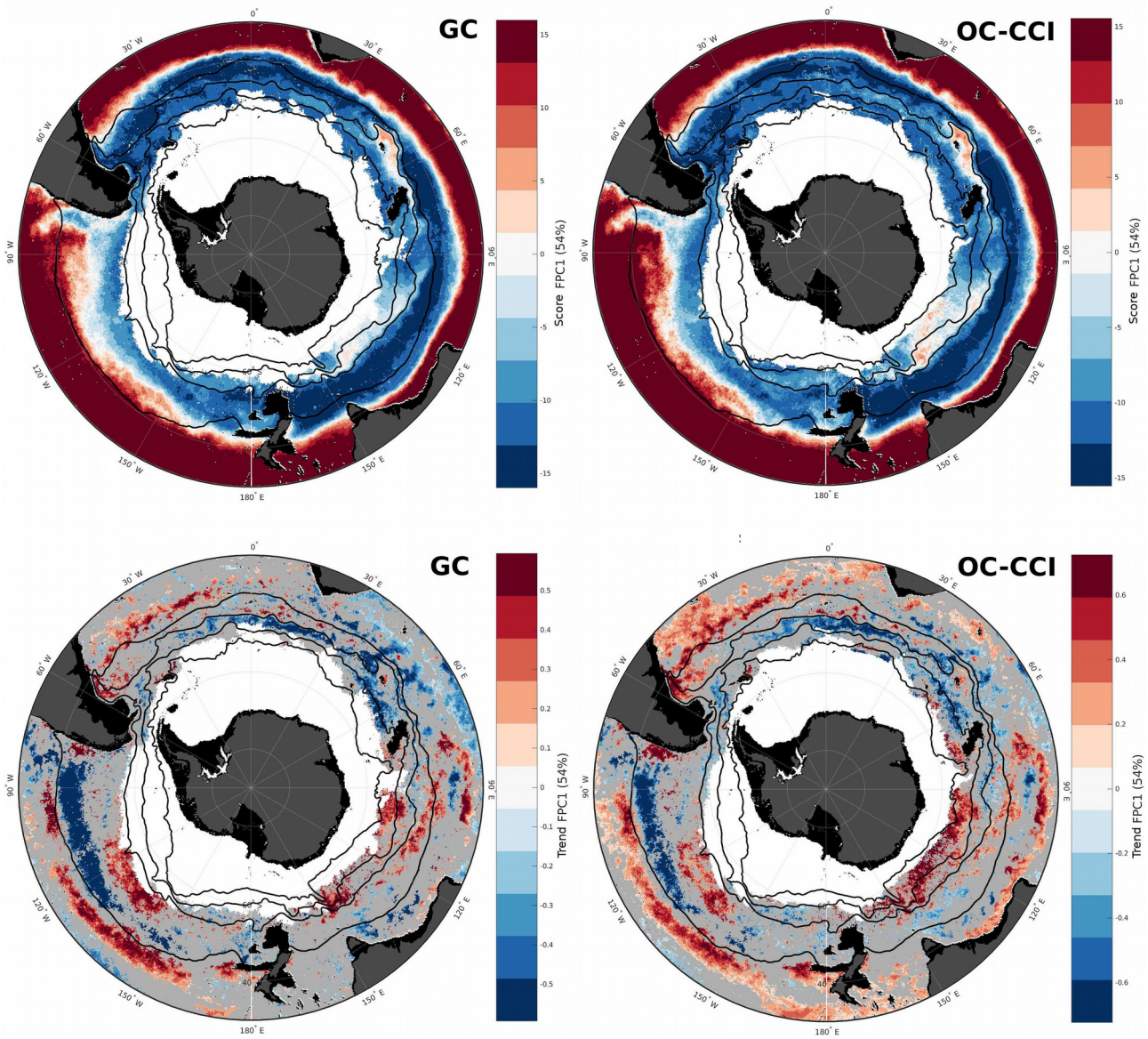


Figure I.18: Comparison between phenology analyses performed with GlobColour and OC-CCI data. Spatial distribution of FPC1 scores (top panels) and trend in FPC1 score for the 1998-2022 period (bottom panels) obtained with GlobColour (GC, left panels) and OC-CCI data (right panels). Black lines represent the major fronts of the SO. Gray areas represent land areas. Black areas represent marine areas with bathymetry shallower than -1000 meters. White areas represent regions with less than 50% valid CHL values in the 1998-2022 time series (see details in caption of Figure I.7).

Text S6. CHL biomass accumulated at each location during the year

To account for the CHL biomass accumulated at each location during the year, the integral over time of each annual cycle (Σ_{CHL}) was computed. The annual cycles used for the computation of Σ_{CHL} were the sum of all the components of the time series. To account for the log-normal distribution of CHL values, the integral over time of each annual cycle (Σ_{CHL}) was log-transformed.

$$\Sigma_{CHL} = \log\left(\int_{seasonal\ cycle} CHL_{MA}(t) + CHL_{Seas}(t) + CHL_{SS}(t)\, dt \right) \quad \text{(I.5)}$$

where CHL_{MA}, CHL_{Seas} and CHL_{SS} are the multi-annual, seasonal, sub-seasonal components of the time series of CHL, respectively (see Section I.3.a).

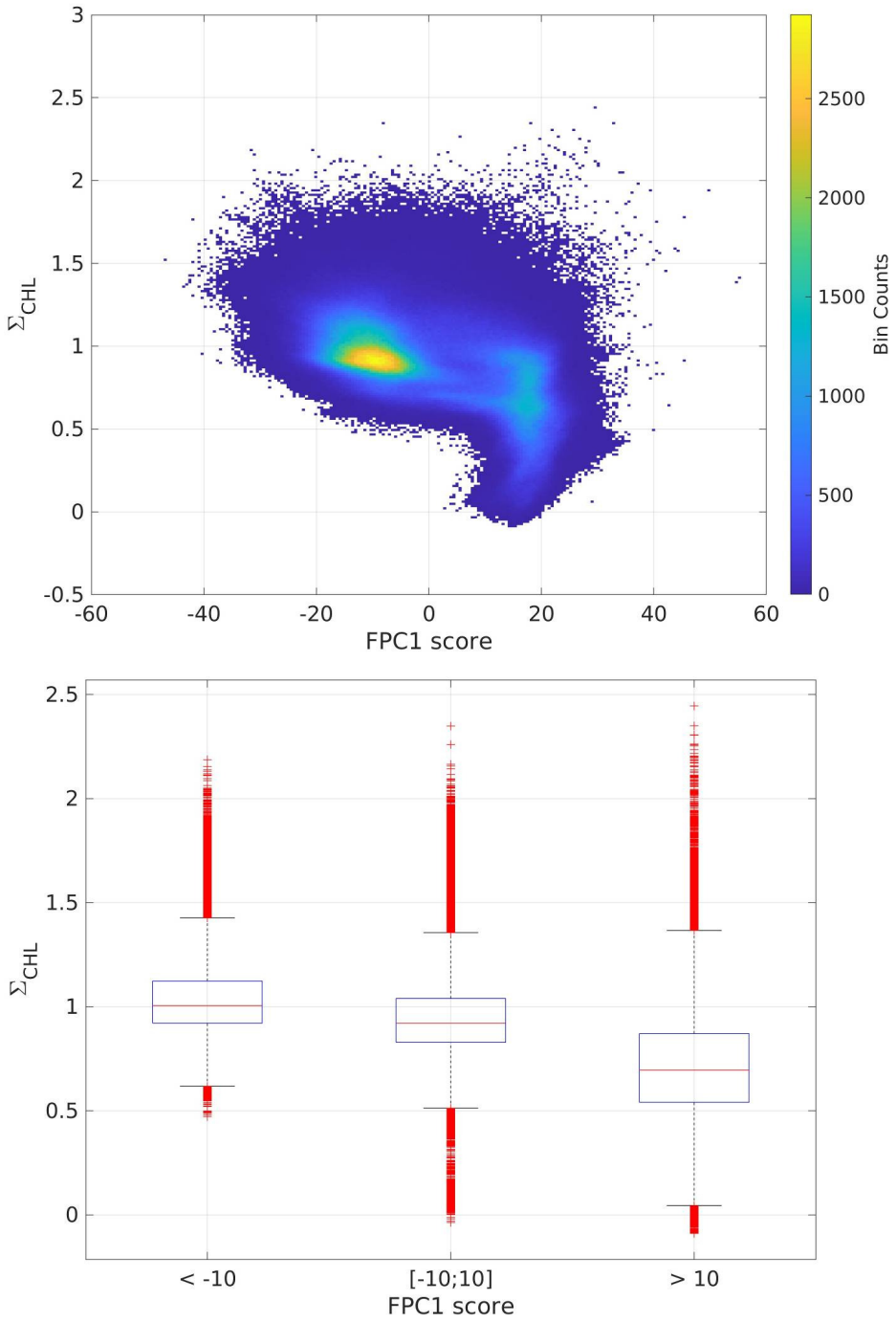


Figure I.19: Comparison between FPC1 score and CHL biomass accumulated during the year. (top) Scatter plot with all points of the study zone (bottom) same data, in the form of a boxplot grouping FPC1 scores into three bins: FPC1 score < -5 (left), -5 <FPC1 score < 5 (middle), 5 < FPC1 score, and associated values of Σ_{CHL}.

Chapter II

A phenological bioregionalization of the Indian Sector of the Southern Ocean

Contents of Chapter II

II.1. Introduction

The Southern Ocean (SO) is a vast, dynamic ocean, that connects with the other oceans of the globe (Marshall & Speer 2012; Talley, 2013; Rintoul, 2018). Its flow, driven by the Antarctic Circumpolar Current (ACC), circles the Antarctic continent, unblocked by continents. The ACC constitutes the main component of the SO circulation scheme (Rintoul et al., 2001). For its high connectivity with the other oceans, the SO has a major role in marine ecosystem functioning on global scales (Murphy et al., 2021). Therefore, understanding how SO biomes are distributed and how they respond to environmental forcing is crucial to assess and manage the impacts of climate change (Sarmiento et al., 2004; Constable et al., 2014; Poloczanska et al., 2016; Henson et al., 2020).

Bioregionalization of the oceanic biomes enables grouping of oceanic zones with similar characteristics. These characteristics may be derived from abiotic (Rosso et al., 2020) or biotic (Fabri-Ruiz et al., 2020) variables, or both (Reygondeau et al., 2013). Longhurst's biogeographical provinces (Longhurst, 2007) are often cited as a reference for classification of the world's oceans into main biomes. Longhurst provinces comprise open-ocean biomes and coastal biomes. Other bioregionalizations only focus on open-ocean areas (Fay & McKinley, 2014). While many bioregionalizations proposed to divide the world's oceans using both abiotic and biotic criteria, other bioregionalizations were purely based on the phenology of remote-sensed chlorophyll-a (CHL) (D'Ortenzio & Ribera d'Alcalà, 2009). Bioregionalization is a powerful tool to examine different biomes and explore the key mechanisms implied in ecosystem functioning (Ardyna et al., 2017). Bioregionalization is also a method to detect changes in time, through examination of variations over time in spatial distribution of described clusters (Fay & Mc Kinley, 2014; Mayot et al., 2016).

Ecosystem dynamics highly rely on primary production pathways, in the sense that different dominant species of primary producers might trigger different trophic chains (Schofield et al., 2017; Henley et al., 2020). Phytoplankton community composition can be described by so-called Phytoplankton Functional Types (PFT). Phytoplankton Size Classes (PSC) are also used to characterize phytoplankton community composition (IOCCG, 2014). A series of PFT (or PSC) models were recently developed to infer phytoplankton community composition-related information from satellite data (Mouw et al., 2017). While some PFT models are purely based on remote sensing reflectance (Uitz et al., 2006; Alvain et al., 2008; El Hourany et al., 2019; Xi et al., 2020), others also include environmental data to derive estimates of PFT concentrations (Brewin et al., 2017; Sun et al., 2023).

At the opposite end of the trophic chain, top predators also provide precious information for characterization of ecosystem dynamics (Cherel and Hobson, 2007; Hindell et al., 2020). Particularly notable since the early 2000s, the development of miniaturized animal-attached data loggers has dramatically increased the capacity for scientists to study animal behavior and derive variables to characterize their environment (Ropert-Coudert & Wilson, 2005; Bograd et al., 2010). Bio-logging techniques can be defined as ‘the use of miniaturized animal-attached tags for logging and/or relaying of data about an animal’s movements, behaviour, physiology and/or environment’ (Rutz and Hays, 2009). Over the past two decades, bio-logging techniques, particularly valuable in remote areas like the SO, have become major tools to monitor marine ecosystems, study ocean dynamics and address conservation issues (Watanabe and Papastamatiou, 2023).

In the present study, we proposed a phenological bioregionalization of the Indian sector of the SO based on satellite CHL. We defined different bioregions in the Indian sector of the SO, characterized by distinct CHL dynamics. We validated our results with previous bioregionalization studies performed at global scale or specific to the SO. We also highlighted the advantages brought by the phenology analysis in determining different bioregions in the SO. Finally, we discussed the potential links that could be established between the defined bioregions and ecosystem functioning with variables related to ecosystem composition and spatial distribution of organisms.

II.2. Data

This section provides detailed information about the data processed in the present chapter and associated methods. Part of the data and methods used in the present study are common to Chapter I (see Sections I.2 and I.3) and Chapter II (present chapter). Therefore, these sections were hereafter summarized and when necessary, references to the corresponding sections in Chapter I were provided for detailed information.

II.2.a. Study period and area of interest

The study period extends from 1998 to 2022. The area of interest of the present study is the Indian sector of the SO. The SO was defined as the oceanic region located South of the 30ºS parallel. In the present study, the Indian sector of the SO was delimited to the west by the 20ºE meridian and to the east by the 150ºE meridian (see supplementary material, Figure II.10).

All data processed in the present study were mapped with a 0.25º x 0.25º resolution (~20 km in the region of the study), which we judged adapted for the purpose of the study.

II.2.b. Coastline and bathymetry data

The coastline and bathymetry data used in the present chapter were identical to those used in Chapter One (see Section I.2.b), namely, the 50 m resolution coastline data product (version 5.1.0) distributed by Natural Earth (available at https://www.naturalearthdata.com/downloads/50m-physical-vectors/) and the ETOPO1 1 Arc-Minute Global Relief Model data from National Centers for Environmental Information at the National Oceanic and Atmospheric Administration (NCEI NOAA; available at https://www.ngdc.noaa.gov/mgg/global/relief/ETOPO1/data/), respectively.

Grid points with bathymetry shallower than -1000 m were not considered in the study. In total, 70 251 grid points (84%) were retained (out of 83 881 grid points in the study zone).

II.2.c. Fronts in the Southern Ocean

The fronts used in the present chapter to delineate water masses boundaries within the SO were identical to those used in Chapter One (see Section I.2.c), namely, the Subantarctic Front (SAF), the Polar Front (PF) and the Southern ACC Front (SACCF) from Park et al. (2019) (available at https://www.seanoe.org/data/00486/59800/) and the the Subtropical Front (STF), based on the definition by Orsi, Whitworth & Nowlin (1995).

II.2.d. Chlorophyll-a concentration (CHL) satellite data

We used 8-day averaged Level 3 CHL products from January 1998 to December 2022 from the CHL product distributed by the Copernicus Marine Service GlobColour project (available at

http://www.globcolour.info/). The Copernicus Marine Service GlobColour product merges the data from multiple sensors. The native resolution of the satellite-derived CHL product is 4 km. We mapped the data with a 0.25º x 0.25º resolution.

The CHL data used in the present chapter was restricted to the southern spring and summer seasons, defined as the period from October to March. Time series with more than 50% missing values over the spring-summer record were removed and the corresponding grid points were not considered in the study. In total, 60 446 grid points (72%) were retained (out of 83 881 grid points in the study zone). The relevance of focusing on CHL production in the SO during spring-summer period is discussed further in this chapter (see Section II.4.d). CHL spring-summer climatology (1998-2022) was computed for each pixel as the mean value of CHL for the period extending from October to March from 1998 to 2022 (see supplementary material Figure II.10).

II.2.e. Phytoplankton Size Classes (PSC) model

We used PSC data from the ecological model developed in Sun et al. (2023). The 17-parameters PSC model developed by Sun et al. (2023) used in the present study is based on ocean-color and Sea Surface Temperature (SST) data. The model provides relative concentration of three PSCs (micro-, nano, picophytoplankton) compared to the total CHL concentration. The PSC model uses [CHL] data from European Space Agency Ocean Color Climate Change Initiative (ESA OC-CCI; available at http://www.oceancolour.org) and monthly composites of Optimal Interpolation Sea Surface Temperature (OISST, version 2) data at 1º resolution distributed by the Physical Sciences Laboratory at the National Oceanic and Atmospheric Administration (PSL NOAA; available at https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html).

II.2.f. Phytoplankton pigment database

Phytoplankton pigment data obtained by High Pressure Liquid Chromatography (HPLC) was used to validate PSC products. To optimally cover the study zone with HPLC samples, we gathered data from the Service d'Analyses de Pigments par HPLC of the Laboratoire d'Océanographie de Villefranche (LOV) at the Institut de la Mer de Villefranche (SAPIGH; https://lov.imev-mer.fr/web/facilities/sapigh/), from SeaWiFS Bio-optical Archive and Storage System distributed by the National Aeronautics and Space Administration (NASA) Ocean Biology Processing Group (OBPG) (SeaBASS; available at https://seabass.gsfc.nasa.gov/) and from the Australian Ocean Data Network (AODN) Portal of Australia's Integrated Marine Observing System (IMOS) (AODN; available at https://portal.aodn.org.au/). The spatial distribution of the HPLC samples included in the present study was provided in supplementary material Figure II.19.

II.2.g. Southern Elephant Seal (SES) dataset

In the present study, we used data from seal-borne data loggers placed on southern elephant seals (SES). The SES data were collected as part of the Système National d'Observation

Mammifères Echantillonneurs du Milieu Océanique (SNO-MEMO) and made freely available by the International Mammals Exploring the Oceans Pole to Pole Consortium and the national programs that contribute to it (MEOP; available at www.meop.net/database/meop-databases).

The so-called bio-logging devices typically record behavioral data (e.g. location, diving depth, prey capture). Bio-logging devices are also useful to characterize the SES's oceanic environment through the measurement of variables like temperature, salinity, light, chlorophyll-a fluorescence. The present study focuses on SRDLs' depth and accelerometry data. The location of the SES dives as well as the bottom depth of each SES dive were computed. The accelerometer data were processed at MEOP following the method described in Viviant et al. (2010) and Gallon et al. (2013) to detect prey capture attempts (PCAs). We mapped the SES data with a 0.25º x 0.25º resolution according to the grid defined in the present section. We computed for each equipped animal (i) the mean number of dives per grid point in the animal trip (further referred to as number of dives), (ii) the maximum depth per dive (further referred to as dive depth), and (iii) the mean number of PCA per dive (further referred to as number of PCA).

The MEOP database is divided into two different datasets according to data resolution. Low-resolution SES data is transmitted by satellite approximately every 15 dives of the animal, with relatively low vertical resolution (~10 m in the surface layer, ~100 m at depth). High-resolution data depend on bio-logging device recovery, has high vertical resolution (1 m) and contains every dive of the animal. High- and low-resolution data hence differ both by their vertical and horizontal resolution. Including datasets with different horizontal resolution was a concern for the counting of number of dives per grid point. High-resolution data introduces a bias compared to low-resolution data by necessarily accounting for more dives per grid point. Yet, high-resolution bio-logging devices are generally used for specific research purposes, and do not equally cover all months of the year (typical deployments are from October to January), nor the SES population (high-resolution devices are mainly placed on female SES) (Guinet et al., 2014). Therefore, for homogeneity of the data examined in the present study, the mean number of dives per grid point in each SES trip at sea was computed with the low-resolution dataset, gathering in total 67 335 SES profiles in the study zone. The high-resolution dataset processed for computation of SES dive depth contained 307 159 SES profiles in the study zone. The PCA data were computed with 97 377 SES profiles (for geographical distribution of SES profiles per dataset, see supplementary material Figures II.21-II.23).

II.3. Methods

II.3.a. CHL Phenology

Phenology analysis of CHL spring-summer cycles

The analysis of CHL phenology was performed following the method described in the previous chapter (see Section I.3.b). In summary, the analysis of CHL phenology consisted in (i) temporal decomposition of CHL time series into multi-annual (CHL_{MA}), seasonal (CHL_{Seas}) and sub-seasonal (CHL_{SS}) components, (ii) normalization of CHL time series, (iii) yearly segmentation of CHL time series, (iv) Functional Principal Component Analysis (FPCA) performed on the seasonal component of CHL time series. In the present chapter, the phenology of CHL was characterized by analyzing the dynamics of CHL during the spring-summer period (October to March).

Reminder of FPCA terminology

The main results of the FPCA are the Functional Principal Components (FPCs). FPCs describe the main modes of variation in CHL spring-summer cycles (FPCs are hereafter indifferently referred to as FPCs, modes of variation and FPCA axes). FPCs were numbered from 1 to 10, in descending order of the associated explained variance. FPCA enables the creation of a new basis defined by the FPCs. Projection of an annual cycle on a FPC is hereafter referred to as its score. Absolute value of the score is not directly interpretable in terms of CHL.

Each grid point is assigned a score each year based on the phenology of the annual cycle recorded at the corresponding location. To each grid point hence corresponds a time series of 24 score values for each FPC. The evolution of CHL annual cycles phenology over time was examined at each grid point using the time series of its FPC scores and classical metrics derived from these, namely mean (hereafter referred to as “climatological score”), and standard deviation.

II.3.b. Clustering

A clustering analysis was performed on the time series. The computed clustering consists in a K-means clustering analysis. The K-means clustering was based both on CHL spring-summer phenology (FPCA analysis) and CHL spring-summer biomass (CHL spring-summer climatology). More precisely, the inputs of the clustering were for each grid point (i) FPC1 score, (ii) standard deviation of FPC1 score ($FPC1_{STD}$), and (iii) the log-transformed spring-summer climatological value of CHL. Prior to performing K-means clustering, all the input data were normalized (mean equal to zero, standard deviation equal to one).

Number of clusters is a user-defined input in the K-means clustering method. The optimum number of clusters was determined using the elbow method. Briefly, the elbow method consists in

successively performing clustering with increasing number of clusters. By assessing at each step within-cluster similarity between points, the elbow method aims at detecting the point from which no more significant improvement is brought by addition of a new cluster. Extended description and graphical support for the method summarized in the present paragraph is provided in supplementary material Text SII.2 and Figure II.14.

Ecosystem-level characteristics of the obtained clusters were derived from PSC and SES data. For each ecosystem variable, Kruskal-Wallis test was performed to test whether the obtained clusters differed in their ecosystem characteristics. A significant result of the Kruskal-Wallis test implies that the distribution of the data of at least one cluster differs from all others.

II.4. Results and Discussion

II.4.a. Phenology analysis

The phenology analysis developed in the present study enables examination of the main modes of variation of CHL spring-summer cycles and of their spatial distribution. Graphical support for the obtained modes of variation, the spatial distribution of FPC scores and standard deviation of FPC scores are provided in supplementary material Figures II.11, II.12, and II.13, respectively.

The FPCs (modes of variation) obtained in the present chapter were at first order equivalent to the FPCs obtained with annual cycles of CHL (Section I.4.a). Briefly, the mode of variation that explains most of the variance (FPC1, 45%) highlights an opposition between CHL cycles with maximum in summer and CHL cycles with minimum in summer (supplementary material Figure II.11, first panel). The second mode of variation (FPC2, 37%) discriminates at finer temporal scale the timing of the summer annual maxima (supplementary material Figure II.11, second panel). From the third axis of the FPCA (FPC3, 12%) and beyond, interpretations of associated modes of variations becomes more delicate because the shapes depicted by the corresponding FPCs are more complex (supplementary material Figure II.11, third to tenth panels).

Additionally, the spatial distribution of FPC scores obtained by analyzing CHL spring-summer phenology (supplementary material Figure II.12) were comparable to those obtained by analyzing CHL annual cycles (Figures I.3-I.4). The first axis (FPC1) mainly depicted a general latitudinal gradient from subtropical to subpolar latitudes. The other FPCs (FPC2 and further) presented more complex distributions with higher zonal diversity.

Thus, the phenology analysis performed in the present study provided similar results to the phenology analysis presented in the previous chapter. However, because focus of the present analysis was placed on spring-summer months, the range of phenologies described by FPC scores did not encompass the same diversity as in when including winter dynamics. For this reason, “spring-summer” FPC1 scores (this chapter) notably appear shifted compared to “annual cycle” FPC1 scores (previous chapter). The shift in FPC1 scores is clearly visible through the larger proportion of red-coloured pixels - notably south of the SAF - in the spatial distribution of spring-summer FPC1 scores (supplementary material Figure II.12) compared to annual cycle FPC1 scores (Figure I.3).

II.4.b. Bioregionalization

Based on the similarity of both their spring-summer CHL phenology and the magnitude of their spring-summer CHL production, grid points were grouped into six clusters (hereafter also referred to as bioregions). The optimal number of clusters was set at six by application of the elbow method to our dataset (for detailed description of the method, see supplementary material Text SII.2 and Figure II.14). Our analysis provides information on CHL biomass (climatological CHL), on CHL phenological type (FPC1), and on the variability of CHL phenology ($FPC1_{STD}$), a proxy for sub-seasonal variability in CHL time series (see Section I.5.b and supplementary material Figure I.15).

It is worth to note that FPC2 scores were not included in the clustering because we judged that the diversity of phenologies depicted by the sole interpretation of FPC1 scores (see Section I.5.a) was sufficient to provide an efficient bioregionalization in the study zone. Additionally, opting for a reduced number of inputs in the K-means analysis enabled a simplified description of the main characteristics of the clusters. Finally, $FPC2_{STD}$ was not included in the clustering inputs because largely redundant with $FPC1_{STD}$ (see Section I.4.c).

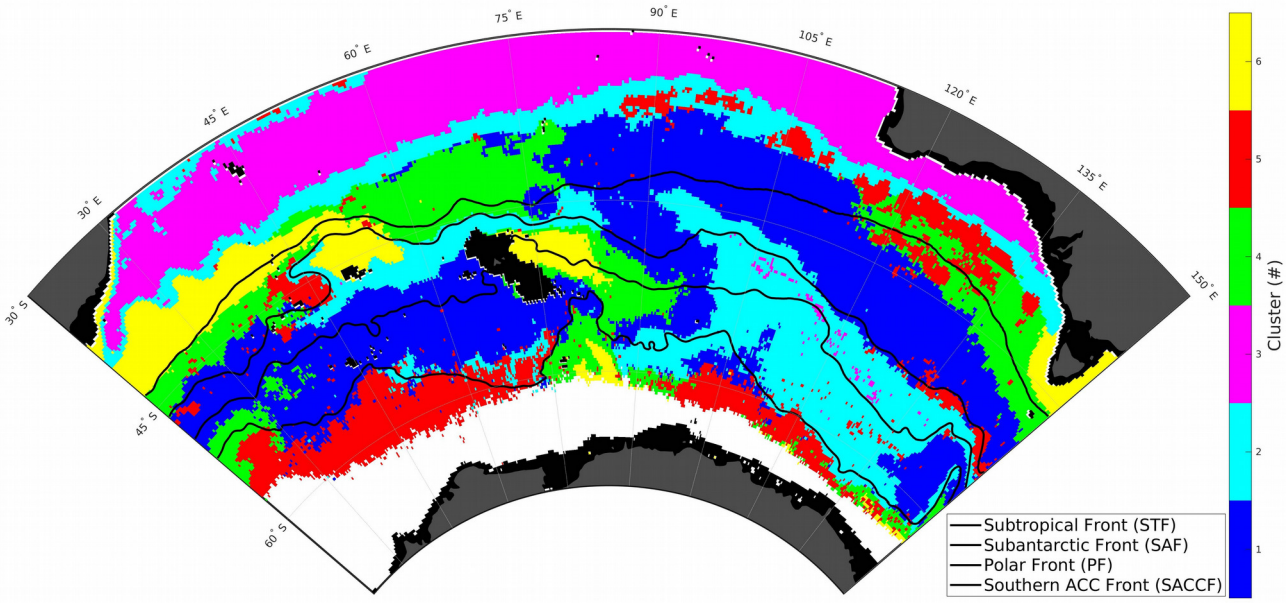


Figure II.1: Phenological bioregionalization of the Indian sector of the Southern Ocean. Spatial distribution of clusters based on CHL biomass and CHL phenology. Black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 spring-summer time series (see Section I.2.d) are represented in white.

The geographical distribution of the six bioregions (Figure II.1) highlights the variety of CHL dynamics in the Indian sector of the SO derived from CHL time series. Based cluster centroid FPC1 score, $FPC1_{STD}$, and spring-summer climatological CHL, the average phenology and amplitude of the spring-summer CHL cycle per bioregion can be reconstructed. Specific characteristics of each bioregion can be hence inferred from cluster centroid data (Figure II.2), as described hereafter. Bioregion numbers were assigned in descending order of geographical extent (number of grid points per cluster, see Figure II.3).

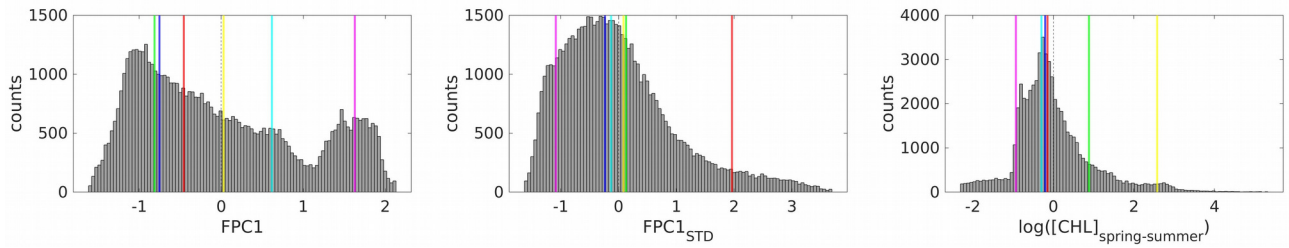


Figure II.2: Clustering inputs and cluster centroids. Histograms of clustering inputs (grey bars) and coordinates of centroid for each clustering input (colored vertical lines). The color used for each cluster is the same as in Figure II.1. Clustering inputs were FPC1 score (FPC1, left panel), standard deviation of FPC1 score ($FPC1_{STD}$, middle panel), and log-transformed spring-summer climatological value of CHL ($log([CHL]_{spring\text{-}summer})$, right panel). Note that the displayed values are normalized values (see Section II.3.b).

The most oligotrophic regions

Bioregion 3 (Figure II.1, magenta) was mainly characterized by a combination of very low CHL production and clearly positive FPC1 score. Grid points of bioregion 3 are associated with typical subtropical phenology (see Section I.4.a). Typical subtropical phenology corresponds to relatively high levels of CHL production in winter, an annual maximum in September, and a rapid decrease of CHL levels after the annual CHL maximum towards an annual CHL minimum in winter. The subtropical phenological type reflects high limitation by nutrients. Moreover, bioregion 3 is associated with very low $FPC1_{STD}$, which reflects low variation in CHL phenology.

Early-bloom, low-productivity regions

Bioregion 2 (Figure II.1, light blue) was the second most represented cluster in the study zone (Figure II.3). Bioregion 2 was characterized by moderate $FPC1_{STD}$, moderate to low CHL production, but a notably high FPC1 score (although still lower than FPC1 scores in bioregion 3). High FPC1 score reflects a relatively early annual maximum, compared to most of the phenologies present south of the STF. The moderate values of $FPC1_{STD}$ bioregion 2. may reflect non-negligible contribution of the the sub-seasonal component to CHL variability.

High productivity regions

Bioregion 6 (Figure II.1, yellow) was the least represented bioregion in the study zone. The main characteristic of grid points in bioregion 6 was high CHL production. Bioregion 6 groups the most productive areas of the Indian sector of the SO (see spring-summer CHL climatology in supplementary material Figure II.10). These areas are mainly located downstream subantarctic islands (e.g. Crozet, Kerguelen) or low bathymetry areas (e.g. southern Kerguelen Plateau), in coastal (e.g. South African coast) and frontal (e.g. northwest Crozet) zones. FPC_{STD} in bioregion 6 revealed that contribution of the sub-seasonal component in bioregion 6 was moderate.

Steady HNLC regions

Bioregion 1 (Figure II.1, dark blue) was the most represented cluster in the study zone. Bioregion 1 was mainly characterized by very negative FPC1 score (Figure II.2), the typical phenology of HNLC waters. As discussed in Chapter I (see Section I.5.a) summer CHL production in HNLC waters is highly limited by by iron delivery. The spatial distribution of bioregion 1 confirms this hypothesis, with grid points corresponding to bioregion 1 being mainly far from any shallow bathymetry areas, or located westward (i.e. “upstream” when considering the flow of the ACC) of those. The average FPC_{STD} in grid points of bioregion 1 is moderate to low.

HNLC regions with influence of sub-seasonal variability

Bioregion 4 (Figure II.1, green) had similar FPC1 score to bioregion 1, which classified bioregion 4 also in HNLC regime. However, the values of $FPC1_{STD}$ in bioregion 4 were on average higher than in bioregion 1, which reflects higher sub-seasonal variability in the grid points belonging to bioregion 4 compared to bioregion 1. Moreover, spring-summer climatological values of CHL were also higher in bioregion 4 than in bioregion 1. Two main patches of grid points belonging to bioregion 4 were located downstream subantarctic islands (namely, Crozet and Kerguelen, see Figure II.1), which reinforces the hypothesis of summer CHL production supported by (sub)mesoscale activity (d’Ovidio et al., 2015).

Highly variable CHL phenology

Finally, bioregion 5 (Figure II.1, red) was mainly characterized by high FPC_{STD}, reflecting highly variable seasonal cycle. CHL phenology bioregion 5 may be highly influenced by variations in CHL at sub-seasonal and inter-annual scales. CHL production in bioregion 5 was moderate. Spatial distribution of grid points belonging to bioregion 5 revealed two main sub-groups. The first sub-group is located in the frontal zone at the interface between the subtropical and the subantarctic phenologies (see Section II.4.a). The second sub-group is located at the sea-ice edge. Grid points of bioregion 5 were grouped because they present highly variable CHL phenology, however the drivers of variability in CHL phenology may be different in each sub-group. While the first sub-group (subtropical frontal) may be subject to higher influence of meandering of the STF, the second one (sea-ice edge sub-group) may be highly dependent on sea ice cover.

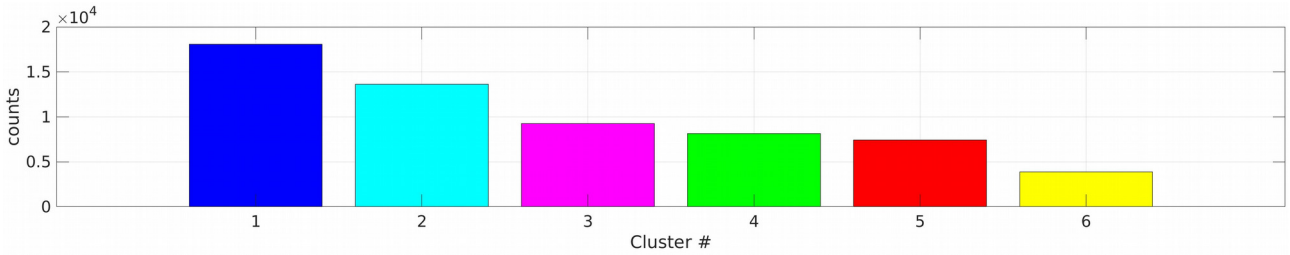


Figure II.3: Number of grid points in each bioregion (cluster). The color used of each cluster is the same as in Figure II.1. In total, 60 446 grid points were included in the study.

Previous studies proposed other bioregionalizations of the SO (Longhurst, 2007; Fay & McKinley, 2014; Ardyna et al., 2017). In the next section we compare our results with these existing bioregionalizations.

II.4.c. Comparison with existing bioregionalizations of the SO

The bioregionalization of the Indian sector of the SO proposed in the present study is in line with other proposed bioregionalizations of the global ocean that globally depict a latitudinal separation of biomes following the main fronts of the SO (Longhurst, 2007; Fay & McKinley, 2014). Although the spatial distribution of the six CHL-derived bioregions described in the present study can obviously not be explained by SO fronts only, clear patterns of SO fronts being boundaries for bioregions are visible in our bioregionalization (Figure II.1). Notably, the SACCF operates in the west part of the study zone a clear delimitation between the area associated with highly variable CHL phenology (bioregion 5) and the HNLC zone (bioregion 4). In addition, the SAF also locates at the interface between bioregions (e.g. bioregions 1 and 2 in the east, bioregions 2 and 4 north of the Kerguelen Plateau).

As previously mentioned, the subtropical zone differs from the subantarctic zone, which itself has different characteristics than the polar zone, making the STF, the SAF and the PF very relevant when spatially analyzing CHL phenology in the SO (see Section I.5.a). However, our analysis provides finer distinction regarding CHL phenology. On top of the meridional classification mainly driven by environmental variables such as sea surface temperature or mixed layer depth, already highlighted in previous studies (Longhurst, 2007; Fay & McKinley, 2014), we highlight in our study clear zonal patterns driven by CHL dynamics (Figure II.1). Including phenology analysis in our classification and focusing on CHL data only was indeed performed with the aim to highlight zonal patterns of CHL productivity.

Bioregionalization of the SO by Ardyna et al., 2017 reveal a global meridional pattern, but also, a clear zonal differentiation throughout the SO is highlighted their study. Their classification,

following the method developed in the Mediterranean Sea in D'Ortenzio & Ribera d'Alcalà (2009), is based on CHL phenology only, with no addition of environmental variables. In our study, also based on CHL data only, we obtained different distribution of the bioregions. Firstly, we used a different number of clusters (seven in Ardyna et al. (2017), six in our study). Secondly, different distribution of the bioregions resulted from the addition of an input related to CHL biomass. Secondly, the addition of a criteria based on the variability of the phenology (namely, $FPC1_{STD}$) in our analysis brings another insight into phenology analysis. The addition of $FPC1_{STD}$ provides information related to sub-seasonal variability (see Section I.5.b and supplementary material Figure I.15).

II.4.d. Use of phenological clusters to derive ecosystem-based analysis

In the present section we discuss the potential of the phenology analysis based on functional data and resulting classification to derive ecosystem-based analysis of the SO. Ecosystem-based analysis is useful to derive relevant proxies characterizing the entire ecosystem, as well as to study the evolution in time of the SO ecosystems.

Focusing on the spring-summer period

Cloud cover is a strong limiting factor for remote-sensing CHL estimates in the SO, especially in winter. Focusing on the spring-summer period enables extension of the validity domain of ocean-colour observations to the south, compared to the analysis presented in the previous chapter (i.e. a reduced number of discarded grid points). The greater number of available observations can be visually assessed by comparing the extent of white-coloured areas (discarded grid points) in spring-summer CHL climatology (supplementary material Figure II.10) and in annual CHL climatology (supplementary material Figure I.7). By extending the geographical extent of included observations, a greater number of ecosystem variables could be included. However, part of the annual CHL production is obviously missed when only considering the spring-summer period only. To investigate the part of the annual CHL production that is concentrated on the spring-summer months, we computed the ratio between spring-summer CHL production and annual CHL production (see supplementary material Figure II.15). In low latitudes, the ratio between spring-summer and annual CHL production falls below 50%, which is in line with FPC1 scores depicting higher CHL levels in winter at subtropical latitudes (see supplementary material Figure II.12). The ratio between spring-summer and annual CHL production increases in high latitudes. South of 45ºS, the ratio between spring-summer and annual CHL production is above 60%. In the present study, important focus is placed on ecosystem dynamics at subantarctic to subpolar latitudes (i.e. south of 45ºS). The main objective of including subtropical latitudes was to constitute a baseline for CHL phenology and CHL production. Therefore, based on the examination of the ratio between spring-summer CHL production and annual CHL production, we consider that focusing on spring-summer months for bioregionalization was relevant and adapted for the purpose of the study. Nevertheless, the ecosystem variables examined in the present section included data from periods covering the entire annual cycle.

Primary production pathways: phytoplankton community composition

Bioregionalization of the SO based on CHL phenology and climatological CHL enables efficient visualization of main phenologies, while keeping a metric related to CHL biomass. The information of bioregion membership was related to phytoplankton size classes (PSC) to infer potential link between amplitude and phenology of CHL production, and phytoplankton community composition. The dynamics of fractions of PSC (pico-, nano-, microphytoplankton) was examined for each cluster (Figures II.4-II.6). Kruskal-Wallis test was performed on the climatological values of PSC fractions and absolute concentrations to test whether the obtained clusters differed in their PSC composition (Section II.3.b). All tests resulted significant (all p values $< 10^{-4}$). In the examination of PSC data, we judged relevant to analyze the entire annual cycle to catch the seasonal dynamics of PSC including winter dynamics. In addition, we judged relevant to analyze fractions of PSC to disentangle CHL biomass dynamics from phytoplankton community composition dynamics. However, the absolute concentrations of PSC were computed and provided in supplementary material Figures II.16-II.18.

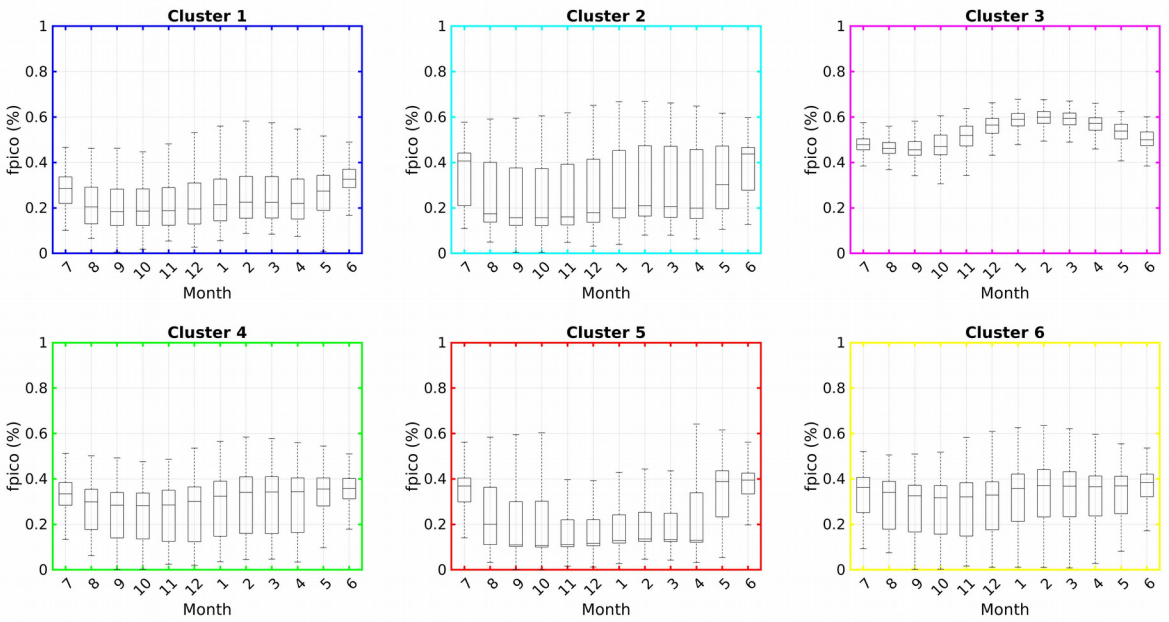


Figure II.4: Fraction of picophytoplankton per bioregion and per month. Cluster number (bioregion) is displayed above each panel and the color of each panel refers to the colors used in Figure II.1. Fraction of picophytoplankton (fpico) is given in percentage of the total CHL. Months in the abscissa are displayed with annual cycle centered on Austral summer (i.e. month of December, number 12). Each black box represents the distribution of the PSC data for the corresponding month of the year. The horizontal black line inside each box represents the median. The top and bottom limits of each box are the 25th and 75th percentiles, respectively. The vertical dashed black lines extending above and below each box represent the full range of non-outlier observations.

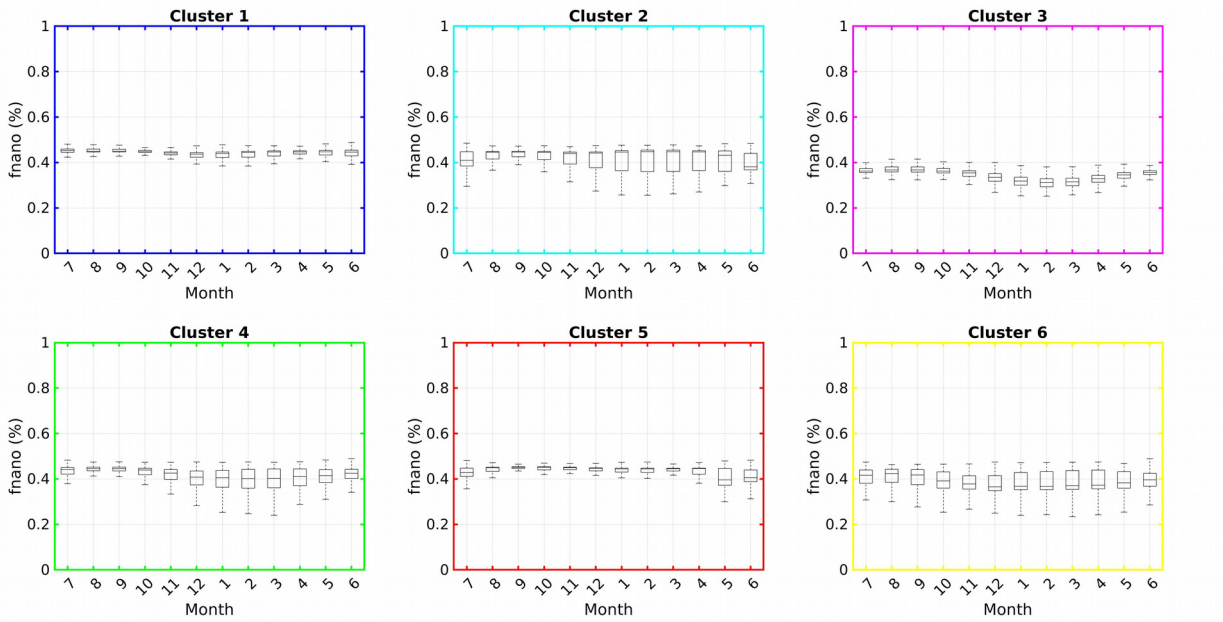


Figure II.5: Fraction of nanophytoplankton per bioregion and per month. Cluster number (bioregion) is displayed above each panel and the color of each panel refers to the colors used in Figure II.1. Fraction of nanophytoplankton (fnano) is given in percentage of the total CHL. Months in the abscissa are displayed with annual cycle centered on Austral summer (i.e. month of December, number 12). Each black box represents the distribution of the PSC data for the corresponding month of the year. The horizontal black line inside each box represents the median. The top and bottom limits of each box are the 25th and 75th percentiles, respectively. The vertical dashed black lines extending above and below each box represent the full range of non-outlier observations.

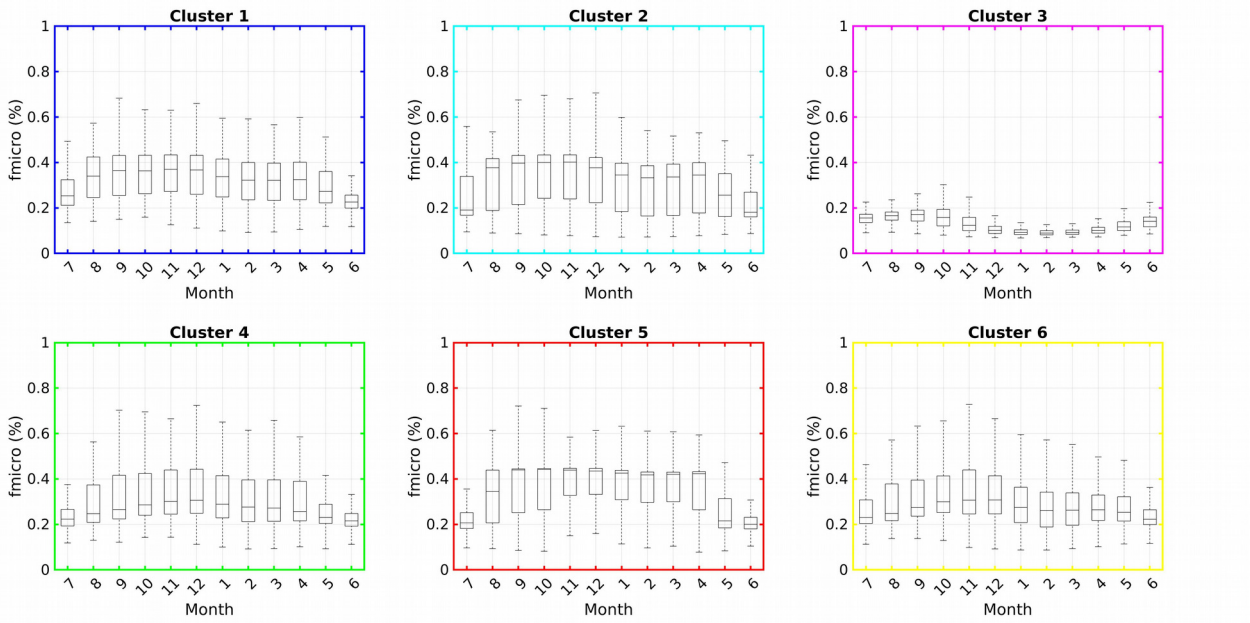


Figure II.6: Fraction of microphytoplankton per bioregion and per month. Cluster number (bioregion) is displayed above each panel and the color of each panel refers to the colors used in Figure II.1. Fraction of microphytoplankton (fmicro) is given in percentage of the total CHL. Months in the abscissa are displayed with annual cycle centered on Austral summer (i.e. month of December, number 12). Each black box represents the distribution of the PSC data for the corresponding month of the year. The horizontal black line inside each box represents the median. The top and bottom limits of each box are the 25th and 75th percentiles, respectively. The vertical dashed black lines extending above and below each box represent the full range of non-outlier observations.

Annual variations in the fraction of the nanophytoplankton class shows relatively low amplitude (fraction of nanophytoplankton generally varied between ~30% and ~45% in all bioregions). In contrast, dynamics of the pico-and microphytoplankton classes are clearly different between bioregions. Bioregion 3 is clearly marked by high proportion of picophytoplankton (> 40%) and a maximum in summer (~60%), while fraction of microphytoplankton in bioregion 3 is low (< 20%) all-year long. Oppositely, fraction of microphytoplankton is above 20% in all other bioregions, with maximum values in bioregion 5 (associated with Highly variable CHL phenology). In all bioregions but bioregion 3, the typical annual cycle seems to describe lower levels of fraction of microphytoplankton in winter and highest levels in summer, while the dynamics of picophytoplankton are opposite. Interestingly, the opposition appears to be especially marked in bioregions 2 and 5.

We demonstrated in the present section that the bioregions defined by CHL phenology and CHL biomass were meaningful regarding phytoplankton community composition and their annual

dynamics. However, we also highlight potential biases of the method regarding the exact rendering of phytoplankton community composition. As such, the low variation in fraction of nanophytoplankton may reflect the difficulties encountered by the PSC model to differentiate the analyzed remote-sensing signal and correctly attribute phytoplankton groups. Another limitation emanates from the fact that in the PSC model used in the present study, the microphytoplankton class is set directly equivalent to the diatom phytoplankton functional type. However the diatom group spans a wide range of cell sizes (IOCCG, 2014). Classifying all diatoms in the microphytoplankton compartment may hence lead to model errors. Additionally, SO high latitude waters are known to have different bio-optical properties (Robinson et al., 2021). The PSC model used in the present study is trained at global scale. Results derived in SO waters should hence be interpreted with caution. To examine the validity of the PSC data used in the present study, we performed a matchup exercise between HPLC samples and PSC model outputs (see supplementary material Figure II.19). In total 676 successful matchups were performed between HPLC data and PSC data (see supplementary material Figure II.20). The HPLC data firstly confirms the north-south gradient in the proportion of pico- and micro phytoplankton size classes (more pico phytoplankton in the north, more micro phytoplankton in the south). The model tends to overestimate phytoplankton concentration in their lower range of values (below ~2.10^{-2} $mg.m^{-3}$). Nonetheless, HPLC data and PSC model data are in good agreement. Finally, it is worth to note that our study highlights the importance of sub-seasonal variations ($FPC1_{STD}$) as a key criteria in the differentiation of CHL phenologies (see Section II.4.b). However, variations in CHL and in environmental conditions (e.g. SST) respond to different time scales of variations (Dunstan et al., 2018). For this reason, “ecological” PFTs (i.e. derived from ocean colour data plus environmental variables, such as SST), might be biased and not fully reliable at small scales (Keerthi et al., 2022). Therefore, the elements listed in the present paragraph validate the model outputs, with some limitations that cannot be overlooked, particularly in the study zone of the present study.

Higher trophic levels: top predator foraging behaviour

The foraging behaviour of a top predator such as the SES was used to examine whether the phenology and composition of the primary producer’s compartment influence higher trophic levels. SESs typically breed at subantarctic latitudes (e.g. Crozet Islands, Kerguelen Archipelago) and forage in the open ocean, between subantarctic latitudes and the Antarctic continent (Guinet et al., 2014). SESs mainly feed on mesopelagic fish (myctophyds, Cherel et al., 2008). Bio-loggers placed on SESs provide useful information to both characterize their behaviour (e.g. localization through GPS, body movement through accelerometry), their abiotic (e.g. temperature, salinity) and biotic environment (e.g. chlorophyll-a fluorescence) (McGovern et al., 2019). We used in the present analysis SES dive statistics (number of dives and dive depth per grid point) as well as movement data (number of PCA) to characterize SES foraging effort and foraging success in the study zone (Section II.2.g). SES foraging behaviour indirectly enables characterization of SES prey distribution and availability. Kruskal-Wallis test was performed on the mean values of (i) number of dives, (ii) dive depth and (iii) number of PCA per grid point to test whether the distribution of these SES data differed between bioregions (Section II.3.b). The three tests resulted significant (p values < 10^{-4}).

Bioregion 3 (subtropical) was almost no represented in the statistical analysis of the section, coherent with the foraging area of the SES mentioned above (see graphical support to geographical distribution of SES profiles per dataset in supplementary material Figures II.21-II.23).

Distribution of the mean number of dives per grid point in each SES trip reveals clear differences between bioregions (Figure II.7). Bioregions 4 and 6 (the most productive bioregions in terms of CHL, see Section II.4.b) were clearly associated with the highest number of dives per grid point, followed by bioregions 2 and 1 (HNLC). The mean number of dives per grid point in bioregion 5 (highly variable CHL phenology) was low. Bioregion 3 (subtropical) was almost no represented in this dataset.

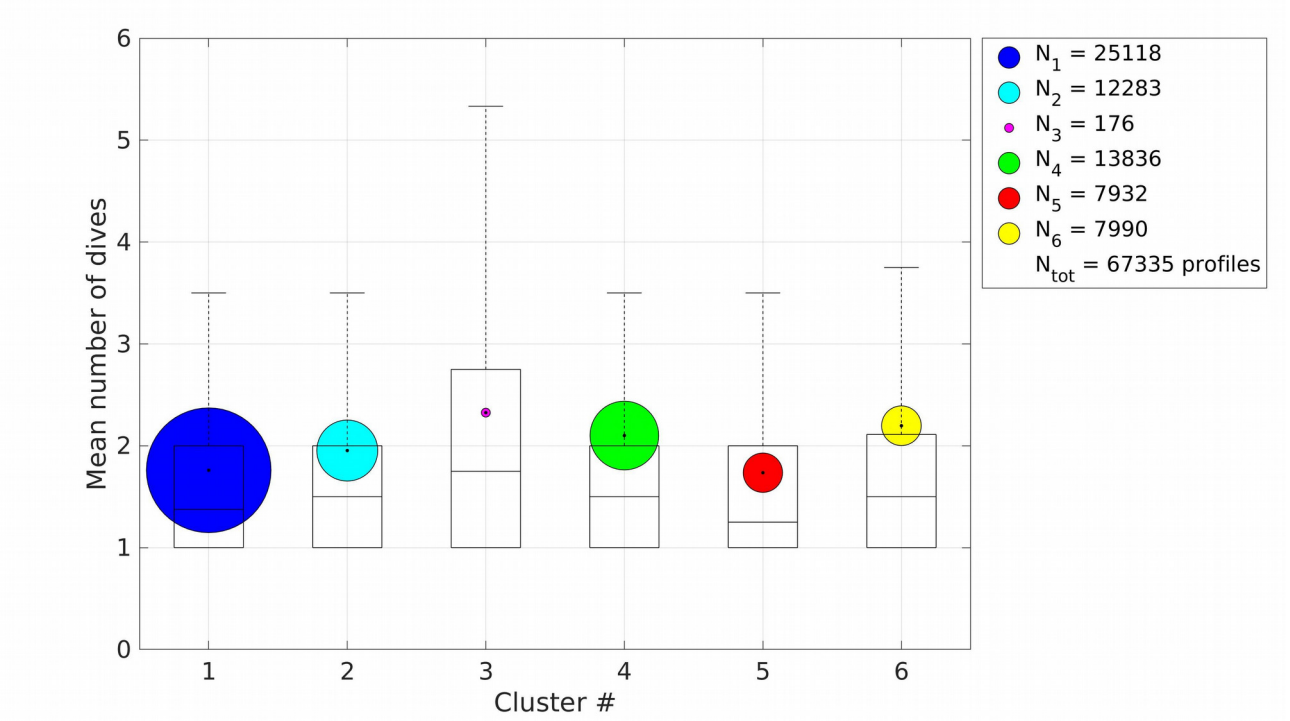


Figure II.7: Number of SES dives per grid point for each bioregion. Each black box represents the distribution of the SES data for the corresponding bioregion. The horizontal black line inside each box represents the median. The top and bottom limits of each box are the 25th and 75th percentiles, respectively. The vertical dashed black lines extending above and below each box represent the full range of non-outlier observations. Black dots in the center of colored circles represent the mean. The color of each circle refers to the colors used in Figure II.1. The size of colored circles is proportional to the number of data points (SES profiles). The number of data points used to compute the distribution of the data in each cluster is indicated in the top-right inset. N_i is the number of data points available in cluster i (i = 1,2,...6). N_{tot} is the total number of SES profiles used to compute number of SES dives per grid point.

Examination of dive depths per cluster suggests better prey accessibility (shallower dives) in bioregion 1 (HNLC) compared to other bioregions (Figure II.8). Distribution of dive depths were spread over a larger interval in clusters 4 and 6, a potential indicator of the larger variety in the environments encountered by SES in these bioregions Oppositely, dive depth distributions in bioregions 2 and 5 were more tightly centered around their median. Excluding bioregion 3 for being probably too fewly represented, bioregion 2 (low CHL productivity) was the bioregion with deepest mean dive depth. Interestingly, bioregions 4 and 6 were associated with mean deeper dive depths than bioregion 1, however the mean number of dives per grid point discussed above suggests that SESs dived with higher frequency in bioregions 4 and 6 than in bioregion 1.

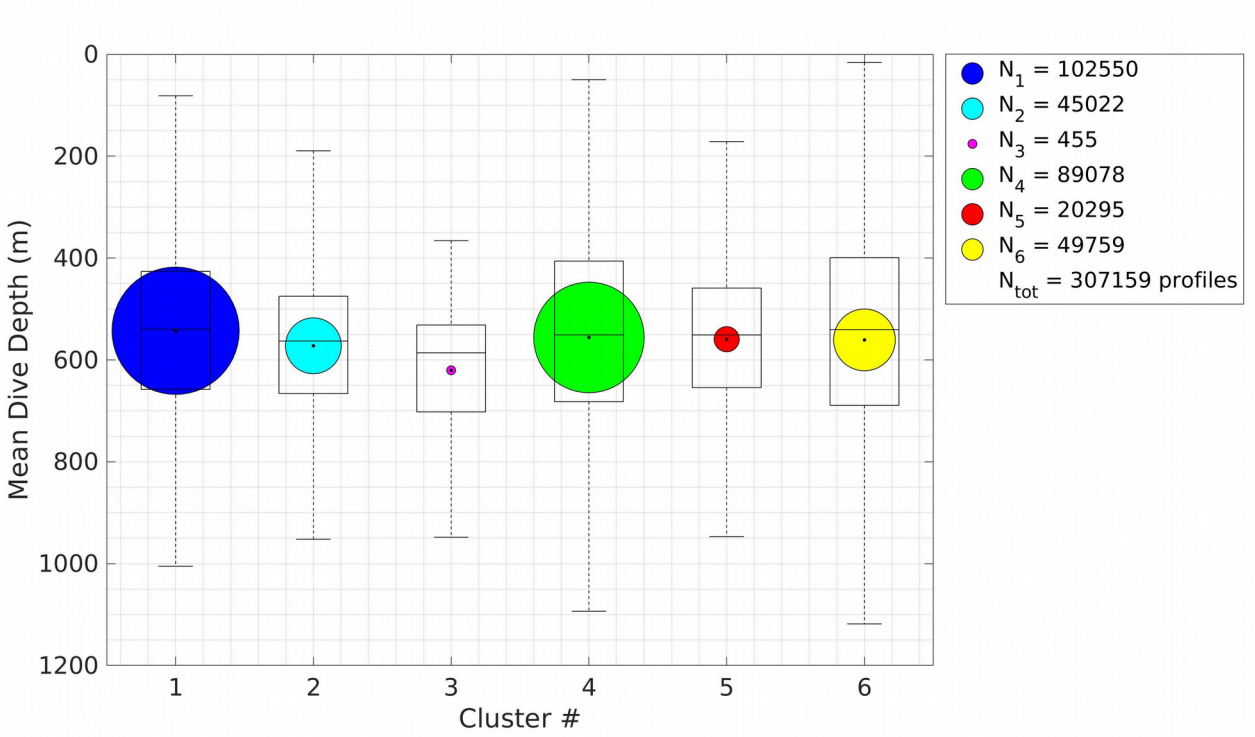


Figure II.8: SES dive depth per grid point for each bioregion. Each black box represents the distribution of the SES data for the corresponding bioregion. The horizontal black line inside each box represents the median. The top and bottom limits of each box are the 25th and 75th percentiles, respectively. The vertical dashed black lines extending above and below each box represent the full range of non-outlier observations. Black dots in the center of colored circles represent the mean. The color of each circle refers to the colors used in Figure II.1. The size of colored circles is proportional to the number of data points (SES profiles). The number of data points used to compute the distribution of the data in each cluster is indicated in the top-right inset. N_i is the number of data points available in cluster i (i = 1,2,...6). N_{tot} is the total number of SES profiles used to compute SES dive depth per grid point.

The number of PCA per bioregion (Figure II.9) supports the opposition between bioregions 4 and 6 on one side and bioregion 1 on the other. In bioregions 4 and 6, deeper dives and higher dive frequency are associated with higher number of PCA. Oppositely, shallower dives and lower dive frequency in bioregion1 are associated with lower number of PCA. The highest mean number of PCA was recorded for bioregion 5, suggesting that areas with highly changing environment - inferred from highly variable CHL phenology - may be favorable to the development of SES's preys. Finally, bioregion 2, previously identified as a bioregion with less frequent and deepeer dives compared to other bioregions, was associated with low number of PCA per dive, and indicator of low foraging success for the SES.

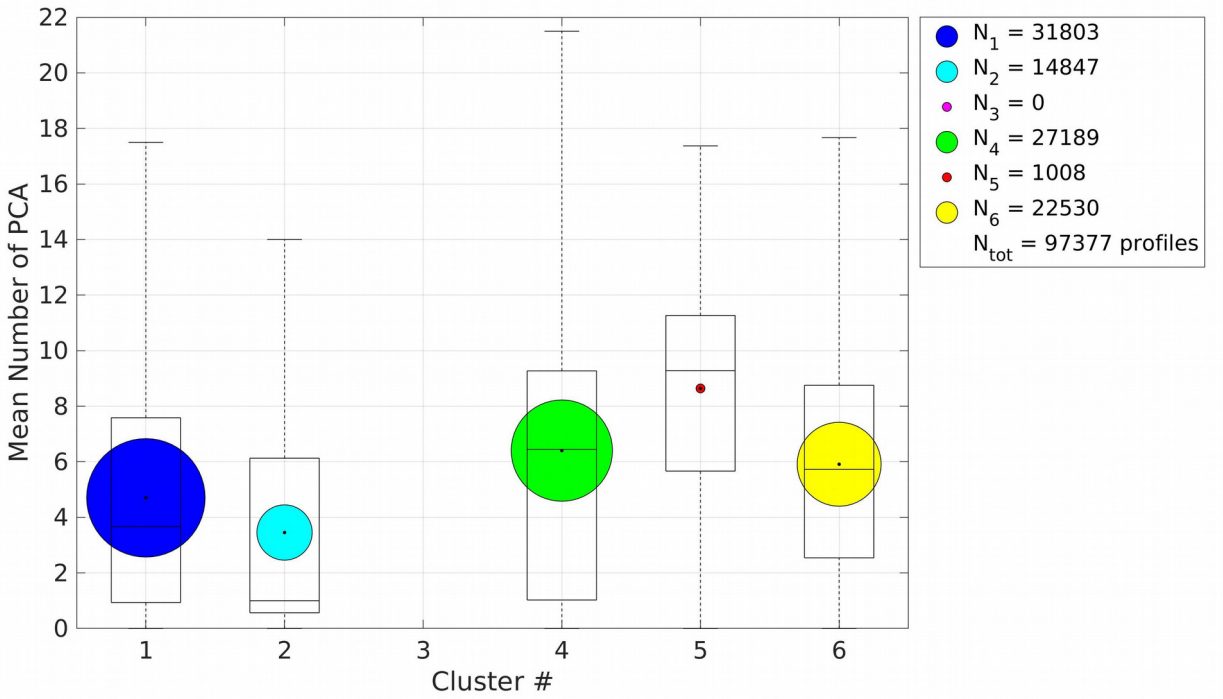


Figure II.9: Number of SES prey capture attempts per grid point for each bioregion. Each black box represents the distribution of the SES data for the corresponding bioregion. The horizontal black line inside each box represents the median. The top and bottom limits of each box are the 25th and 75th percentiles, respectively. The vertical dashed black lines extending above and below each box represent the full range of non-outlier observations. Black dots in the center of colored circles represent the mean. The color of each circle refers to the colors used in Figure II.1. The size of colored circles is proportional to the number of data points (SES profiles). The number of data points used to compute the distribution of the data in each cluster is indicated in the top-right inset. N_i is the number of data points available in cluster i (i = 1,2,...6). N_{tot} is the total number of SES profiles used to compute number of SES prey capture attempts (PCA) per grid point.

It is worth to note that the sampling performed with SES bio-logging data has biases, as mentioned previously (Section II.2.g and supplementary material Figures II.21-II.23), due to the specific scientific questions bio-logging devices are supposed to address. Bio-logging devices deployment spans a large range of diversified objectives, from identification of important foraging areas (Hindell et al., 2016; Mestre et al., 2020), to collection of oceanographic data (Siegelman et al. 2020). For this reason, the sampling is inherently biased because equipped animals are not randomly chosen. However, even if the sample of SES profiles processed in the present study could to some extent be considered as biased, its large size underlines the fact that it contains useful information for ecosystem-derived metrics. Therefore, by combining the information provided by clustering analysis based on CHL data with bio-logging-derived SES foraging behaviour data, we clearly highlighted in the present section bioregions with specific properties in terms of higher-trophic level ecosystem dynamics.

II.5. Conclusion

In the present study, we proposed a bioregionalization of the Indian sector of the SO from a combination of biomass- and phenology-based criteria, all based on the sole use of remote sensing CHL data. Six clusters with distinct characteristics were described, from oligotrophic subtropical open-ocean biomes, to typical subpolar HNLC biomes. Compared to existing global bioregionalizations of the ocean (Longhurst, 2007; Fay & McKinley, 2014), our study brings a closer insight into primary production dynamics and enables zonal differentiation. In addition to CHL biomass, studying CHL phenology proved efficient in revealing large zones in the Indian sector of the SO with distinct seasonal patterns. Also, the addition of a criteria depicting the variability in CHL phenology over time enables finer discrimination of main seasonal patterns driving CHL dynamics in the Indian sector of the SO.

Establishing a link between CHL production, phenology, and phytoplankton community composition is not trivial. As discussed in Chapter I (see Section I.5.c), interpretation of variations in CHL is subject to limitations due to a series of both biological and instrumental factors. However we suggest through our study that distinct phenologies could be associated with distinct phytoplankton community composition (Henson et al., 2017). Investigating the link between CHL phenology and phytoplankton functional types (or phytoplankton size classes) in different bioregions of the SO, as proposed in the present study, is essential for assessing the future impacts of climate change on SO primary production and its implications for ecosystem functioning and carbon export.

Moreover, we suggest in our study that potential changes in terms of primary production could propagate throughout the trophic network towards higher trophic levels. Observing mid-trophic levels in the ocean is a real challenge which requires intense sampling. A good indicator of fish abundance can be derived from a top predator’s foraging success, measurable through bio-logging techniques (Goulet et al., 2020, Tournier et al., 2021). Also, the foraging behaviour of a top predator provides useful information to characterize prey distribution. The examination of SES foraging behaviour in relation to SO bioregions based on CHL characteristics is proposed as a way to study the propagation of primary production dynamics throughout the ecosystem.

Grouping and combining indicators derived from various trophic levels is essential to describe SO ecosystem functioning and define main SO biomes. Furthermore, assessing the sensitivity of SO ecosystems to environmental changes through ecosystem-derived metrics is crucial to establish sustainable resource management strategies in the SO (Chown & Brooks, 2019; Hindell et al., 2020). In the present study, we established clear links between primary production dynamics ecosystem composition. Moreover, in the previous chapter, we discussed the importance of sub-seasonal variations in the control of annual CHL dynamics (see Section I.5.b). As a consequence of the results obtained in Chapter I and Chapter II, intensified sampling of metrics characterizing CHL production in the SO emerges as a crucial necessity to better understand the processes that drive primary production in the SO. This issue is addressed in Chapter III.

Supplementary material for Chapter II

Text SII.1. CHL spring-summer climatology for the Indian sector of the SO

CHL spring-summer climatology (1998-2022) for the Indian sector of the SO was computed for each pixel as the mean value of CHL for the period extending from October to March from 1998 to 2022. CHL spring-summer climatology was derived from GlobColour 8-day averaged Level 3 product and computed with a 0.25º x 0.25º resolution.

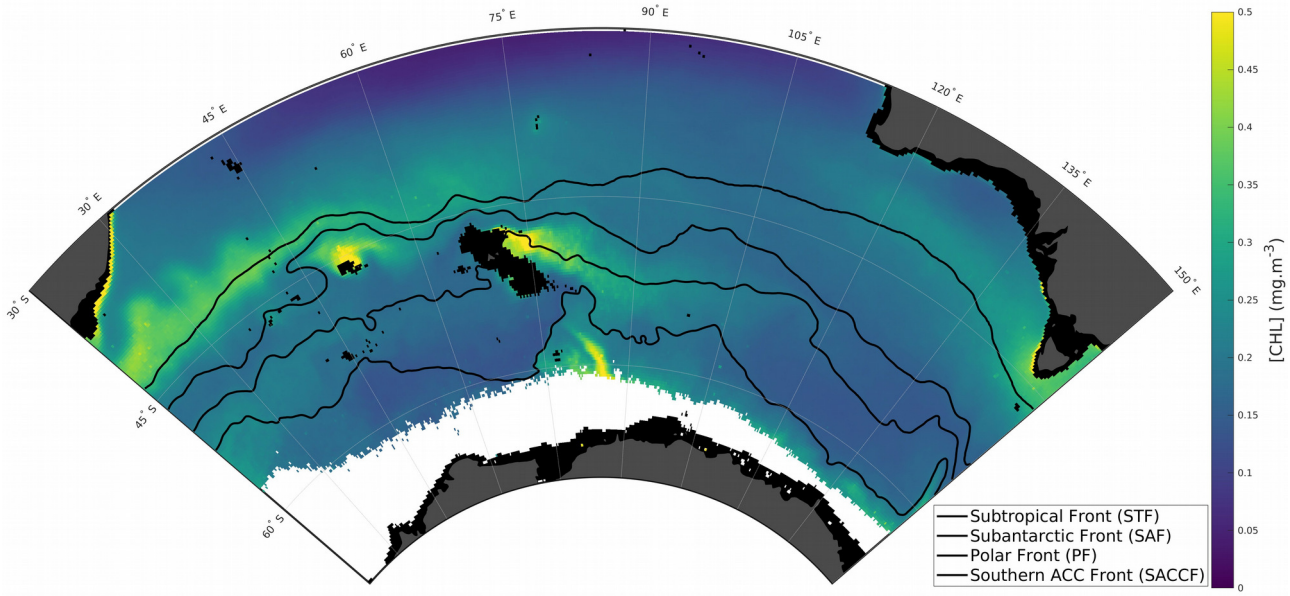


Figure II.10: CHL spring-summer climatology (1998-2022) in the Indian sector of the Southern Ocean derived from GlobColour 8-day averaged Level 3 product with 0.25º x 0.25º resolution. Black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 spring-summer time series (see Section I.2.d) are represented in white.

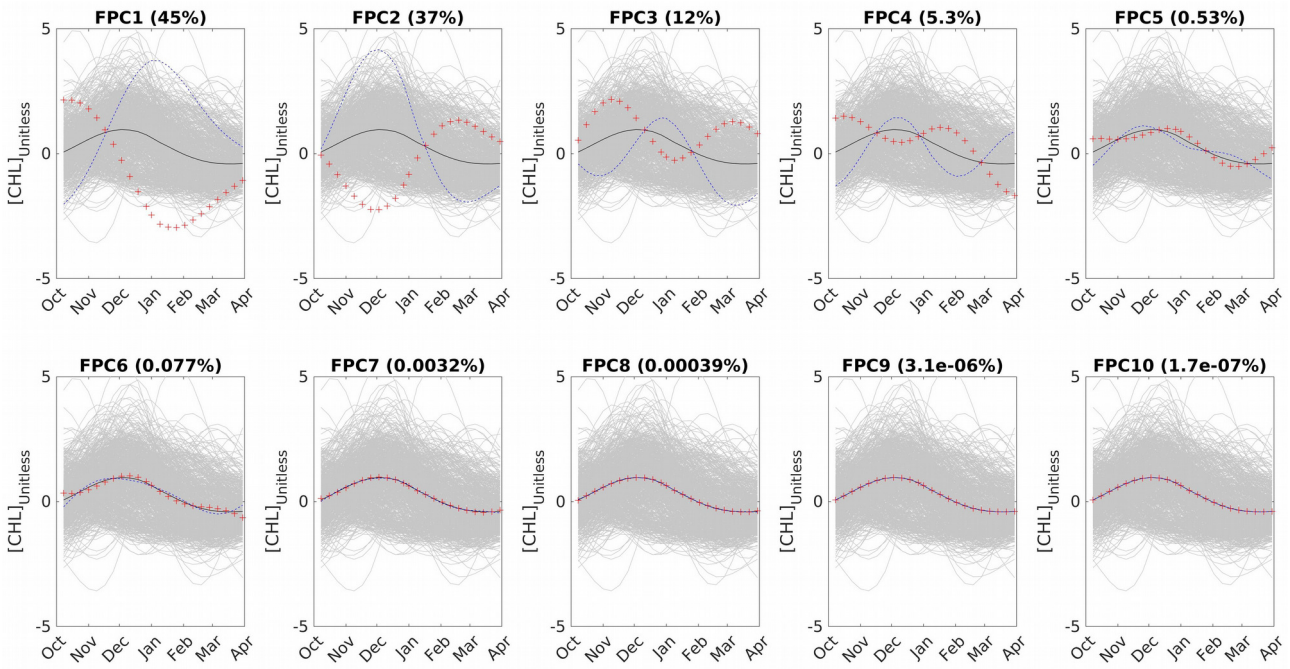


Figure II.11: Ten first FPCs (modes of variation) obtained after running FPCA for CHL spring-summer cycles (see Section II.3.a). The value in brackets above each panel is the percentage of variance explained by the corresponding FPC. FPC1-5 (FPC6-10) were represented in the top (bottom) panels. In each panel, the blue (red) curve depicted by "-" ("+") symbols shows a typical phenology associated with a negative (positive) FPC score. The continuous black line displayed in every panels correspond to the average phenology in the study zone (Indian sector of the Southern Ocean)

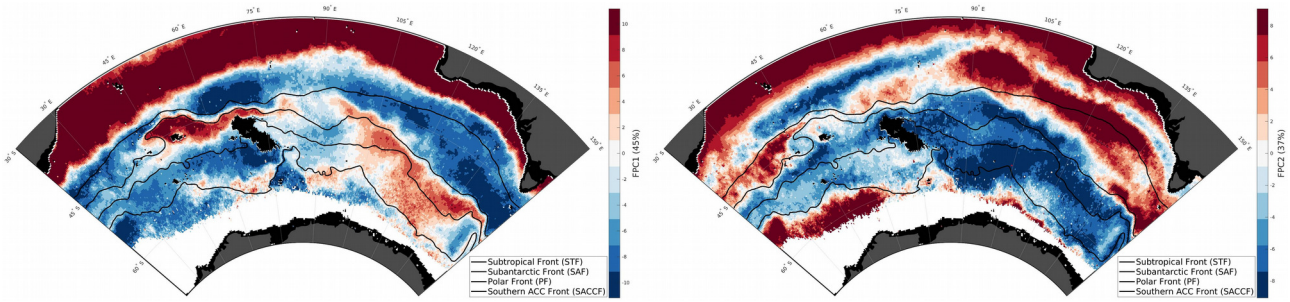


Figure II.12: Spatial distribution of FPC climatological scores. Spatial distributions of FPC1 (left panel) and FPC2 (right panel) climatological scores (see Section II.3.a) were represented with the same color code as in Figure II.11. Black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 spring-summer time series (see Section I.2.d) are represented in white.

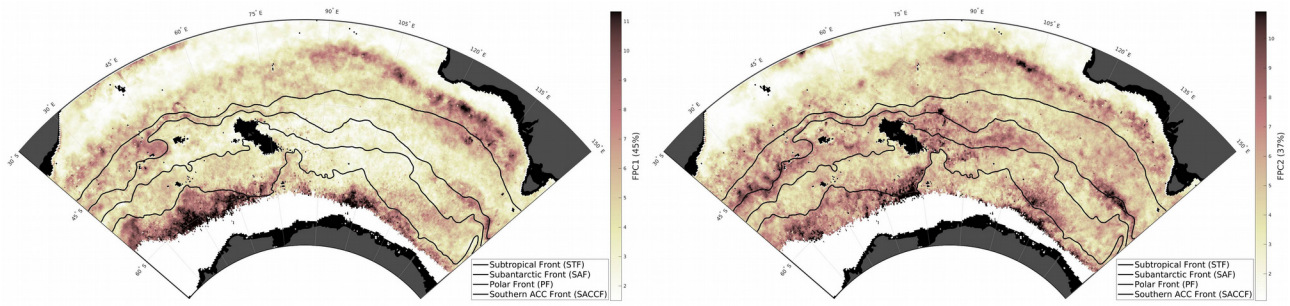


Figure II.13: Spatial distribution of standard deviation of FPC scores. The standard deviation of the time series of FPC scores (see Section II.3.a) were represented for the first ($FPC1_{STD}$, left panel) and the second ($FPC2_{STD}$, right panel) axes of the FPCA. Black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 spring-summer time series (see Section I.2.d) are represented in white.

Text SII.2. Optimum number of clusters

Number of clusters is a user-defined input in the K-means clustering method. The optimum number of clusters was determined using the elbow method. Briefly, the elbow method consists in successively performing clustering with increasing number of clusters. At each step, each point is attributed a cluster. Each cluster has a centroid (cluster mean). At each step, the within-cluster sum of squares (WCSS) is computed as the sum of squares of the distance of each data point from the center point of the cluster it belongs to. Euclidian distance was used to calculate WCSS. A proxy for within-cluster similarity between points, WCSS decreases with increasing number of clusters. The elbow method consists in detecting the point from which no more significant improvement is brought by addition of a new cluster in terms of within-cluster similarity between points. We applied the elbow method for number of clusters varying from one to twelve. For robustness of the results, we performed 30 iterations of the clustering at each step of the method (i.e. for each tested number of clusters, from one to twelve). The optimum number of clusters in the present analysis was fixed at six. Graphical support for the method described in the present paragraph is provided in Figure II.14.

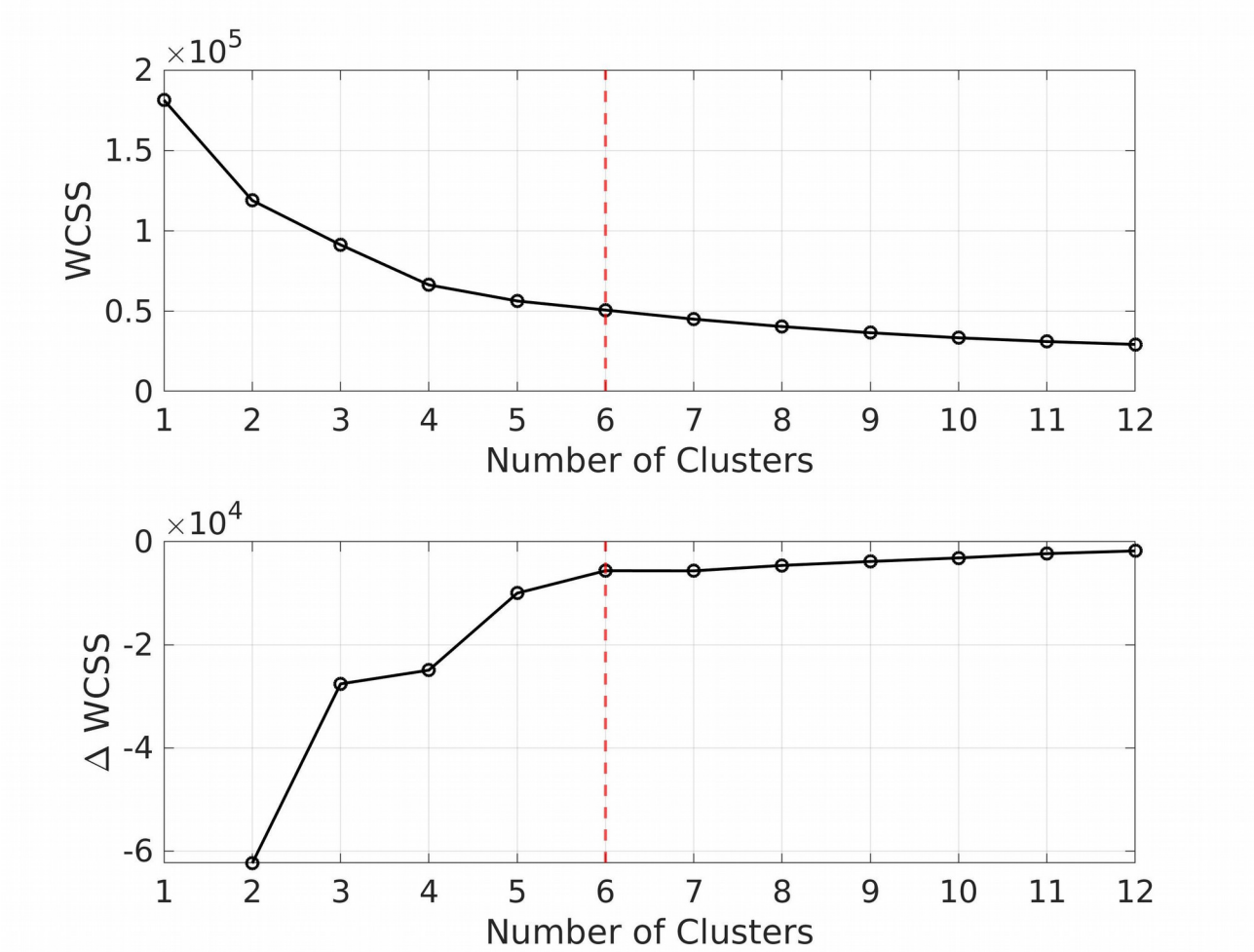


Figure II.14: Determination of optimum number of clusters with elbow method. (top panel) Within-cluster sum of squares (WCSS) as a function of number of clusters and (bottom panel) ΔWCSS represents the difference in WCSS between step n (n clusters) and step n-1 (n-1 clusters). The vertical red dashed line represents the optimum number of clusters (six) fixed for the present study.

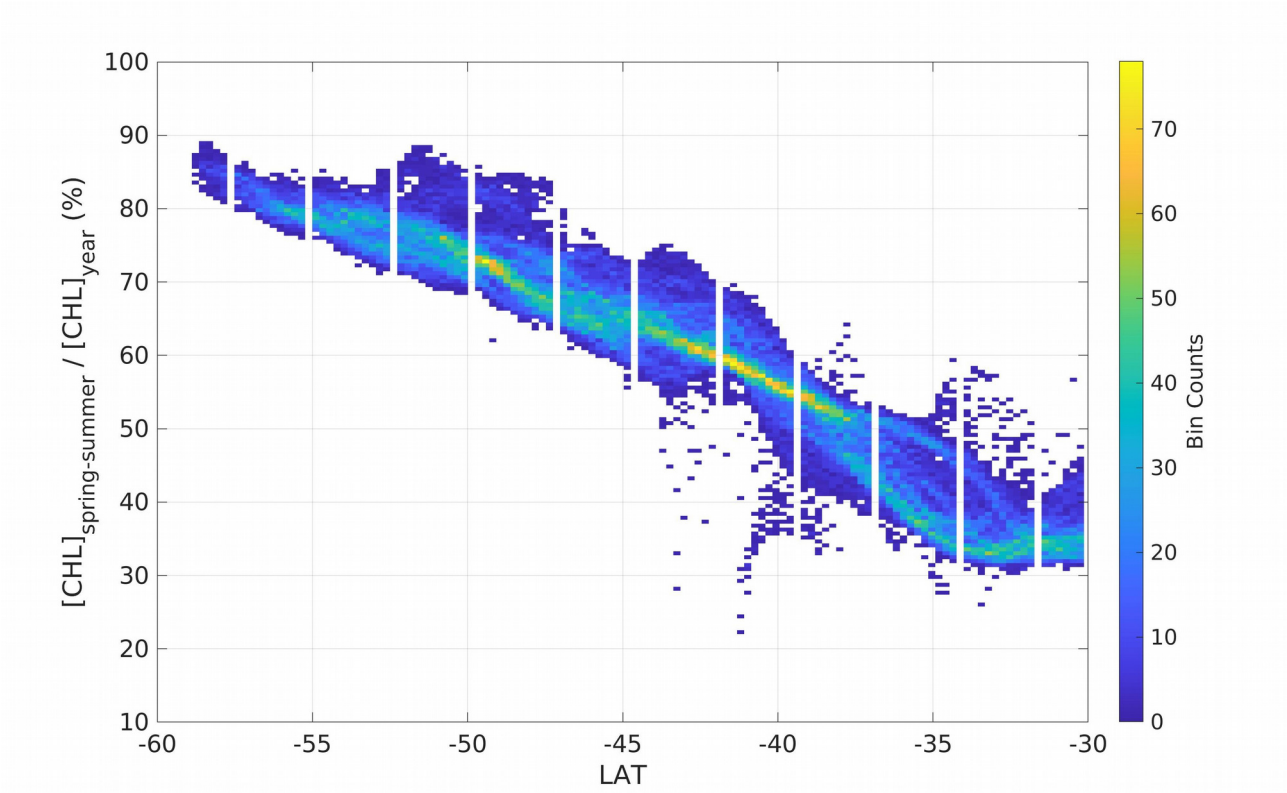


Figure II.15: Proportion of CHL production in the spring-summer period. Mean spring-summer integrated CHL ($[CHL]_{spring-summer}$) and mean annual integrated CHL ($[CHL]_{year}$) were computed for each grid point and represented as a function of latitude. The proportion of CHL production in the spring-summer period is the ratio between $[CHL]_{spring-summer}$ and $[CHL]_{year}$ and was expressed as percentage of $[CHL]_{year}$.

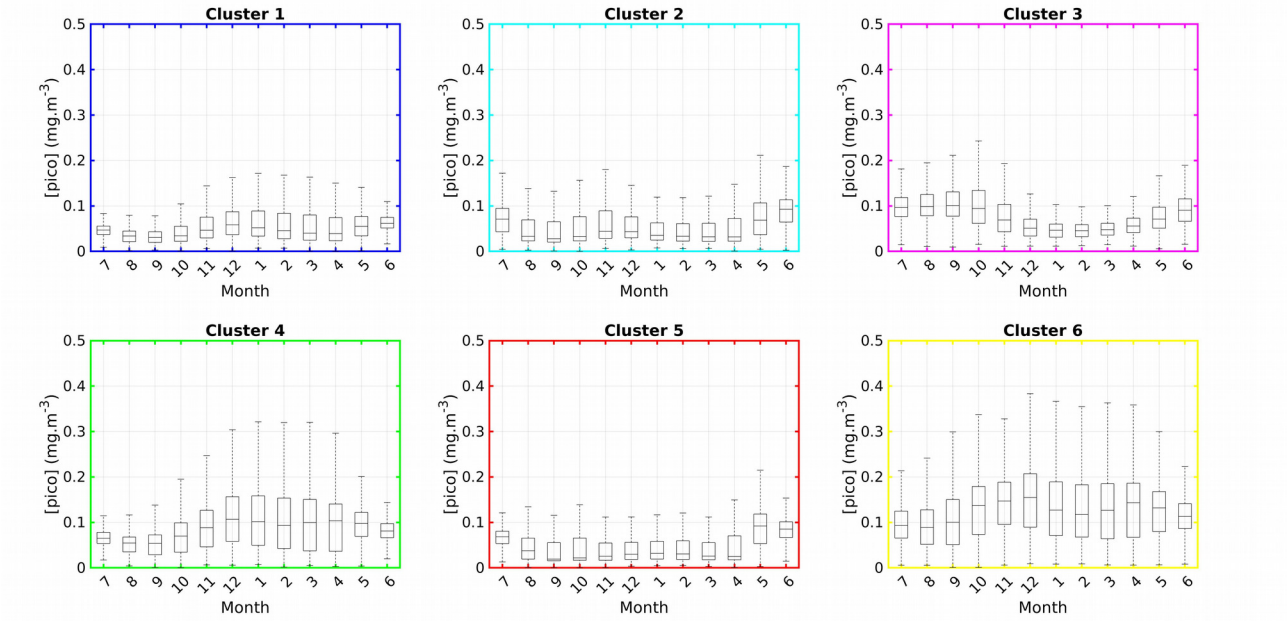


Figure II.16: Concentration of picophytoplankton per bioregion and per month. Cluster number (bioregion) is displayed above each panel and the color of each panel refers to the colors used in Figure II.1. Concentration of picophytoplankton ([pico]) is given in mg.m^{-3}. Months in the abscissa are displayed with annual cycle centered on Austral summer (i.e. month of December, number 12). Each black box represents the distribution of the PSC data for the corresponding month of the year. The horizontal black line inside each box represents the median. The top and bottom limits of each box are the 25th and 75th percentiles, respectively. The vertical dashed black lines extending above and below each box represent the full range of non-outlier observations.

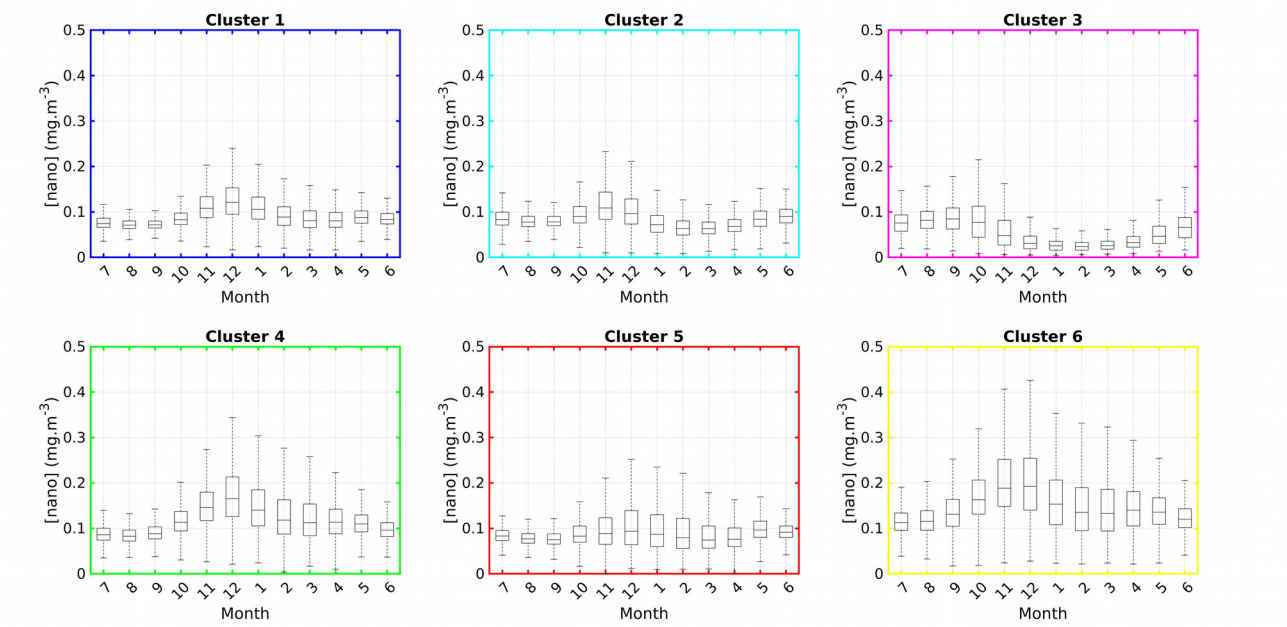


Figure II.17: Concentration of nanophytoplankton per bioregion and per month. Cluster number (bioregion) is displayed above each panel and the color of each panel refers to the colors used in Figure II.1. Concentration of nanophytoplankton ([nano]) is given in mg.m^{-3}. Months in the abscissa are displayed with annual cycle centered on Austral summer (i.e. month of December, number 12). Each black box represents the distribution of the PSC data for the corresponding month of the year. The horizontal black line inside each box represents the median. The top and bottom limits of each box are the 25th and 75th percentiles, respectively. The vertical dashed black lines extending above and below each box represent the full range of non-outlier observations.

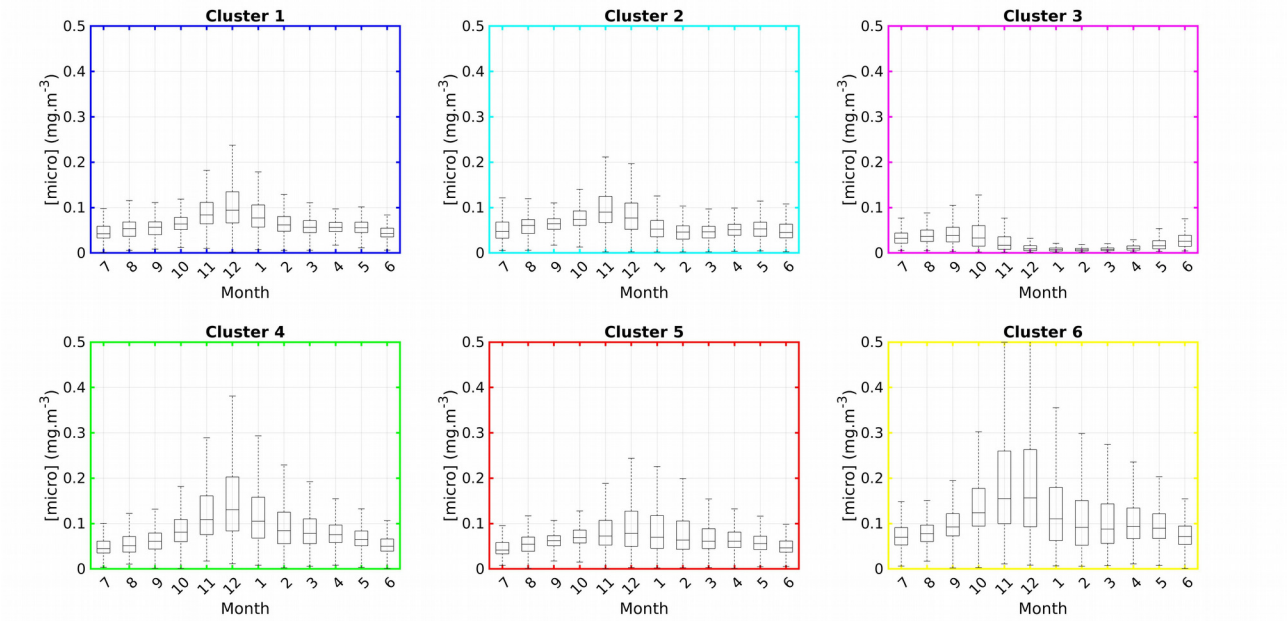


Figure II.18: Concentration of microphytoplankton per bioregion and per month. Cluster number (bioregion) is displayed above each panel and the color of each panel refers to the colors used in Figure II.1. Concentration of microphytoplankton ([micro]) is given in mg.m^{-3}. Months in the abscissa are displayed with annual cycle centered on Austral summer (i.e. month of December, number 12). Each black box represents the distribution of the PSC data for the corresponding month of the year. The horizontal black line inside each box represents the median. The top and bottom limits of each box are the 25th and 75th percentiles, respectively. The vertical dashed black lines extending above and below each box represent the full range of non-outlier observations.

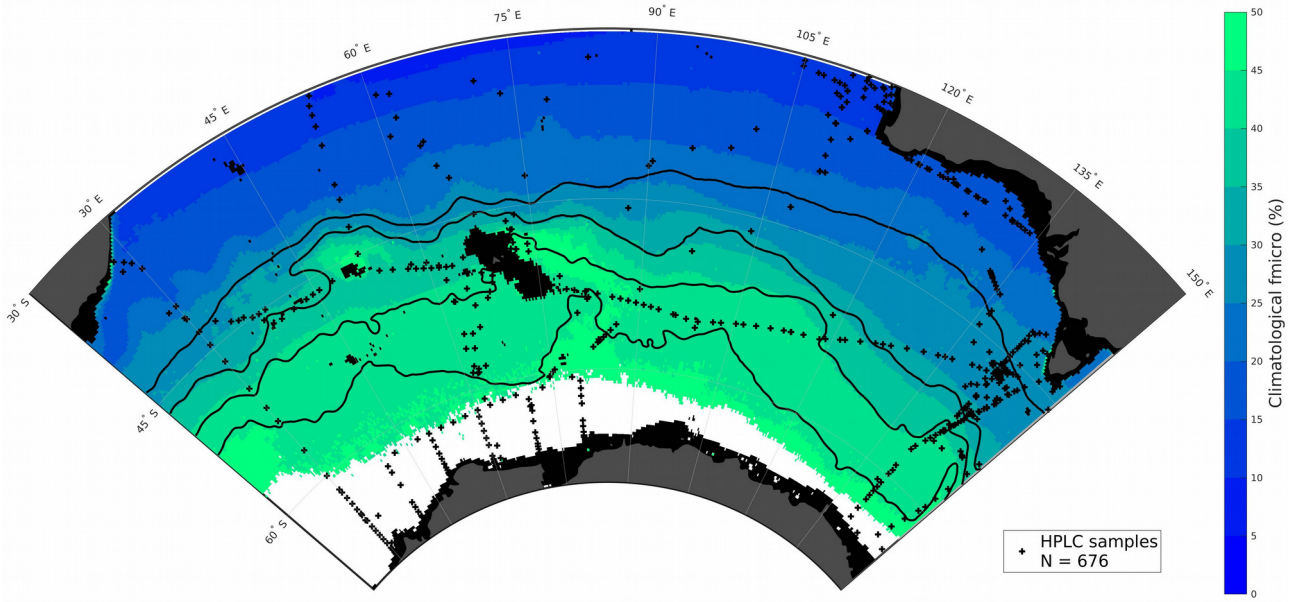


Figure II.19: Spatial distribution of HPLC samples. Pigment data obtained from HPLC was used to validate PSC products (Section II.2.f). In total, N = 676 HPLC samples were included in the validation exercise. The background map is the climatological fraction of microphytoplankton (fmicro) in the study zone, computed as the mean value of fmicro from 1998 to 2022. Black lines represent the major fronts of the SO defined in Section I.2.c (from north to south): Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF) and Southern ACC Front (SACCF). The areas filled with gray color are land areas. The areas filled with black color depict marine areas with bathymetry shallower than -1000 meters. Pixels with less than 50% valid CHL values in the 1998-2022 spring-summer time series (see Section I.2.d) are represented in white.

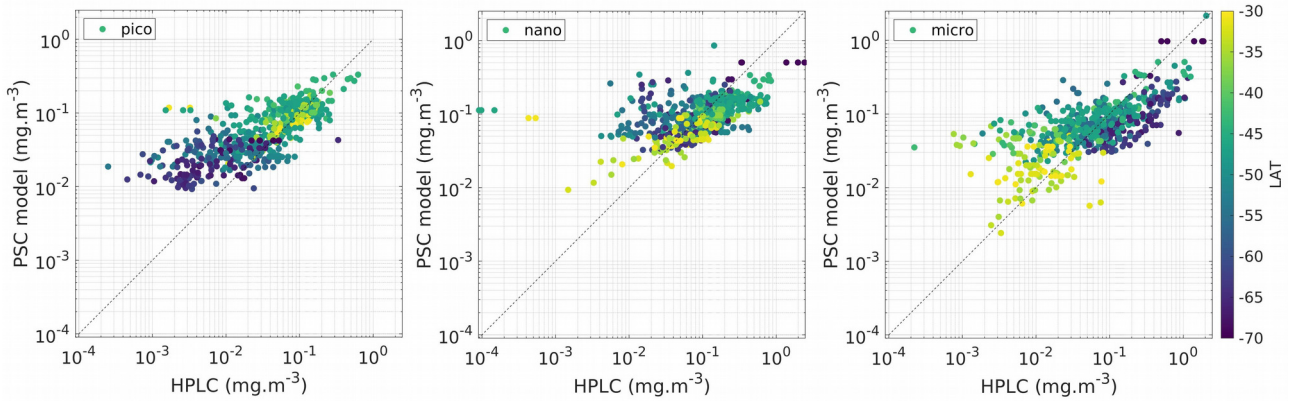


Figure II.20: Validation of PSC model with HPLC data. Comparison of concentrations in pico- (left panel), nano- (middle panel) and microphytoplankton (right panel) from HPLC samples with concomitant model estimate. Abscissa are HPLC data (see Section II.2.f) and ordinates are PSC model data (see Section II.2.e). The color of the points represent the latitude of the samples , from high latitudes (dark blue) to low latitudes (yellow).The dashed black line represents the 1:1 line. Note that the data is represented in log scale.

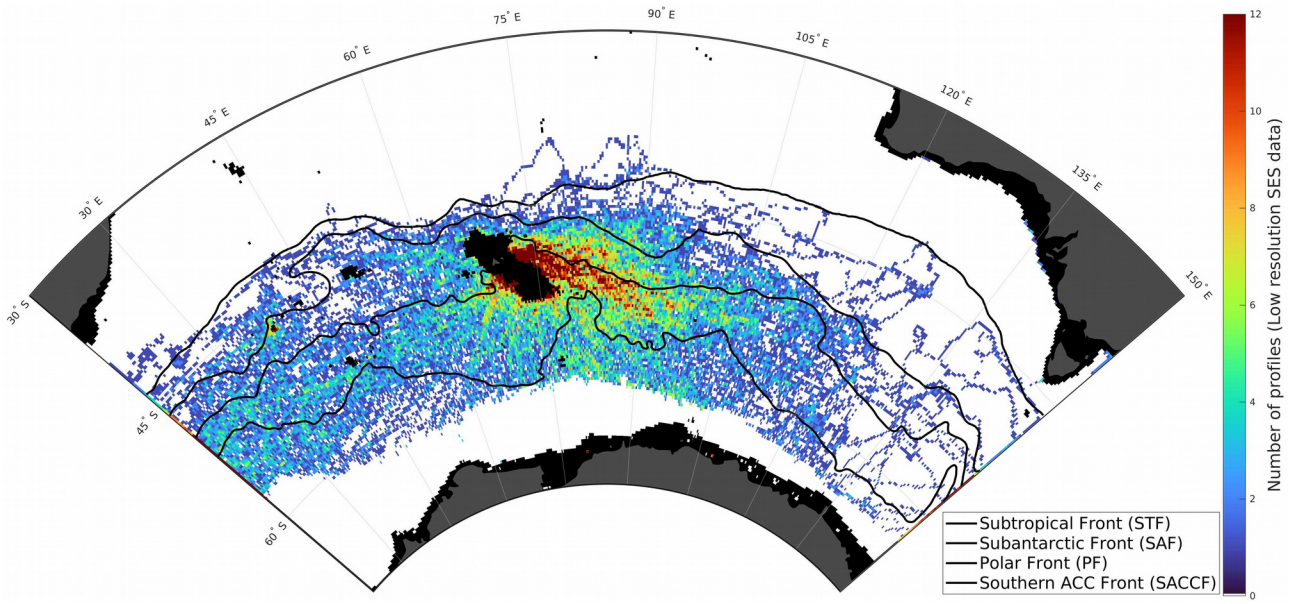


Figure II.21: Geographical distribution of SES profiles used to compute number of SES dives per grid point. In total, 67 335 SES profiles were processed to compute number of SES dives per grid point.

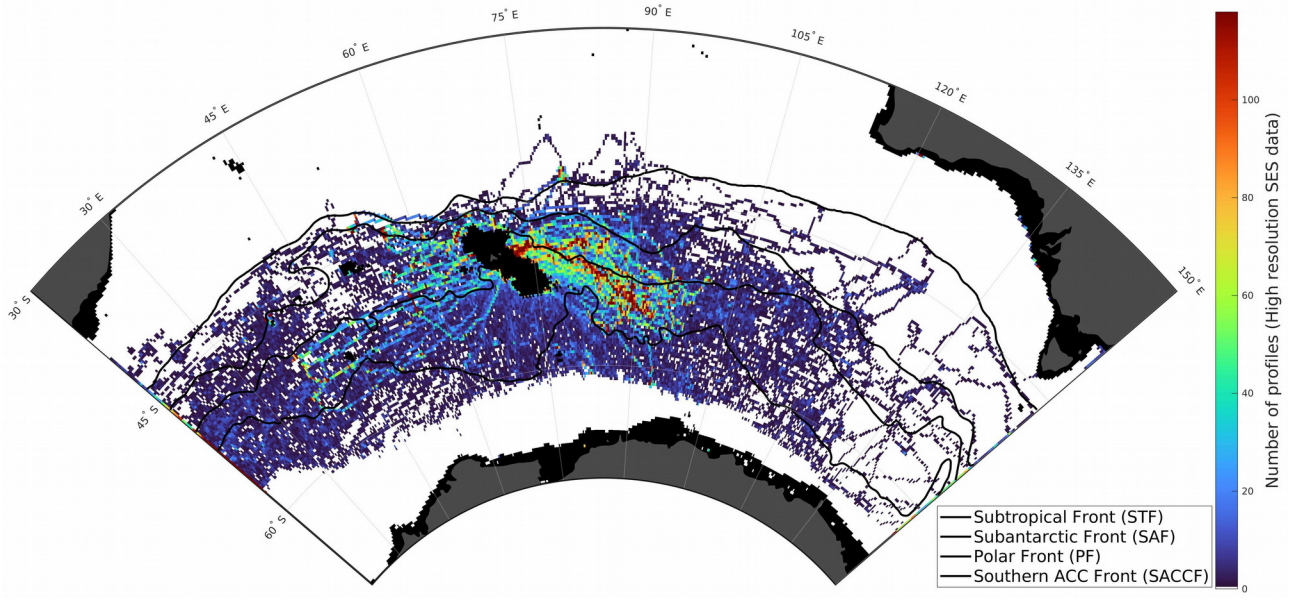


Figure II.22: Geographical distribution of SES profiles used to compute SES dive depth per grid point. In total, 307 159 SES profiles were processed to compute SES dive depth per grid point.

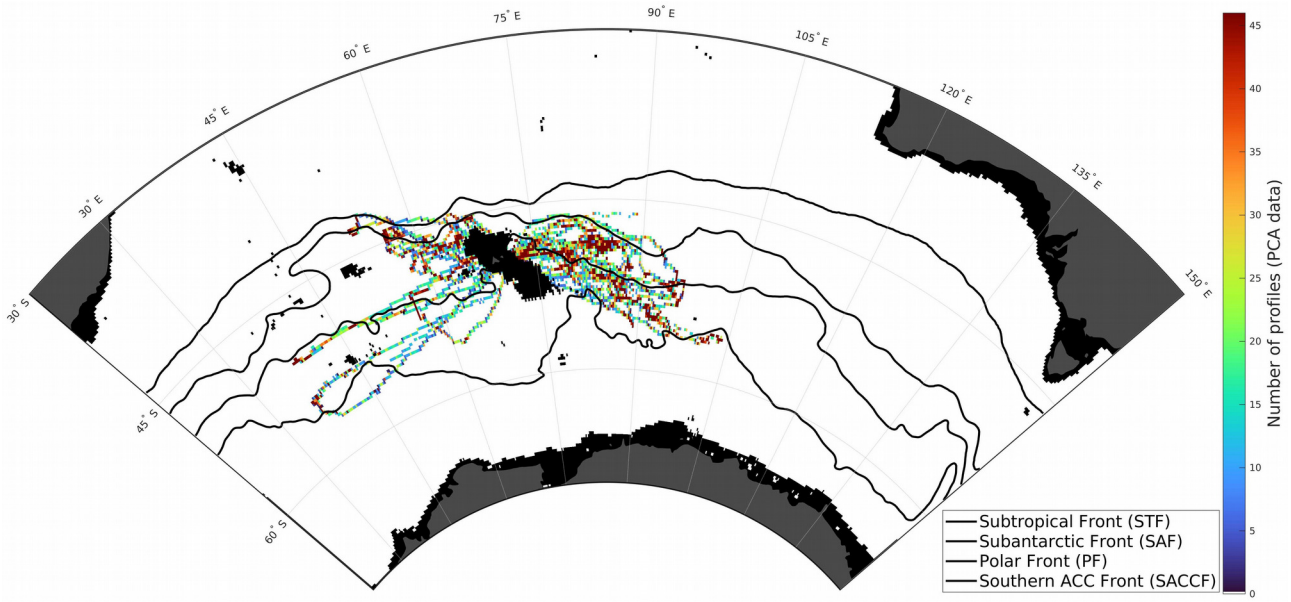
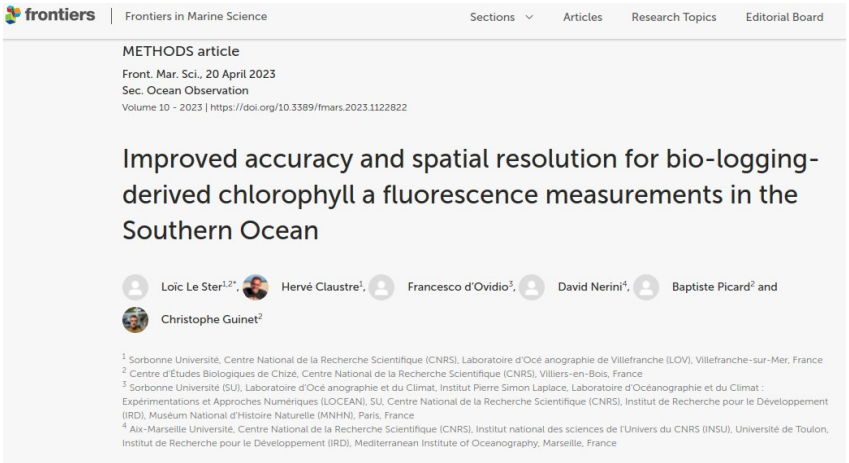


Figure II.23: Geographical distribution of SES profiles used to compute number of SES PCA per grid point. In total, 97 377 SES profiles were processed to compute number of SES PCA per grid point.

Chapter III

Bio-logging data as a tool for the studying of (sub)mesoscale bio-physical interactions in the open ocean

Published in Frontiers in Marine Science in April 2023



https://doi.org/10.3389/fmars.2023.1122822

Contents of Chapter III

Improved accuracy and spatial resolution for bio-logging-derived chlorophyll a fluorescence measurements in the Southern Ocean

Loïc Le Ster[1,2 *], Hervé Claustre[1], Francesco d'Ovidio[3], David Nerini[4], Baptiste Picard[2], Christophe Guinet[2]

[1] Sorbonne Université, CNRS, Laboratoire d'Océanographie de Villefranche, LOV, 06230 Villefranche-sur-Mer, France
[2] Centre d'Études Biologiques de Chizé, CNRS, 79360 Villiers-en-Bois, France
[3] Sorbonne Université, Laboratoire d'Océanographie et du Climat, Institut Pierre Simon Laplace (LOCEAN, SU/CNRS/IRD/MNHN), 75252 Paris Cedex 05, France
[4] Aix-Marseille Université, CNRS/INSU, Université de Toulon, IRD, Mediterranean Institute of Oceanography, UM 110, 13288 Marseille, France

*** Correspondence:**
Loïc Le Ster
loic.lester@imev-mer.fr

Keywords: Chla fluorescence, Southern Ocean, Sensor calibration, Submesoscale, Southern elephant seal, bio-logging tag

Abstract

The ocean's meso- and submeso-scales (1-100 km, days to weeks) host features like filaments and eddies that have a key structuring effect on phytoplankton distribution, but that due to their ephemeral nature, are challenging to observe. This problem is exacerbated in regions with heavy cloud coverage and/or difficult access like the Southern Ocean, where observations of phytoplankton distribution by satellite are sparse, manned campaigns costly, and automated devices limited by power consumption. Here, we address this issue by considering high-resolution *in-situ* data from 18 bio-logging devices deployed on southern elephant seals (*Mirounga leonina)* in the Kerguelen Islands between 2018 and 2020. These devices have submesoscale-resolving capabilities of light profiles due to the high spatio-temporal frequency of the animals' dives (on average 1.1 +- 0.6 km between consecutive dives, up to 60 dives per day), but observations of fluorescence are much coarser due to power constraints. Furthermore, the chlorophyll a concentrations derived from the (uncalibrated) bio-logging devices' fluorescence sensors lack a common benchmark to properly qualify the data and allow comparisons of observations. By proposing a method based on functional data analysis, we show that a reliable predictor of chlorophyll a concentration can be constructed from light profiles (14 686 in our study). The combined use of light profiles and matchups with satellite ocean-color data enable effective (1) homogenization then calibration of the bio-logging devices' fluorescence data and (2) filling of the spatial gaps in coarse-grained fluorescence sampling. The developed method improves the spatial resolution of the chlorophyll a field description from ~30 km to ~12 km. These results open the way to empirical study of the coupling between physical forcing and biological response at submesoscale in the Southern Ocean, especially useful in the context of upcoming high-resolution ocean-circulation satellite missions.

III.1. Introduction

Primary producers are key elements in the structuring of marine food webs and their distribution in the ocean largely drives ecosystem dynamics (Lévy et al., 2018; Henley et al., 2020). Primary production also plays a critical role in biogeochemical cycles given its involvement in CO_2 sequestration through the process of the biological carbon pump (DeVries et al., 2012; Siegel 2014; Boyd et al., 2019). Resolving phytoplankton distribution in the ocean is however a challenging issue due to the extreme heterogeneity of environmental conditions and the short time scales of events relative to phytoplankton growth. From the mesoscale (O(100 km)) to the submesoscale (O(10 km)), the oceanic landscape is shaped by dynamic processes such as filaments or eddies which directly impact phytoplankton distribution (Mahadevan, 2016; Lévy et al., 2018). Complex shapes and high patchiness, observable from space with ocean-color radiometry, result from these processes (d'Ovidio et al., 2010; Lehahn et al., 2018), with consequences on the variability of the associated biogeochemical processes (Resplandy et al., 2009).

Remote sensing of ocean-color enables monitoring of the distribution of chlorophyll a (Chla hereafter), a proxy for phytoplankton biomass, with the advantage of providing a synoptic view of the processes occurring at the surface of the ocean. Yet the reflectance signal upcoming from the ocean surface is subject to obstruction by clouds or masking by sea ice (at high latitudes), which requires coupling satellite data with *in-situ* sampling. Furthermore, the critical need for collecting *in-situ* data is reinforced by the fact that the vertical distribution of Chla escapes remote detection. Indeed, ocean-color measurements are restricted to the near-surface. Satellite observations consequently only include part of the productive layer and omit potential subsurface features (e.g. deep chlorophyll maxima, see Baldry et al., 2020; Cornec et al., 2021).

While the mesoscale is quite well covered by current satellite observations of physical dynamics coupled with *in-situ* platforms sampling biogeochemical variables (McGillicuddy, 2016), recent missions like the Surface Water and Ocean Topography (SWOT) mission enable access to spatial scales down to 15-30 km (Morrow et al., 2019) but there is no *in-situ* counterpart to support the remotely-sensed observations (d'Ovidio et al., 2019). Phytoplankton distribution at submesoscale is hence inadequately resolved due to the gap between satellite observations and *in-situ* data.

One region where an enhanced submesoscale observation of phytoplankton distribution would be particularly valuable is the Southern Ocean (SO). Considered as a main contributor to global air-sea CO_2 exchange (Ardyna et al., 2017; Devries et al., 2019; Bushinsky et al., 2019), the SO hosts a large variety of ecosystems, from unicellular organisms up to charismatic megafauna, that rely greatly on ocean biogeochemistry (Deppeler and Davidson, 2017; Henley et al., 2020). In addition to displaying marked seasonal and regional features (Blain et al., 2008; Deppeler and Davidson, 2017), the spatio-temporal variability of phytoplankton concentration in the SO is subject to the heavy structuring effect of the (sub)mesoscale (Bachman et al., 2017) and is strongly influenced by sub-seasonal forcings (Prend et al., 2022). Monitoring the distribution of phytoplankton at such short spatial and temporal scales is therefore crucial. However the monitoring

of primary production in the SO through *in-situ* sampling by research vessels is highly limited by harsh meteorological conditions and by the presence of sea ice. In addition, satellite-based observations in the SO are frequently restricted by cloud coverage. As a result, despite the preeminent position of the SO in the Earth's climate system and ecosystem functioning, it remains undersampled compared to other ocean basins.

The limitations associated with research vessel-based sampling in the SO lead to opting for autonomous measuring platforms like AUVs (Autonomous Underwater Vehicles). However, both the large extent and the remoteness of the zone highly constrain any AUV deployment and recovery. Nonetheless, large efforts have been made in the past two decades to increase the number of autonomous platforms monitoring the SO through the measurement of biogeochemical variables (Chai et al., 2020). While Biogeochemical-Argo (BGC-Argo) floats enable the sampling of a region over several years (Claustre et al., 2020), gliders (Testor et al., 2019) and marine mammals equipped with bio-logging devices (Blain et al., 2013; Guinet et al., 2013; Treasure et al., 2017) are more suitable for the observation of short-lived (sub)mesoscale processes. Gliders are indisputably a powerful tool for characterizing phytoplankton distribution at these scales due to the high spatio-temporal density they can achieve in the sampling (0.5–6 km, 0.5–6 h between 2 vertical profiles, Rudnick et al., 2016a; Testor et al., 2019). However, despite some examples of successful glider deployments in the SO providing an insight into phytoplankton distribution at high resolution (e.g. Alex Kahl et al., 2010), high-frequency data remain rare in the SO because of the deployment constraints mentioned above. By comparison, bio-logging devices mounted on deep-diving animals such as southern elephant seals (*Mirounga leonina*, SES hereafter) offer the possibility of acquiring as many as 60 profiles per day at depths regularly exceeding 500 m (Siegelman et al., 2019). Bio-logging devices hence have the potential to address the (sub)mesoscale sampling issue in zones as remote and turbulent as the SO.

The Satellite Relayed Data Logger (SRDL, see Boehme et al., 2009) developed by the Sea Mammal Research Unit (SMRU, UK) is a bio-logging device designed for marine mammals like the SES. SRDLs commonly include a Conductivity, Temperature and Depth (CTD) sensor head. Optionally, SRDLs may include a light sensor, and a fluorometer to measure Chla fluorescence (Fluo hereafter). SRDLs can also act as high-frequency sampling loggers which need to be recovered when the SESs are back ashore in order to obtain access to the data. The present study focuses on SRDLs (referred to as "tags" hereafter) measuring light (L hereafter) and Fluo. Although High Pressure Liquid Chromatography (HPLC) is the reference technique for accurate Chla concentration estimates ([Chla], $mg.m^{-3}$) (Wright et al., 1991; Ras et al., 2008), it requires the collection of water samples whereas fluorometers provide a real-time estimate of *in-situ* [Chla]. Due to their ease of use and the relative simplicity of their integration, fluorescence-based sensors have recently been largely implemented in autonomous platforms. As a consequence, Fluo has become a universal standard variable for the estimate of [Chla]. The measurement of Fluo is based on the optical properties of the Chla photosynthetic pigments present in the sampled water volume (Lorenzen, 1966; Huot and Babin, 2010; Roesler and Barnard, 2013). Fluo is at first order proportional to [Chla]. However, more precise examinations of the Fluo signal reveal that the relationship between the observed fluorescence and the actual phytoplankton biomass can vary

according to phytoplankton community composition, physiological factors or light conditions (Serôdio and Lavaud, 2011; Xing et al., 2012; Roesler et al., 2017; Schallenberg, et al., 2022). Fluo is hence an imperfect proxy which does not straightforwardly reflect phytoplankton concentration, but remains to date the best means to obtain widespread estimates of *in-situ* [Chla].

Previous studies have already shown that Fluo data quality can be enhanced by the use of concomitant radiometric measurements, and vice versa (Morel and Maritorena, 2001; Morel et al., 2007; Xing et al., 2011). These methods rely on the hypothesis that light absorption in the water column is mainly due to the presence of phytoplankton. Such a hypothesis is commonly made for oceanic waters with no direct terrestrial influence, classified as “case 1” waters (Morel and Prieur, 1977; Morel, 1988). Based on the same hypothesis but more specifically for inference purposes in the framework of functional data analysis, it has been proved that the vertical diffuse attenuation coefficient for L (K_L hereafter) can be a good predictor for Fluo (Bayle et al., 2015). In the present study we propose to exploit the predictive capabilities of a linear functional model (LFM) similar to the one described by Bayle et al. (2015) (who limited their analysis to the inference of low resolution Fluo data) to adjust the (uncalibrated) Fluo data provided by multiple tags (18 in our study). The tags’ intercalibration does not resolve the issue of the absolute Fluo to [Chla] conversion. Consequently, following the merging of all the intercalibrated tags, we have selected [Chla] estimated by ocean-color radiometry as the benchmark to carry out absolute Fluo calibration.

Another key issue regarding Fluo measurements in the context of bio-logging is related to energy consumption. Not only is energy consumption a major concern for autonomous platforms in general but Fluo measurements are also particularly energy-demanding as they rely on an active optical sensor. The issue is especially critical for bio-logging tags due to the reduced size of such loggers and therefore the highly limited volume of their batteries. A trade-off between the vertical and temporal resolutions of the acquisitions is necessary to best optimize battery lifetime. As a consequence, despite the high sampling resolution of the tags enabled by SES diving behaviour (up to ~60 dives per day), suitable for the observation of submesoscale processes, the spatial resolution of the SRDL’s Fluo measurements was reduced (~4 profiles per day) and becomes insufficient for a proper description of phytopankton distribution at that scale. To address this observation gap, a LFM was designed to infer [Chla] from K_L and increase the resolution of the [Chla] field description towards the submesoscale.

To summarize, the method developed in the present study aims at enhancing the quality of the [Chla] estimates provided by a set of multiple SRDLs in terms of accuracy and horizontal resolution through the use of K_L derived from vertical light profiles, combined with satellite estimates of [Chla]. The objective of the present study is to propose and to validate a method based on bio-logging SRDL data to retrieve (1) a calibrated measurement of *in-situ* [Chla] (2) at submesoscale (O(10 km)). The main steps of the method described in the present study are summarized in Figure III.1. An application of the method is presented with the data from two tags deployed in the Kerguelen Islands region.

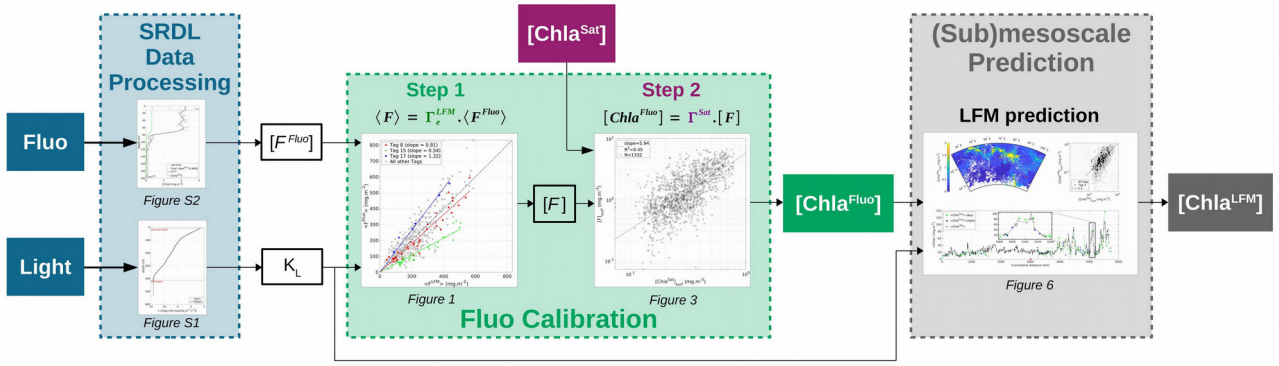


Figure III.1: Flowchart summarizing the main steps of the method described in the present study. The method aims at enhancing the quality of the in-situ [Chla] estimates provided by a set of multiple bio-logging devices (SRDLs) in terms of accuracy and horizontal resolution. Based on SRDL data (Fluo *and L in-situ measurements, blue boxes) combined with concomitant satellite estimates of [Chla] ([Chla^{Sat}], purple box), the method firstly enables the constitution of calibrated datasets of in-situ [Chla] estimates ([Chla^{Fluo}], green box). The method then extends the description of the [Chla] field to (sub)mesoscale ([Chla^{LFM}], gray box) through the use of K_L as a predictor for [Chla].*

III.2. Material and Methods

III.2.a. Tag data

The present analysis is based on the data from 18 tags deployed on female SESs in the Kerguelen Islands region. The study area is located in the Indian sector of the SO and extends from 43ºS to 62ºS and from 35ºE to 101ºE. The 18 tags were deployed during the SESs' post-breeding foraging trip, which occurs from October of year *N* to January of year *N* + 1. In the present study, post-breeding deployments from years 2018 to 2020 were analyzed, totaling 89 197 vertical profiles (for detailed metadata per tag, see supplementary material, Table III.2).

The tags were glued on the fur of the SESs' head using a two component industrial epoxy (see McMahon et al. 2008; Boehme et al. 2009 for animal capture and tag attachment details). After their post-breeding foraging trip, the female SESs were located, recaptured and the tags were retrieved. The tags measure and record pressure (dbar), temperature (ºC), salinity (dimensionless), L (µmol quanta.m^{-2}.s^{-1}) and Fluo (mg.m^{-3}) at 0.5 Hz (note that the different variables used in this study and their associated symbols, definitions and units are detailed in Table 1). The tags' sensor data were continuously sampled during the SESs' trip, with the exception of the Fluo sensor, which was intermittently switched off to save battery power (see details in Section 2.2.3). The archived time series were processed to produce only one vertical profile per dive, corresponding to the ascent phase of the dive, starting from the deepest part of the dive (down to ~1000 m depth), up to the surface. The SRDL data were interpolated at 1 m resolution. The vertical resolution of 1 m is consistent with the sampling rate of the tags (0.5 Hz) and the vertical speed of the SES during the ascent (~1.5 m.s^{-1}, see Richard et al., 2014; McGovern et al., 2019).

For each surfacing phase (i.e. each time the animal emerges to breathe), when available, the location of the animal was recorded, using by default the Argos satellite system, operated by Collecte Localisation Satellites (CLS). When no positioning was transmitted, the location was a posteriori estimated by linearly interpolating the trajectory of the animal. During the interpolation process, the horizontal speed of the animal was taken into account to ensure the spatio-temporal coherency of the location data. Ten of the studied SESs were also equipped with a biometric sonar and movement tag (DTAG). The DTAG placed on the animal's head picks up the GPS position using the Snapshot GPS acquisition algorithm (Goulet et al., 2019) and enables a more accurate GPS positioning of the profiles than with the Argos system. The positioning accuracy is 2-3 km for Argos, ~50 m for GPS (see Dragon et al., 2012; Irvine et al., 2020).

To avoid the influence of coastal waters and specifically focus on open-ocean so-called case 1 waters (Morel and Prieur, 1977; Morel, 1988), only profiles for which the seabed was deeper than

-1 500 m were kept. The ocean bathymetry data was based on ETOPO1 1 Arc-Minute Global Relief Model data from NOAA National Centers for Environmental Information and downloaded from https://www.ngdc.noaa.gov/mgg/global/relief/ETOPO1/data/. Following the filtering of the profiles according to the bathymetry criterion, the analysis included 63 791 light profiles and 4 404 Fluo profiles (see Table III.1 for a summary of the number of selected profiles).

Table III.1: Dataset selection criteria and number of selected PAR profiles

Filtering criterion	**Number of profiles**
All profiles	89 197
Bathymetry > 1 500 m	63 791
Day profiles	39 595
Depth interval ([Z_{inf} ; Z_{sup}])	14 686
Concomitant L + Fluo	1387

III.2.b. Data processing

Light profiles and derived quantities

The light sensor embedded in the SRDL is a Hamamatsu S1227-1010BR photodiode (340-1000 nm spectral response range, 100 mm^2 effective photosensitive area). The photodiode points to the right side of the animal with a 90º angle compared to the frontward axis of the animal. The SRDL light sensor provides an estimate of the diffused light level in the animal's environment (L, expressed in µmol quanta.m^{-2}.s^{-1}). The vertical profiles of light ranged from the maximum diving depth of the animal up to the surface. The processing steps of the raw vertical profiles of light include (for detailed description of the processing steps and graphical support, see supplementary material, Text S1 and Figure III.9): detection of the dark depth; dark-offset correction; removal of saturated values at the surface; application of a piecewise cubic polynomial fit. The applied piecewise polynomial fit is constrained, so that L monotonously decreases with depth. The vertical diffuse attenuation coefficient for L (K_L , m^{-1}) was derived from the processed light profiles. Vertical profiles of K_L were defined with the same vertical resolution as light profiles (i.e. 1 m) and computed as follows:

$$K_L(z) = \frac{d}{dz}\left(\log\left(L(z)\right)\right) \quad \text{(III.1)}$$

where *z* refers to the depth of the measurement.

The present analysis focuses on daylight periods only. According to the location and time of each profile, the solar angle was computed and only light profiles with positive solar angle (i.e. above the horizon) were retained. Profiles with no location available (23%) were still examined to

recover the day/night information from the mean surface values of L. This was enabled by the significant difference observed in the mean surface values of L between day (35 µmol quanta.m^{-2}.s^{-1} +-14) and night (0.65 µmol quanta.m^{-2}.s^{-1} +-3.7). As a result, 39 395 day profiles (62%) were retained after the filtering of the light profiles based on the daylight period criterion (see Table III.1).

Following the processing of the raw light profiles, the euphotic depth (Z_{eu}) and the penetration depth (Z_{pd}) were computed. Z_{eu} is defined as the depth at which L is reduced to 1% of its value just below the surface. Z_{eu} was only computed for light profiles with no sensor saturation in the surface layer. Z_{pd} (also called first optical depth) characterizes the thickness of the superficial layer of the ocean "seen" by satellites and was defined as Z_{eu}/4.6 (Gordon and McCluney, 1975; Morel, 1988).

Temperature, salinity and mixed layer depth

SRDLs carry a Conductivity, Temperature and Depth (CTD) sensor head. Temperature (T) and salinity (S) profiles, defined from the maximum diving depth of the animal up to the ocean surface, were quality controlled and corrected to prevent density inversions according to the algorithm proposed in Siegelman et al. (2019). Density profiles were computed based on temperature and salinity profiles. The mixed layer depth (Z_{MLD}) was computed using a density threshold of 0.03 kg.m^{-3} with respect to a near-surface value at 10 m depth (de Boyer Montégut et al., 2004).

Fluorescence profiles and derived quantities

SRDLs also include a fluorometer (Valeport Hyperion 470 nm/696 nm emission/reception) that sample Fluo at 0.5 Hz. However, to optimize tags's energy consumption, their Fluo sampling resolution was reduced so that the onset of the fluorescence sensor was triggered only every ~15 dives and Fluo was only sampled during the ascending phase of the dives from Z_{inf} = 200 m to the surface. Accordingly, the SRDLs performed around four fluorescence profiles every 24 hours. The processing steps of the raw vertical profiles of Fluo include (for detailed description of the processing steps and graphical support, see supplementary material, Text S2 and Figure III.10): dark-offset correction; Non-Photochemical Quenching (NPQ) correction; spikes smoothing with a piecewise cubic polynomial fit. Finally, the smoothed, dark- and NPQ-corrected Fluo data (hereafter denoted [F^{Fluo}]) were converted into [Chla]. The actual Chla concentration derived from [F^{Fluo}] ([$Chla^{Fluo}$] hereafter), was obtained by applying a calibration coefficient to the [F^{Fluo}] data. A specific calibration coefficient was computed for each tag, based on both K_L and the comparison of *in-situ* data with concomitant satellite-based [Chla] observations (see details in Section 2.4).

A series of metrics was computed from the vertical profiles of [Chla^{Fluo}]. Defined for each profile, the metrics were (see Figure III.2):

- <Chla^{Fluo}>, the water-column integrated value of [Chla^{Fluo}] defined as

$$\left\langle Chla^{Fluo} \right\rangle = \int_{Z_{inf}}^{0} \left[Chla^{Fluo} \right](z)\, dz \tag{III.2}$$

- [Chla^{Fluo}]$_{max}$, the maximum value of [Chla^{Fluo}]

- Z_{max}^{Fluo} , the depth where [Chla^{Fluo}] (z) = [Chla^{Fluo}]$_{max}$,

- [Chla^{Fluo}]$_{surf}$, the surface value of [Chla^{Fluo}] defined as

$$\left[Chla^{Fluo} \right]_{surf} = \overline{\left[Chla^{Fluo} \right]_{0 \leqslant z \leqslant Z_{pd}}} \tag{III.3}$$

Additionally, the percentage of Chla within the mixed layer was defined as

$$\left\langle Chla^{Fluo} \right\rangle_{\%ML} = \frac{100}{\left\langle Chla^{Fluo} \right\rangle} \int_{Z_{MLD}}^{0} \left[Chla^{Fluo} \right](z)\, dz \tag{III.4}$$

When relevant, the same metrics were computed for any other variable defined on the vertical in the present study (e.g. [F^{Fluo}]) with the same notations.

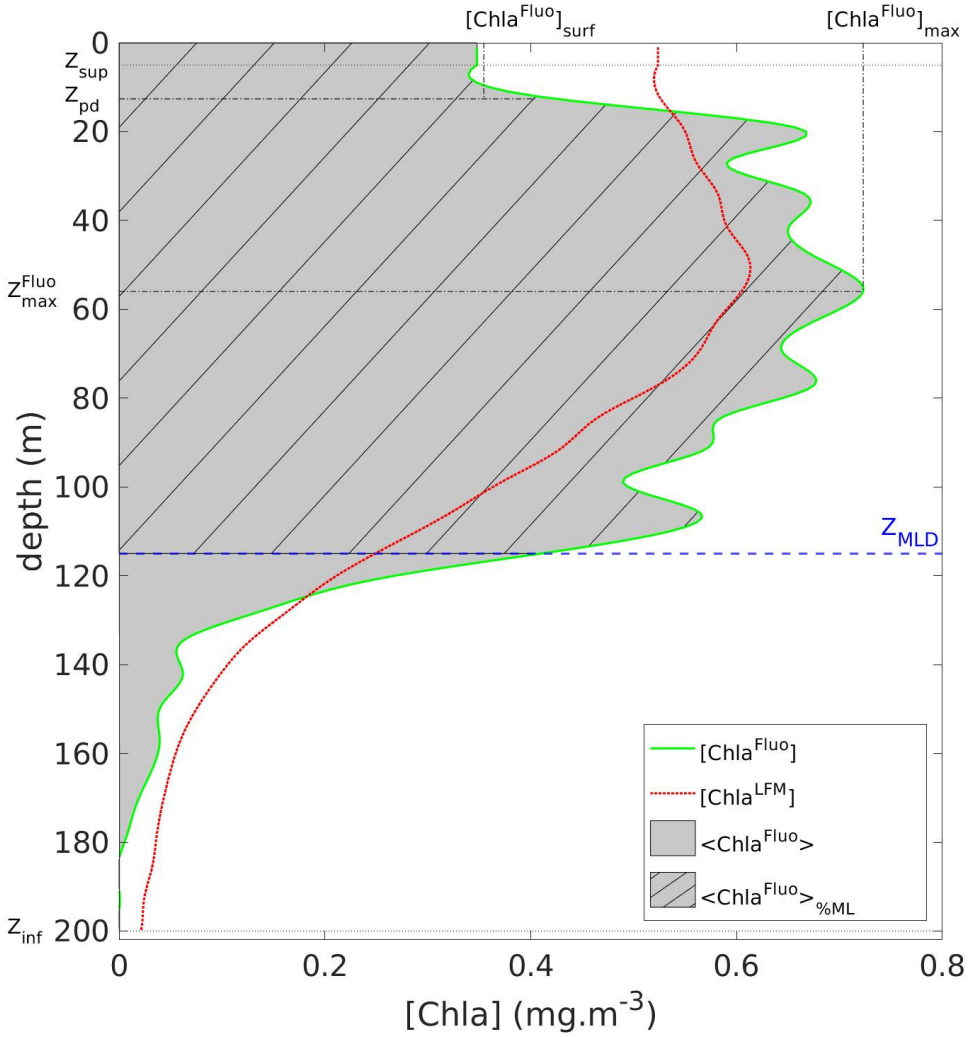


Figure III.2: Graphical representation of the metrics defined on a vertical profile of [Chla^{Fluo}]: <Chla^{Fluo}>, [Chla^{Fluo}]$_{max}$, Z_{max}^{Fluo} , [Chla^{Fluo}]$_{surf}$ (see Section 2.2.3). The solid green line represents the [Chla^{Fluo}] data (for detailed information about the Fluo data processing, see Section 2.2.3 and supplementary material, Text S2 and Figure S2). The gray area represents <Chla^{Fluo}>, the vertically-integrated amount of Chla. The dashed blue line represents Z_{MLD}. The dashed area materializes <Chla^{Fluo}>$_{\%ML}$, the proportion of <Chla^{Fluo}> located above Z_{MLD}. The red dotted line represents the corresponding predicted [Chla^{LFM}] profile (see Section 2.5). When necessary, the metrics were identically computed on the vertical profiles of [F^{Fluo}], [F], [F^{LFM}] and [Chla^{LFM}].

III.2.c. Linear Functional Model principle

The Linear Functional Model (LFM) developed in the present study is a statistical model based on Functional Data Analysis (Ramsay and Silverman, 1997). The LFM was used as an inference tool to predict [Chla] from K_L. The model is constructed from a statistical sample composed of concomitant vertical profiles of K_L (predictor) and [Chla] (observations), following the method described in Bayle et al. (2015). With the statistical sample at hand, the LFM is designed to minimize the error between model predictions and observations. The functional approach, by handling the vertical profiles as functional variables (i.e. curves), presents the advantage of integrating the shape of the profiles in the analysis (for detailed information about the construction of the model, see Bayle et al., 2015 and supplementary material, Text S4 and Figures III.12-III.13).

The statistical sample was composed of vertical profiles continuously defined on a depth interval ranging at least from Z_{sup} = 5 m to Z_{inf} = 200 m. The predicted profiles were defined on the same depth interval. The main limiting factor for the determination of Z_{sup} was the recurrent saturation of the light sensor at the surface (39% of the light profiles were saturated down to at least 5 m depth, see Table III.1 and supplementary material, Figure III.9). The choice of Z_{inf} was determined by the maximum depth of the Fluo measurements. For a proper interpretation of the results, the values in the upper 0-5 m layer of the predicted profiles were extrapolated from $z = Z_{sup}$ to the surface with their value at $z = Z_{sup}$. $[F^{Fluo}]$ and $[Chla^{Fluo}]$ values issued from Fluo measurements in the 0-Z_{sup} layer, meanwhile, were generally available. Following the filtering of the profiles based on the depth-interval criterion, the dataset contains 14 686 light profiles (from which K_L is derived, see Equation III.1), which includes 1 387 concomitant L and Fluo (i.e. $[F^{Fluo}]$ or $[Chla^{Fluo}]$) profiles. The number of L profiles following the application of the successive selection criteria is summarized in Table III.1.

In the present study, the LFM approach was used to predict either $[F^{Fluo}]$ or $[Chla^{Fluo}]$ from K_L, with different objectives, described hereafter (see Sections 2.4 and 2.5).

III.2.d. Fluo calibration

The calibration procedure applied to the *in-situ* $[F^{Fluo}]$ data aims at ensuring (1) the interoperability of the tags through the intercalibration of the Fluo sensors and (2) the consistency of the outputs of the model developed in the present study in terms of absolute values of *in-situ* [Chla] compared to satellite estimates. This two-step sequence is described hereafter (see flowchart, Figure III.1).

Step 1: LFM-based (relative) calibration

The predictive capabilities of the LFM approach were first exploited to intercalibrate the Fluo sensors. The LFM-based intercalibration step consists in predicting [F^{Fluo}] from K_L with a model that merges observations from all the tags. The predicted variable is hereafter denoted [F^{LFM}]. Within this step, the focus is not on the reconstruction of vertical profiles of [F^{Fluo}] but the intended goal is to examine and quantify the relative biases between the Fluo sensors. Consequently, rather than retrieving the parametric definition of the K_L-to-[F^{Fluo}] functional relationship, the comparison between <F^{LFM}> (predictions) and <F^{Fluo}> (observations) enables derivation of a correction factor, proper to each tag, that addresses inter-tag variability, with K_L as a common benchmark. The choice of <F^{LFM}> (i.e. indirectly, K_L) as the reference variable for Fluo intercalibration is discussed further (see Section 4.1). The [F^{Fluo}] data of Tag *e* were re-calibrated with the correction factor Γ_e^{LFM} so that

$$<F> = \Gamma_e^{LFM} . <F^{Fluo}> \quad \text{(III.5)}$$

where <F> is the re-calibrated <F^{Fluo}> data and *e* refers to the tag number (see list of tags by tag number in supplementary material, Table III.2).

In practical terms, the coefficient Γ_e^{LFM} (unitless) is the slope of the linear regression between the values of <F^{Fluo}> and the corresponding predicted values of <F^{LFM}> for Tag *e*. To compute the Γ_e^{LFM} coefficients, the sample of concomitant [F^{Fluo}] and K_L observations was merged and randomly split into two subsets: 70% of the profiles were used to construct the LFM (970 profiles). The remaining 30% (417 profiles) were used to evaluate the Γ_e^{LFM} coefficients. The sample used to evaluate the Γ_e^{LFM}coefficients was hence independent from the statistical sample used to construct the model. A bootstrap procedure was performed to gain robustness in the determination of the calibration coefficients: the LFM-based calibration was repeated one thousand times with a different random sampling at each iteration. At each iteration, a Γ_e^{LFM} calibration coefficient was calculated. Finally, the Γ_e^{LFM} calibration coefficient retained for Tag *e* corresponds to the median value of the Γ_e^{LFM} coefficients iteratively calculated for Tag *e*.

Step 2: Satellite-based (absolute) calibration

In a second phase of the calibration procedure, a single calibration factor Γ^{Sat} common to all tags was computed. Γ^{Sat} was based on ocean-color data as a benchmark to convert [F] into [$Chla^{Fluo}$]. Surface measurements of *in-situ* [F] ($[F]_{surf}$) of all the tags were merged and compared to the corresponding satellite-derived estimates of surface [Chla] ($[Chla^{Sat}]_{surf}$).

Matchups between satellite and *in-situ* data were performed following the procedure described in Bailey and Werdell (2006) using normalized satellite remote sensing reflectance (Rrs) daily Level-3 (L3) products from multiple sensors, with a 4 km resolution. The use of more stringent matchup protocols improves the quality of the matchup exercise (Concha et al., 2021), but

critically decreases the number of matchups (Haëntjens et al. 2017; Xi et al. 2020; Terrats et al. 2020), especially in the SO where cloud cover is a strong limiting factor. The narrow time window defined in Bailey and Werdell (2006) (+-3 h) was widened to a 24-hour time window (corresponding to the maximum temporal resolution of L3 ocean-color products). Haëntjens et al. (2017) show that expanding the temporal window from +-3 h to a 24-hour window increases the number of matchups, without significantly impacting the quality of the matchups. Accordingly, the matchup protocol used in the present study was based on the averaged data of a 3 x 3 pixel box centered on *in-situ* measurement with a 1-day time window. Satellite-derived estimates of surface [Chla] were obtained from the Copernicus Marine Service's GlobColour data archive (http://www.globcolour.info/). The coefficient Γ^{Sat} was defined as the slope of the linear regression between $[F]_{surf}$ and $[Chla^{Sat}]_{surf}$. Finally, the satellite-corrected data $[Chla^{Fluo}]$ was defined as follows

$$[Chla^{Fluo}] = \Gamma^{Sat} \, . \, [F] \qquad \text{(III.6)}$$

III.2.e. [Chla] prediction

Following the calibration procedure described in the previous section, the LFM approach was applied for inference purposes. A new LFM was designed to infer $[Chla^{Fluo}]$ from K_L. Being constructed with $[Chla^{Fluo}]$ profiles, the resulting LFM model hence inherently contains the calibration of the $[F^{Fluo}]$ data (i.e. Γ_e^{LFM} and Γ^{Sat} coefficients). The output variable is denoted $[Chla^{LFM}]$ (see flowchart, Figure III.1). The objective of the prediction phase is to increase the spatial resolution of the $[Chla^{Fluo}]$ field description with $[Chla^{LFM}]$. The prediction is made on the basis of the 14 686 available light profiles (see Table III.1). Prior to the retrieval of the $[Chla^{Fluo}]$ field at (sub)mesoscale, the performance of the LFM was assessed.

Performance assessment

The sample of concomitant $[Chla^{Fluo}]$ and K_L observations was likewise randomly split into two subsets: 70% of the profiles for the construction of the LFM (970 profiles) and 30% to assess the performance of the LFM (validation sample, 417 profiles). Assessments regarding LFM prediction error and model performance were carried out on the validation sample by comparing the metrics previously defined for $[Chla^{Fluo}]$ (see Section 2.2.3) with the same metrics derived from the predicted $[Chla^{LFM}]$ profiles, namely $<Chla^{LFM}>$, $[Chla^{LFM}]_{surf}$, $[Chla^{LFM}]_{max}$, and Z_{max}^{LFM} .

(Sub)mesoscale prediction

Following the assessment of the model itself via the analysis performed on the validation sample, the LFM was constructed with the entire sample of concomitant $[Chla^{Fluo}]$ and K_L profiles (1 387 profiles, see Table III.1). Subsequently, a prediction exercise was carried out with the 14 686 available K_L profiles (i.e. including K_L profiles which were not associated with any $[Chla^{Fluo}]$ data). The aim of the prediction exercise is to predict [Chla] at (sub)mesoscale.

III.2.f. Spectrum analysis

To further assess the improvement brought by the LFM in terms of spatial resolution, the variance spectra of the tags signals were analyzed. For a surface tracer measured with variable *V* (e.g. $<Chla^{Fluo}>$), the variance spectrum is calculated as the Fourier transform of the squared spatial anomaly of *V*. As a result, the power spectrum for variable *V* along the trajectory of an equipped animal depicts the energy of the signal as a function of the spatial frequency in the horizontal plane defined by the ocean surface. The variable on the horizontal axis of the computed power spectra is called wave number (m^{-1}). Increasing wave numbers correspond to smaller spatial scales.

In the present study, the variance spectra of $<Chla^{Fluo}>$ and $<Chla^{LFM}>$ were computed for each tag and compared. An additional variable (hereafter denoted $dark^{KL}$) was included in the spectrum analysis, defined as the vertical diffuse attenuation coefficient for L, restricted only to the dark signal of L (see Section 2.2.1 and supplementary material, Text S1 and Figure III.9). For each profile, $dark^{KL}$ corresponds to the mean vertical diffuse attenuation coefficient for L (i.e. the derivative of the log-transformed light profile) from the dark depth to the bottom of the dive. While $dark^{KL}$ contains no useful information for the inference of <Chla>, it was included in the spectrum analysis as a benchmark in terms of spectral behavior. Since $dark^{KL}$ is computed from dark noise, the corresponding power spectrum theoretically depicts the behavior of pure instrumental noise. The comparison with $dark^{KL}$ offers a means to determine if the observed tracers ($<Chla^{Fluo}>$ or $<Chla^{LFM}>$) indeed contain useful signals and depict coherent structures, or conversely, behave like noise.

III.3. Results

III.3.a. Calibration

Step 1: LFM-based tag intercalibration

The LFM-based tag intercalibration was performed based on the comparison of $<F^{Fluo}>$ and $<F^{LFM}>$ for the validation sample. The calibration coefficients were obtained after one thousand iterations of the LFM-based calibration procedure (see Section 2.4). An illustration of one iteration is presented in Figure III.3. Graphical examination of the residuals reveals that they are organized and persistent for a given tag, thus confirming the relevance of the intercalibration method. Z_{max}^{Fluo} values ranged from 0.29 for Tag 2 to 1.55 for Tag 17 (see complete list of Γ_e^{LFM} values in supplementary material, Table III.3). Small samples have the highest variability because they are not always well represented with the random sampling. Essentially, gaining robustness in the determination of the coefficient is the reason for performing a bootstrap procedure with one thousand iterations of the random sampling. The LFM-based calibration procedure enables computing of the [F] data, which ensures the interoperability of all the tags.

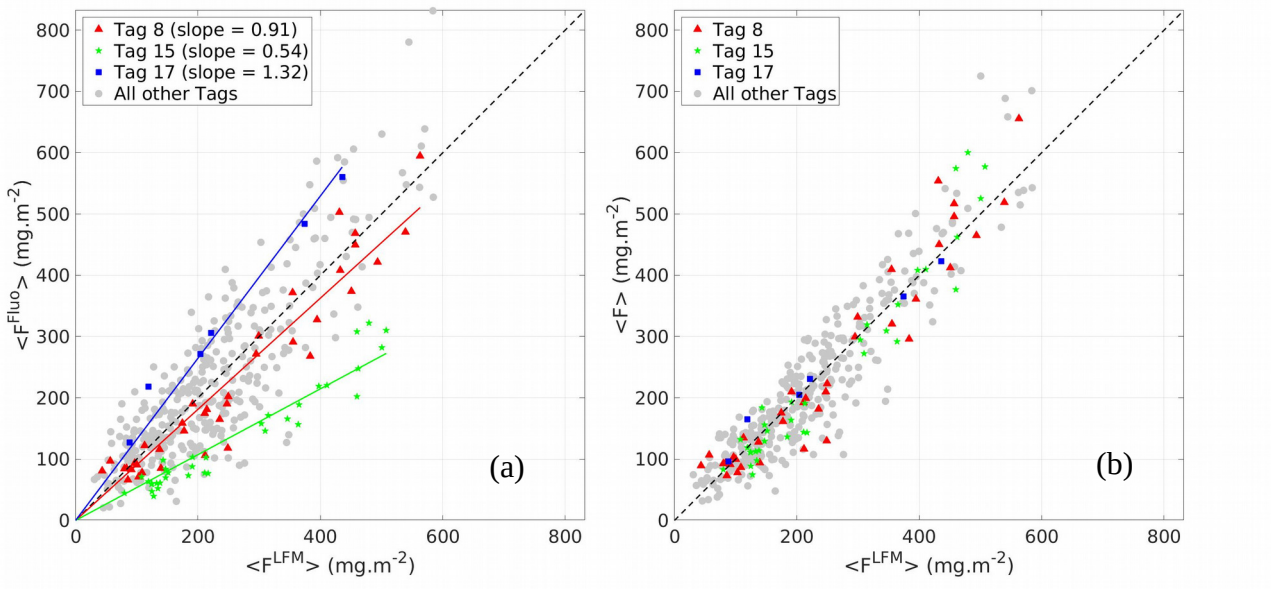


Figure III.3: Comparison between $<F^{Fluo}>$ and $<F^{LFM}>$ (a) before the LFM-based intercalibration step (b) after the LFM-based intercalibration step. Three tags are highlighted (red triangle, green dots and blue squares, for tags 8, 15 and 17, respectively). Data from all other tags are displayed by gray dots. The black dashed line represents the 1:1 reference line.

Step 2: Satellite-based calibration

The satellite-based calibration procedure was performed after the inter-tag calibration and by merging the [F] data of all the tags. The merging of the tags after the intercalibration step strengthens the power of the satellite-based calibration. Among the 5 791 [F] profiles available, 1 332 successful matchups (23%) were achieved (Table III.1). Γ^{Sat} = 5.9 was obtained from the slope of the linear regression between $[F]_{surf}$ and $[Chla^{Sat}]_{surf}$. The regression had a satisfactory significance level (F-test, p-value < 10^{-15}).

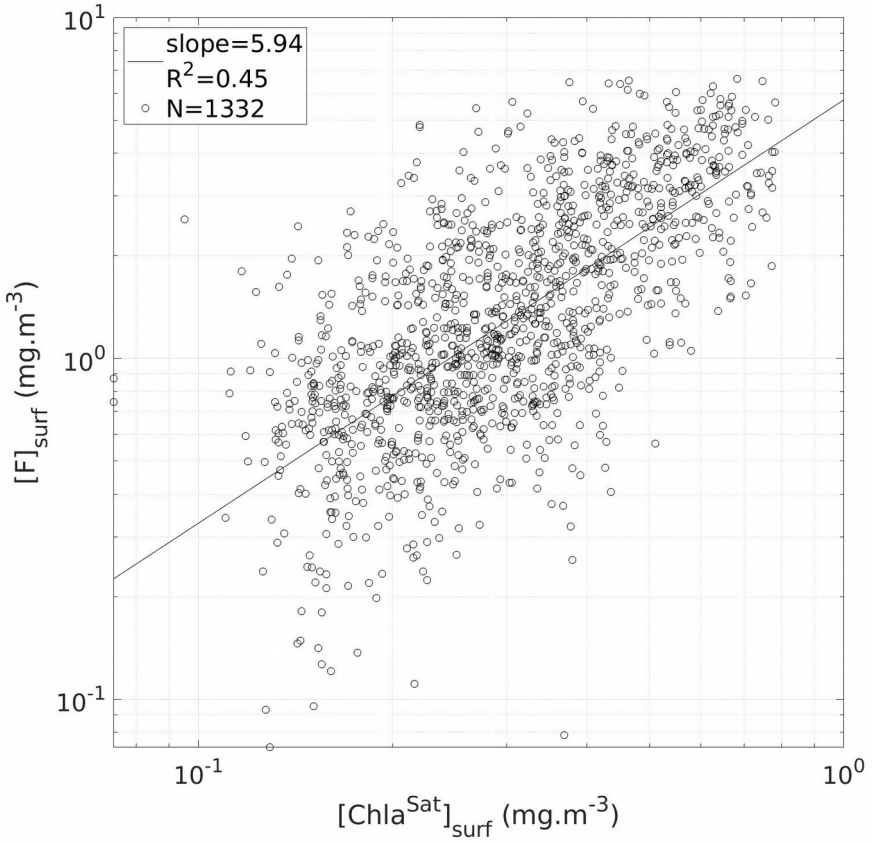


Figure III.4: Illustration of the second step of the [F^{Fluo}] data calibration (see Section 2.4). Estimates of surface [Chla] derived from intercalibrated [F] data ($[F]_{surf}$) compared to satellite estimates of surface [Chla] ($[Chla^{Sat}]_{surf}$). The black circles represent the sample points (in total: 1 332 matchups are displayed) and the black line materializes the linear regression of all the sample points (slope = 5.95; R^2 = 0.45; N = 1 332 matchups).

III.3.b. LFM Assessment

Following the conversion of the homogenized variable [F] into [Chla^{Fluo}] with the Γ^{Sat} correction factor (Equation III.6), a new LFM model was constructed on the basis of concomitant [Chla^{Fluo}] and K_L profiles, with [Chla^{LFM}] as the output variable. For assessment purposes, the statistical sample of 1 387 concomitant [Chla^{Fluo}] and K_L profiles was randomly split (see Section 2.5), so that the LFM was constructed with 70% of the statistical sample and assessed with the remaining 30% (validation sample). The performance of the model was assessed on the validation sample through examination of the metrics defined in Section 2.2.3 (see Figure III.2), for both [Chla^{Fluo}] and [Chla^{LFM}] (Figure III.5).

<Chla>

The predicted <Chla^{LFM}> differs very little from the targeted <Chla^{Fluo}> (on average 0.9% +-21.1). Within a factor Γ^{Sat}, the performance of the model in predicting <Chla^{Fluo}> both in terms of accuracy and precision is exemplified in Figure III.3. The sound agreement between <F^{Fluo}> and <F^{LFM}> observed in the intercalibration phase firstly confirms the inference capabilities of the model in terms of accuracy (Figure III.3a). Additionally, by correcting the inter-tag variability, the LFM-based calibration procedure (Equation III.5) inherently increases the precision of the predictions regarding the estimation of the water-column integrated Chla biomass (Figure III.3b). Finally, no further change in terms of accuracy and precision is implied by applying the Γ_e^{LFM} factor, common to all tags, to obtain [Chla^{Fluo}] from [F] (Equation III.6).

<Chla>$_{\%ML}$

The distribution of the Chla biomass in the vertical is further investigated with the variable <Chla>$_{\%ML}$. <Chla>$_{\%ML}$ represents the ratio between the Chla content in the 0-Z_{MLD} layer and the total Chla content in the water column (<Chla>). The slope of the linear regression between <Chla^{LFM}>$_{\%ML}$ and <Chla^{Fluo}>$_{\%ML}$ reveals that the model renders, with a satisfactory preciseness, the proportion of the vertically-integrated Chla amount located above and below Z_{MLD} (slope = 0.92, R^2 = 0.89). On average, the LFM underestimates <Chla>$\%_{ML}$ by only 4.8% (+-7.2).

[Chla]$_{surf}$, [Chla]$_{max}$ and Z$_{max}$

The ability of the model to retrieve the exact vertical distribution of Chla is examined with variables [Chla]$_{surf}$, [Chla]$_{max}$ and $Z_{max}^{(c)}$. [Chla^{LFM}]$_{surf}$ and [Chla^{LFM}]$_{max}$ are compliant with the corresponding observations of [Chla^{Fluo}]$_{surf}$ and [Chla^{Fluo}]$_{max}$, although slightly underestimated (Figure III.5b-c). The retrieval of Z_{max} (Figure III.5d) is however not satisfying and in some way reveals the limits of the LFM. The poor correlation between Z_{max}^{Fluo} and Z_{max}^{LFM} highlights the weak accuracy of the LFM for retrieving the exact vertical structure of the [Chla] profile.

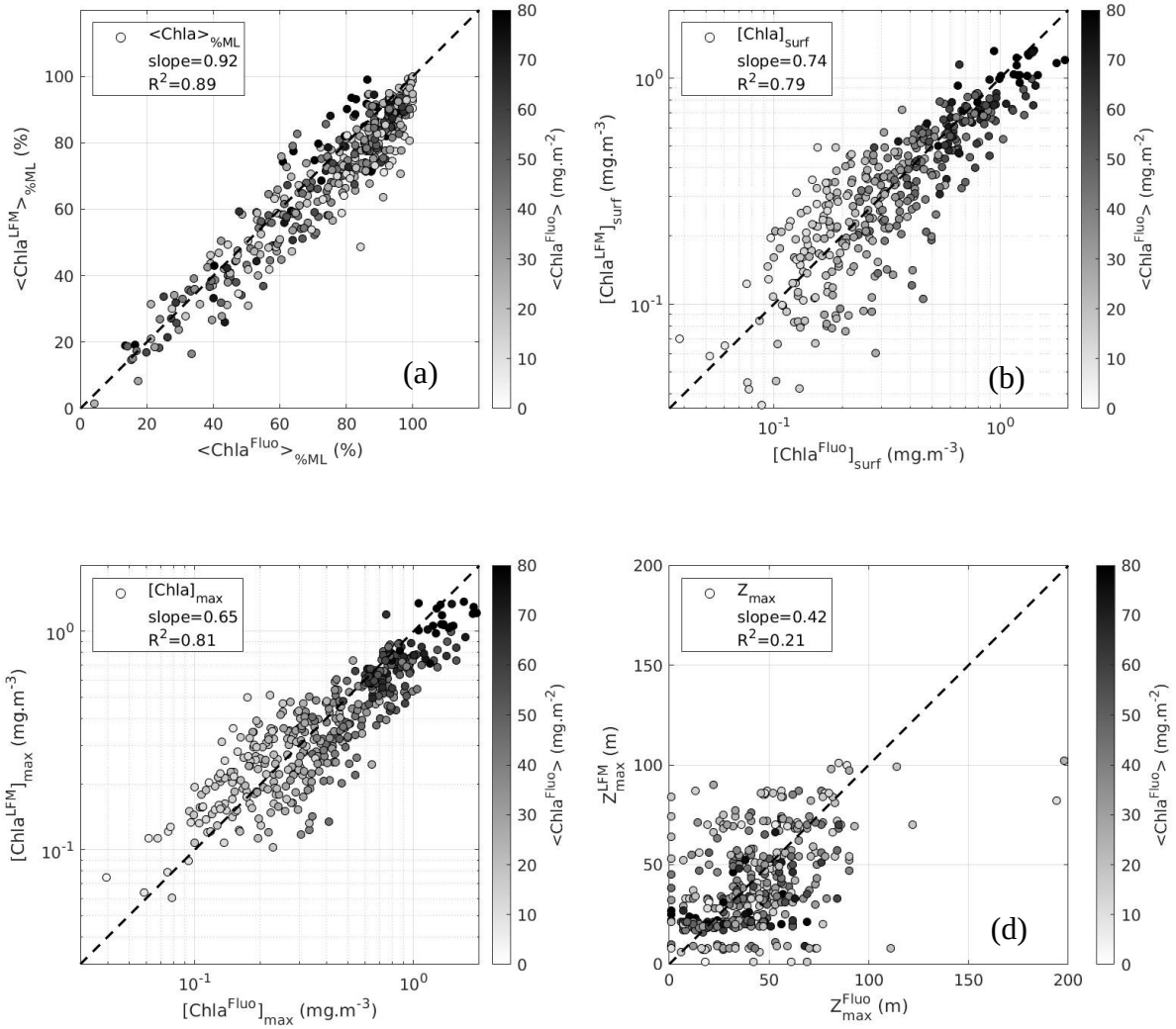


Figure III.5: Assessment of the LFM performance on the validation sample. Values in the horizontal axis are metrics derived from the observations ([Chla^{fluo}]). Values in the vertical axis are metrics derived from the predictions ([Chla^{LFM}]). The metrics examined to assess the performance of the LFM are (a) [Chla^{Fluo}]$_{\%ML}$, the percentage of Chla in the mixed layer (b) the surface value of [Chla], [Chla^{Fluo}]$_{surf}$ (c) the maximum value of [Chla], [Chla^{Fluo}]$_{max}$ and (d) Z_{max}^{Fluo} , the depth of the maximum value of [Chla]. The dashed black lines in each plot represent the 1:1 reference line. The metrics (R^2, slope) associated with the linear regression performed between model predictions and observations are indicated in each plot.

These results of the model assessment lead to the conclusion that the LFM performs quite well in detecting the amount of Chla in a given profile. The rough vertical distribution of [Chla] in relation to the location of Z_{MLD} is also well achieved by the model. However metrics on the vertical such as Z_{max} are not accurately rendered and present high variability.

III.3.c. Validation with satellite data

Following the assessment of the LFM performance, the model was constructed with all available profiles of $[Chla^{Fluo}]$ and K_L (1 387 profiles, see Table III.1). A prediction exercise was then carried out with all the available K_L profiles (derived from the 14 686 light profiles).

The validity of the LFM predictions was tested through examination of the satellite matchups corresponding to the predicted $[Chla^{LFM}]$ profiles (see Section 2.4 for the matchup procedure), i.e. by comparing $[Chla^{LFM}]_{surf}$ with the co-located $[Chla^{Sat}]_{surf}$ estimates. In total, 3 320 successful matchups were achieved (23%). As a direct consequence of the calibration procedure previously performed on the $[F^{Fluo}]$ data (see Section 3.2), the slope factor of the linear regression of $[Chla^{LFM}]_{surf}$ with $[Chla^{Sat}]_{surf}$ (slope = 1.01) confirms the consistency of the model outputs in relation to satellite-derived estimates of [Chla] (Figure III.6). The compliance of the predicted values with the corresponding $[Chla^{Sat}]_{surf}$ estimates validates the LFM predictions. However, a clear divergence is noticeable for low values of $[Chla^{LFM}]_{surf}$ ($[Chla^{LFM}]_{surf} < 0.1$ mg.m^{-3}). To further investigate the validity of the low values of $[Chla^{LFM}]_{surf}$, the available concomitant values of $[Chla^{Fluo}]_{surf}$ were examined and compared to $[Chla^{Sat}]_{surf}$ (the 14 686 K_L profiles of the prediction exercise include the sample of concomitant $[Chla^{Fluo}]$ and K_L profiles, i.e. 1 387 $[Chla^{LFM}]$ profiles with an available concomitant value of $[Chla^{Fluo}]$). The comparison of $[Chla^{Fluo}]_{surf}$ values with the corresponding $[Chla^{Sat}]_{surf}$ estimates (N = 329 successful matchups), also represented in Figure III.6, reveals a similar divergence to that observed for $[Chla^{LFM}]$ below ~ 0.1 mg.m^{-3} . The difference between low values of $[Chla^{LFM}]_{surf}$ and the corresponding $[Chla^{Sat}]_{surf}$ estimates is therefore attributable to the discrepancy between *in-situ* and satellite measurements of [Chla] (discussed further, see Sections 4.2 and 4.3.2) rather than to a model deviation. Low values of $[Chla^{LFM}]_{surf}$ are hence valid in the sense that they match with the targeted $[Chla^{Fluo}]_{surf}$.

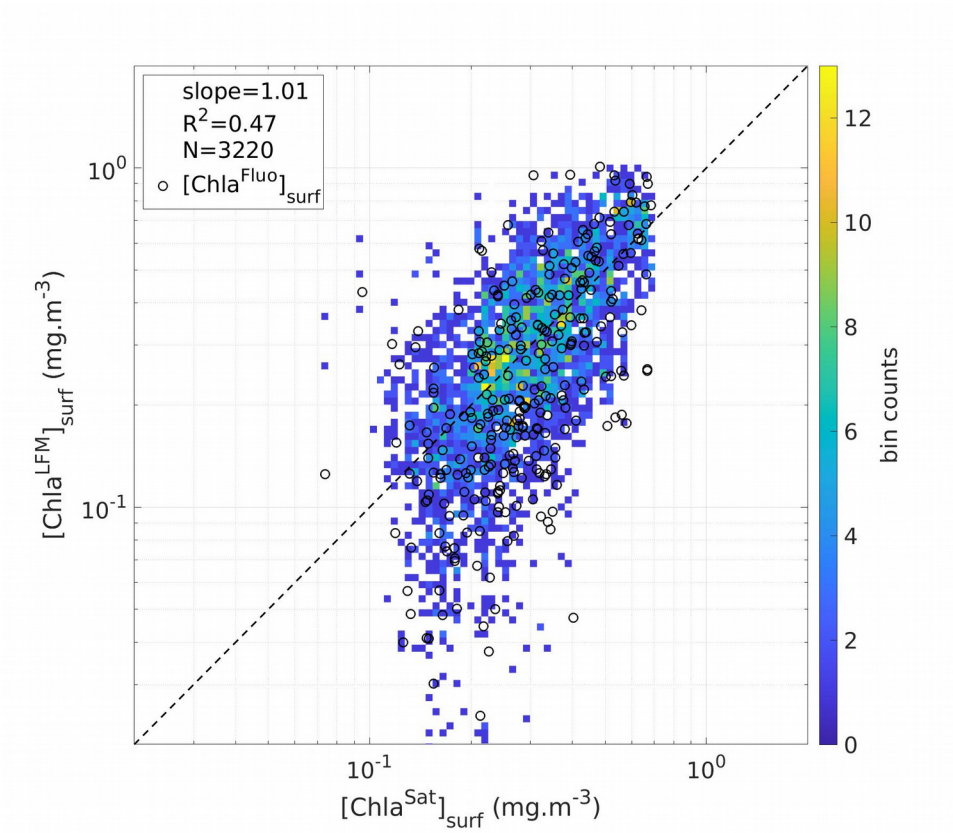


Figure III.6: Estimates of surface [Chla] derived from LFM predictions ($[Chla^{LFM}]_{surf}$) compared to satellite estimates of surface [Chla] ($[Chla^{Sat}]_{surf}$). The color of the pixels represent the density of the points in the plot (in total: 3 320 matchups are displayed). The dashed black line materializes the linear regression including all matched samples of $[Chla^{LFM}]_{surf}$ and $[Chla^{Sat}]_{surf}$ (slope = 1.01; R^2 = 0.47; N = 3 320 matchups). The black circles represent the available concomitant values of $[Chla^{Fluo}]_{surf}$ (N = 329 successful matchups within the 1 387 $[Chla^{Fluo}]$ profiles used to construct the model, see Section 3.3).

III.3.d. Application: (sub)mesoscale retrieval

Transect of a SES equipped in Kerguelen

A subset of the SES dataset corresponding to a single individual transect (Tag 3) is shown in Figure III.7. The transect is 5 746 km long and covers 70 days at sea, between 25-Oct-2019 and 02-Jan-2020 (Figure III.7a). The transect comprises 234 profiles of $[Chla^{Fluo}]$ and 879 light profiles. One notable detail regarding the transect of Tag 3 is the malfunctioning of the fluorescence sensor during a certain period of the deployment. As a consequence, no Fluo data were available for the

time interval extending from 09-Nov-2019 to 02-Dec-2019 (over the same period, 342 T, S and light profiles were sampled by the SRDL).

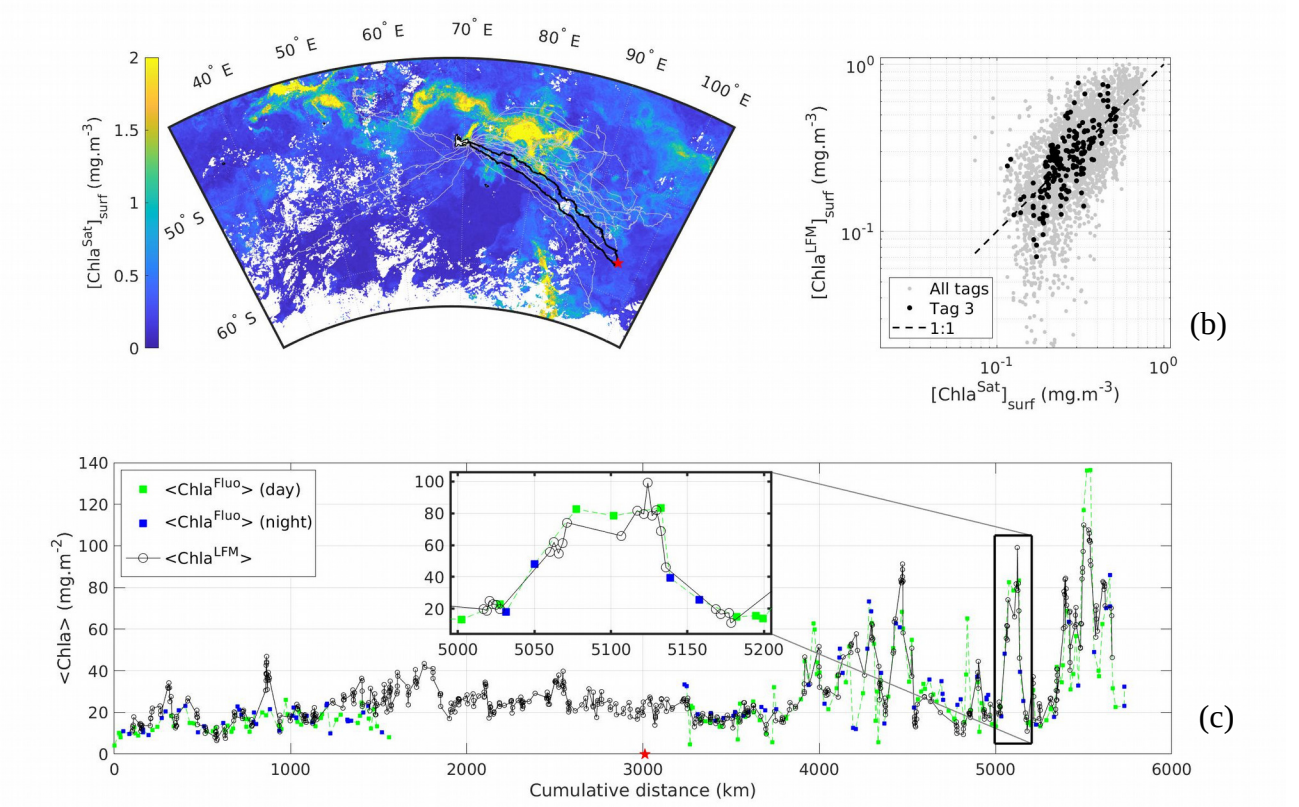


Figure III.7: Transect of a SES equipped in Kerguelen in October 2019 with SRDL referred to as Tag 3 including (a) a map of the trajectory described by the SES from 25-Oct-2019 to 02-Jan-2020 departed from- and arrived at Kerguelen (b) comparison of $[Chla^{LFM}]_{surf}$ with $[Chla^{Sat}]_{surf}$ for Tag 3 and (c) <Chla> as measured (<$Chla^{Fluo}$>) and predicted (<$Chla^{LFM}$>) along the transect of Tag 3, where green (blue) squares represent the <$Chla^{Fluo}$> data measured during the day (night) and black circles represent <$Chla^{LFM}$>. The inset in (c) highlights the data on a section of the animal transect (~200 km). The red star in (a) and (c) represents the furthest location from Kerguelen in the trajectory of the animal equipped with Tag 3. Light gray dots in (a) and (b) represent the data of all the other tags included in the present study. The dashed black line in (b) represents the 1:1 reference line. The background map in (a) is the climatology of $[Chla^{Sat}]_{surf}$ derived from GlobColour computed for each pixel as the mean value of $[Chla^{Sat}]_{surf}$ during the month of November 2019.

The comparison of <$Chla^{Fluo}$> and <$Chla^{LFM}$> along the animal's trajectory clearly reveals that the <Chla> signal is well captured by the model. K_L-based LFM predictions faithfully reproduce [$Chla^{Fluo}$] observations at water-column level. Additionally, as previously observed when merging all the tags (Figure III.6) the general strong correlation between $[Chla^{Sat}]_{surf}$ and $[Chla^{LFM}]_{surf}$

(Figure III.7b) confirms the validity of the LFM predictions. [$Chla^{LFM}$] data are especially valuable during the period for which no Fluo data were available due to the malfunctioning of the fluorescence sensor: the missing block (23-days long) of [$Chla^{Fluo}$] data was retrieved thanks to the [$Chla^{LFM}$] predictions. Inversely, [$Chla^{Fluo}$] estimates are available at night when no [$Chla^{LFM}$] profiles could be derived from variable L (see inset in Figure III.7c). Both situations emphasize the complementary assets of [$Chla^{Fluo}$] and [$Chla^{LFM}$] estimates. Due to the fact that <$Chla^{LFM}$> estimates are available at a higher spatial resolution than that of <$Chla^{Fluo}$> (see inset in Figure III.7c) during daylight periods, LFM predictions performed between two consecutive [$Chla^{Fluo}$] profiles enable the scale of the observations to be refined. The gain relative to the spatial resolution of <$Chla^{LFM}$> estimates compared to <$Chla^{Fluo}$> is examined hereafter.

Variance spectra

The variance spectra of <$Chla^{Fluo}$> and <$Chla^{LFM}$> were computed for each tag included in the present study (see Section 2.6). A specific focus was placed on Tag 11 for which the recordings of both Fluo and L were continuous (uninterrupted) during the studied transect. The variance spectrum of $dark^{KL}$ for Tag 11 was also computed. The highlighted transect corresponding to Tag 11 is 2 231 km long and covers 43 days at sea (between 19-Oct-2018 and 30-Nov-2018). The transect comprises 209 profiles of [$Chla^{Fluo}$] and 851 light profiles.

(a)

The variance spectra of both observations and predictions of <Chla> along the transect of Tag 11 were compared (Figure III.8). The extension of the <$Chla^{LFM}$> signal towards the (sub)mesoscale is clearly visible through comparison of the variance spectra of <$Chla^{Fluo}$> and <$Chla^{LFM}$>. While the smallest spatial scale reached with [$Chla^{Fluo}$] observations is ~21 km in the example of Tag 11, the spectrum of the LFM predictions extends to a spatial scale of ~2 km.

On the interval where both signals are defined, a clear energy decay is visible in the variance spectra of <$Chla^{Fluo}$> and <$Chla^{LFM}$>, following a power-law behavior in k^{-a} (where k is the wave number, and $-a$ the spectral slope on a log-log plot). For scales larger than ~21 km, the spectral slope of the <$Chla^{LFM}$> signal ($k^{-1.17}$) is in line with the spectral slope of <$Chla^{Fluo}$> ($k^{-1.22}$), attesting to the good agreement between observations and model predictions (for interpretation of the spectral slopes, see Section 4.3.4). From ~21 km to ~7 km, the <$Chla^{LFM}$> signal similarly follows a power-law behavior with a spectral slope equal to -2.3 (i.e. a steeper decrease than at larger spatial scales). From ~7 km down to ~2 km the spectral slope is -5.2, but this part of the spectrum appears to be much noisier.

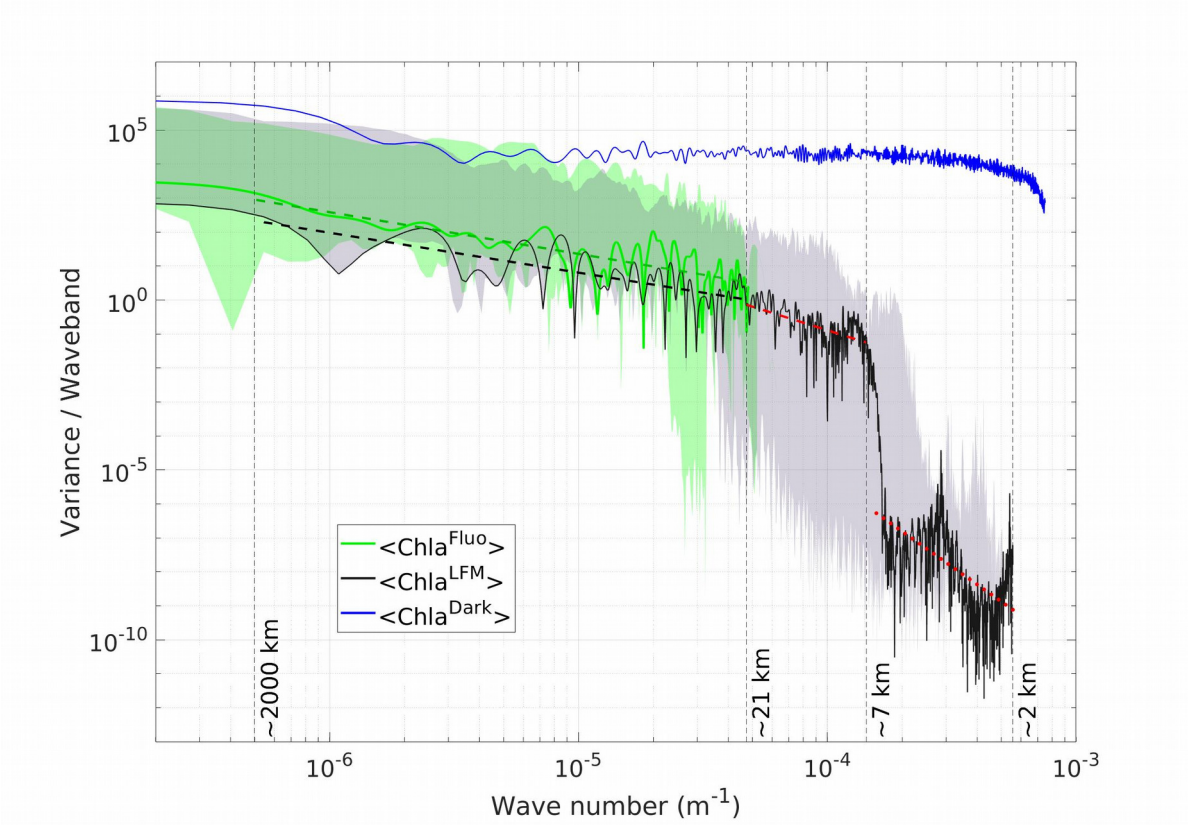


Figure III.8: Variance spectrum of <Chla> from observations and model predictions along the trajectory of Tag 11. Increasing wave numbers correspond to smaller spatial scales (see Section 2.6). The green (black) solid line represents the variance per waveband of the <Chla^{Fluo}> (<Chla^{LFM}>) signal. The blue solid line represents the variance spectrum of dark^{KL}. The green (gray) shaded area represents the envelope of the <Chla^{Fluo}> (<Chla^{LFM}>) spectra. The envelope encompasses the minimum and the maximum variances per waveband obtained in the dataset of the 18 SES tags included in the present study. *The green (black) dashed line represents the linear regression of the variance spectra of <Chla^{Fluo}> (<Chla^{LFM}>) on the spatial scale interval between ~2 000 km and ~21 km, with a spectral slope equal to -1.22 (-1.17). The dashed (dotted) red line represents the linear regression of the variance spectrum of <Chla^{LFM}> on the spatial scale interval between ~21 km and ~7 km (~7 km and ~2 km).* The *exact corresponding wave numbers are 5 10^{-7} m^{-1} (~2 000 km), 4.7 10^{-5} m^{-1} (~21 km), 1.4 10^{-4} m^{-1} (~7 km) and 5 10^{-4} m^{-1} (~2 km). (For the determination of the thresholds used for the piecewise linear regressions calculated on the variance spectra of <Chla^{Fluo}> and <Chla^{LFM}>, see supplementary material, Text S7 and Figures S8-S9).*

The spectrum of dark^{KL} derived from the dark noise of L, computed for Tag 11 (see Section 2.6), was added in Figure III.8 to illustrate the spectral characteristics of dark noise. The variance spectrum of dark^{KL} is almost flat for wave numbers smaller than ~3 10^{-4} m^{-1}, meaning that all spatial frequencies larger than ~3 km are equally represented within the dark^{KL} signal. Such a spectrum is

coherent with the definition of pure noise and contrasts with the power-law behavior of $<Chla^{Fluo}>$ and $<Chla^{LFM}>$.

For spatial scales smaller than ~21 km, where only $<Chla^{LFM}>$ is defined, two distinct wave number intervals are clearly discernible, separated by a pronounced drop in the energy of the signal, located at around ~7 km in the case of Tag 11. The signal gets much noisier after the energy drop. A similar behavior was observed for every tag included in the present study, namely a pronounced energy drop materializing a spatial scale threshold below which the spectrum follows a power-law behavior, and above which the signal loses coherency (i.e. gets much noisier). It is consequently reasonable to consider that the interpretations of the structures depicted by the $<Chla^{LFM}>$ signal are valid up to the scale of the energy-drop threshold and should be discarded for higher spatial frequencies. In the case of Tag 11 (specifically highlighted in Figure III.8), the spatial resolution of the observations hence extends from ~21 km with Fluo to at least ~7 km with LFM predictions. The energy-drop threshold was different for each tag. The shaded gray area in Figure III.8 represents the envelope of the $<Chla^{LFM}>$ spectra, encompassing the minimum and the maximum variances per waveband obtained in the dataset of the 18 SES tags included in the present study. The envelope of the $<Chla^{LFM}>$ spectra reveals that the spatial scale of the energy drop spreads from ~30 km to ~4 km (corresponding to wave numbers of ~$3\ 10^{-5}$ m^{-1} and ~$2.5\ 10^{-4}$ m^{-1}, respectively).

Similarly, a comparable energy drop was observed in the $<Chla^{Fluo}>$ spectrum for some, though not all, of the tags included in the present study (see green shaded area in Figure III.8). The energy drop in the $<Chla^{Fluo}>$ spectra occurred at larger spatial scales than in the $<Chla^{LFM}>$ spectra, starting from ~45 km.

III.4. Discussion

III.4.a. Exploiting light-Fluo synergies through LFM: data intercalibration and homogenization

Including data from multiple tags in a study raises the issue of the intercalibration of the fluorescence sensors, a critical point when [Chla] estimates are to be derived from Fluo measurements. Evidences of inter-sensor variability have already been pointed out for different fluorometer models (e.g. Guinet et al., 2013; Xing et al., 2014; Keates et al., 2020), emphasizing the necessity of homogenizing the Fluo data from one tag to another before any further analysis. A common benchmark which provides an absolute Fluo to [Chla] conversion theoretically resolves the inter-tag calibration issue. However, the fluorescence sensors embedded in the tags examined in the present study did not undergo any *in-situ* calibration process and not all of them could be successfully independently calibrated with concomitant satellite estimates of [Chla] (see Section 4.2).

The lack of any direct comparative benchmark for Fluo led to an investigation of the *in-situ* data concomitantly sampled by the SRDLs in order to best take advantage of them. Accordingly, the first step of the Fluo calibration procedure relies on variable K_L as a common variable to all the tags and exploits the predictive assets of the LFM to intercalibrate the Fluo sensors. It has been previously demonstrated (Morel, 1988; Morel and Maritorena, 2001) that the optical properties of open-ocean waters (so-called case 1 waters) are essentially driven by their phytoplankton content (depicted by the concentration in Chla) and their associated living or inanimate materials (heterotrophic organisms, including bacteria; various debris; and excreted organic matter). Such relationships between the water-column algal content and optical properties have also been used to deeper examine the data acquired by electronic tags deployed on pelagic animals (Teo et al., 2009; Jaud et al., 2012; Bayle et al., 2015). Here, we further exploit the synergy between K_L and Fluo by proposing a method to make the Fluo data from different (and not intercalibrated) tags inter-comparable. As K_L depicts gradients (derivative) of light in the water column rather than absolute light levels, it is much less dependent on the light sensor design, calibration or drift (e.g. biofouling). Therefore, K_L is a highly robust measurement that can potentially serve as a reference measurement for long-term observations like those obtained via autonomous platforms (floats, gliders) or animals.

In this context, a method based on an analytical relationship linking the diffuse attenuation for downward irradiance K_d to [Chla] (Morel and Maritorena, 2001) was first proposed in Xing et al. (2011) to take advantage of fluoresence profiles acquired by BGC-Argo floats simultaneously with radiometric profiles of downward irradiance. The relationship between [Chla] and downward irradiance is investigated at three specific wavelengths and enables the calibration of Fluo data in

terms of [Chla] as well as the handling of any potential drift of the Fluo sensor over time. In our study, the SRDL provides a measurement on a large spectral interval (340 to 1000 nm, see Section 2.2.1). Additionally the quality of the radiometric measurements performed with the SRDL is not comparable to those of BGC-Argo floats, for which the verticality of the sensor as well at the exact time of the sampling are monitored to optimize the quality of the radiometry data (see Organelli et al. 2016; Xing et al. 2011). As a consequence, we judged relevant to adopt a less strictly analytical method for the matching between K_L and [Chla] but instead, encompass the variability in the entire visible spectrum with a shape-based approach (LFM). The LFM imposes no a priori model regarding the relationship between K_L and [Chla] during the construction of the model, leaving the calibration of Fluo in terms of [Chla] for a later stage of the procedure. Accordingly, the first step of the Fluo calibration procedure solely relies on variable K_L as a common variable to all the tags and exploits the predictive assets of the LFM to intercalibrate the Fluo sensors.

Finally, it is worth noting that the power and robustness of the LFM depend on the size of the statistical sample used to construct the model (for results regarding the robustness of the model in relation to the composition of the statistical sample used to construct the model, see supplementary material, Text S5 and Figure III.14). The amount of data available to feed the model is limited by the fact that K_L is only exploitable during daytime (in the present study, only light profiles associated with a positive solar angle were selected). During daytime, the determination of K_L is partially influenced by the solar angle (Morel et al., 2007). Nevertheless, the model's accuracy appears not to be influenced by solar angle as the prediction error presented a similar dispersion for all positive values of the solar angle (see supplementary material, Text S6. and Figure III.15).

III.4.b. Matching *in-situ* and satellite measurements for absolute calibration

A per-tag satellite-based calibration procedure comparing surface measurements of Fluo with concomitant satellite estimates of [Chla] is an alternative way to convert [F^{Fluo}] into actual [Chla] (e.g. Lavigne et al., 2012; Terrats et al., 2020), in particular when no pre-deployment HPLC [Chla] data are available for the calibration of the tags. However, although surface Fluo measurements could be quite successfully matched with satellite data for some of the tags, others critically lacked sufficient satellite coverage to permit trustworthy calibration. The per-tag satellite matchup procedure hence could not be generalized to all the tags. Therefore the LFM-based step discussed above was an essential requirement to correct for inter-tag variability and to render all the [F^{Fluo}] data interoperable, thus constituting a homogeneous data base from which absolute calibration (i.e. conversion from [F] into [Chla], see flowchart, Figure III.1) could subsequently be established. The merging of all the intercalibrated tags indeed reinforces the quality of the comparison with satellite data and increases the robustness of the calibration procedure.

The second step of the Fluo calibration therefore converts the fluorescence signal into an estimate of [Chla], based on concomitant satellite measurements. Satellite-based ocean-color algorithms for the retrieval of [Chla] do not perform equally in all regions of the globe (Szeto et al., 2011). Specifically for the SO, the need for having regionally-tuned algorithms is a matter of debate, with some arguing that satellite-derived [Chla] is underestimated by a 2-3 factor (Johnson et al., 2013; Guinet et al., 2013), while others reporting that standard algorithms for the global ocean perform well (Haëntjens et al., 2017). The standard satellite [Chla] product used here (namely, the Copernicus Marine Service's GlobColour ocean-color data) is a global scale product which we consider to be adapted in the context of the main study purposes. Furthermore, merging all the tags' data prior to the satellite calibration was relevant because all the tags were deployed in the same region, namely the Kerguelen Islands, which reinforces the interoperable nature of the various tag observations.

III.4.c. Assessment of the method

Retrieval of the vertical distribution of Chla

Following the constitution of calibrated [Chla] datasets, the LFM was developed such that the modeled [$Chla^{LFM}$] matched as closely as possible the targeted [$Chla^{Fluo}$]. The retrieval of <$Chla^{Fluo}$> was generally well achieved with the LFM predictions (see Section 3.2). The model however lacks accuracy along the vertical dimension. In the SO, Z_{MLD} is a major driver in the vertical distribution of Chla. For a large majority of profiles, the mixed layer contains most of the Chla biomass and [Chla] is homogeneous within this layer (Cornec et al., 2021). Yet a remainder of the vertically-integrated Chla can be present below Z_{MLD}. Since the model also proved to successfully estimate <$Chla^{LFM}$>$_{\%ML}$ (Figure III.5a), an estimate of this remainder is achievable. Here, we make the hypothesis that this remainder may not be present deeper than 1.5 times Z_{eu}. As a consequence, by associating <$Chla^{LFM}$> predictions with <$Chla^{LFM}$>$_{\%ML}$, Z_{MLD} and Z_{eu}, an insight into the distribution of Chla in the vertical is still achievable even in the absence of properly reconstructed vertical profiles.

Fluo uncertainty

The LFM is based on estimates derived from fluorescence measurements, with the aim of reproducing them. Nevertheless, uncertainties exist regarding the algal biomass estimates derived from fluorescence measurements. The relationship between Chla fluorescence and actual Chla concentration is in particular governed by the fluorescence quantum yield (expressed as: mole emitted photons (mole of absorbed photons)$^{-1}$) which depends on many factors, including phytoplankton community composition, photo-physiological as well as nutrient status (Roesler et al., 2017; Schallenberg, et al., 2022). The tags included in the present study were all deployed from the Kerguelen Islands, but the trajectories of the equipped animals spread in the ocean from East to West of the Kerguelen Plateau. The Kerguelen plateau region is highly contrasted between the iron-limited western part and the iron-fertilized eastward zone (Blain et al., 2008). The merging of [F^{Fluo}]

data of all the tags as part of the Fluo calibration procedure could thus be in some ways questionable. However, the intercalibration coefficients derived from the LFM predictions were examined and no significant difference was observed between East and West of the Kerguelen Plateau (Mann-Whitney-Wilcoxon test, p-value of 0.63).

Model - observation discrepancies

The LFM is based on the assumption that the presence of phytoplankton is the main source of light attenuation in the water column, as commonly hypothesized when studying oceanic case 1 waters. Accordingly, a minimum bathymetry criterion was fixed to avoid dealing with coastal waters (depth > 1 500 m). However, persistent overestimates of the Chla content are observed in some short portions of the predicted signal (Figure III.7c). Such deviations are of special interest for analyzing specific issues or limitations regarding the LFM. Generally however, locally persistent deviations may result from small-scale variations in the bio-optical properties of the corresponding water masses. These could be due the presence of other covarying substances contributing to light attenuation and/or affecting the fluorescence signal (Bricaud et al., 1998; Loisel et al., 2002; Bellacicco et al., 2019). Such small-scale variations were however not investigated in the present study.

Towards filling [Chla] observational gaps at the (sub)mesoscale

One of the potentially interesting outcomes of the development of the present method is that it becomes obvious that measurements of light rather than fluorescence represent a cost-effective alternative for filling [Chla] observational gaps. The observation gap originates from the limitations relative to the power consumption of SRDLs mounted on SESs, which impedes Fluo sampling at a scale compatible with (sub)mesoscale observations. By contrast, measurements of light require less energy and can be performed at much higher spatio-temporal resolution. Comparing the variance spectra of both observations and predictions (Figure III.8) corroborates the gain brought by the LFM in terms of spatial resolution. The gain in terms of spatial resolution is also visible on the transect presented in Figure III.7c. In the case of the SRDLs, $<Chla^{LFM}>$ is the only variable defined at (sub)mesoscale and higher to describe phytoplankton dynamics, opening up a new way to fill the (sub)mesoscale observational gap.

The analysis of the variance spectrum of $<Chla^{LFM}>$ was further used as a means to validate the consistency of the predictions at different spatial scales. At spatial scales where both the measured ($<Chla^{Fluo}>$) and the predicted ($<Chla^{LFM}>$) signals were available (e.g. from ~2 000 km to ~21 km in the case of Tag 11), the predictions were validated by the similarity between the spectral slopes of $<Chla^{Fluo}>$ and $<Chla^{LFM}>$ (Section 3.4.2). At spatial scales where $<Chla^{Fluo}>$ was not defined (e.g. smaller than ~21 km in the case of Tag 11), the variance spectra of $dark^{KL}$ and $<Chla^{LFM}>$ resulted in clearly distinct shapes (Figure III.8), hence ensuring that $<Chla^{LFM}>$ predictions did not result from pure observational noise.

The predicted signal therefore constitutes a useful signal with a coherent energy decrease across the observed spatial scales. From large-scale (~2 000 km) to (sub)mesoscale (O(10-100 km)), the energy decay of the spatial variance observing a power-law behavior in k^{-a} (for the definition of wave number k and spectral slope $-a$, see Section 3.4.2) is consistent with the expected behavior of a tracer such as [Chla] (Bracco et al., 2009; Lévy et al., 2018). Furthermore, spectral slopes become more negative as the wave number increases (i.e. the spectrum has a steeper decrease at smaller spatial scales), from $\sim k^{-1}$ at mesoscale to $\sim k^{-2}$ at submesoscale (see Section 3.4.2). These slopes are highly consistent with the expected decay slopes of a signal depicting phytoplankton distribution at (sub)mesoscale (Martin and Srokosz, 2002; Callies and Ferrari, 2013; Van Gennip et al., 2016).

The energy-drop threshold materializing the validity domain of the predictions (see Section 3.4.2) corresponded for each tag to twice the mean spatial frequency of the tags' light measurements (for detailed theoretical interpretation of the validity domain, see supplementary material, Text S7 and Figures III.16-III.17). The energy-drop threshold is hence directly dependent on the inherent properties of the corresponding transect. In the dataset of the 18 SES tags included in the present study, the mean distance between two valid consecutive Fluo profiles is 14.9 km +- 4.1, whereas for light profiles this distance is reduced to 5.9 km +- 3.1. As a result, while on average the SRDL Fluo measurements enable observation of phytoplankton dynamics at spatial scales up to ~30 km, LFM predictions extend the spatial scale of the observations up to ~12 km. In the present study, the gain enabled by the use of $<Chla^{LFM}>$ as a proxy for $<Chla^{Fluo}>$ is on average a factor 2.8 +- 0.9 towards finer observation scales.

Although light measurements are obviously restricted to the daytime period it is worth noting that in high-latitude environments with extended day lengths during the productive season, light measurements provided by SES tags might represent a unique tool to better address (sub)mesoscale coupling between physical forcing and biological response. However, independently of the bio-physical processes occurring along the transect of the tag, the spatial resolution of the light measurements performed by the tag directly depends on the horizontal speed of the SES (i.e. the distance between the consecutive dives of the SES, see previous paragraph), and also quite frequently, on the quality and validity of the measurements. For example, in the present study, the saturation issue (see Section 2.2.1 and supplementary material, Figure III.9) clearly lowers the spatial resolution achieved by the light measurements. Therefore, our recommendation regarding SRDLs is to implement a less sensitive light sensor to limit sensor saturation under high light levels (mainly occurring around noon, in the surface layer). Theoretically, if the light sensor is free from the saturation issue and all daylight profiles can hence be included in the LFM, the gain in terms of spatial resolution could potentially reach a factor 9, meaning that the method developed in the present study would enable observation of phytoplankton dynamics at a scale of ~3-4 km.

III.5. Conclusion and perspectives

The present study highlights the benefits of using the LFM both to homogenize the Fluo data from different sensors and to infer the Chla content in the water column. The interest of a model such as the LFM using K_L to describe the dynamics of Chla along the trajectory of an equipped SES stands out especially for a device with severe power consumption constraints such as the SRDL. The substantially low mean error associated with LFM predictions (see Section 3.2) emphasizes the accuracy of the LFM-based method for retrieving the variability of the <Chla> field and extending the spatio-temporal scale of observations (see Section 4.3.4). While the sole use of fluorescence measurements might not be sufficient to access (sub)mesoscale processes, the finer horizontal resolution achievable with LFM predictions unlocks the (sub)mesoscale observation gap of SRDLs.

Examples of SES foraging behavior being influenced by the environmental oceanographic conditions at (sub)mesoscale have already been described (Campagna et al., 2006; Della Penna et al., 2015; Siegelman et al., 2019). In parallel, while recent missions like SWOT aim at describing the ocean surface dynamics at an unprecedented resolution (15-30 km, see Morrow et al., 2019), the use of $<Chla^{LFM}>$ as a proxy for $<Chla^{Fluo}>$ enables the resolution of *in-situ* biological tracers to be aligned with the spatial scales targeted in such recent missions. Because primary production is largely driven by ephemeral physical processes occurring from the mesoscale (O(100 km)) to the submesoscale (O(10 km)), *in-situ* information at such scales is critical to describe phytoplankton dynamics (Mahadevan, 2016; McGillicuddy, 2016; Lévy et al., 2018). The improvements brought by the LFM in terms of spatial scales hence contribute key elements for deepening study of the coupling between phytoplankton distribution and the ocean's physical structure at (sub)mesoscale (including Lagrangian studies, Lehahn et al., 2018), and also provide novel data for studying SES behavior and the horizontal exploration of the ocean by such marine predators.

In this way, the dynamics of phytoplankton along the trajectories of SESs are optimally described by merging satellite-calibrated $[Chla^{Fluo}]$ data derived from the tags' fluorescence measurements (covering both day and night periods) with $[Chla^{LFM}]$ estimates (only available during daylight hours), which improves the spatio-temporal resolution of the data. In the present study, the data from tag deployment campaigns performed between 2018 and 2020 were included. The predictive capabilities of the LFM can possibly be extended towards a larger range of tags, e.g. tags deployed in the past measuring light but not Fluo. Indeed, bio-logging data have proved to be a considerable source of *in-situ* data at (sub)mesoscale in the SO in the past two decades. Every light profile, despite not providing truly reliable metrics on the vertical, is useful, under the method developed in the present study, to feed biogeochemical models with an estimate of the vertically-integrated Chla amount and the proportion of the Chla amount present within the mixed layer. Datasets comprising hundreds of thousands of *in-situ* vertical profiles sampled by equipped animals in the SO hence constitute a possible insight into the ocean subsurface to extend the quasi-synoptic - but surface-only - vision provided by satellite data. These numerous profiles are potentially highly valuable data for developing our knowledge about [Chla] variability at different spatio-temporal scales in the under-sampled SO, from short-lived processes to decadal variability.

Acknowledgments

This work was partly funded by the CNES and REFINE (European Research Council, Grant agreement 834177) project. L.L.S. is supported by a joint CNES-CNRS doctoral grant. The elephant seal work was supported as part of the SNO-MEMO and by the CNES-TOSCA project Elephant seals as Oceanographic Samplers of submesoscale features led by C. Guinet with support of the French Polar Institute (programmes 109 and 1201). This research was carried out, in part, at the Centre d'Études Biologiques de Chizé, under a contract with the Centre National d'Études Spatiales (CNES) and at the Laboratoire d'Océanographie de Villefranche.

Data availability

The marine mammal data were collected as part of the Système National d'Observation Mammifères Echantillonneurs du Milieu Océanique (SNO-MEMO) and made freely available by the International MEOP Consortium and the national programs that contribute to it. The MEOP data are available at www.meop.net/database/meop-databases. The computer codes produced for the present study are available at https://doi.org/10.17882/93268. The GlobColor products were produced and distributed by the Copernicus Marine Service with support from CNES and ACRI-ST, and are available at http://marine.copernicus.eu/services-portfolio/access-to-products/.

Author contributions

Conceived and designed the study, L.L.S, H.C and C.G.; B.P. helped with the processing of the MEOP data. L.L.S.and D.N. conceived the functional model and analyzed the model performance. L.L.S., H.C., F.d.O. and C.G. wrote the manuscript. H.C., F.d.O. and C.G. helped with analyzing the results. L.L.S., H.C., F.d.O., D.N., B.P. and C.G. reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Supplementary Material for Chapter III

Supplementary Material for

Improved accuracy and spatial resolution for bio-logging-derived chlorophyll a fluorescence measurements in the Southern Ocean

Loïc Le Ster[1,2*], Hervé Claustre[1], Francesco d'Ovidio[3], David Nerini[4], Baptiste Picard[2], Christophe Guinet[2]

[1] Sorbonne Université, CNRS, Laboratoire d'Océanographie de Villefranche, LOV, 06230 Villefranche-sur-Mer, France
[2] Centre d'Études Biologiques de Chizé, CNRS, 79360Villiers-en-Bois, France
[3] Sorbonne Université, Laboratoire d'Océanographie et du Climat, Institut Pierre Simon Laplace (LOCEAN, SU/CNRS/IRD/MNHN), 75252 Paris Cedex 05, France
[4] Aix-Marseille Université, CNRS/INSU, Université de Toulon, IRD, Mediterranean Institute of Oceanography, UM 110, 13288 Marseille, France

*** Correspondence:** loic.lester@imev-mer.fr

Contents of this file

Introduction

This supporting information:
- Presents the metadata of the tags included in the present study (Table S1);
- Gives the calibration coefficients (unitless) from LFM-based tag inter-calibration for all the tags included in the present study (Table S2);
- Provides graphical support for the processing of the light (Text S1, Figure S1) and Fluo (Text S2, Figure S2) data;
- Provides graphical support for the examination of the consistency of the NPQ correction applied in the present study (Text S3, Figure S3);
- Provides detailed information about the adjustment of the LFM parameters (Text S4), namely the number of basis functions (Figure S4) and the smoothing parameter (Figure S5);
- Describes the bagging procedure performed to test the robustness of the LFM against modifications of the statistical sample used to construct the LFM (Text S5, Figure S6);
- Presents the distribution of the LFM prediction error according to the solar angle (Text S6, Figure S7);
- Provides extended interpretation of the variance spectra through the assessment of the upper limit of the observable wave number domain enabled by the predicted variable $<Chla^{LFM}>$(Text S7, Figures S8-S9).

Table III.2: Summary of tag metadata

Tag Number	SMRU Tag Reference	Date Start	Date End	Min Latitude	Max Latitude	Min Longitude	Max Longitude	Number of Profiles
1	ct152-059-19	29-Oct-2019	28-Dec-2019	51ºS	47ºS	63ºE	72ºE	3 012
2	ct152-060-19	26-Oct-2019	12-Nov-2019	52ºS	49ºS	70ºE	80ºE	1 266
3	ct152-873-BAT-19	25-Oct-2019	02-Jan-2020	58ºS	49ºS	70ºE	97ºE	4 365
4	ct152-876-BAT-19	29-Oct-2019	05-Jan-2020	55ºS	48ºS	70ºE	95ºE	5 113
5	ct152-943-19	23-Oct-2019	06-Jan-2020	54ºS	48ºS	70ºE	89ºE	5 452
6	ct159-873-BAT2-19	22-Oct-2020	03-Jan-2021	54ºS	45ºS	70ºE	90ºE	5 087
7	ct159-881-BAT2-19	23-Oct-2020	30-Dec-2020	57ºS	49ºS	70ºE	93ºE	4 447
8	ct159-943-BAT-19	21-Oct-2020	06-Jan-2021	54ºS	49ºS	70ºE	101ºE	4 642
9	ct159-F905-BAT2-19	23-Oct-2020	09-Jan-2021	62ºS	48ºS	70ºE	95ºE	5 780
10	ft22-686-18	21-Oct-2018	02-Jan-2019	50ºS	43ºS	55ºE	71ºE	4 501
11	ft22-873-18	19-Oct-2018	08-Jan-2019	56ºS	49ºS	55ºE	71ºE	6 028
12	ft22-874-18	21-Oct-2018	05-Jan-2019	57ºS	49ºS	35ºE	71ºE	6 430
13	ft22-875-18	21-Oct-2018	09-Jan-2019	57ºS	48ºS	70ºE	96ºE	4 999
14	ft22-876-18	20-Oct-2018	08-Jan-2019	51ºS	48ºS	70ºE	96ºE	6 720
15	ft22-878-18	22-Oct-2018	01-Jan-2019	53ºS	46ºS	70ºE	87ºE	4 834
16	ft22-879-18	21-Oct-2018	08-Jan-2019	55ºS	49ºS	70ºE	90ºE	5 554
17	ft22-881-18	25-Oct-2018	06-Jan-2019	53ºS	49ºS	70	97	5 473
18	ft22-882-18	21-Oct-2018	31-Dec-2018	54ºS	46ºS	51	71	5 494

Table III.3: Calibration coefficients (unitless) from LFM-based tag inter-calibration

Tag Number	**Median Γ_e^{LFM} coefficient (1 000 iterations)**	**Std**
1	1.00	0.09
2	0.29	0.19
3	1.05	0.25
4	1.01	0.20
5	0.90	0.11
6	1.24	0.09
7	1.02	0.12
8	0.94	0.05
9	0.88	0.07
10	0.81	0.06
11	1.15	0.07
12	1.25	0.09
13	1.38	0.13
14	1.39	0.13
15	0.65	0.05
16	1.18	0.13
17	1.55	0.74
18	0.79	0.09

Text S1. Processing of the light data

The following steps describe the processing of the raw light vertical profiles (for graphical support, see Figure S1).

(i) A correction was applied to the light profiles on a per-profile basis according to the method described in Organelli et al. (2016) to deal with deep values of measured radiometric profiles and eliminate sensor noise. The method consists in detecting, for each profile, the depth below which no more information is present in the signal, called “dark depth”. Data below the detected “dark depth” may contain only pure noise and were discarded from the analysis. The median value of L below the dark depth is called “dark value” and characterizes the noise level of the sensor. The offset due to the noise present in the signal (i.e. dark value) was subtracted from the raw signal in order to systematically set to zero the light values corresponding to depths greater than the dark depth (see IOCCG Protocol Series, 2019). The dark value offset for a given sensor may change with time due to sensor drift and because environmental conditions can have an impact on the sensor response (mainly temperature). For the analysis to be robust against such changes, the dark values time series of dark values were computed on a daily basis for each tag by averaging the dark values of the 5 deepest profiles per day. It is important to note that the main variable derived from L in the present study (namely K_L) is not sensitive to the offset of the light signal. Absolute light values were only required for the correction of the Fluo profiles (see Section 2.2.3) and were picked up in relatively shallow portions of the profile where the dark value was very small in relation to the light signal itself (<<1%).

(ii) Vertical profiles of light were often subject to sensor saturation in the near-surface. Saturated portions of the light signal (i.e. where $\frac{dL(z)}{dz}=0$) were discarded.

(iii) To ensure physical and optical consistency of the light data in relation to classic radiometry sensing issues such as the passage of a cloud or a wave focusing at surface (see for example Mueller et al., 2003; Xing et al., 2011; Organelli et al., 2016; IOCCG Protocol Series, 2019), as well as animal head movements, a constrained fit was applied to the light profiles so that L monotonously decreases with depth. The fit was performed on the log-transformed vertical profile of light and defined as a piecewise cubic polynomial (see details in Section 2.3.1).

(iv) K_L vertical profiles were computed as the derivative of the fitted (log-transformed) light profiles. Vertical profiles of K_L were defined with the same vertical resolution as light profiles (i.e. 1 m).

$$K_L(z)=\frac{d}{dz}\left(\log\left(L(z)\right)\right)$$

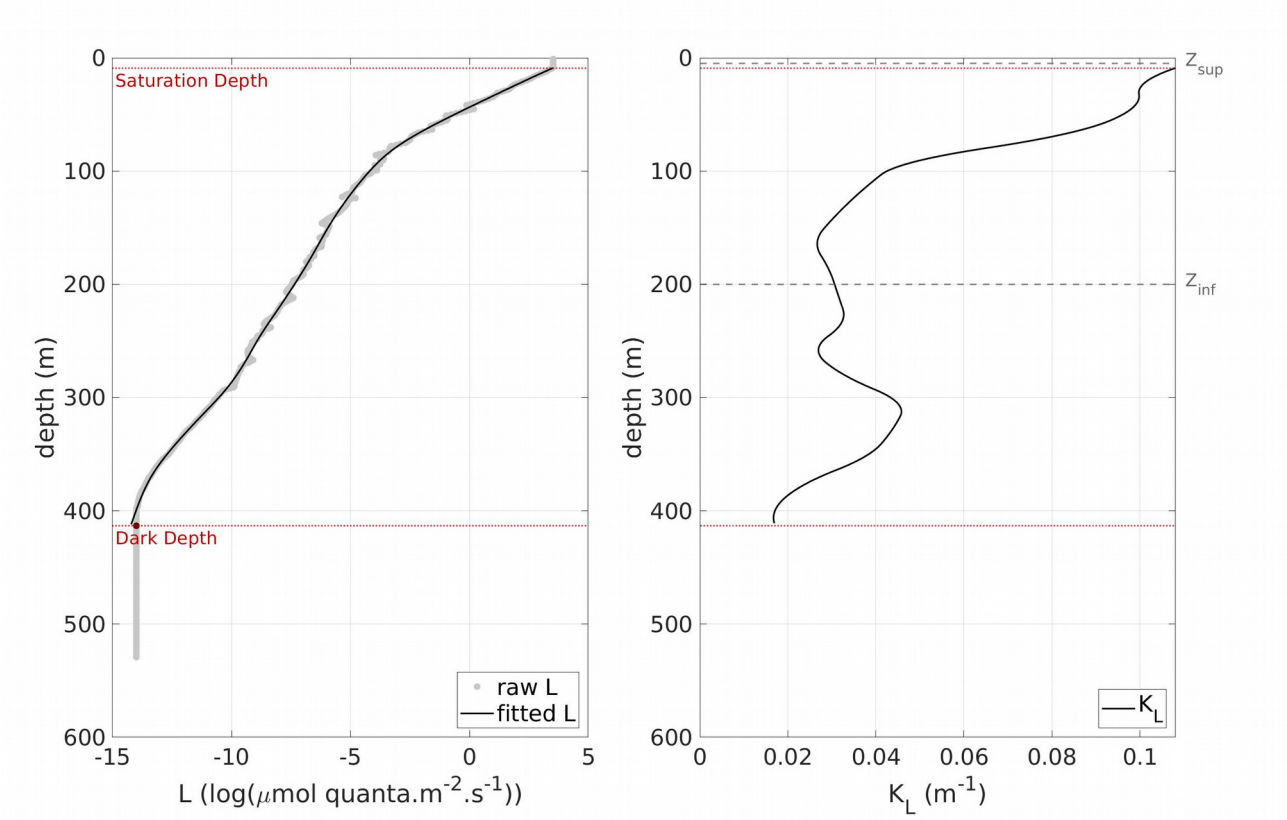


Figure III.9: Example of light (L) and corresponding K_L vertical profiles summarizing the processing of the irradiance data. (a) The raw light profile is represented with the gray dots, the black continuous line represents the fit applied to the raw L data, the upper and lower red dotted lines represent the Saturation Depth and the Dark Depth, respectively. The sections of the light profile located above the Saturation Depth or below the Dark Depth were discarded. (b) The corresponding K_L profile profile is represented with the continuous black line, the upper and lower gray dashed lines represent the shallower (Z_{sup}) and deeper (Z_{inf}) boundaries of the LFM prediction. Note that the profile of the present figure was only displayed to illustrate the processing of the light data but was not included in the LFM because the Saturation Depth (9 m) was deeper than Z_{sup} = 5 m.

Text S2. . Processing of the Fluo data

The raw Fluo vertical profiles were processed and quality controlled using the following procedure (for graphical support, see supplementary material, Figure S2):

(i) The raw Fluo was firstly corrected to remove instrumental noise, namely the fluorescence dark signal ($dark^{Fluo}$). A single value of $dark^{Fluo}$ was computed for every tag. The procedure described in Schmechtig et al. (2014) (originally designed for BGC-Argo floats) that deals with non-zero values of Fluo observed at depth was applied to SRDL Fluo data. However, because SRDLs do not sample Fluo as deep as BGC-Argo floats do, Schmechtig et al. (2014) was adapted to match SRDL data characteristics. The intensity of the fluorescence dark signal was calculated on a per-profile basis by computing the median value of Fluo between 190 m and 200 m. For a given tag, $dark^{Fluo}$ corresponds to the median value of all the successive fluorescence dark signal values computed in the tag's time series. The fluorescence dark signal was subsequently subtracted from the Fluo signal for the entire tag data set. Given the distribution of Z_{eu} and Z_{MLD}, computing $dark^{Fluo}$ between 190 m and 200 m was appropriate and suitable for the studied zone (on average, Z_{MLD} = 78 +- 37 m and Z_{eu} = 84 +- 19 m).

(ii) Vertical profiles of Fluo were corrected for the Non-Photochemical Quenching (NPQ) effect according to the method described in Xing et al. (2018). The method relies on the identification of a so-called "NPQ layer" from the surface to the shallowest value between Z_{MLD} and Z_{NPQ}. Z_{NPQ} corresponds to the depth above which the light signal is greater than a given threshold, fixed at 15 µmol quanta.m^{-2}.s^{-1}. For each profile, the maximum value of Fluo encountered within the "NPQ layer" was thus extrapolated up to the surface.

(iii) The dark- and NPQ-corrected Fluo profiles were fitted to smooth any spikes. The fit is similar to the fit applied to light profiles (i.e. a piecewise cubic polynomial, except that no monotonicity constraint was imposed on the Fluo fit).

(iv) Finally, the smoothed, dark- and NPQ-corrected Fluo data (hereafter denoted [F^{Fluo}]) were converted into [Chla]. [$Chla^{Fluo}$], the actual Chla concentration derived from [F^{Fluo}], was obtained by applying a calibration coefficient to the [F^{Fluo}] data. A calibration coefficient is by default provided by manufacturer Valeport, based on factory calibration tests performed on samples with known concentrations of [Chla]. However, in the present study, a specific calibration coefficient was computed for each tag, based on both K_L and the comparison of *in-situ* data with concomitant satellite-based [Chla] observations (see details in Section 2.3.2).

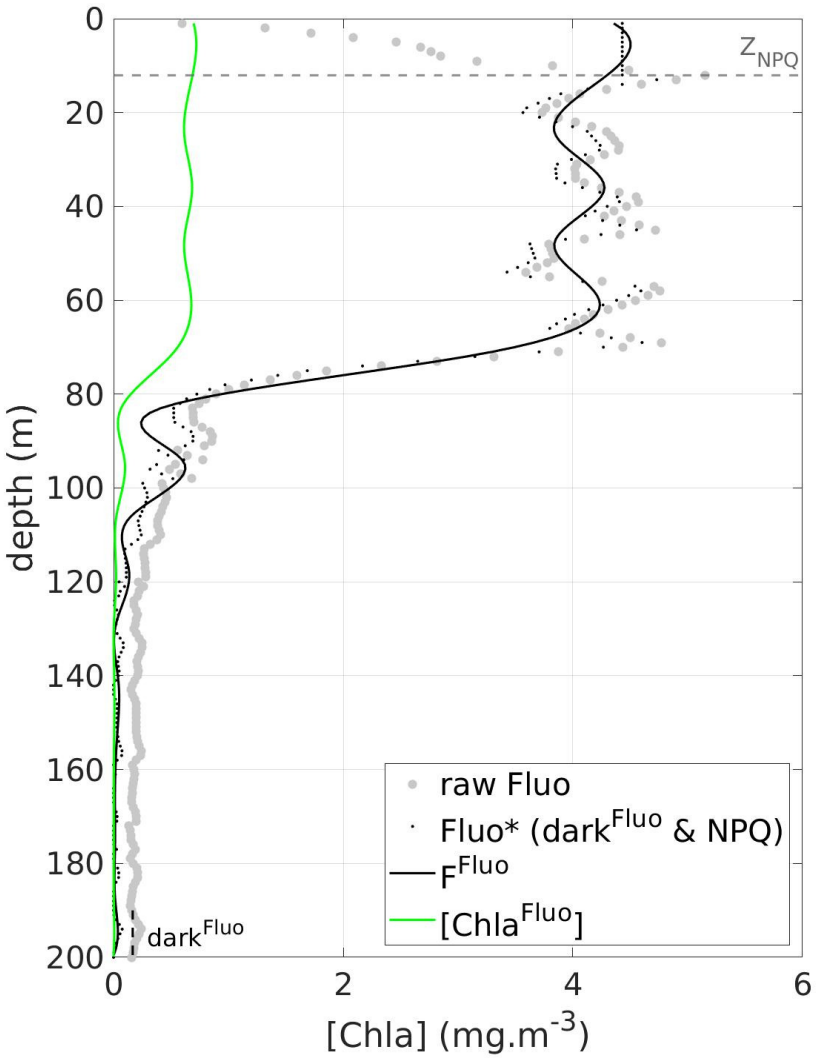


Figure III.10: Example of Fluo vertical profile summarizing the processing of the Fluo data. The gray dots represent the raw Fluo data (Fluo), the thin black dots represent the Fluo data corrected from the fluorescence dark signal (dark^{Fluo}) and from the NPQ (Non-Photochemical Quenching), denoted Fluo, the continuous black line represents F^{Fluo}, and finally, the continuous green line represents the [Chla^{Fluo}] data obtained after calibration. Note that the profile displayed in Figure S2 corresponds to the profile displayed in Figure S1 (i.e. was concomitantly sampled).*

Text S3. NPQ correction

Vertical profiles of Fluo were corrected for the Non-Photochemical Quenching (NPQ) effect according to the method described in Xing et al. (2018). The method relies on the identification of a so-called "NPQ layer", which takes into account Z_{MLD} and the depth above which the light signal is greater than a threshold fixed at 15 µmol quanta.m^{-2}.s^{-1} (see Text S2 and Figure S2). Xing et al. (2018) use the downwelling irradiance to determine the light threshold. Given that the light sensor of the SRDL does not strictly measure the downwelling irradiance but a diffused light level in the animal's environment, the threshold value of 15 µmol quanta.m^{-2}.s^{-1} used in the present study for the quenching correction needs to be verified. For this purpose, we compared the values of $<Chla^{Fluo}>$ between day and night profiles to assess the consistency of the NPQ correction. The distribution of the values differ very few between day and night, which reinforces the validity of the quenching correction applied to the Fluo profiles in the present study.

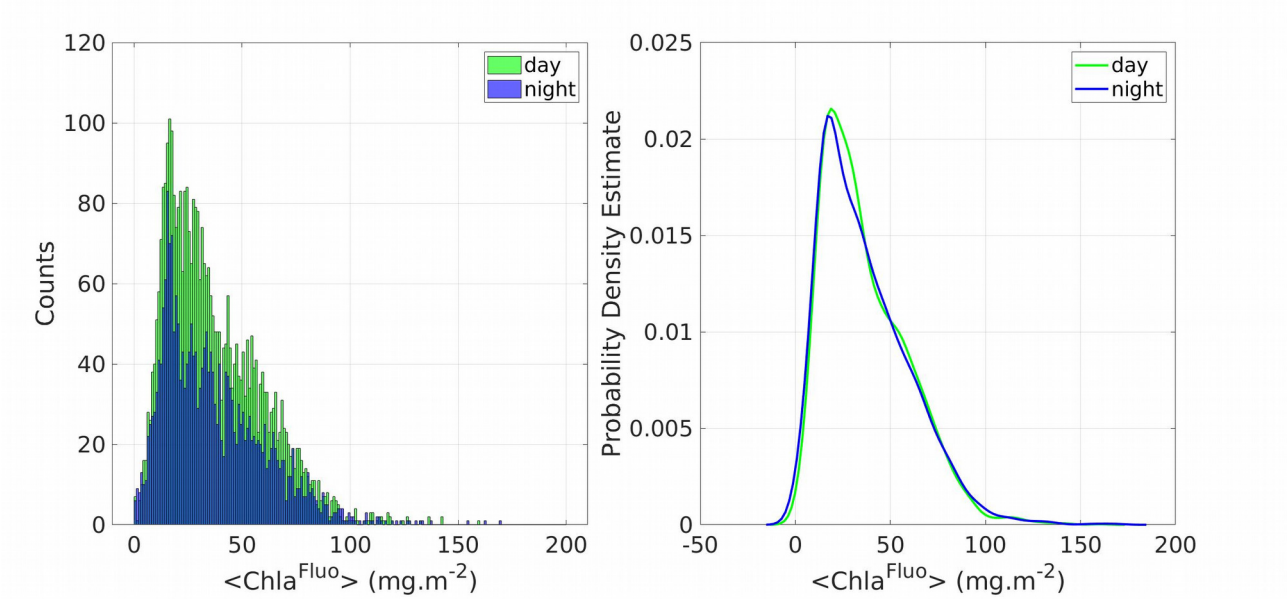


Figure III.11: Comparison of water-column integrated values of [$Chla^{Fluo}$] between day and night. The histogram of the values of $<Chla^{Fluo}>$ are shown in (a) and the estimates of the probability density function are shown in (b). In both panels, green color refers to day profiles and blue color refers to night profiles.

Text S4. Number of basis functions and Smoothing parameter

To develop the LFM, the handled variables (i.e. vertical profiles of K_L and $[Chla]^{Fluo}$) are converted into functional objects. In a functional space, vertical profiles are not classically defined as vectors expressed along a geo-referenced vertical axis (of *j* discrete depth values), but instead, they are expressed as a linear combination of *m* basis functions ($m < j$). The basis functions are the same for all profiles, meaning that the functional basis is unique for all the profiles. The *m* ($m = 30$ in our study) basis functions define the reference frame of the study. The *m* profiles' coefficients can be considered as their coordinates in that frame.

The basis functions in the present study are B-spline functions. B-spline functions are piecewise polynomials of degree *n* (in the present study, $n = 3$) with continuity and derivability constraints on some well-suited depths called knots. We chose a uniform distribution of the fit knots along the vertical. The regularity of the B-spline fit is ensured by a smoothing parameter (λ) that enables dampening high local variations in the vertical dimension.

The fit performed on the raw vertical profiles of L and Fluo (see Sections 2.2.1 and 2.2.3) therefore has two functions: (1) smoothing the data with, eventually, some constraints applied to the fitted data (e.g. monotonicity for L) and (2) converting the vertical profiles into functional objects, i.e. expression of each vertical profile as a linear combination of the *m* B-spline basis functions.

Once placed in a functional workspace, many classic statistical tools can be adapted to functional data (see Ramsay and Silverman, 1997). In the present study, the core operator of the LFM is obtained after computing a Functional Principal Component Analysis (FPCA) based on the *m* coefficients (i.e. coordinates in the functional space) of the predictor and the predicted variables.

From the FPCA, a linear functional relationship is derived between the functional objects associated with the vertical profiles of K_L (the derivative of the log-transformed light profile, also expressed as a functional object) and $[Chla]^{Fluo}$.

The (invertible) linear functional relationship on which the LFM relies is subsequently used as an inference tool to predict $[Chla]^{Fluo}$ from K_L.

Converting vertical profiles into functional objects using a B-spline basis is achieved when fitting the data with piecewise polynomials of degree *n* (in the present study, $n = 3$), with specific constraints imposed to the fit (Meyer 2012), such as the smoothness of the curve or monotonicity.

The smoothing parameter (λ) enables to constrain the distance between raw data and the functional fit. Although the LFM models used in the present study were developed following precisely the same method as described in Bayle et al. 2015, the vertical profiles processed in the present study have a much higher vertical resolution (1 m resolution in the present study versus 10 m in Bayle et al. 2015). The number *m* of B-spline basis functions had consequently to be re-adjusted, as well as the smoothing parameter λ. The optimization of *m* and λ is addressed in this section.

The optimum number of basis functions was m = 30 and the optimum value for the smoothing parameter was λ = 0.03. The knots of the fit were uniformly distributed along the vertical.

The prediction error was defined as the difference between [Chla^{LFM}] and [Chla^{Fluo}]. Subsequently, the root-mean-square error of the LFM (RMSE^{LFM}) was defined based on the prediction performed on the independent validation sample (see Sections 2.4 and 2.5) by positively summing the prediction error at every point (i.e. every depth of every profile of the validation sample) and then dividing by the number of profiles included in the prediction exercise. RMSE^{LFM} (expressed in mg.m^{-2}) is positive and characterizes the distance between the prediction [Chla^{LFM}] and the targeted [Chla^{Fluo}]. RMSE^{LFM} is explicitly defined as follows:

$$RMSE^{LFM} = \frac{1}{n_{val}} \sum_{i=1}^{n_{test}} \int_{z=Z_{inf}}^{z=0} \sqrt{\left([Chla^{LFM}]^{(i)}(z) - [Chla^{Fluo}]^{(i)}(z)\right)^2}\, dz \qquad \text{(III.7)}$$

Where $V^{(i)}$ is the i^{th} profile of variable V in the validation sample, n_{val} is the number of validation profiles (n_{val} = 417) and z is the depth of the measurement (Z_{inf} < z < 0 m).

Similarly, we define RMSE^{Fluo}, the root-mean-square error of the fit applied to the Fluo data (see Section 2.2.3) by positively summing the difference between the raw Fluo* data (the dark- and NPQ-corrected Fluo data, see Figure S2) and F^{Fluo}, the functional fit of Fluo*, at every point (i.e. every depth of every profile of the validation sample) and then dividing by the number of profiles included in the prediction exercise. RMSE^{Fluo} (expressed in mg.m^{-2}) is positive and characterizes the distance between the smoothed Fluo* data and the raw Fluo data. RMSE^{Fluo} is explicitly defined as follows:

$$RMSE^{Fluo} = \frac{1}{n_{test}} \sum_{i=1}^{n_{test}} \int_{z=Z_{inf}}^{z=0} \sqrt{\left(FChla^{Fluo(i)}(z) - Fluo*^{(i)}(z)\right)^2}\, dz \qquad \text{(III.8)}$$

The number m of basis functions was adjusted by testing equally spaced values of m ranging from 5 to 50 and computing RMSE^{LFM} for each resulting model, based on the prediction performed on the independent validation sample. The effect of the number m of basis functions on RMSE^{LFM} is displayed in Figure S4. The lowest RMSE^{LFM} value corresponds to m = 5. However the associated value of RMSE^{Fluo} is at its highest. This means that when the number of basis functions is reduced (m = 5-6), although [Chla^{LFM}] appears to be quite close to [Chla^{Fluo}], the Fluo fit (F^{Fluo}) is coarse and far away from the Fluo* data. It is hence preferable to consider higher values of m so that the Fluo* is not too degraded when computing F^{Fluo}. For m > 12, RMSE^{LFM} decreases continuously as m increases, however it appears clearly that the gain becomes negligible for m > 30. One idea beneath the model is to reduce as much as possible the dimension of the problem, hence the number m of parameters necessary to express the handled objects, while minimizing the prediction error.

As a result, a functional space defined with m = 30 basis functions was found to be the best compromise to jointly minimize $RMSE^{Fluo}$, $RMSE^{LFM}$ and model complexity.

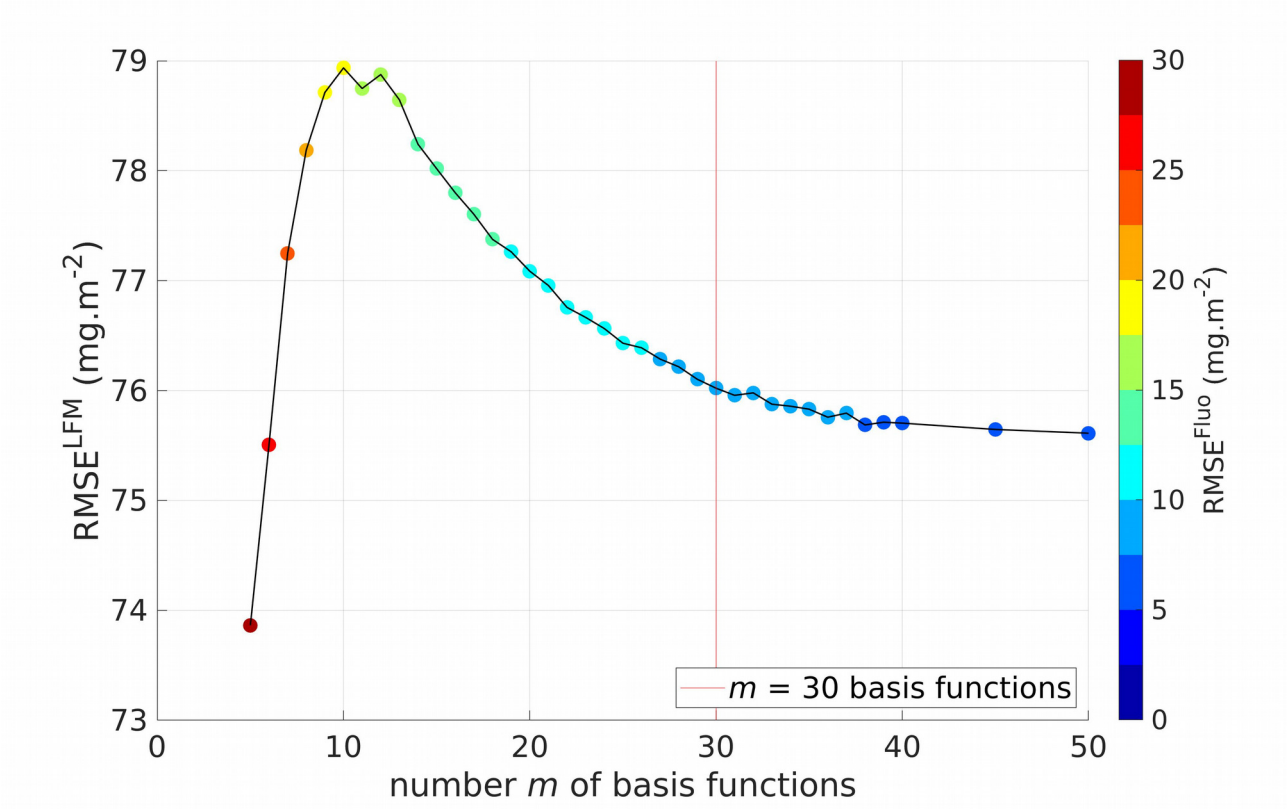


Figure III.12: Root-mean-square error of the LFM ($RMSE^{LFM}$) according to number m of basis functions. The continuous red vertical line represents the number of basis functions selected for the model (m = 30). The color of the points refer to the value of $RMSE^{Fluo}$.

The smoothing parameter λ was adjusted by testing values of λ equally spaced on a logarithmic scale from 10^{-4} to $10^{1.5}$ (i.e. approximately from 0.0001 to 30). The testing of λ was performed with a fixed number of basis functions m = 30, as previously adjusted. A high smoothing parameter means that the fit will be highly constrained and thus local variations of the raw data will be highly smoothed. While the RMSE seems to be quite constant for low values of λ, it increases rapidly for λ > 1, after reaching an optimum around 0.03 (Figure S5). For this reason, we chose λ = 0.03.

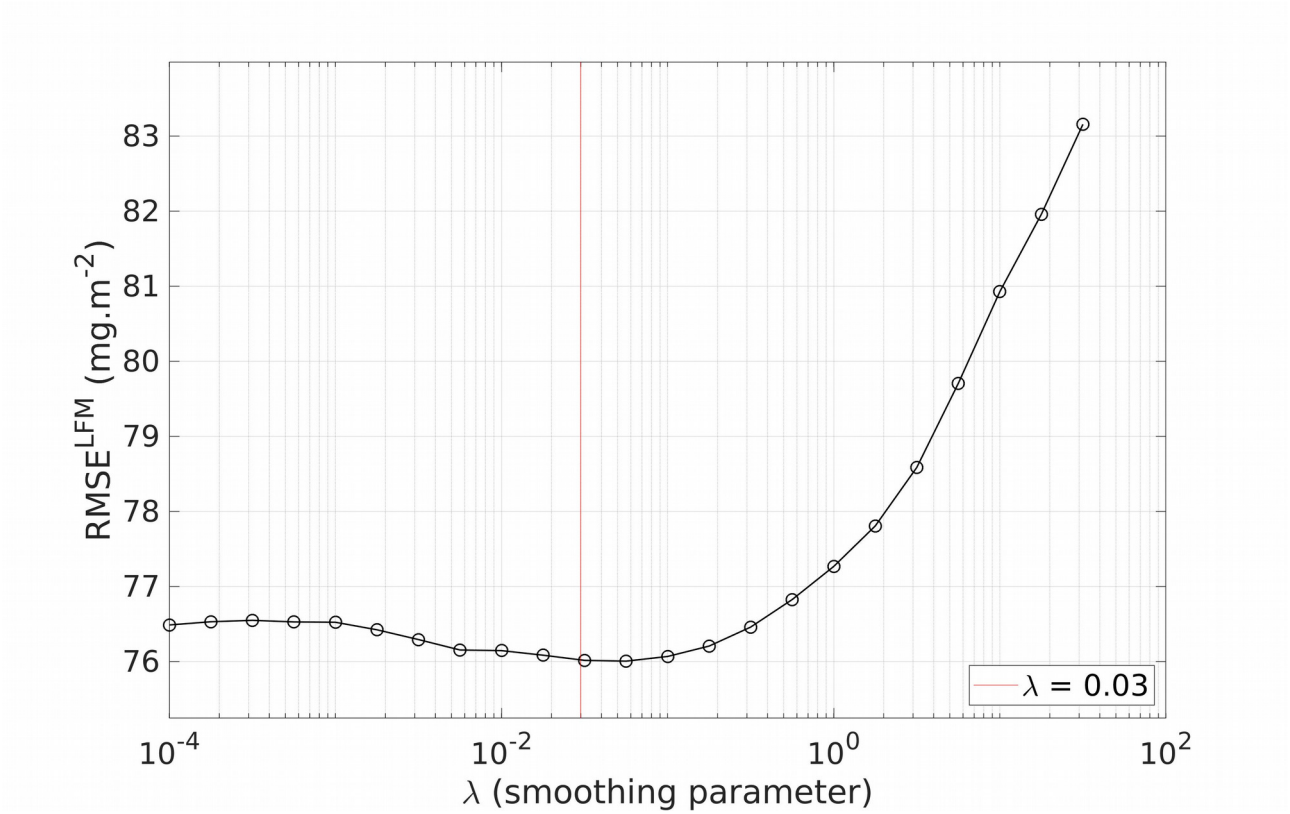


Figure III.13: Root-mean-square error of the LFM ($RMSE^{LFM}$) according to smoothing parameter λ. The continuous red vertical line represents the value selected for the model (λ = 0.03).

Text S5. Bagging procedure

To test the robustness of the model against modifications of the statistical sample used to construct the LFM, a bagging (bootstrap aggregating) validation procedure was carried out. In addition to testing the sensibility of the model to slight changes in the composition of the statistical sample of concomitant [$Chla^{Fluo}$] and K_L profiles , the bagging procedure, by averaging numerous versions of a given model, can potentially bring a substantial gain in accuracy to the bagged model (Breiman 1996, Nerini and Ghattas 2007). The bagging procedure consists in conducting a series of N iterations, where at each step a new statistical sample is created from the existing one by randomly sampling the set of available concomitant [$Chla^{Fluo}$] and K_L profiles (1 387 profiles, see Table 2). The modification of the statistical sample used to construct the LFM creates a new version of the LFM. The sampling is performed with replacement so that (i) the sample has always the same number of profiles but (ii) the set of selected profiles slightly differs from one iteration to the next. At step j ($1 \leqslant j \leqslant N$), the procedure is repeated j times so that j versions of the LFM are produced, each one corresponding to a different sample. A unique model LFM(j) is computed as the average of the j versions of the LFM created at step j. At each iteration of the bagging procedure, model LFM(j) was applied to the validation sample and $RMSE^{LFM}(j)$ was computed. The same dataset split as previously used was kept for the bagging (70% for the construction of the model / 30% for validation). N = 500 iterations of this procedure were carried out.

The dependency of the LFM relatively to the composition of the statistical sample used to construct the model (970 profiles) was examined through the $RMSE^{LFM}$ resulting from the bagging procedure. The values of $RMSE^{LFM}(j)$ computed at each iteration of the bagging procedure (N = 500 iterations in total) were reported in Figure S6. The convergence of the model in terms of prediction error primarily testifies to the robustness of the aggregated model against the presence of potential outlier profiles in the statistical sample (assuming that the sample is large enough). But mostly, the gain in accuracy enabled by the bagging procedure of less than 2% after 500 iterations points out the initial stability inherent to the LFM method developed in this study.

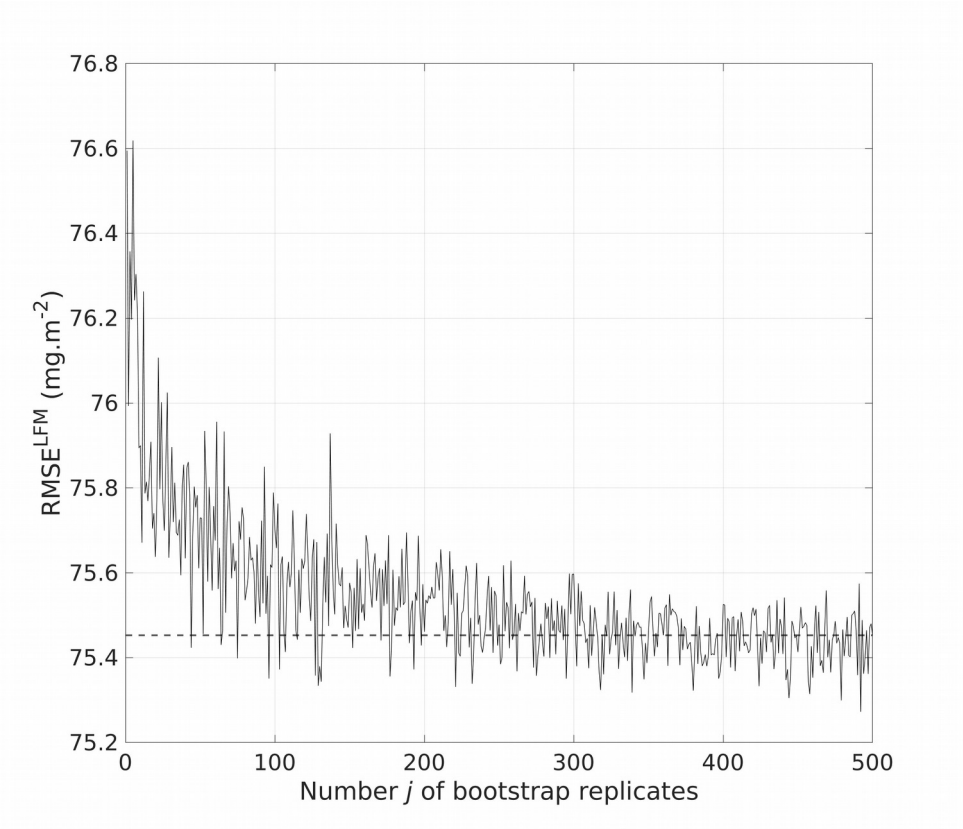


Figure III.14: $RMSE^{LFM}$ of the bagged predictions based on the bootstrap replicates of the statistical sample of 970 (70%) concomitant $[Chla^{Fluo}]$ and K_L profiles, with N = 500 iterations. The dashed line represents the convergence value of $RMSE^{LFM}$, computed as the median value of $RMSE^{LFM}(j)$ between the 300^{th} and 500^{th} iterations.

Text S6. Distribution of the prediction error according to the solar angle

The prediction error, computed as the difference (positive of negative) between $<Chla^{LFM}>$ and $<Chla^{Fluo}>$, was computed on the sample used to assess the LFM performance (validation sample, n_{val} = 417 profiles, see Section 2.3.3). The distribution of the prediction error was examined in Figure S7. The solar angle intervals were 0º to 15º, 15º to 30º, 30º to 45º, and higher than 45º. The prediction error presented a similar distribution for all positive values of the solar angle (note that no predictions were computed for negative solar angles, i.e. at night, see Section 2.2.1). The significance of the difference between the distributions of the prediction error for different intervals of solar angle was tested. The four distributions corresponding to the four solar angle intervals were not significantly different from one another (Mann-Whitney-Wilcoxon test, lowest p-value of 0.13).

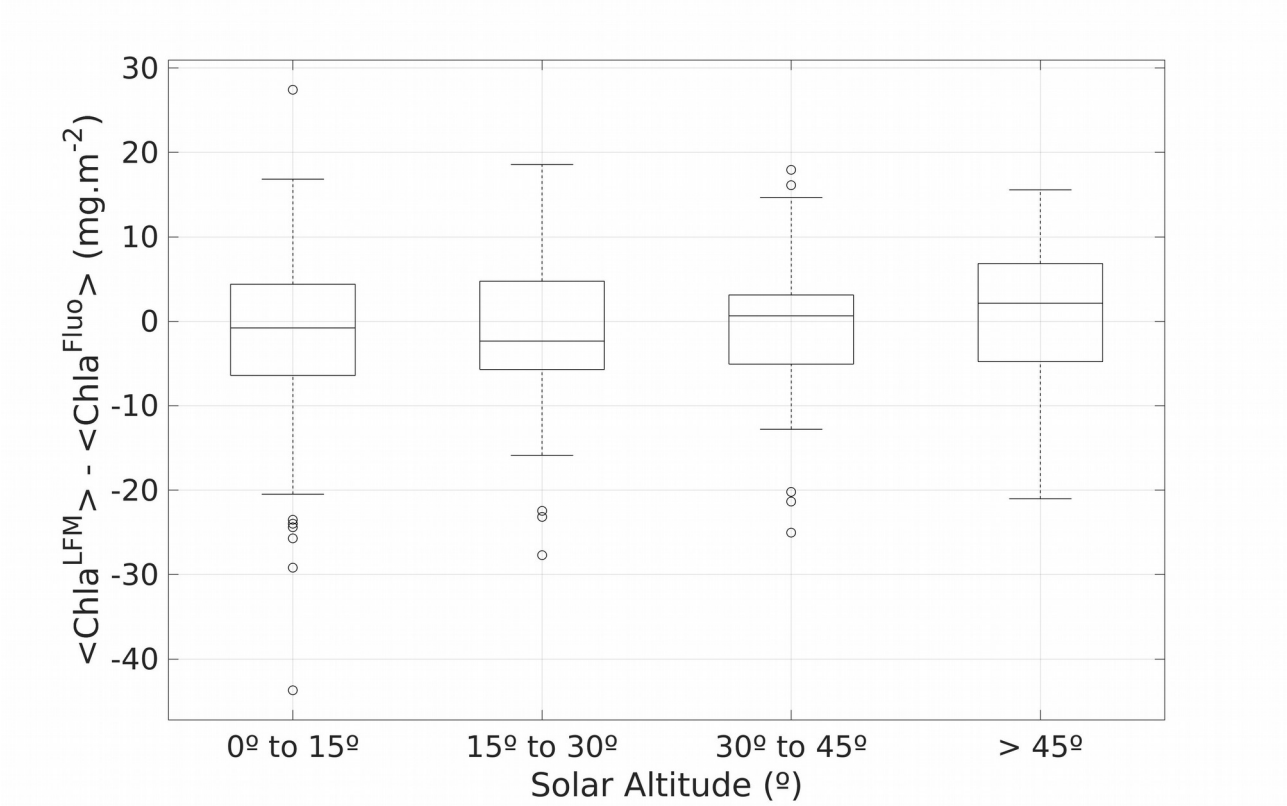


Figure III.15: Distribution of the prediction error for each (positive) solar angle intervals (0º to 15º, 15º to 30º, 30º to 45º and higher than 45º). For each box (i.e. each solar angle interval) the upper and lower limits of the box represent the 25th and 75th percentiles, respectively, the central line in the box represents the median value, the whiskers extend from the minimum to the minimum values of the data (excluding outliers) and the black circles represent the outliers.

Text S7. Assessing the upper limit of the observable wave number domain

In the present study, the variance spectra of $<Chla^{Fluo}>$ and $<Chla^{LFM}>$ were computed for each tag and compared. Prior to the calculation of the Fourier transform, a symmetric Hamming window was applied to the analyzed signals. A pronounced drop in the energy of the $<Chla^{LFM}>$ signal was observed for each tag at a different spatial scale (see Section 3.4.2), marking a threshold above which the spectrum was much noisier and lost the clear power-law behavior described for larger spatial scales (see Section 4.3.4). It is consequently reasonable to consider that the interpretations of the structures depicted by the $<Chla^{LFM}>$ signal are valid up to the scale of the energy-drop threshold and should be discarded for higher wave numbers.

The spatial scale corresponding to the energy-drop threshold ($f_{E\text{-}drop}$) was computed for each tag. For this purpose, the variance spectrum of $<Chla^{LFM}>$ was smoothed using B-splines functions. More precisely, the smoothed variance spectrum was computed independently for each tag and was defined as a piecewise cubic polynomial. The smoothed variance spectrum of $<Chla^{LFM}>$ is hence derivable and $f_{E\text{-}drop}$ corresponds to the minimum of its derivative (i.e. the wave number for which the variance spectrum is the steepest). An example with Tag 11 is shown in Figure 8. The consistency of the obtained value of $f_{E\text{-}drop}$ was visually checked for the 18 tags.

In order to characterize the spatial frequency of the light sampling of each tag, f_{mean} was defined as the inverse of the mean distance between two consecutive light profiles and was computed for each equipped SES.

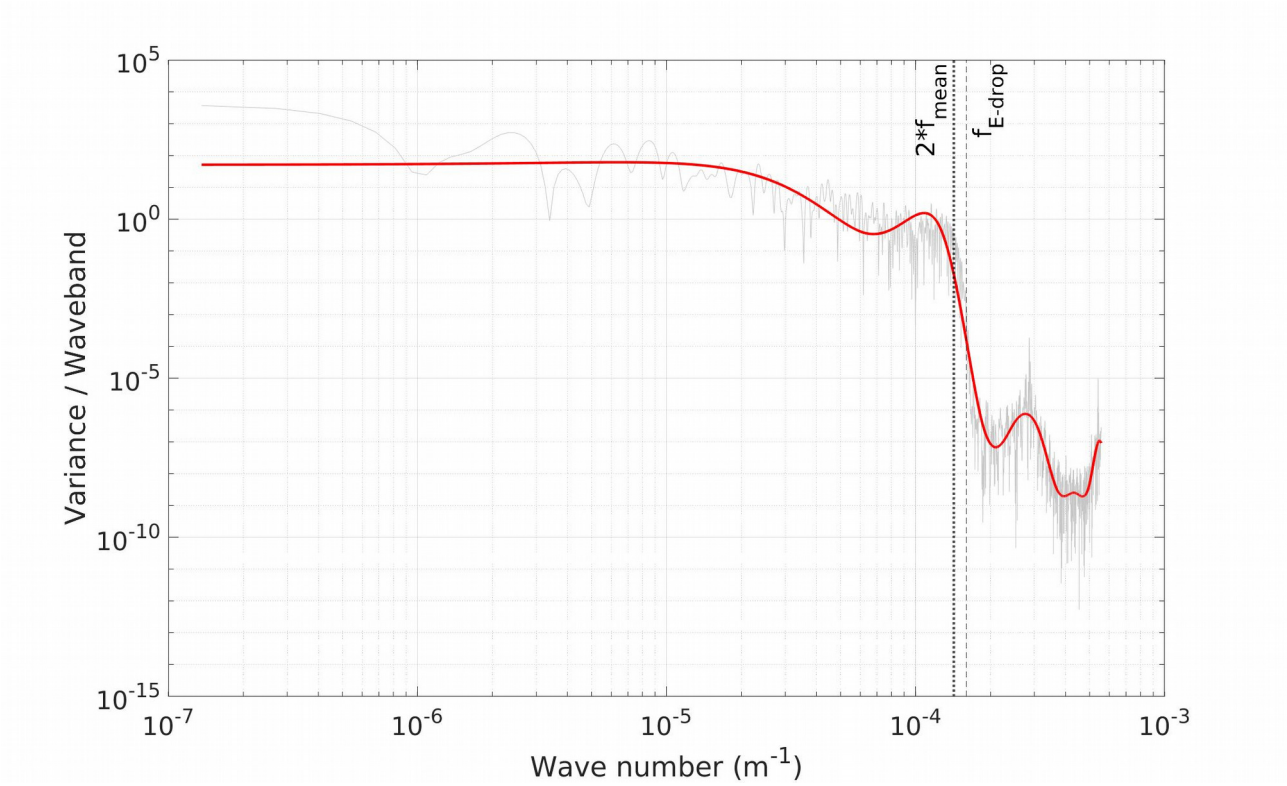


*Figure III.16: Computation of $f_{E\text{-}drop}$ and comparison with f_{mean} through the example of Tag 11. The continuous gray (red) line represents the (smoothed) variance spectrum of <Chla^{LFM}>. The black dashed vertical line represents $f_{E\text{-}drop}$ and the black dotted vertical line represents twice f_{mean}. In the present example, $f_{E\text{-}drop}$ = 1.6 10^{-4} m^{-1}, which corresponds to a spatial scale of ~6.2 km and 2*f_{mean} = 1.4 10^{-4} m^{-1}, which corresponds to a spatial scale of ~7.0 km.*

As exemplified in Figure S8 for Tag 11, $f_{E\text{-}drop}$ was subsequently compared to f_{mean} for all the tags (Figure S9). The spatial scale of the <Chla^{LFM}> signal energy drop appeared to be closely linked to the spatial frequency of the light sampling of each tag (i.e. the mean distance between two consecutive valid vertical profiles of light) and corresponds to twice the mean spatial frequency of the SRDL light measurements that were included in the model (see Figure S9, slope = 1.07, R^2 = 0.99). This result is to be linked with the Nyquist-Shannon theorem, that relates the sampling frequency of a signal with the maximum observable frequency in the sampled signal (Nyquist, 1928; Shannon, 1949; Jerri, 1977). The Nyquist-Shannon theorem states that for the reconstruction of a periodic signal with no loss, the sampling frequency must be at least twice the maximum frequency of the signal to be reconstructed. Here, the signal to be reconstructed is the <Chla> field along the trajectory of the animal and <Chla^{LFM}>, based on the vertical profiles of light sampled during the trip of the animal, is equivalent to a spatial subsampling of the <Chla> field. Interpreting reversely the Nyquist-Shannon theorem leads to the conclusion that it may be reasonable to give a much lower confidence level to the method when trying to detect structures in the <Chla^{LFM}> signal at a scale smaller than twice the mean spatial frequency of the SRDL light measurements (< ~7 km in the case of Tag 11).

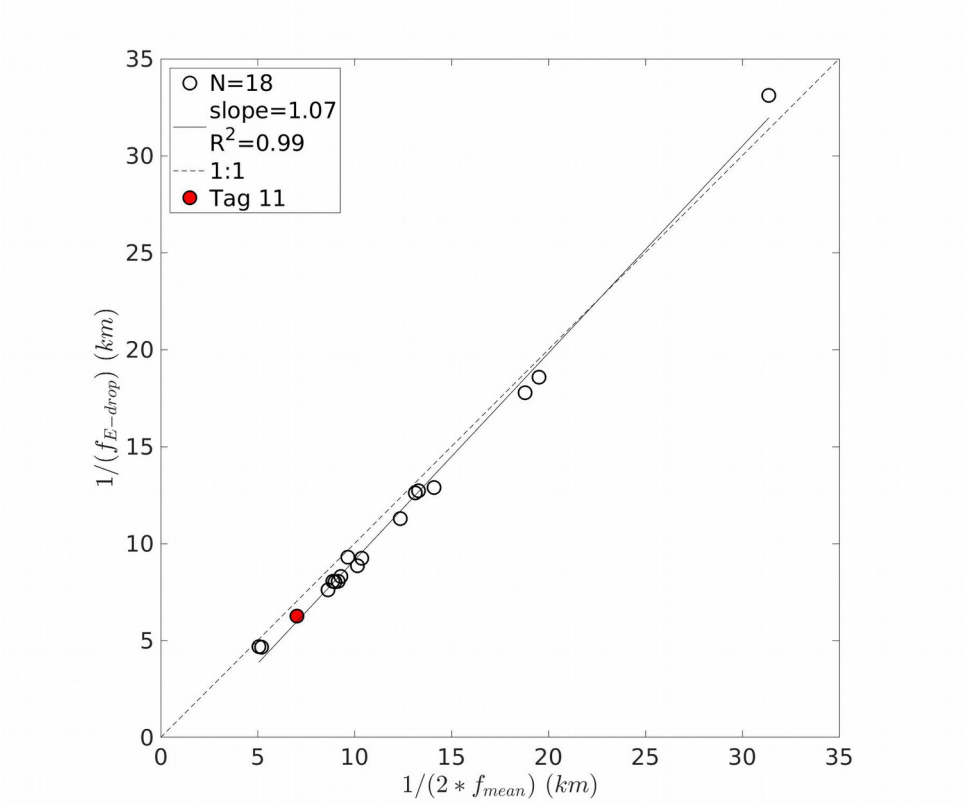


Figure III.17: Energy-drop threshold ($f_{E\text{-}drop}$) versus mean distance between two consecutive valid light profiles (f_{mean}). The black circles represent the data from the 18 tags included in the present study, the continuous black line materializes the linear regression of all the sample points (slope = 1.07; R^2 = 0.99; N = 18 points) and the black dashed line represents the 1:1 reference line. The red point represents the data from Tag 11, highlighted in Figure S8.

The threshold materializing the validity domain of the predictions directly depends on the inherent properties of the corresponding transect. Indeed, independently of the bio-physical processes occurring along the transect of the tag, the spatial scale of the energy drop is inherent to the spatial resolution of the measurements performed by the tag, which directly depends on the horizontal speed of the SES and also quite frequently, on the quality and validity of the measurements (e.g. the saturation issue clearly lowers the spatial resolution achieved by the light measurements). The mean distance between two valid consecutive Fluo profiles is 14.9 km +- 4.1 and this distance is reduced to 5.9 km +- 3.1 for light profiles. As a result, while on average the SRDL Fluo measurements enable observation of phytoplankton dynamics at spatial scales up to ~30 km, LFM predictions extend the spatial scale of the observations up to ~12 km. In the present study, the gain enabled by the use of $<Chla^{LFM}>$ as a proxy for $<Chla^{Fluo}>$ is on average a factor 2.8 +- 0.9 towards finer observation scales.

Chapter IV

General Conclusion

Contents of General Conclusion

IV.1. Reminder of the objectives and main results

The main objective of the present thesis was to examine the linkages between CHL dynamics and SO processes, as well as their influence on ecosystem functioning, through analysis of variations in chlorophyll-a fluorescence at suitable spatio-temporal scales. Throughout the chapters of the present thesis, we focused on CHL dynamics in the SO, from large scale and long-term changes, to small-scale, short-term variations. Firstly, we took advantage of two decades of ocean-colour data to examine at large spatial scale the time scales of variation in phytoplankton concentration in the SO (Chapter I). Then, through a regional study focused on the Indian sector of the SO, we assessed the propagation of variations in primary production throughout the ecosystem (Chapter II). Finally, we developed a method aimed at providing key data to understand how external forcing conditions phytoplankton growth at fine spatio-temporal scale (Chapter III). The objectives and main results of each chapter are detailed below.

The study presented in Chapter I aimed at disentangling the time scales of variation in satellite-based estimates of chlorophyll-a concentration (CHL) by decomposing CHL time series into a sum of components. Each component corresponded to a different time scale of variation, namely multi-annual, seasonal and sub-seasonal. Firstly, we investigated multi-annual variations in CHL in the SO. We demonstrated that the confidence level associated with estimates of long-term CHL biomass trends in the SO is low. The difficulty in obtaining long-term time series of CHL biomass estimates from space was probably highly linked to ocean-colour sensor inter-calibration issues. We then focused on smaller time scales of variation through examination of CHL phenology. To study CHL phenology, we developed a method focused on the seasonal component of CHL time series and based on normalized CHL. We described CHL dynamics in the SO at seasonal to sub-seasonal time scales. Our findings were validated with regional case studies from previous studies. Moreover, the joined regional case studies brought insights into the physical processes supporting CHL production in the different zones of the SO highlighted in our study. Finally, by studying the different components of CHL time series, we highlighted an overlap in their spectrum, resulting in the influence of one component on the others. Notably, we showed that besides seasonal cycles, sub-seasonal variations are determinant in shaping annual cycles of CHL, giving high importance to (sub)mesoscale processes.

In Chapter II, we focused on the Indian sector of the SO to examine linkages between primary production dynamics and ecosystem functioning. We took benefit of the phenology analysis developed in Chapter I to define coherent bioregions in relation to primary production. We proposed a bioregionalization of the Indian sector of the SO based on CHL biomass and CHL phenology to investigate at regional scale the influence of CHL dynamics on ecosystem functioning. We thus defined six bioregions in the study zone. Then, we studied ecosystem characteristics of bioregions and highlighted significant differences between those. Ecosystem-level characteristics were derived (i) from model data providing information on phytoplankton community composition and (ii) from

bio-logging data providing information on the foraging behaviour of a top predator in the SO. Firstly, we analyzed the dynamics of each bioregion in terms of phytoplankton community composition. We demonstrated that the distribution of phytoplankton size classes (PSC) were significantly different between bioregions. To summarize, the most productive regions were mainly associated with microphytoplankton, sub-tropical regions had proportionally more picophytoplankton, while nanophytoplankton was relatively ubiquitous. Then, we used indicators derived from tracking data of a top predator - the southern elephant seal (*Mirounga leonina*) - to infer potential influence of variations in primary production on higher trophic levels. We showed that elephant seal foraging behaviour was significantly different between bioregions, with productive regions being associated with less accessible preys but higher foraging success. Therefore, by merging ecosystem variables characterizing CHL biomass, CHL phenology, phytoplankton community composition and top predator behaviour, the analysis of Chapter II opens the way to ecosystem-level analysis of the SO.

In Chapter III, we proposed a tool for applications linked to the study of interactions between physical and biological processes at (sub)mesoscale. We developed a method based on vertical profiles of light obtained from bio-logging data to derive estimates of water-colum integrated CHL. The estimates of water-colum integrated CHL were calibrated with satellite CHL estimates. The developed method was presented as a cost-effective, energy-efficient way to obtain high-resolution *in-situ* CHL data in the SO. In addition, the proposed method suggests that systematic implementation of light sensors on autonomous platforms measuring chlorophyll-a fluorescence is a way to ensure sensor interoperability. We demonstrated the potential of the study with an application of the method on the transect of a southern elephant seal departed from the Kerguelen Archipelago. Application of the method enabled improvement of the horizontal resolution of CHL data from ~21 km to ~7 km . Observing both physics (e.g. sea surface height) and biology (e.g. phytoplankton concentration) the SO at fine spatial (submesoscale) and temporal (sub-seasonal) scales is necessary to identify the processes supporting primary production in the euphotic zone. Therefore, by enabling examination CHL dynamics at submesoscale, the proposed method contributes to filling the observation gap in the SO and highlights the promising potential of data derived from bio-logging devices placed on marine mammals to study polar environments.

IV.2. General discussion

IV.2.a. Context

The key (and threatened) function of the Southern Ocean in the climate system

A vast, dynamic region of the world's oceans, the SO is central to climate studies. The SO is an important sink of CO_2 (Mikaloff Fletcher et al., 2006; Gruber et al., 2019a), driven by complex spatio-temporal patterns of the air-sea CO_2 flux (Landschützer et al., 2015, 2016). Within the fraction of anthropogenic emissions sequestered by the world's oceans - estimated ~25% (Khatiwala et al., 2009) - an important part (~40%) is attributed to the SO (Khatiwala et al., 2009; Gruber et al., 2019b; DeVries, 2022; Friedlingstein et al., 2022). Furthermore, the world's oceans are a major contributor of the Earth's energy budget since ~90% of the excess of heat is estimated to be absorbed by the ocean (Church et al., 2011; Levitus et al., 2012; Trenberth et al., 2014; IPCC, 2013; von Schuckmann et al., 2016). Furthermore, the SO circulation scheme is a key element in the Earth climate system, notably because it connects the main ocean basins of the world (Talley et al., 2011; Marshall & Speer 2012; Talley, 2013; Rintoul, 2018).

In the context of climate change, SO circulation is affected by both warming (Toggweiler & Russell, 2008; Shi et al., 2021), and modified wind regime (Böning et al., 2008; Beech et al., 2020; Behrens & Bostock, 2023). These changes result in modified air-sea CO_2 fluxes (Séférian et al., 2012; Bushinsky et al., 2019) and affect biological activity (Lovenduski & Gruber, 2005; Chapman et al., 2020). Moreover, the SO is subject to increasing direct pressure from human activity, including increased marine traffic, intensification of fishing effort and development of tourism (Chown et al., 2012; Morley et al., 2020; Grant et al., 2021), which strongly threatens SO ecosystems (Pinkerton et al., 2021; Volzke et al. 2021; Johnston et al., 2022).

A crucial need to monitor phytoplankton in the SO

Therefore, monitoring phytoplankton dynamics in the SO is crucial because primary production in the SO mitigates climate change and supports SO ecosystems. Firstly, phytoplankton dynamics play a key role in SO processes that mitigate global warming, notably through the essential function primary producers play in the biological carbon pump (BCP). The BCP is the process by which organic matter is transported from the surface ocean to the ocean interior (Siegel et al., 2023). By photosynthesis, phytoplankton organisms transform inorganic carbon into organic carbon. Phytoplankton growth predominantly occurs in the well-lit surface ocean. Mainly induced by gravity, ocean mixing or vertical migration of animals, a series of processes are responsible for the transport of organic carbon from the surface to depth, where it is sequestered away from the

atmosphere on decadal to millennial timescales (Boyd et al., 2019a). Through the biological carbon pump, the biogenic carbon pool initiated by phytoplankton growth was identified as a major contributor to the oceanic carbon sink in the SO (Huang et al., 2023). In a lesser extent, phytoplankton organisms were also shown to have an effect on the ocean heat balance (Sathyendranath et al., 1991; Asselot et al., 2022). Secondly, besides its role in mitigating global warming, phytoplankton is the base of SO marine food webs (Boyd et al., 2019b). Formation of organic matter by autotroph organisms is the initial trigger for the transfer of energy to higher trophic levels. At every level (from phytoplankton to zooplankton, fish and SO charismatic megafauna), the response of SO ecosystems to climate change highly depends on future SO primary production (Deppeler and Davidson, 2017; Henley et al., 2020). As such, proxies of phytoplankton abundance, distribution and diversity are hence considered essential ocean variables (EOV) in the frame of ocean biodiversity and ecosystem monitoring (Constable et al., 2016; Miloslavich et al., 2018).

To respond to future needs in ocean management, we must address challenges related to the monitoring of phytoplankton according to three main pillars: observe, explain, model. These challenges are described in the following section.

IV.2.b. Limitations of the study and perspectives

Observing: unavoidable biases

Significant efforts are made to monitor SO primary production dynamics through *in-situ* water sampling (e.g. Nardelli et al., 2023), using proxies like chlorophyll-a fluorescence (e.g. Chai et al., 2020) or analyzing ocean-colour data (e.g. Henson et al., 2017). These efforts are needed to assess the response of SO ecosystems to global change. However, although the techniques and proxies used to characterize primary production are constantly developed and improved (e.g. Claustre et al., 2020), objective interpretation of observed variations in phytoplankton are compromised by fundamental issues related to available observations.

Remote-sensing-based observations of phytoplankton use natural light (Clarke et al., 1970). Phytoplankton dynamics as seen from space are hence necessarily biased because only sensed during day time. This implies an observational bias both on daily (day versus night) and seasonal (towards summer) timescales. Moreover, ocean-colour data is only reliably obtained under clear-sky conditions, which, by undersampling other (harsher) weather conditions, sums an additional bias to satellite observations. Another bias of satellite observations is that they only reflect surface dynamics (Gordon and McCluney, 1975) and do not render the vertical structure of phytoplankton. However, variations in surface chlorophyll do not necessarily accurately reflect its vertically-integrated biomass (Sallée et al., 2008; Uitz et al., 2009; Baldry et al., 2020). Additionally, as discussed in Chapter I, long-term satellite time series are complex to interpret due to ocean-colour

sensor inter-calibration issues (Pope et al., 2017; Henson et al., 2017). Finally, despite many efforts in developing regionally-tuned algorithms (Szeto et al., 2011; Johnson et al., 2013; Robinson et al., 2021), it remains essential to collect more *in-situ* samples to adequately tune SO-specific satellite products (Pope et al., 2017).

Pairing remote-sensing observations with *in-situ* time series is therefore necessary, but the latter also raise their share of issues. While fixed SO moorings like the Southern Ocean Time Series Observatory moored south of Tasmania (SOTS, imos.org.au/facilities/deepwatermoorings/sots) or the KERFIX mooring located southwest of the Kerguelen Islands (see Jeandel et al., 1998), offer the possibility to acquire high-resolution data for time periods generally exceeding seasonal timescales, they lack geographical coverage and the representativeness of their measurements must be discussed. Oppositely, autonomous free-drifting platforms like BGC-Argo floats (Chai et al., 2020) can capture phytoplankton dynamics on larger spatial scales and for several years, but their measurements do mix time and space variability. While remotely-controlled gliders enable to partially keep control on the location of the measurements, they are rapidly subject to limitations in energy (Testor et al., 2019). Power consumption usually constrains duration of glider developments from a few days to several months, depending on the sampling strategy (e.g. du Plessis et al., 2019). A potential way to overcome the constraint related to power limitation for sensor displacement is to use animal-borne sensors (Harcourt et al., 2019, McMahon et al., 2021). However, due to the reduced size of the devices, bio-logging deployments remain strongly limited by energy consumption, with notable impact on the number of measured variables and on the sampling frequency. Finally, sampling onboard research vessels remain the best way to obtain reliable and accurate *in-situ* estimates of phytoplankton abundance and diversity. However, the seasonal bias previously mentioned is also present in oceanographic campaign data due to logistical constraints and meteorologic conditions in the SO, especially harsh in winter.

Explaining: the difficulties in identifying underlying processes

Identifying the processes that support phytoplankton growth is key to assess the response of SO primary production to climate change. However, direct attribution of a given process to an increase or a decrease in phytoplankton biomass, or a change in its composition, is complex because our observations - on top of being biased (see previous section) - inherently contain various timescales of variation. As discussed in Chapter I, the growth and fate of phytoplankton in the SO can respond to a large range of processes on sub-seasonal time scales. Phytoplankton growth in the SO can be triggered (Thomalla et al., 2015; Pellichero et al., 2020) or maintained (Carranza & Gille 2015; d'Ovidio et al. 2015) by short-term, submesoscale features. Similarly, the fate of phytoplankton was shown to be driven by diverse processes, including physical advection, downward transport and grazing, each one responding to different timescales (Moreau et al., 2020). Phytoplankton growth can also be enhanced by punctual events like atmospheric dust deposition (Duce & Tindale, 1991) or, conversely, limited by intense submesoscale activity (du Plessis et al. 2019). These sub-seasonal variations strongly influence phytoplankton phenology (Prend et al.,

2022), together with the natural seasonal cycles of environmental factors such as mixed layer depth (Swart et al., 2015). On seasonal timescales, the annual occurrence of deep winter mixing drives the amount of dissolved iron available to phytoplankton in the surface layer, and consequently, has also an important role in phytoplankton phenology (Mtshali et al., 2019). On sub-seasonal to inter-annual timescales, climate indices like Southern Annular Mode (SAM) influence phytoplankton growth (Lovenduski and Gruber 2005). On inter-annual (Vantrepotte and Mélin, 2011) to decadal (Gregg e al., 2003) timescales and above, climate change comes into play and must be taken into account to assess long-term changes in phytoplankton dynamics (Deppeler & Davidson, 2017; Pinkerton et al., 2021; Henson et al., 2021).

Besides the complexity that resides in disentangling the timescales of variation in phytoplankton concentration to link it with physical, biogeochemical or biological processes, fluorescence-based proxies are subject to biases, that are inherent to the measurement technique. These biases lead the observers to question their data: is our observation a true signal, is it an artefact due to our instrument, are we misinterpreting our instrument's signal? Autonomous platforms and satellites commonly use chlorophyll-a fluorescence to characterize phytoplankton abundance, diversity and distribution. However, fluorescence remains an imperfect proxy for phytoplankton concentration, notably because sensitive to variations in phytoplankton species composition and physiological state (Bricaud et al., 2004). Photoacclimation of phytoplankton cells can modify the fluorescence-to-chlorophyll ratio, both in low (Cornec et al., 2021) and high (Xing et al., 2018) light conditions. Physiological changes in phytoplankton due to photoacclimation can occur on timescales of seconds (Li et al., 2009). This implies that day and night, deep and shallow, fluorescence data won't be interpreted the same way (Doblin et al., 2011). Furthermore, the bias induced by variations in phytoplankton species composition has a notable impact on the study of long-term variations in phytoplankton concentration, notably for studies based on satellite estimates of chlorophyll-a concentration (Sun et al., 2023).

Remote-sensing estimates of phytoplankton concentration remain a powerful tool to study long-term changes at wide spatial scale. Nonetheless, emergence times of climate-driven changes are potentially larger than space-borne time series of ocean-colour data, especially in the SO (Henson et al., 2016, 2017). Time series of ~25 years as examined in the present study might be too short to distinguish climate change-driven trends from background natural variability (Henson et al., 2016, 2017). As a consequence, to enlarge our vision on the occurring changes and foster our comprehension of the Earth system, there is an evident necessity to develop suitable biogeochemical models, based on our (partial) understanding of the interactions between physical, biogeochemical and biological processes in the ocean.

Modeling: the fine-scale challenge

The use of models is a powerful tool to isolate the response of ecosystems to specific drivers (Asselot et al., 2022), study a region where observations are scarce (Taylor et al., 2013) or project into future climate scenarios (Boyd & Doney, 2002; Frölicher et al., 2015; Behrens & Bostock, 2023). In our study, we highlighted the importance of fine-scale, short-term processes. The resolution of the sampling has an influence on the observed variability, as demonstrated both for observations (Hales & Takahashi, 2004; Monteiro et al., 2015; Kaufman et al., 2018) and models (Lévy et al., 2014; Aumont et al., 2015). The impact of model resolution on our ability to assess or not the effect of a range of processes emphasizes the need to develop and better constrain high-resolution biogeochemical models capable of resolving fine-scale (submesoscale), short-term (< sub-seasonal) variability (Lévy et al., 2018).

However, the complexity of climate system mechanisms is far from simple to model. Our lack of understanding of the full range of processes that drive a key biogeochemical process such as the BCP (see Section IV.2.a) is revealed by the little agreement of climate models in assessing the sensitivity of SO BCP to climate change (Henson et al., 2022; Thomalla et al., 2023). In addition, uncertainties associated with projections of the Coupled Model Intercomparison Project Phase 6 (CMIP6) models in terms of primary production for the 21st century are large, with some models projecting positive anomalies relative to the reference period, while other projecting negative ones (Kwiatkowski et al., 2020). Therefore, divergence between climate models highlight the complexity of the climate system and our difficulty to accurately model its functioning.

Which ocean management for sustainable future?

As discussed in the previous paragraphs, SO equilibrium is maintained through a complex network of interactions between physical, biogeochemical and biological processes (Lévy et al., 2018; Boyd et al., 2019; Siegel et al., 2023). The processes affecting the state of the SO can even be extended to socio-economic, cultural and ethical questions (Murphy et al., 2021), but these aspects were out of the scope of the present manuscript. To resolve the interactions between SO processes, we need to observe and model the Earth system at suitable spatio-temporal scales, according to the question at hand. Understanding biogeochemical cycles and ecosystem dynamics in the SO is necessary to support sustainable ocean management policies (Murphy & Hofmann, 2012).

Future conservation of the Antarctic raises important challenges, emanating from increasing threats on SO ecosystems due to climate change. In addition, increasing stress due to human activity in the region - including marine resource harvesting, tourism and growing pressure for mineral and hydrocarbon resource - results in a crucial need for sustainable, internationally-approved, ocean management policies (Chown and Brooks, 2019). Set by the Antarctic Treaty System, the Antarctic regulatory framework promotes “peaceful purposes” and international cooperation in the Antarctic

(Berkman et al., 2012). However, these principles are threatened by the sometimes divergent interests of participating (or more problematically: non-participating) nations (Chown et al., 2012).

SO ecosystems are changing (Constable et al., 2014; Rogers et al., 2020). Mitigation of human-induced impacts on Antarctic ecosystems, and notably interaction with fisheries (Ropert-Coudert et al., 2018), require identification of areas of ecological significance (Hindell et al., 2020) to protect key regions of the SO under the status of marine protected areas (e.g. Douglass et al., 2014), as defined by the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR; ccamlr.org). Identification of these areas of ecological significance (e.g. using animal tracking data, see Hindell et al., 2003; Ropert-Coudert et al., 2018; Hindell et al., 2020) require continuous effort of the scientific community in such an extreme environment as the SO. However, long-term time series are critical for the monitoring and assessment of the impact of environmental change on SO species (e.g. Jenouvrier et al., 2018). It is therefore essential to extend the spatial and temporal resolution of biotic and abiotic variables measured in the SO through continued sampling effort to ensure sustainable management of human activities in the SO. More generally, management of human activities in relation to their environment appears as a necessary - yet apparently not obvious to all parties - condition for sustainable human presence on Earth.

References

Abbott, M. R., Richman, J. G., Letelier, R. M., and Bartlett, J. S. (2000). The spring bloom in the Antarctic Polar Frontal Zone as observed from a mesoscale array of bio-optical sensors. *Deep Sea Research Part II: Topical Studies in Oceanography* 47, 3285–3314. doi: 10.1016/S0967-0645(00)00069-2.

Allison, L. C., Johnson, H. L., Marshall, D. P., and Munday, D. R. (2010). Where do winds drive the Antarctic Circumpolar Current? *Geophys. Res. Lett.* 37, n/a-n/a. doi: 10.1029/2010GL043355.

Alvain, S., Moulin, C., Dandonneau, Y., and Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: A satellite view. *Global Biogeochem. Cycles* 22, n/a-n/a. doi: 10.1029/2007GB003154.

Ardyna, M., Claustre, H., Sallée, J.-B., D'Ovidio, F., Gentili, B., van Dijken, G., et al. (2017). Delineating environmental control of phytoplankton biomass and phenology in the Southern Ocean: Phytoplankton Dynamics in the SO. *Geophys. Res. Lett.* 44, 5016–5024. doi: 10.1002/2016GL072428.

Ardyna, M., Lacour, L., Sergi, S., d'Ovidio, F., Sallée, J.-B., Rembauville, M., et al. (2019). Hydrothermal vents trigger massive phytoplankton blooms in the Southern Ocean. *Nat Commun* 10, 2451. doi: 10.1038/s41467-019-09973-6.

Arrigo, K. R., van Dijken, G. L., and Bushinsky, S. (2008). Primary production in the Southern Ocean, 1997–2006. *J. Geophys. Res.* 113, C08004. doi: 10.1029/2007JC004551.

Arteaga, L. A., Boss, E., Behrenfeld, M. J., Westberry, T. K., and Sarmiento, J. L. (2020). Seasonal modulation of phytoplankton biomass in the Southern Ocean. *Nat Commun* 11, 5364. doi: 10.1038/s41467-020-19157-2.

Arteaga, L., Pahlow, M., and Oschlies, A. (2014). Global patterns of phytoplankton nutrient and light colimitation inferred from an optimality-based model. *Global Biogeochem. Cycles* 28, 648–661. doi: 10.1002/2013GB004668.

Asselot, R., Lunkeit, F., Holden, P. B., and Hense, I. (2022). Climate pathways behind phytoplankton-induced atmospheric warming. *Biogeosciences* 19, 223–239. doi: 10.5194/bg-19-223-2022.

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., and Gehlen, M. (2015). PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geosci. Model Dev.* 8, 2465–2513. doi: 10.5194/gmd-8-2465-2015.

Bachman, S. D., Taylor, J. R., Adams, K. A., and Hosegood, P. J. (2017). Mesoscale and Submesoscale Effects on Mixed Layer Depth in the Southern Ocean. *Journal of Physical Oceanography* 47, 2173–2188. doi: 10.1175/JPO-D-17-0034.1.

Bailey, S. W., and Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment* 102, 12–23. doi: 10.1016/j.rse.2006.01.015.

Baldry, K., Strutton, P. G., Hill, N. A., and Boyd, P. W. (2020). Subsurface Chlorophyll-a Maxima in the Southern Ocean. *Front. Mar. Sci.* 7, 671. doi: 10.3389/fmars.2020.00671.

Banse, K., and English, D. (1997). Near-surface phytoplankton pigment from the Coastal Zone Color Scanner in the Subantarctic region southeast of New Zealand. *Mar. Ecol. Prog. Ser.* 156, 51–66. doi: 10.3354/meps156051.

Barbraud, C., and Weimerskirch, H. (2006). Antarctic birds breed later in response to climate change. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6248–6251. doi: 10.1073/pnas.0510397103.

Barreau, E., Ropert-Coudert, Y., Delord, K., Barbraud, C., and Kato-Ropert, A. (2019). Scale matters: sea ice and breeding success of Adélie penguins. *Polar Biol* 42, 1405–1410. doi: 10.1007/s00300-019-02531-2.

Baudena, A., Ser-Giacomi, E., D'Onofrio, D., Capet, X., Cotté, C., Cherel, Y., et al. (2021). Fine-scale structures as spots of increased fish concentration in the open ocean. *Sci Rep* 11, 15805. doi: 10.1038/s41598-021-94368-1.

Bayle, S., Monestiez, P., Guinet, C., and Nerini, D. (2015). Moving toward finer scales in oceanography: Predictive linear functional model of Chlorophyll a profile from light data. *Progress in Oceanography* 134, 221–231. doi: 10.1016/j.pocean.2015.02.001.

Beaulieu, C., Henson, S. A., Sarmiento, J. L., Dunne, J. P., Doney, S. C., Rykaczewski, R. R., et al. (2013). Factors challenging our ability to detect long-term trends in ocean chlorophyll. *Biogeosciences* 10, 2711–2724. doi: 10.5194/bg-10-2711-2013.

Beech, N., Rackow, T., Semmler, T., Danilov, S., Wang, Q., and Jung, T. (2022). Long-term evolution of ocean eddy activity in a warming world. *Nat. Clim. Chang.* 12, 910–917. doi: 10.1038/s41558-022-01478-3.

Behrens, E., and Bostock, H. (2023). The Response of the Subtropical Front to Changes in the Southern Hemisphere Westerly Winds—Evidence From Models and Observations. *JGR Oceans* 128. doi: 10.1029/2022JC019139.

Bellacicco, M., Cornec, M., Organelli, E., Brewin, R. J. W., Neukermans, G., Volpe, G., et al. (2019). Global Variability of Optical Backscattering by Non algal particles From a Biogeochemical Argo Data Set. *Geophys. Res. Lett.* 46, 9767–9776. doi: 10.1029/2019GL084078.

Bender, N. A., Crosbie, K., and Lynch, H. J. (2016). Patterns of tourism in the Antarctic Peninsula region: a 20-year analysis. *Antarctic Science* 28, 194–203. doi: 10.1017/S0954102016000031.

Berkman, P. A., Lang, M. A., and Walton, D. W. H. (2012). Science diplomacy: Antarctica, science and governance of international spaces. *Antarctic Science* 24, 428–429. doi: 10.1017/S095410201200051X.

Blain, S., Quéguiner, B., Armand, L., Belviso, S., Bombled, B., Bopp, L., et al. (2007). Effect of natural iron fertilization on carbon sequestration in the Southern Ocean. *Nature* 446, 1070–1074. doi: 10.1038/nature05700.

Blain, S., Quéguiner, B., and Trull, T. (2008). The natural iron fertilization experiment KEOPS (KErguelen Ocean and Plateau compared Study): An overview. *Deep Sea Research Part II: Topical Studies in Oceanography* 55, 559–565. doi: 10.1016/j.dsr2.2008.01.002.

Blain, S., Renaut, S., Xing, X., Claustre, H., and Guinet, C. (2013). Instrumented elephant seals reveal the seasonality in chlorophyll and light-mixing regime in the iron-fertilized Southern Ocean. *Geophys. Res. Lett.* 40, 6368–6372. doi: 10.1002/2013GL058065.

Blain, S., Tréguer, P., Belviso, S., Bucciarelli, E., Denis, M., Desabre, S., et al. (2001). A biogeochemical study of the island mass effect in the context of the iron hypothesis: Kerguelen Islands, Southern Ocean. *Deep Sea Research Part I: Oceanographic Research Papers* 48, 163–187. doi: 10.1016/S0967-0637(00)00047-9.

Boehme, L., Lovell, P., Biuw, M., Roquet, F., Nicholson, J., Thorpe, S. E., et al. (2009). Technical Note: Animal-borne CTD-Satellite Relay Data Loggers for real-time oceanographic data collection. *Ocean Sci.* 5, 685–695. doi: 10.5194/os-5-685-2009.

Bograd, S., Block, B., Costa, D., and Godley, B. (2010). Biologging technologies: new tools for conservation. Introduction. *Endang. Species. Res.* 10, 1–7. doi: 10.3354/esr00269.

Böning, C. W., Dispert, A., Visbeck, M., Rintoul, S. R., and Schwarzkopf, F. U. (2008). The response of the Antarctic Circumpolar Current to recent climate change. *Nature Geosci* 1, 864–869. doi: 10.1038/ngeo362.

Bourgeois, T., Goris, N., Schwinger, J., and Tjiputra, J. F. (2022). Stratification constrains future heat and carbon uptake in the Southern Ocean between 30°S and 55°S. *Nat Commun* 13, 340. doi: 10.1038/s41467-022-27979-5.

Boyd, P., Kawaguchi, S., Strzepek, R., Meiners, K., Hallegraeff, G., McMinn, A., et al. (2019a). *Report Card: Potential Tipping Points for life in the Southern Ocean*. Antarctic Climate & Ecosystems Cooperative Research Centre (ACE CRC) doi: 10.25959/100.00031768.

Boyd, P., LaRoche, J., Gall, M., Frew, R., and McKay, R. M. L. (1999). Role of iron, light, and silicate in controlling algal biomass in subantarctic waters SE of New Zealand. *J. Geophys. Res.* 104, 13395–13408. doi: 10.1029/1999JC900009.

Boyd, P. W., Claustre, H., Levy, M., Siegel, D. A., and Weber, T. (2019b). Multi-faceted particle pumps drive carbon sequestration in the ocean. *Nature* 568, 327–335. doi: 10.1038/s41586-019-1098-2.

Boyd, P. W., Crossley, A. C., DiTullio, G. R., Griffiths, F. B., Hutchins, D. A., Queguiner, B., et al. (2001). Control of phytoplankton growth by iron supply and irradiance in the subantarctic Southern Ocean: Experimental results from the SAZ Project. *J. Geophys. Res.* 106, 31573–31583. doi: 10.1029/2000JC000348.

Boyd, P. W., and Doney, S. C. (2002). Modelling regional responses by marine pelagic ecosystems to global climate change. *Geophys. Res. Lett.* 29, 53-1-53–4. doi: 10.1029/2001GL014130.

Boyd, P. W., and Ellwood, M. J. (2010). The biogeochemical cycle of iron in the ocean. *Nature Geosci* 3, 675–682. doi: 10.1038/ngeo964.

Boyd, P. W., Ellwood, M. J., Tagliabue, A., and Twining, B. S. (2017). Biotic and abiotic retention, recycling and remineralization of metals in the ocean. *Nature Geosci* 10, 167–173. doi: 10.1038/ngeo2876.

Boyd, P. W., and Law, C. S. (2001). The Southern Ocean Iron RElease Experiment (SOIREE)—introduction and summary. *Deep Sea Research Part II: Topical Studies in Oceanography* 48, 2425–2438. doi: 10.1016/S0967-0645(01)00002-9.

Boyd, P. W., Strzepek, R., Chiswell, S., Chang, H., DeBruyn, J. M., Ellwood, M., et al. (2012). Microbial control of diatom bloom dynamics in the open ocean. *Geophys. Res. Lett.* 39. doi: 10.1029/2012GL053448.

Bracco, A., Clayton, S., and Pasquero, C. (2009). Horizontal advection, diffusion, and plankton spectra at the sea surface. *J. Geophys. Res.* 114, C02001. doi: 10.1029/2007JC004671.

Bradford Grieve, J. M., Chang, F. H., Gall, M., Pickmere, S., and Richards, F. (1997). Size fractionated phytoplankton standing stocks and primary production during austral winter and spring 1993 in the Subtropical Convergence region near New Zealand. *New Zealand Journal of Marine and Freshwater Research* 31, 201–224. doi: 10.1080/00288330.1997.9516759.

Breiman, L. (1996). Bagging predictors. *Mach Learn* 24, 123–140. doi: 10.1007/BF00058655.

Brewin, R. J. W., Ciavatta, S., Sathyendranath, S., Jackson, T., Tilstone, G., Curran, K., et al. (2017). Uncertainty in Ocean-Color Estimates of Chlorophyll for Phytoplankton Groups. *Front. Mar. Sci.* 4. doi: 10.3389/fmars.2017.00104.

Bricaud, A., Claustre, H., Ras, J., and Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations. *J. Geophys. Res.* 109, C11010. doi: 10.1029/2004JC002419.

Bricaud, A., Morel, A., Babin, M., Allali, K., and Claustre, H. (1998). Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models. *J. Geophys. Res.* 103, 31033–31044. doi: 10.1029/98JC02712.

Brody, S. R., Lozier, M. S., and Dunne, J. P. (2013). A comparison of methods to determine phytoplankton bloom initiation. *J. Geophys. Res. Oceans* 118, 2345–2357. doi: 10.1002/jgrc.20167.

Bushinsky, S. M., Landschützer, P., Rödenbeck, C., Gray, A. R., Baker, D., Mazloff, M. R., et al. (2019). Reassessing Southern Ocean Air Sea CO_2 Flux Estimates With the Addition of Biogeochemical Float Observations. *Global Biogeochemical Cycles* 33, 1370–1388. doi: 10.1029/2019GB006176.

Butler, W. L. (1978). Energy Distribution in the Photochemical Apparatus of Photosynthesis. *Annu. Rev. Plant. Physiol.* 29, 345–378. doi: 10.1146/annurev.pp.29.060178.002021.

Callies, J., and Ferrari, R. (2013). Interpreting Energy and Tracer Spectra of Upper-Ocean Turbulence in the Submesoscale Range (1–200 km). *Journal of Physical Oceanography* 43, 2456–2474. doi: 10.1175/JPO-D-13-063.1.

Campagna, C., Piola, A. R., Rosa Marin, M., Lewis, M., and Fernández, T. (2006). Southern elephant seal trajectories, fronts and eddies in the Brazil/Malvinas Confluence. *Deep Sea Research Part I: Oceanographic Research Papers* 53, 1907–1924. doi: 10.1016/j.dsr.2006.08.015.

Carranza, M. M., and Gille, S. T. (2015). Southern Ocean wind-driven entrainment enhances satellite chlorophyll-a through the summer. *J. Geophys. Res. Oceans* 120, 304–323. doi: 10.1002/2014JC010203.

Carranza, M. M., Gille, S. T., Franks, P. J. S., Johnson, K. S., Pinkel, R., and Girton, J. B. (2018). When Mixed Layers Are Not Mixed. Storm Driven Mixing and Bio optical Vertical Gradients in Mixed Layers of the Southern Ocean. *JGR Oceans* 123, 7264–7289. doi: 10.1029/2018JC014416.

Cavanagh, R. D., Melbourne-Thomas, J., Grant, S. M., Barnes, D. K. A., Hughes, K. A., Halfter, S., et al. (2021). Future Risk for Southern Ocean Ecosystem Services Under Climate Change. *Front. Mar. Sci.* 7, 615214. doi: 10.3389/fmars.2020.615214.

Chai, F., Johnson, K. S., Claustre, H., Xing, X., Wang, Y., Boss, E., et al. (2020). Monitoring ocean biogeochemistry with autonomous platforms. *Nat Rev Earth Environ* 1, 315–326. doi: 10.1038/s43017-020-0053-y.

Chapman, C. C., Lea, M.-A., Meyer, A., Sallée, J.-B., and Hindell, M. (2020). Defining Southern Ocean fronts and their influence on biological and physical processes in a changing climate. *Nat. Clim. Chang.* 10, 209–219. doi: 10.1038/s41558-020-0705-4.

Cherel, Y., Ducatez, S., Fontaine, C., Richard, P., and Guinet, C. (2008). Stable isotopes reveal the trophic position and mesopelagic fish diet of female southern elephant seals breeding on the Kerguelen Islands. *Mar. Ecol. Prog. Ser.* 370, 239–247. doi: 10.3354/meps07673.

Cherel, Y., and Hobson, K. (2007). Geographical variation in carbon stable isotope signatures of marine predators: a tool to investigate their foraging areas in the Southern Ocean. *Mar. Ecol. Prog. Ser.* 329, 281–287. doi: 10.3354/meps329281.

Chisholm, S. W. and Morel, F. M. M. (1991). Conference: American Society of Limnology and Oceanography (ASLO) symposium on what controls phytoplankton production in nutrient-rich areas of the open sea, San Marcos, CA (United States), 22-24 Feb 1991. in.

Chiswell, S. M., Bradford-Grieve, J., Hadfield, M. G., and Kennan, S. C. (2013). Climatology of surface chlorophyll *a* , autumn-winter and spring blooms in the southwest Pacific Ocean. *J. Geophys. Res. Oceans* 118, 1003–1018. doi: 10.1002/jgrc.20088.

Chown, S. L., and Brooks, C. M. (2019). The State and Future of Antarctic Environments in a Global Context. *Annu. Rev. Environ. Resour.* 44, 1–30. doi: 10.1146/annurev-environ-101718-033236.

Chown, S. L., Clarke, A., Fraser, C. I., Cary, S. C., Moon, K. L., and McGeoch, M. A. (2015). The changing form of Antarctic biodiversity. *Nature* 522, 431–438. doi: 10.1038/nature14505.

Chown, S. L., Lee, J. E., Hughes, K. A., Barnes, J., Barrett, P. J., Bergstrom, D. M., et al. (2012). Challenges to the Future Conservation of the Antarctic. *Science* 337, 158–159. doi: 10.1126/science.1222821.

Church, J. A., White, N. J., Konikow, L. F., Domingues, C. M., Cogley, J. G., Rignot, E., et al. (2011). Revisiting the Earth's sea-level and energy budgets from 1961 to 2008. *Geophys. Res. Lett.* 38, n/a-n/a. doi: 10.1029/2011GL048794.

Clarke, G. L., Ewing, G. C., and Lorenzen, C. J. (1970). Spectra of Backscattered Light from the Sea Obtained from Aircraft as a Measure of Chlorophyll Concentration. *Science* 167, 1119–1121. doi: 10.1126/science.167.3921.1119.

Claustre, H., Johnson, K. S., and Takeshita, Y. (2020). Observing the Global Ocean with Biogeochemical-Argo. *Annu. Rev. Mar. Sci.* 12, 23–48. doi: 10.1146/annurev-marine-010419-010956.

Cole, H., Henson, S., Martin, A., and Yool, A. (2012). Mind the gap: The impact of missing data on the calculation of phytoplankton phenology metrics. *J. Geophys. Res.* 117, n/a-n/a. doi: 10.1029/2012JC008249.

Concha, J. A., Bracaglia, M., and Brando, V. E. (2021). Assessing the influence of different validation protocols on Ocean Colour match-up analyses. *Remote Sensing of Environment* 259, 112415. doi: 10.1016/j.rse.2021.112415.

Constable, A. J., Costa, D. P., Schofield, O., Newman, L., Urban, E. R., Fulton, E. A., et al. (2016). Developing priority variables ("ecosystem Essential Ocean Variables" — eEOVs) for observing dynamics and change in Southern Ocean ecosystems. *Journal of Marine Systems* 161, 26–41. doi: 10.1016/j.jmarsys.2016.05.003.

Constable, A. J., Melbourne-Thomas, J., Corney, S. P., Arrigo, K. R., Barbraud, C., Barnes, D. K. A., et al. (2014). Climate change and Southern Ocean ecosystems I: how changes in physical habitats directly affect marine biota. *Glob Change Biol* 20, 3004–3025. doi: 10.1111/gcb.12623.

Cornec, M., Claustre, H., Mignot, A., Guidi, L., Lacour, L., Poteau, A., et al. (2021). Deep Chlorophyll Maxima in the Global Ocean: Occurrences, Drivers and Characteristics. *Global Biogeochem Cycles* 35. doi: 10.1029/2020GB006759.

Cravatte, S., Serazin, G., Penduff, T., and Menkes, C. (2021). Imprint of chaotic ocean variability on transports in the southwestern Pacific at interannual timescales. *Ocean Sci.* 17, 487–507. doi: 10.5194/os-17-487-2021.

d'Ovidio, F., De Monte, S., Alvain, S., Dandonneau, Y., and Lévy, M. (2010). Fluid dynamical niches of phytoplankton types. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18366–18370. doi: 10.1073/pnas.1004620107.

d’Ovidio, F., Della Penna, A., Trull, T. W., Nencioli, F., Pujol, M.-I., Rio, M.-H., et al. (2015). The biogeochemical structuring role of horizontal stirring: Lagrangian perspectives on iron delivery downstream of the Kerguelen Plateau. *Biogeosciences* 12, 5567–5581. doi: 10.5194/bg-12-5567-2015.

d’Ovidio, F., Pascual, A., Wang, J., Doglioli, A. M., Jing, Z., Moreau, S., et al. (2019). Frontiers in Fine-Scale in situ Studies: Opportunities During the SWOT Fast Sampling Phase. *Front. Mar. Sci.* 6, 168. doi: 10.3389/fmars.2019.00168.

De Boer, A. M., Graham, R. M., Thomas, M. D., and Kohfeld, K. E. (2013). The control of the Southern Hemisphere Westerlies on the position of the Subtropical Front: The Subtropical Front and Wind Stress. *J. Geophys. Res. Oceans* 118, 5669–5675. doi: 10.1002/jgrc.20407.

de Boyer Montégut, C. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res.* 109, C12003. doi: 10.1029/2004JC002378.

Della Penna, A., De Monte, S., Kestenare, E., Guinet, C., and d’Ovidio, F. (2015). Quasi-planktonic behavior of foraging top marine predators. *Sci Rep* 5, 18063. doi: 10.1038/srep18063.

Demarcq, H., Reygondeau, G., Alvain, S., and Vantrepotte, V. (2012). Monitoring marine phytoplankton seasonality from space. *Remote Sensing of Environment* 117, 211–222. doi: 10.1016/j.rse.2011.09.019.

Deppeler, S. L., and Davidson, A. T. (2017). Southern Ocean Phytoplankton in a Changing Climate. *Front. Mar. Sci.* 4. doi: 10.3389/fmars.2017.00040.

DeVries, T. (2022). The Ocean Carbon Cycle. *Annu. Rev. Environ. Resour.* 47, 317–341. doi: 10.1146/annurev-environ-120920-111307.

DeVries, T., Le Quéré, C., Andrews, O., Berthet, S., Hauck, J., Ilyina, T., et al. (2019). Decadal trends in the ocean carbon sink. *Proc. Natl. Acad. Sci. U.S.A.* 116, 11646–11651. doi: 10.1073/pnas.1900371116.

DeVries, T., Primeau, F., and Deutsch, C. (2012). The sequestration efficiency of the biological pump. *Geophys. Res. Lett.* 39, n/a-n/a. doi: 10.1029/2012GL051963.

Doblin, M. A., Petrou, K. L., Shelly, K., Westwood, K., Van Den Enden, R., Wright, S., et al. (2011). Diel variation of chlorophyll-a fluorescence, phytoplankton pigments and productivity in the Sub-Antarctic and Polar Front Zones south of Tasmania, Australia. *Deep Sea Research Part II: Topical Studies in Oceanography* 58, 2189–2199. doi: 10.1016/j.dsr2.2011.05.021.

D’Ortenzio, F., and Ribera d’Alcalà, M. (2009). On the trophic regimes of the Mediterranean Sea: a satellite analysis. *Biogeosciences* 6, 139–148. doi: 10.5194/bg-6-139-2009.

Douglass, L. L., Turner, J., Grantham, H. S., Kaiser, S., Constable, A., Nicoll, R., et al. (2014). A Hierarchical Classification of Benthic Biodiversity and Assessment of Protected Areas in the Southern Ocean. *PLoS ONE* 9, e100551. doi: 10.1371/journal.pone.0100551.

Dragon, A., Bar-Hen, A., Monestiez, P., and Guinet, C. (2012). Comparative analysis of methods for inferring successful foraging areas from Argos and GPS tracking data. *Mar. Ecol. Prog. Ser.* 452, 253–267. doi: 10.3354/meps09618.

Du Plessis, M., Swart, S., Ansorge, I. J., and Mahadevan, A. (2017). Submesoscale processes promote seasonal restratification in the Subantarctic Ocean. *J. Geophys. Res. Oceans* 122, 2960–2975. doi: 10.1002/2016JC012494.

Du Plessis, M., Swart, S., Ansorge, I. J., Mahadevan, A., and Thompson, A. F. (2019). Southern Ocean Seasonal Restratification Delayed by Submesoscale Wind–Front Interactions. *Journal of Physical Oceanography* 49, 1035–1053. doi: 10.1175/JPO-D-18-0136.1.

Duce, R. A., and Tindale, N. W. (1991). Atmospheric transport of iron and its deposition in the ocean. *Limnol. Oceanogr.* 36, 1715–1726. doi: 10.4319/lo.1991.36.8.1715.

Dunstan, P. K., Foster, S. D., King, E., Risbey, J., O'Kane, T. J., Monselesan, D., et al. (2018). Global patterns of change and variation in sea surface temperature and chlorophyll a. *Sci Rep* 8, 14624. doi: 10.1038/s41598-018-33057-y.

Eayrs, C., Li, X., Raphael, M. N., and Holland, D. M. (2021). Rapid decline in Antarctic sea ice in recent years hints at future change. *Nat. Geosci.* 14, 460–464. doi: 10.1038/s41561-021-00768-3.

Edwards, A. M., Platt, T., and Sathyendranath, S. (2004). The high-nutrient, low-chlorophyll regime of the ocean: limits on biomass and nitrate before and after iron enrichment. *Ecological Modelling* 171, 103–125. doi: 10.1016/j.ecolmodel.2003.06.001.

Edwards, M., and Richardson, A. J. (2004). Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430, 881–884. doi: 10.1038/nature02808.

El Hourany, R., Abboud Abi Saab, M., Faour, G., Aumont, O., Crépon, M., and Thiria, S. (2019). Estimation of Secondary Phytoplankton Pigments From Satellite Observations Using Self Organizing Maps (SOMs). *J. Geophys. Res. Oceans* 124, 1357–1378. doi: 10.1029/2018JC014450.

Fabri Ruiz, S., Danis, B., Navarro, N., Koubbi, P., Laffont, R., and Saucède, T. (2020). Benthic ecoregionalization based on echinoid fauna of the Southern Ocean supports current proposals of Antarctic Marine Protected Areas under IPCC scenarios of climate change. *Glob Change Biol* 26, 2161–2180. doi: 10.1111/gcb.14988.

Fay, A. R., and McKinley, G. A. (2014). Global open-ocean biomes: mean and temporal variability. *Earth Syst. Sci. Data* 6, 273–284. doi: 10.5194/essd-6-273-2014.

Ferreira, A. S., Visser, A. W., MacKenzie, B. R., and Payne, M. R. (2014). Accuracy and precision in the calculation of phenology metrics. *J. Geophys. Res. Oceans* 119, 8438–8453. doi: 10.1002/2014JC010323.

Forcada, J., Trathan, P. N., Reid, K., Murphy, E. J., and Croxall, J. P. (2006). Contrasting population changes in sympatric penguin species in association with climate warming. *Global Change Biology* 12, 411–423. doi: 10.1111/j.1365-2486.2006.01108.x.

Forcén-Vázquez, A., Williams, M. J. M., Bowen, M., Carter, L., and Bostock, H. (2017). Campbell Plateau: A major control on the SW Pacific sector of theSouthern Ocean circulation. In situ Observations/Temperature, Salinity and Density Fields/All Depths/Deep Seas: Southern Ocean doi: 10.5194/os-2017-36.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., et al. (2022). Global Carbon Budget 2022. *Earth Syst. Sci. Data* 14, 4811–4900. doi: 10.5194/essd-14-4811-2022.

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., and Winton, M. (2015). Dominance of the Southern Ocean in Anthropogenic Carbon and Heat Uptake in CMIP5 Models. *Journal of Climate* 28, 862–886. doi: 10.1175/JCLI-D-14-00117.1.

Fyfe, J. C. (2006). Southern Ocean warming due to human influence. *Geophys. Res. Lett.* 33, L19701. doi: 10.1029/2006GL027247.

Gallon, S., Bailleul, F., Charrassin, J.-B., Guinet, C., Bost, C.-A., Handrich, Y., et al. (2013). Identifying foraging events in deep diving southern elephant seals, Mirounga leonina, using acceleration data loggers. *Deep Sea Research Part II: Topical Studies in Oceanography* 88–89, 14–22. doi: 10.1016/j.dsr2.2012.09.002.

Gille, S. T. (2008). Decadal-Scale Temperature Trends in the Southern Hemisphere Ocean. *Journal of Climate* 21, 4749–4765. doi: 10.1175/2008JCLI2131.1.

Godard, M., Manté, C., Guinet, C., Picard, B., and Nerini, D. (2020). Diving Behavior of Mirounga leonina: A Functional Data Analysis Approach. *Front. Mar. Sci.* 7, 595. doi: 10.3389/fmars.2020.00595.

Gordon, H. R., and McCluney, W. R. (1975). Estimation of the Depth of Sunlight Penetration in the Sea for Remote Sensing. *Appl. Opt.* 14, 413. doi: 10.1364/AO.14.000413.

Goulet, P., Guinet, C., Campagna, C., Campagna, J., Tyack, P. L., and Johnson, M. (2020). Flash and grab: deep-diving southern elephant seals trigger anti-predator flashes in bioluminescent prey. *Journal of Experimental Biology* 223, jeb222810. doi: 10.1242/jeb.222810.

Goulet, P., Guinet, C., Swift, R., Madsen, P. T., and Johnson, M. (2019). A miniature biomimetic sonar and movement tag to study the biotic environment and predator-prey interactions in aquatic animals. *Deep Sea Research Part I: Oceanographic Research Papers* 148, 1–11. doi: 10.1016/j.dsr.2019.04.007.

Grant, S. M., Waller, C. L., Morley, S. A., Barnes, D. K. A., Brasier, M. J., Double, M. C., et al. (2021). Local Drivers of Change in Southern Ocean Ecosystems: Human Activities and Policy Implications. *Front. Ecol. Evol.* 9, 624518. doi: 10.3389/fevo.2021.624518.

Gregg, W. W., Conkright, M. E., Ginoux, P., O'Reilly, J. E., and Casey, N. W. (2003). Ocean primary production and climate: Global decadal changes. *Geophys. Res. Lett.* 30. doi: 10.1029/2003GL016889.

Gregg, W. W., Rousseaux, C. S., and Franz, B. A. (2017). Global trends in ocean phytoplankton: a new assessment using revised ocean colour data. *Remote Sensing Letters* 8, 1102–1111. doi: 10.1080/2150704X.2017.1354263.

Gruber, N., Clement, D., Carter, B. R., Feely, R. A., Van Heuven, S., Hoppema, M., et al. (2019a). The oceanic sink for anthropogenic CO_2 from 1994 to 2007. *Science* 363, 1193–1199. doi: 10.1126/science.aau5153.

Gruber, N., Landschützer, P., and Lovenduski, N. S. (2019b). The Variable Southern Ocean Carbon Sink. *Annu. Rev. Mar. Sci.* 11, 159–186. doi: 10.1146/annurev-marine-121916-063407.

Guieu, C., Bonnet, S., Petrenko, A., Menkes, C., Chavagnac, V., Desboeufs, K., et al. (2018). Iron from a submarine source impacts the productive layer of the Western Tropical South Pacific (WTSP). *Sci Rep* 8, 9075. doi: 10.1038/s41598-018-27407-z.

Guinet, C., Vacquié-Garcia, J., Picard, B., Bessigneul, G., Lebras, Y., Dragon, A., et al. (2014). Southern elephant seal foraging success in relation to temperature and light conditions: insight into prey distribution. *Mar. Ecol. Prog. Ser.* 499, 285–301. doi: 10.3354/meps10660.

Guinet, C., Xing, X., Walker, E., Monestiez, P., Marchand, S., Picard, B., et al. (2013). Calibration procedures and first dataset of Southern Ocean chlorophyll a profiles collected by elephant seals equipped with a newly developed CTD-fluorescence tags. *Earth Syst. Sci. Data* 5, 15–29. doi: 10.5194/essd-5-15-2013.

Gutiérrez Rodríguez, A., Safi, K., Fernández, D., Forcén Vázquez, A., Gourvil, P., Hoffmann, L., et al. (2020). Decoupling Between Phytoplankton Growth and Microzooplankton Grazing Enhances Productivity in Subantarctic Waters on Campbell Plateau, Southeast of New Zealand. *J. Geophys. Res. Oceans* 125. doi: 10.1029/2019JC015550.

Hadfield, M. G., Rickard, G. J., and Uddstrom, M. J. (2007). A hydrodynamic model of Chatham Rise, New Zealand. *New Zealand Journal of Marine and Freshwater Research* 41, 239–264. doi: 10.1080/00288330709509912.

Haëntjens, N., Boss, E., and Talley, L. D. (2017). Revisiting Ocean Color algorithms for chlorophyll a and particulate organic carbon in the Southern Ocean using biogeochemical floats. *JGR Oceans* 122, 6583–6593. doi: 10.1002/2017JC012844.

Hales, B., and Takahashi, T. (2004). High-resolution biogeochemical investigation of the Ross Sea, Antarctica, during the AESOPS (U. S. JGOFS) Program. *Global Biogeochem. Cycles* 18, n/a-n/a. doi: 10.1029/2003GB002165.

Hammond, M. L., Beaulieu, C., Henson, S. A., and Sahu, S. K. (2020). Regional surface chlorophyll trends and uncertainties in the global ocean. *Sci Rep* 10, 15273. doi: 10.1038/s41598-020-72073-9.

Harcourt, R., Sequeira, A. M. M., Zhang, X., Roquet, F., Komatsu, K., Heupel, M., et al. (2019). Animal-Borne Telemetry: An Integral Component of the Ocean Observing Toolkit. *Front. Mar. Sci.* 6, 326. doi: 10.3389/fmars.2019.00326.

Hauck, J., and Völker, C. (2015). Rising atmospheric CO_2 leads to large impact of biology on Southern Ocean CO_2 uptake via changes of the Revelle factor. *Geophys. Res. Lett.* 42, 1459–1464. doi: 10.1002/2015GL063070.

Hauck, J., Völker, C., Wolf Gladrow, D. A., Laufkötter, C., Vogt, M., Aumont, O., et al. (2015). On the Southern Ocean CO_2 uptake and the role of the biological carbon pump in the 21st century. *Global Biogeochem. Cycles* 29, 1451–1470. doi: 10.1002/2015GB005140.

Henley, S. F., Cavan, E. L., Fawcett, S. E., Kerr, R., Monteiro, T., Sherrell, R. M., et al. (2020). Changing Biogeochemistry of the Southern Ocean and Its Ecosystem Implications. *Front. Mar. Sci.* 7, 581. doi: 10.3389/fmars.2020.00581.

Henson, S. A., Beaulieu, C., Ilyina, T., John, J. G., Long, M., Séférian, R., et al. (2017a). Rapid emergence of climate change in environmental drivers of marine ecosystems. *Nat Commun* 8, 14682. doi: 10.1038/ncomms14682.

Henson, S. A., Beaulieu, C., and Lampitt, R. (2016). Observing climate change trends in ocean biogeochemistry: when and where. *Glob Change Biol* 22, 1561–1571. doi: 10.1111/gcb.13152.

Henson, S. A., Cael, B. B., Allen, S. R., and Dutkiewicz, S. (2021). Future phytoplankton diversity in a changing climate. *Nat Commun* 12, 5372. doi: 10.1038/s41467-021-25699-w.

Henson, S. A., Cole, H. S., Hopkins, J., Martin, A. P., and Yool, A. (2017b). Detection of climate change-driven trends in phytoplankton phenology. *Glob Change Biol* 24, e101–e111. doi: 10.1111/gcb.13886.

Henson, S. A., Laufkötter, C., Leung, S., Giering, S. L. C., Palevsky, H. I., and Cavan, E. L. (2022). Uncertain response of ocean biological carbon export in a changing world. *Nat. Geosci.* 15, 248–254. doi: 10.1038/s41561-022-00927-0.

Hillenbrand, C.-D., and Cortese, G. (2006). Polar stratification: A critical view from the Southern Ocean. *Palaeogeography, Palaeoclimatology, Palaeoecology* 242, 240–252. doi: 10.1016/j.palaeo.2006.06.001.

Hindell, M. A., Bradshaw, C. J. A., Sumner, M. D., Michael, K. J., and Burton, H. R. (2003). Dispersal of female southern elephant seals and their prey consumption during the austral summer: relevance to management and oceanographic zones. *J Appl Ecology* 40, 703–715. doi: 10.1046/j.1365-2664.2003.00832.x.

Hindell, M. A., McMahon, C. R., Bester, M. N., Boehme, L., Costa, D., Fedak, M. A., et al. (2016). Circumpolar habitat use in the southern elephant seal: implications for foraging success and population trajectories. *Ecosphere* 7. doi: 10.1002/ecs2.1213.

Hindell, M. A., Reisinger, R. R., Ropert-Coudert, Y., Hückstädt, L. A., Trathan, P. N., Bornemann, H., et al. (2020). Tracking of marine predators to protect Southern Ocean ecosystems. *Nature* 580, 87–92. doi: 10.1038/s41586-020-2126-y.

Hogg, A. McC., Penduff, T., Close, S. E., Dewar, W. K., Constantinou, N. C., and Martínez Moreno, J. (2022). Circumpolar Variations in the Chaotic Nature of Southern Ocean Eddy Dynamics. *JGR Oceans* 127. doi: 10.1029/2022JC018440.

Huang, Y., Fassbender, A. J., and Bushinsky, S. M. (2023). Biogenic carbon pool production maintains the Southern Ocean carbon sink. *Proc. Natl. Acad. Sci. U.S.A.* 120, e2217909120. doi: 10.1073/pnas.2217909120.

Hughes, K. A., Pescott, O. L., Peyton, J., Adriaens, T., Cottier Cook, E. J., Key, G., et al. (2020). Invasive non native species likely to threaten biodiversity and ecosystems in the Antarctic Peninsula region. *Glob Change Biol* 26, 2702–2716. doi: 10.1111/gcb.14938.

Huot, Y., and Babin, M. (2010). "Overview of Fluorescence Protocols: Theory, Basic Concepts, and Practice," in *Chlorophyll a Fluorescence in Aquatic Sciences: Methods and Applications*, eds. D. J. Suggett, O. Prášil, and M. A. Borowitzka (Dordrecht: Springer Netherlands), 31–74. doi: 10.1007/978-90-481-9268-7_3.

Hutchins, D. A., Sedwick, P. N., DiTullio, G. R., Boyd, P. W., Quéguiner, B., Griffiths, F. B., et al. (2001). Control of phytoplankton growth by iron and silicic acid availability in the subantarctic Southern Ocean: Experimental results from the SAZ Project. *J. Geophys. Res.* 106, 31559–31572. doi: 10.1029/2000JC000333.

International Ocean Colour Coordinating Group (IOCCG) (2014). Phytoplankton Functional Types from Space. doi: 10.25607/OBP-106.

International Ocean Colour Coordinating Group (IOCCG) (2019). Ocean Optics and Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation, Volume 3.0: Protocols for Satellite Ocean Colour Data Validation: In Situ Optical Radiometry. International Ocean Colour Coordinating Group (IOCCG) doi: 10.25607/OBP-691.

IPCC (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*.

Irvine, L. M., Winsor, M. H., Follett, T. M., Mate, B. R., and Palacios, D. M. (2020). An at-sea assessment of Argos location accuracy for three species of large whales, and the effect of deep-diving behavior on location error. *Anim Biotelemetry* 8, 20. doi: 10.1186/s40317-020-00207-x.

Jaud, T., Dragon, A.-C., Garcia, J. V., and Guinet, C. (2012). Relationship between Chlorophyll a Concentration, Light Attenuation and Diving Depth of the Southern Elephant Seal Mirounga leonina. *PLoS ONE* 7, e47444. doi: 10.1371/journal.pone.0047444.

Jeandel, C., Ruiz-Pino, D., Gjata, E., Poisson, A., Brunet, C., Charriaud, E., et al. (1998). KERFIX, a time-series station in the Southern Ocean: a presentation. *Journal of Marine Systems* 17, 555–569. doi: 10.1016/S0924-7963(98)00064-5.

Jenouvrier, S., Desprez, M., Fay, R., Barbraud, C., Weimerskirch, H., Delord, K., et al. (2018). Climate change and functional traits affect population dynamics of a long lived seabird. *J Anim Ecol* 87, 906–920. doi: 10.1111/1365-2656.12827.

Jerri, A. J. (1977). The Shannon sampling theorem—Its various extensions and applications: A tutorial review. *Proc. IEEE* 65, 1565–1596. doi: 10.1109/PROC.1977.10771.

Ji, R., Edwards, M., Mackas, D. L., Runge, J. A., and Thomas, A. C. (2010). Marine plankton phenology and life history in a changing climate: current research and future directions. *Journal of Plankton Research* 32, 1355–1368. doi: 10.1093/plankt/fbq062.

Jickells, T. D., An, Z. S., Andersen, K. K., Baker, A. R., Bergametti, G., Brooks, N., et al. (2005). Global Iron Connections Between Desert Dust, Ocean Biogeochemistry, and Climate. *Science* 308, 67–71. doi: 10.1126/science.1105959.

Johnson, R., Strutton, P. G., Wright, S. W., McMinn, A., and Meiners, K. M. (2013). Three improved satellite chlorophyll algorithms for the Southern Ocean. *J. Geophys. Res. Oceans* 118, 3694–3703. doi: 10.1002/jgrc.20270.

Johnston, N. M., Murphy, E. J., Atkinson, A., Constable, A. J., Cotté, C., Cox, M., et al. (2022). Status, Change, and Futures of Zooplankton in the Southern Ocean. *Front. Ecol. Evol.* 9, 624692. doi: 10.3389/fevo.2021.624692.

Jönsson, B. F., Salisbury, J., Atwood, E. C., Sathyendranath, S., and Mahadevan, A. (2023). Dominant timescales of variability in global satellite chlorophyll and SST revealed with a MOving Standard deviation Saturation (MOSS) approach. *Remote Sensing of Environment* 286, 113404. doi: 10.1016/j.rse.2022.113404.

J.W. Brewin, R., Ciavatta, S., Sathyendranath, S., Skákala, J., Bruggeman, J., Ford, D., et al. (2019). The Influence of Temperature and Community Structure on Light Absorption by Phytoplankton in the North Atlantic. *Sensors* 19, 4182. doi: 10.3390/s19194182.

Kahl, L. A., Schofield, O., and Fraser, W. R. (2010). Autonomous Gliders Reveal Features of the Water Column Associated with Foraging by Adelie Penguins. *Integrative and Comparative Biology* 50, 1041–1050. doi: 10.1093/icb/icq098.

Kaufman, D. E., Friedrichs, M. A. M., Hemmings, J. C. P., and Smith Jr., W. O. (2018). Assimilating bio-optical glider data during a phytoplankton bloom in the southern Ross Sea. *Biogeosciences* 15, 73–90. doi: 10.5194/bg-15-73-2018.

Keates, T. R., Kudela, R. M., Holser, R. R., Hückstädt, L. A., Simmons, S. E., and Costa, D. P. (2020). Chlorophyll fluorescence as measured in situ by animal-borne instruments in the northeastern Pacific Ocean. *Journal of Marine Systems* 203, 103265. doi: 10.1016/j.jmarsys.2019.103265.

Keerthi, M. G., Lévy, M., and Aumont, O. (2021). Intermittency in phytoplankton bloom triggered by modulations in vertical stability. *Sci Rep* 11, 1285. doi: 10.1038/s41598-020-80331-z.

Keerthi, M. G., Levy, M., Aumont, O., Lengaigne, M., and Antoine, D. (2020). Contrasted Contribution of Intraseasonal Time Scales to Surface Chlorophyll Variations in a Bloom and an Oligotrophic Regime. *JGR Oceans* 125. doi: 10.1029/2019JC015701.

Keerthi, M. G., Prend, C. J., Aumont, O., and Lévy, M. (2022). Annual variations in phytoplankton biomass driven by small-scale physical processes. *Nat. Geosci.* 15, 1027–1033. doi: 10.1038/s41561-022-01057-3.

Khatiwala, S., Primeau, F., and Hall, T. (2009). Reconstruction of the history of anthropogenic CO2 concentrations in the ocean. *Nature* 462, 346–349. doi: 10.1038/nature08526.

Kim, Y. S., and Orsi, A. H. (2014). On the Variability of Antarctic Circumpolar Current Fronts Inferred from 1992–2011 Altimetry. *Journal of Physical Oceanography* 44, 3054–3071. doi: 10.1175/JPO-D-13-0217.1.

Korb, R. E., and Whitehouse, M. (2004). Contrasting primary production regimes around South Georgia, Southern Ocean: large blooms versus high nutrient, low chlorophyll waters. *Deep Sea Research Part I: Oceanographic Research Papers* 51, 721–738. doi: 10.1016/j.dsr.2004.02.006.

Krumhardt, K. M., Long, M. C., Sylvester, Z. T., and Petrik, C. M. (2022). Climate drivers of Southern Ocean phytoplankton community composition and potential impacts on higher trophic levels. *Front. Mar. Sci.* 9, 916140. doi: 10.3389/fmars.2022.916140.

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., et al. (2020). Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. *Biogeosciences* 17, 3439–3470. doi: 10.5194/bg-17-3439-2020.

Landschützer, P., Gruber, N., and Bakker, D. C. E. (2016). Decadal variations and trends of the global ocean carbon sink: DECADAL AIR-SEA CO_2 FLUX VARIABILITY. *Global Biogeochem. Cycles* 30, 1396–1417. doi: 10.1002/2015GB005359.

Landschützer, P., Gruber, N., Haumann, F. A., Rödenbeck, C., Bakker, D. C. E., Van Heuven, S., et al. (2015). The reinvigoration of the Southern Ocean carbon sink. *Science* 349, 1221–1224. doi: 10.1126/science.aab2620.

Lannuzel, D., Bowie, A. R., Remenyi, T., Lam, P., Townsend, A., Ibisanmi, E., et al. (2011). Distributions of dissolved and particulate iron in the sub-Antarctic and Polar Frontal Southern Ocean (Australian sector). *Deep Sea Research Part II: Topical Studies in Oceanography* 58, 2094–2112. doi: 10.1016/j.dsr2.2011.05.027.

Lavigne, H., D'Ortenzio, F., Claustre, H., and Poteau, A. (2012). Towards a merged satellite and in situ fluorescence ocean chlorophyll product. *Biogeosciences* 9, 2111–2125. doi: 10.5194/bg-9-2111-2012.

Le Quéré, C., Rödenbeck, C., Buitenhuis, E. T., Conway, T. J., Langenfelds, R., Gomez, A., et al. (2007). Saturation of the Southern Ocean CO_2 Sink Due to Recent Climate Change. *Science* 316, 1735–1738. doi: 10.1126/science.1136188.

Lehahn, Y., d'Ovidio, F., and Koren, I. (2018). A Satellite-Based Lagrangian View on Phytoplankton Dynamics. *Annu. Rev. Mar. Sci.* 10, 99–119. doi: 10.1146/annurev-marine-121916-063204.

Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., et al. (2012). World ocean heat content and thermosteric sea level change (0-2000 m), 1955-2010. *Geophys. Res. Lett.* 39, n/a-n/a. doi: 10.1029/2012GL051106.

Lévy, M., Ferrari, R., Franks, P. J. S., Martin, A. P., and Rivière, P. (2012). Bringing physics to life at the submesoscale. *Geophys. Res. Lett.* 39, n/a-n/a. doi: 10.1029/2012GL052756.

Lévy, M., Franks, P. J. S., and Smith, K. S. (2018). The role of submesoscale currents in structuring marine ecosystems. *Nat Commun* 9, 4758. doi: 10.1038/s41467-018-07059-3.

Lévy, M., Resplandy, L., and Lengaigne, M. (2014). Oceanic mesoscale turbulence drives large biogeochemical interannual variability at middle and high latitudes. *Geophys. Res. Lett.* 41, 2467–2474. doi: 10.1002/2014GL059608.

Li, Z., Wakao, S., Fischer, B. B., and Niyogi, K. K. (2009). Sensing and Responding to Excess Light. *Annu. Rev. Plant Biol.* 60, 239–260. doi: 10.1146/annurev.arplant.58.032806.103844.

Lim, H., Dunne, J. P., Stock, C. A., and Kwon, M. (2022). Attribution and Predictability of Climate Driven Variability in Global Ocean Color. *JGR Oceans* 127. doi: 10.1029/2022JC019121.

Llort, J., Lévy, M., Sallée, J.-B., and Tagliabue, A. (2015). Onset, intensification, and decline of phytoplankton blooms in the Southern Ocean. *ICES Journal of Marine Science* 72, 1971–1984. doi: 10.1093/icesjms/fsv053.

Loeb, V. J., and Santora, J. A. (2015). Climate variability and spatiotemporal dynamics of five Southern Ocean krill species. *Progress in Oceanography* 134, 93–122. doi: 10.1016/j.pocean.2015.01.002.

Loisel, H., Nicolas, J.-M., Deschamps, P.-Y., and Frouin, R. (2002). Seasonal and inter-annual variability of particulate organic matter in the global ocean. *Geophys. Res. Lett.* 29, 49-1-49–4. doi: 10.1029/2002GL015948.

Long, M. C., Stephens, B. B., McKain, K., Sweeney, C., Keeling, R. F., Kort, E. A., et al. (2021). Strong Southern Ocean carbon uptake evident in airborne observations. *Science* 374, 1275–1280. doi: 10.1126/science.abi4355.

Longhurst, A. R. (2007). *Ecological Geography of the Sea*. Elsevier doi: 10.1016/B978-0-12-455521-1.X5000-1.

Lorenzen, C. J. (1966). A method for the continuous measurement of in vivo chlorophyll concentration. *Deep Sea Research and Oceanographic Abstracts* 13, 223–227. doi: 10.1016/0011-7471(66)91102-8.

Lovenduski, N. S., and Gruber, N. (2005). Impact of the Southern Annular Mode on Southern Ocean circulation and biology. *Geophys. Res. Lett.* 32, L11603. doi: 10.1029/2005GL022727.

Lucas, M., Seeyave, S., Sanders, R., Mark Moore, C., Williamson, R., and Stinchcombe, M. (2007). Nitrogen uptake responses to a naturally Fe-fertilised phytoplankton bloom during the 2004/2005 CROZEX study. *Deep Sea Research Part II: Topical Studies in Oceanography* 54, 2138–2173. doi: 10.1016/j.dsr2.2007.06.017.

Mahadevan, A. (2016). The Impact of Submesoscale Physics on Primary Productivity of Plankton. *Annu. Rev. Mar. Sci.* 8, 161–184. doi: 10.1146/annurev-marine-010814-015912.

Marshall, J., and Speer, K. (2012). Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geosci* 5, 171–180. doi: 10.1038/ngeo1391.

Martin, A. P., and Srokosz, M. A. (2002). Plankton distribution spectra: inter-size class variability and the relative slopes for phytoplankton and zooplankton. *Geophys. Res. Lett.* 29, 66-1-66–4. doi: 10.1029/2002GL015117.

Martin, J. H., Fitzwater, S. E., and Gordon, R. M. (1990). Iron deficiency limits phytoplankton growth in Antarctic waters. *Global Biogeochem. Cycles* 4, 5–12. doi: 10.1029/GB004i001p00005.

Mayot, N., D'Ortenzio, F., Ribera d'Alcalà, M., Lavigne, H., and Claustre, H. (2016). Interannual variability of the Mediterranean trophic regimes from ocean color satellites. *Biogeosciences* 13, 1901–1917. doi: 10.5194/bg-13-1901-2016.

McClain, C. R. (2009). A Decade of Satellite Ocean Color Observations. *Annu. Rev. Mar. Sci.* 1, 19–42. doi: 10.1146/annurev.marine.010908.163650.

McCormack, S. A., Melbourne-Thomas, J., Trebilco, R., Griffith, G., Hill, S. L., Hoover, C., et al. (2021). Southern Ocean Food Web Modelling: Progress, Prognoses, and Future Priorities for Research and Policy Makers. *Front. Ecol. Evol.* 9, 624763. doi: 10.3389/fevo.2021.624763.

McGillicuddy, D. J. (2016). Mechanisms of Physical-Biological-Biogeochemical Interaction at the Oceanic Mesoscale. *Annu. Rev. Mar. Sci.* 8, 125–159. doi: 10.1146/annurev-marine-010814-015606.

McGovern, K., Rodríguez, D., Lewis, M., and Davis, R. (2019). Classification and behavior of free-ranging female southern elephant seal dives based on threedimensional movements and video-recorded observations. *Mar. Ecol. Prog. Ser.* 620, 215–232. doi: 10.3354/meps12936.

McMahon, C. R., Field, I. C., Bradshaw, C. J. A., White, G. C., and Hindell, M. A. (2008). Tracking and data–logging devices attached to elephant seals do not affect individual mass gain or survival. *Journal of Experimental Marine Biology and Ecology* 360, 71–77. doi: 10.1016/j.jembe.2008.03.012.

McMahon, C. R., Roquet, F., Baudel, S., Belbeoch, M., Bestley, S., Blight, C., et al. (2021). Animal Borne Ocean Sensors – AniBOS – An Essential Component of the Global Ocean Observing System. *Front. Mar. Sci.* 8, 751840. doi: 10.3389/fmars.2021.751840.

Mélin, F. (2016). Impact of inter-mission differences and drifts on chlorophyll- *a* trend estimates. *International Journal of Remote Sensing* 37, 2233–2251. doi: 10.1080/01431161.2016.1168949.

Meredith, M. P., and Hogg, A. M. (2006). Circumpolar response of Southern Ocean eddy activity to a change in the Southern Annular Mode. *Geophys. Res. Lett.* 33, L16608. doi: 10.1029/2006GL026499.

Mestre, J., Authier, M., Cherel, Y., Harcourt, R., McMahon, C. R., Hindell, M. A., et al. (2020). Decadal changes in blood δ^{13} C values, at-sea distribution, and weaning mass of southern elephant seals from Kerguelen Islands. *Proc. R. Soc. B.* 287, 20201544. doi: 10.1098/rspb.2020.1544.

Meyer, M. C. (2012). Constrained penalized splines. *Can J Statistics* 40, 190–206. doi: 10.1002/cjs.10137.

Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Doney, S. C., Dutkiewicz, S., Gerber, M., et al. (2006). Inverse estimates of anthropogenic CO2 uptake, transport, and storage by the ocean. *Global Biogeochem. Cycles* 20, n/a-n/a. doi: 10.1029/2005GB002530.

Miloslavich, P., Bax, N. J., Simmons, S. E., Klein, E., Appeltans, W., Aburto Oropeza, O., et al. (2018). Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. *Global Change Biology* 24, 2416–2433. doi: 10.1111/gcb.14108.

Monteiro, P. M. S., Gregor, L., Lévy, M., Maenner, S., Sabine, C. L., and Swart, S. (2015). Intraseasonal variability linked to sampling alias in air-sea CO_2 fluxes in the Southern Ocean: Alias in Air-Sea CO_2 Fluxes. *Geophys. Res. Lett.* 42, 8507–8514. doi: 10.1002/2015GL066009.

Moreau, S., Boyd, P. W., and Strutton, P. G. (2020). Remote assessment of the fate of phytoplankton in the Southern Ocean sea-ice zone. *Nat Commun* 11, 3108. doi: 10.1038/s41467-020-16931-0.

Moreau, S., Mostajir, B., Bélanger, S., Schloss, I. R., Vancoppenolle, M., Demers, S., et al. (2015). Climate change enhances primary production in the western Antarctic Peninsula. *Glob Change Biol* 21, 2191–2205. doi: 10.1111/gcb.12878.

Morel, A. (1988). Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). *J. Geophys. Res.* 93, 10749. doi: 10.1029/JC093iC09p10749.

Morel, A., and Gentili, B. (1996). Diffuse reflectance of oceanic waters III Implication of bidirectionality for the remote-sensing problem. *Appl. Opt.* 35, 4850. doi: 10.1364/AO.35.004850.

Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., and 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. doi: 10.1016/j.rse.2007.03.012.

Morel, A., and Maritorena, S. (2001). Bio-optical properties of oceanic waters: A reappraisal. *J. Geophys. Res.* 106, 7163–7180. doi: 10.1029/2000JC000319.

Morel, A., and Prieur, L. (1977). Analysis of variations in ocean color1: Ocean color analysis. *Limnol. Oceanogr.* 22, 709–722. doi: 10.4319/lo.1977.22.4.0709.

Morel, A. Y., and Gordon, H. R. (1980). Report of the working group on water color. *Boundary-Layer Meteorol* 18, 343–355. doi: 10.1007/BF00122030.

Morley, S. A., Abele, D., Barnes, D. K. A., Cárdenas, C. A., Cotté, C., Gutt, J., et al. (2020). Global Drivers on Southern Ocean Ecosystems: Changing Physical Environments and Anthropogenic Pressures in an Earth System. *Front. Mar. Sci.* 7, 547188. doi: 10.3389/fmars.2020.547188.

Morrison, A. K., Waugh, D. W., Hogg, A. McC., Jones, D. C., and Abernathey, R. P. (2022). Ventilation of the Southern Ocean Pycnocline. *Annu. Rev. Mar. Sci.* 14, 405–430. doi: 10.1146/annurev-marine-010419-011012.

Morrow, R., Fu, L.-L., Ardhuin, F., Benkiran, M., Chapron, B., Cosme, E., et al. (2019). Global Observations of Fine-Scale Ocean Surface Topography With the Surface Water and Ocean Topography (SWOT) Mission. *Front. Mar. Sci.* 6, 232. doi: 10.3389/fmars.2019.00232.

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. J. W., Bricaud, A., et al. (2017). A Consumer's Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean. *Front. Mar. Sci.* 4. doi: 10.3389/fmars.2017.00041.

Mtshali, T. N., Horsten, N. R., Thomalla, S. J., Ryan Keogh, T. J., Nicholson, S. A., Roychoudhury, A. N., et al. (2019). Seasonal Depletion of the Dissolved Iron Reservoirs in the Sub Antarctic Zone of the Southern Atlantic Ocean. *Geophysical Research Letters* 46, 4386–4395. doi: 10.1029/2018GL081355.

Mueller, J. L., Morel, A., Frouin, R., Davis, C., Arnone, R., Carder, K., et al. (2003). Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4.Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4. Volume III: Radiometric Measurements and Data Analysis Protocols. doi: 10.25607/OBP-62.

Murphy, E. J., and Hofmann, E. E. (2012). End-to-end in Southern Ocean ecosystems. *Current Opinion in Environmental Sustainability* 4, 264–271. doi: 10.1016/j.cosust.2012.05.005.

Murphy, E. J., Johnston, N. M., Hofmann, E. E., Phillips, R. A., Jackson, J. A., Constable, A. J., et al. (2021). Global Connectivity of Southern Ocean Ecosystems. *Front. Ecol. Evol.* 9, 624451. doi: 10.3389/fevo.2021.624451.

Murphy, R. J., Pinkerton, M. H., Richardson, K. M., Bradford Grieve, J. M., and Boyd, P. W. (2001). Phytoplankton distributions around New Zealand derived from SeaWiFS remotely sensed ocean colour data. *New Zealand Journal of Marine and Freshwater Research* 35, 343–362. doi: 10.1080/00288330.2001.9517005.

Nardelli, S. C., Gray, P. C., Stammerjohn, S. E., and Schofield, O. (2023). Characterizing coastal phytoplankton seasonal succession patterns on the West Antarctic Peninsula. *Limnology & Oceanography* 68, 845–861. doi: 10.1002/lno.12314.

Nerini, D., and Ghattas, B. (2007). Classifying densities using functional regression trees: Applications in oceanology. *Computational Statistics & Data Analysis* 51, 4984–4993. doi: 10.1016/j.csda.2006.09.028.

Nerini, D., Monestiez, P., and Manté, C. (2010). Cokriging for spatial functional data. *Journal of Multivariate Analysis* 101, 409–418. doi: 10.1016/j.jmva.2009.03.005.

Nyquist, H. (1928). Certain Topics in Telegraph Transmission Theory. *Trans. Am. Inst. Electr. Eng.* 47, 617–644. doi: 10.1109/T-AIEE.1928.5055024.

Organelli, E., Claustre, H., Bricaud, A., Schmechtig, C., Poteau, A., Xing, X., et al. (2016). A Novel Near-Real-Time Quality-Control Procedure for Radiometric Profiles Measured by Bio-Argo Floats: Protocols and Performances. *Journal of Atmospheric and Oceanic Technology* 33, 937–951. doi: 10.1175/JTECH-D-15-0193.1.

Orsi, A. H., Whitworth, T., and Nowlin, W. D. (1995). On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers* 42, 641–673. doi: 10.1016/0967-0637(95)00021-W.

Park, Y. H., Park, T., Kim, T. W., Lee, S. H., Hong, C. S., Lee, J. H., et al. (2019). Observations of the Antarctic Circumpolar Current Over the Udintsev Fracture Zone, the Narrowest Choke Point in the Southern Ocean. *J. Geophys. Res. Oceans* 124, 4511–4528. doi: 10.1029/2019JC015024.

Pauthenet, E., Roquet, F., Madec, G., and Nerini, D. (2017). A Linear Decomposition of the Southern Ocean Thermohaline Structure. *Journal of Physical Oceanography* 47, 29–47. doi: 10.1175/JPO-D-16-0083.1.

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, 2535–2552. doi: 10.1175/JPO-D-19-0120.1.

Pellichero, V., Boutin, J., Claustre, H., Merlivat, L., Sallée, J., and Blain, S. (2020). Relaxation of Wind Stress Drives the Abrupt Onset of Biological Carbon Uptake in the Kerguelen Bloom: A Multisensor Approach. *Geophys. Res. Lett.* 47. doi: 10.1029/2019GL085992.

Pezzulli, S., Stephenson, D. B., and Hannachi, A. (2005). The Variability of Seasonality. *Journal of Climate* 18, 71–88. doi: 10.1175/JCLI-3256.1.

Pinkerton, M. H., Boyd, P. W., Deppeler, S., Hayward, A., Höfer, J., and Moreau, S. (2021). Evidence for the Impact of Climate Change on Primary Producers in the Southern Ocean. *Front. Ecol. Evol.* 9, 592027. doi: 10.3389/fevo.2021.592027.

Pinkerton, M., Richardson, K., Boyd, P., Gall, M., Zeldis, J., Oliver, M., et al. (2005). Intercomparison of ocean colour band-ratio algorithms for chlorophyll concentration in the Subtropical Front east of New Zealand. *Remote Sensing of Environment* 97, 382–402. doi: 10.1016/j.rse.2005.05.004.

Planquette, H., Statham, P. J., Fones, G. R., Charette, M. A., Moore, C. M., Salter, I., et al. (2007). Dissolved iron in the vicinity of the Crozet Islands, Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* 54, 1999–2019. doi: 10.1016/j.dsr2.2007.06.019.

Platt, T., Fuentes-Yaco, C., and Frank, K. T. (2003). Spring algal bloom and larval fish survival. *Nature* 423, 398–399. doi: 10.1038/423398b.

Pollard, R., Sanders, R., Lucas, M., and Statham, P. (2007). The Crozet Natural Iron Bloom and Export Experiment (CROZEX). *Deep Sea Research Part II: Topical Studies in Oceanography* 54, 1905–1914. doi: 10.1016/j.dsr2.2007.07.023.

Pollard, R. T., Salter, I., Sanders, R. J., Lucas, M. I., Moore, C. M., Mills, R. A., et al. (2009). Southern Ocean deep-water carbon export enhanced by natural iron fertilization. *Nature* 457, 577–580. doi: 10.1038/nature07716.

Poloczanska, E. S., Burrows, M. T., Brown, C. J., García Molinos, J., Halpern, B. S., Hoegh-Guldberg, O., et al. (2016). Responses of Marine Organisms to Climate Change across Oceans. *Front. Mar. Sci.* 3. doi: 10.3389/fmars.2016.00062.

Pope, A., Wagner, P., Johnson, R., Shutler, J. D., Baeseman, J., and Newman, L. (2017). Community review of Southern Ocean satellite data needs. *Antarctic Science* 29, 97–138. doi: 10.1017/S0954102016000390.

Prend, C. J., Gille, S. T., Talley, L. D., Mitchell, B. G., Rosso, I., and Mazloff, M. R. (2019). Physical Drivers of Phytoplankton Bloom Initiation in the Southern Ocean's Scotia Sea. *JGR Oceans* 124, 5811–5826. doi: 10.1029/2019JC015162.

Prend, C. J., Keerthi, M. G., Lévy, M., Aumont, O., Gille, S. T., and Talley, L. D. (2022). Sub Seasonal Forcing Drives Year To Year Variations of Southern Ocean Primary Productivity. *Global Biogeochemical Cycles* 36. doi: 10.1029/2022GB007329.

R. Rintoul, S., W. Hughes, C., and Olbers, D. (2001). "Chapter 4.6 The antarctic circumpolar current system," in *International Geophysics* (Elsevier), 271–XXXVI. doi: 10.1016/S0074-6142(01)80124-8.

Racault, M.-F., Le Quéré, C., Buitenhuis, E., Sathyendranath, S., and Platt, T. (2012). Phytoplankton phenology in the global ocean. *Ecological Indicators* 14, 152–163. doi: 10.1016/j.ecolind.2011.07.010.

Ramsay, J. O., and Silverman, B. W. (1997). *Functional Data Analysis*. New York, NY: Springer New York doi: 10.1007/978-1-4757-7107-7.

Ramsay, J. O., and Silverman, B. W. (2005). *Functional Data Analysis*. New York, NY: Springer New York doi: 10.1007/b98888.

Ras, J., Claustre, H., and Uitz, J. (2008). Spatial variability of phytoplankton pigment distributions in the Subtropical South Pacific Ocean: comparison between in situ and predicted data. *Biogeosciences* 5, 353–369. doi: 10.5194/bg-5-353-2008.

Resplandy, L., Lévy, M., d'Ovidio, F., and Merlivat, L. (2009). Impact of submesoscale variability in estimating the air-sea CO2 exchange: Results from a model study of the POMME experiment. *Global Biogeochem. Cycles* 23, n/a-n/a. doi: 10.1029/2008GB003239.

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., and Maury, O. (2013). Dynamic biogeochemical provinces in the global ocean. *Global Biogeochem. Cycles* 27, 1046–1058. doi: 10.1002/gbc.20089.

References

Richard, G., Vacquié-Garcia, J., Jouma’a, J., Picard, B., Génin, A., Arnould, J. P. Y., et al. (2014). Variation in body condition during the post-moult foraging trip of southern elephant seals and its consequences on diving behaviour. *Journal of Experimental Biology*, jeb.088542. doi: 10.1242/jeb.088542.

Rintoul, S. (2000). Southern Ocean currents and climate. *PPRST*, 41–50. doi: 10.26749/rstpp.133.3.41.

Rintoul, S. R. (2011). “The Southern Ocean in the Earth System,” in *Science diplomacy*

Ropert-Coudert, Y., and Wilson, R. P. (2005). Trends and perspectives in animal-attached remote sensing. *Frontiers in Ecology and the Environment* 3, 437–444. doi: 10.1890/1540-9295(2005)003[0437:TAPIAR]2.0.CO;2.

Roquet, F., Park, Y.-H., Guinet, C., Bailleul, F., and Charrassin, J.-B. (2009). Observations of the Fawn Trough Current over the Kerguelen Plateau from instrumented elephant seals. *Journal of Marine Systems* 78, 377–393. doi: 10.1016/j.jmarsys.2008.11.017.

Rosenzweig, C., Casassa, G., Karoly, D. J., Imeson, A., Liu, C., Menzel, A., et al. (2007). Assessment of observed changes and responses in natural and managed systems. doi: 10.5167/UZH-33180.

Rosso, I., Mazloff, M. R., Talley, L. D., Purkey, S. G., Freeman, N. M., and Maze, G. (2020). Water Mass and Biogeochemical Variability in the Kerguelen Sector of the Southern Ocean: A Machine Learning Approach for a Mixing Hot Spot. *JGR Oceans* 125. doi: 10.1029/2019JC015877.

Rudnick, D. L., Davis, R. E., and Sherman, J. T. (2016). Spray Underwater Glider Operations. *Journal of Atmospheric and Oceanic Technology* 33, 1113–1122. doi: 10.1175/JTECH-D-15-0252.1.

Rutz, C., and Hays, G. C. (2009). New frontiers in biologging science. *Biol. Lett.* 5, 289–292. doi: 10.1098/rsbl.2009.0089.

Ryan-Keogh, T. J., Thomalla, S. J., Monteiro, P. M. S., and Tagliabue, A. (2023). Multidecadal trend of increasing iron stress in Southern Ocean phytoplankton. *Science* 379, 834–840. doi: 10.1126/science.abl5237.

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., et al. (2004). The Oceanic Sink for Anthropogenic CO $_2$. *Science* 305, 367–371. doi: 10.1126/science.1097403.

Sadai, S., Condron, A., DeConto, R., and Pollard, D. (2020). Future climate response to Antarctic Ice Sheet melt caused by anthropogenic warming. *Sci. Adv.* 6, eaaz1169. doi: 10.1126/sciadv.aaz1169.

Sallée, J. B., Speer, K., and Morrow, R. (2008). Response of the Antarctic Circumpolar Current to Atmospheric Variability. *Journal of Climate* 21, 3020–3039. doi: 10.1175/2007JCLI1702.1.

Sallée, J.-B. (2018). Southern Ocean Warming. *Oceanog* 31. doi: 10.5670/oceanog.2018.215.

Sallée, J.-B., Llort, J., Tagliabue, A., and Lévy, M. (2015). Characterization of distinct bloom phenology regimes in the Southern Ocean. *ICES Journal of Marine Science* 72, 1985–1998. doi: 10.1093/icesjms/fsv069.

Sanial, V., Van Beek, P., Lansard, B., d'Ovidio, F., Kestenare, E., Souhaut, M., et al. (2014). Study of the phytoplankton plume dynamics off the Crozet Islands (Southern Ocean): A geochemical-physical coupled approach. *J. Geophys. Res. Oceans* 119, 2227–2237. doi: 10.1002/2013JC009305.

Sarmiento, J. L., Hughes, T. M. C., Stouffer, R. J., and Manabe, S. (1998). Simulated response of the ocean carbon cycle to anthropogenic climate warming. *Nature* 393, 245–249. doi: 10.1038/30455.

Sarmiento, J. L., Slater, R., Barber, R., Bopp, L., Doney, S. C., Hirst, A. C., et al. (2004). Response of ocean ecosystems to climate warming. *Global Biogeochem. Cycles* 18, n/a-n/a. doi: 10.1029/2003GB002134.

Sathyendranath, S., Brewin, R. J. W., Jackson, T., Mélin, F., and Platt, T. (2017). Ocean-colour products for climate-change studies: What are their ideal characteristics? *Remote Sensing of Environment* 203, 125–138. doi: 10.1016/j.rse.2017.04.017.

Sathyendranath, S., Gouveia, A. D., Shetye, S. R., Ravindran, P., and Platt, T. (1991). Biological control of surface temperature in the Arabian Sea. *Nature* 349, 54–56. doi: 10.1038/349054a0.

Saulquin, B., Fablet, R., Mangin, A., Mercier, G., Antoine, D., and Fanton d'Andon, O. (2013). Detection of linear trends in multisensor time series in the presence of autocorrelated noise: Application to the chlorophyll-a SeaWiFS and MERIS data sets and extrapolation to the incoming Sentinel 3-OLCI mission. *J. Geophys. Res. Oceans* 118, 3752–3763. doi: 10.1002/jgrc.20264.

Schallenberg, C., Strzepek, R. F., Bestley, S., Wojtasiewicz, B., and Trull, T. W. (2022). Iron Limitation Drives the Globally Extreme Fluorescence/Chlorophyll Ratios of the Southern Ocean. *Geophysical Research Letters* 49. doi: 10.1029/2021GL097616.

Schofield, O., Saba, G., Coleman, K., Carvalho, F., Couto, N., Ducklow, H., et al. (2017). Decadal variability in coastal phytoplankton community composition in a changing West Antarctic Peninsula. *Deep Sea Research Part I: Oceanographic Research Papers* 124, 42–54. doi: 10.1016/j.dsr.2017.04.014.

Schwartz, M. D. (2013). "Introduction," in *Phenology: An Integrative Environmental Science*, ed. M. D. Schwartz (Dordrecht: Springer Netherlands), 1–5. doi: 10.1007/978-94-007-6925-0_1.

Sedwick, P. N., Blain, S., Quéguiner, B., Griffiths, F. B., Fiala, M., Bucciarelli, E., et al. (2002). Resource limitation of phytoplankton growth in the Crozet Basin, Subantarctic Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* 49, 3327–3349. doi: 10.1016/S0967-0645(02)00086-3.

Sedwick, P. N., DiTullio, G. R., Hutchins, D. A., Boyd, P. W., Griffiths, F. B., Crossley, A. C., et al. (1999). Limitation of algal growth by iron deficiency in the Australian Subantarctic Region. *Geophys. Res. Lett.* 26, 2865–2868. doi: 10.1029/1998GL002284.

Séférian, R., Iudicone, D., Bopp, L., Roy, T., and Madec, G. (2012). Water Mass Analysis of Effect of Climate Change on Air–Sea CO_2 Fluxes: The Southern Ocean. *J. Climate* 25, 3894–3908. doi: 10.1175/JCLI-D-11-00291.1.

Serôdio, J., and Lavaud, J. (2011). A model for describing the light response of the nonphotochemical quenching of chlorophyll fluorescence. *Photosynth Res* 108, 61–76. doi: 10.1007/s11120-011-9654-0.

Shannon, C. E. (1949). Communication in the Presence of Noise. *Proc. IRE* 37, 10–21. doi: 10.1109/JRPROC.1949.232969.

Shi, J.-R., Talley, L. D., Xie, S.-P., Peng, Q., and Liu, W. (2021). Ocean warming and accelerating Southern Ocean zonal flow. *Nat. Clim. Chang.* 11, 1090–1097. doi: 10.1038/s41558-021-01212-5.

Shi, J.-R., Xie, S.-P., and Talley, L. D. (2018). Evolving Relative Importance of the Southern Ocean and North Atlantic in Anthropogenic Ocean Heat Uptake. *Journal of Climate* 31, 7459–7479. doi: 10.1175/JCLI-D-18-0170.1.

Shiskin, J., Young, A. H., and Musgrave, J. C. (1967). The X-11 variant of Census Method II Seasonal Adjustment Programme. *U.S. Bureau of the Census* Tech. Paper No. 15.

Siegel, D. A., Buesseler, K. O., Doney, S. C., Sailley, S. F., Behrenfeld, M. J., and Boyd, P. W. (2014). Global assessment of ocean carbon export by combining satellite observations and food-web models. *Global Biogeochem. Cycles* 28, 181–196. doi: 10.1002/2013GB004743.

Siegel, D. A., DeVries, T., Cetinić, I., and Bisson, K. M. (2023). Quantifying the Ocean's Biological Pump and Its Carbon Cycle Impacts on Global Scales. *Annu. Rev. Mar. Sci.* 15, 329–356. doi: 10.1146/annurev-marine-040722-115226.

Siegelman, L., Klein, P., Rivière, P., Thompson, A. F., Torres, H. S., Flexas, M., et al. (2020). Enhanced upward heat transport at deep submesoscale ocean fronts. *Nat. Geosci.* 13, 50–55. doi: 10.1038/s41561-019-0489-1.

Siegelman, L., O'Toole, M., Flexas, M., Rivière, P., and Klein, P. (2019a). Submesoscale ocean fronts act as biological hotspot for southern elephant seal. *Sci Rep* 9, 5588. doi: 10.1038/s41598-019-42117-w.

Siegelman, L., Roquet, F., Mensah, V., Rivière, P., Pauthenet, E., Picard, B., et al. (2019b). Correction and Accuracy of High- and Low-Resolution CTD Data from Animal-Borne Instruments. *Journal of Atmospheric and Oceanic Technology* 36, 745–760. doi: 10.1175/JTECH-D-18-0170.1.

Sigmond, M., Reader, M. C., Fyfe, J. C., and Gillett, N. P. (2011). Drivers of past and future Southern Ocean change: Stratospheric ozone versus greenhouse gas impacts. *Geophys. Res. Lett.* 38, n/a-n/a. doi: 10.1029/2011GL047120.

Sokolov, S., and Rintoul, S. R. (2009a). Circumpolar structure and distribution of the Antarctic Circumpolar Current fronts: 1. Mean circumpolar paths. *J. Geophys. Res.* 114, C11018. doi: 10.1029/2008JC005108.

Sokolov, S., and Rintoul, S. R. (2009b). Circumpolar structure and distribution of the Antarctic Circumpolar Current fronts: 2. Variability and relationship to sea surface height. *J. Geophys. Res.* 114, C11019. doi: 10.1029/2008JC005248.

Steiner, N. S., Bowman, J., Campbell, K., Chierici, M., Eronen-Rasimus, E., Falardeau, M., et al. (2021). Climate change impacts on sea-ice ecosystems and associated ecosystem services. *Elementa: Science of the Anthropocene* 9, 00007. doi: 10.1525/elementa.2021.00007.

Suggett, D., Moore, C., Hickman, A., and Geider, R. (2009). Interpretation of fast repetition rate (FRR) fluorescence: signatures of phytoplankton community structure versus physiological state. *Mar. Ecol. Prog. Ser.* 376, 1–19. doi: 10.3354/meps07830.

Sun, X., Brewin, R. J. W., Sathyendranath, S., Dall'Olmo, G., Airs, R., Barlow, R., et al. (2023). Coupling ecological concepts with an ocean-colour model: Phytoplankton size structure. *Remote Sensing of Environment* 285, 113415. doi: 10.1016/j.rse.2022.113415.

Swart, N. C., Gille, S. T., Fyfe, J. C., and Gillett, N. P. (2018). Recent Southern Ocean warming and freshening driven by greenhouse gas emissions and ozone depletion. *Nature Geosci* 11, 836–841. doi: 10.1038/s41561-018-0226-1.

Swart, S., Thomalla, S. J., and Monteiro, P. M. S. (2015). The seasonal cycle of mixed layer dynamics and phytoplankton biomass in the Sub-Antarctic Zone: A high-resolution glider experiment. *Journal of Marine Systems* 147, 103–115. doi: 10.1016/j.jmarsys.2014.06.002.

Szeto, M., Werdell, P. J., Moore, T. S., and Campbell, J. W. (2011). Are the world's oceans optically different? *J. Geophys. Res.* 116, 2011JC007230. doi: 10.1029/2011JC007230.

Tagliabue, A., Bopp, L., Dutay, J.-C., Bowie, A. R., Chever, F., Jean-Baptiste, P., et al. (2010). Hydrothermal contribution to the oceanic dissolved iron inventory. *Nature Geosci* 3, 252–256. doi: 10.1038/ngeo818.

Tagliabue, A., Sallée, J.-B., Bowie, A. R., Lévy, M., Swart, S., and Boyd, P. W. (2014). Surface-water iron supplies in the Southern Ocean sustained by deep winter mixing. *Nature Geosci* 7, 314–320. doi: 10.1038/ngeo2101.

Talley, L. (2013). Closure of the Global Overturning Circulation Through the Indian, Pacific, and Southern Oceans: Schematics and Transports. *oceanog* 26, 80–97. doi: 10.5670/oceanog.2013.07.

Talley, L. D., Pickard, G. L., Emery, W. J., and Swift, J. H. (2011). "Global Circulation and Water Properties," in *Descriptive Physical Oceanography* (Elsevier), 473–511. doi: 10.1016/B978-0-7506-4552-2.10014-9.

Taylor, M. H., Losch, M., and Bracher, A. (2013). On the drivers of phytoplankton blooms in the Antarctic marginal ice zone: A modeling approach. *J. Geophys. Res. Oceans* 118, 63–75. doi: 10.1029/2012JC008418.

Teo, S., Kudela, R., Rais, A., Perle, C., Costa, D., and Block, B. (2009). Estimating chlorophyll profiles from electronic tags deployed on pelagic animals. *Aquat. Biol.* 5, 195–207. doi: 10.3354/ab00152.

Terrats, L., Claustre, H., Cornec, M., Mangin, A., and Neukermans, G. (2020). Detection of Coccolithophore Blooms With BioGeoChemical Argo Floats. *Geophys. Res. Lett.* 47. doi: 10.1029/2020GL090559.

Testor, P., de Young, B., Rudnick, D. L., Glenn, S., Hayes, D., Lee, C. M., et al. (2019). OceanGliders: A Component of the Integrated GOOS. *Front. Mar. Sci.* 6, 422. doi: 10.3389/fmars.2019.00422.

Thomalla, S. J., Du Plessis, M., Fauchereau, N., Giddy, I., Gregor, L., Henson, S., et al. (2023). Southern Ocean phytoplankton dynamics and carbon export: insights from a seasonal cycle approach. *Phil. Trans. R. Soc. A.* 381, 20220068. doi: 10.1098/rsta.2022.0068.

Thomalla, S. J., Fauchereau, N., Swart, S., and Monteiro, P. M. S. (2011). Regional scale characteristics of the seasonal cycle of chlorophyll in the Southern Ocean. *Biogeosciences* 8, 2849–2866. doi: 10.5194/bg-8-2849-2011.

Thomalla, S. J., Racault, M.-F., Swart, S., and Monteiro, P. M. S. (2015). High-resolution view of the spring bloom initiation and net community production in the Subantarctic Southern Ocean using glider data. *ICES Journal of Marine Science* 72, 1999–2020. doi: 10.1093/icesjms/fsv105.

Thompson, D. W. J., and Wallace, J. M. (2000). Annular Modes in the Extratropical Circulation. Part I: Month-to-Month Variability. *J. Climate* 13, 1000–1016. doi: 10.1175/1520-0442(2000)013<1000:AMITEC>2.0.CO;2.

Toggweiler, J. R., and Russell, J. (2008). Ocean circulation in a warming climate. *Nature* 451, 286–288. doi: 10.1038/nature06590.

Toggweiler, J. R., and Samuels, B. (1995). Effect of drake passage on the global thermohaline circulation. *Deep Sea Research Part I: Oceanographic Research Papers* 42, 477–500. doi: 10.1016/0967-0637(95)00012-U.

Tournier, M., Goulet, P., Fonvieille, N., Nerini, D., Johnson, M., and Guinet, C. (2021). A novel animal-borne miniature echosounder to observe the distribution and migration patterns of intermediate trophic levels in the Southern Ocean. *Journal of Marine Systems* 223, 103608. doi: 10.1016/j.jmarsys.2021.103608.

Treasure, A., Roquet, F., Ansorge, I., Bester, M., Boehme, L., et al. (2017). Marine Mammals Exploring the Oceans Pole to Pole: A Review of the MEOP Consortium. *Oceanog.* 30, 132–138. doi: 10.5670/oceanog.2017.234.

Trenberth, K. E., Fasullo, J. T., and Balmaseda, M. A. (2014). Earth's Energy Imbalance. *Journal of Climate* 27, 3129–3144. doi: 10.1175/JCLI-D-13-00294.1.

Uchida, T., Balwada, D., P. Abernathey, R., A. McKinley, G., K. Smith, S., and Lévy, M. (2020). Vertical eddy iron fluxes support primary production in the open Southern Ocean. *Nat Commun* 11, 1125. doi: 10.1038/s41467-020-14955-0.

Uitz, J., Claustre, H., Griffiths, F. B., Ras, J., Garcia, N., and Sandroni, V. (2009). A phytoplankton class-specific primary production model applied to the Kerguelen Islands region (Southern Ocean). *Deep Sea Research Part I: Oceanographic Research Papers* 56, 541–560. doi: 10.1016/j.dsr.2008.11.006.

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *J. Geophys. Res.* 111, C08005. doi: 10.1029/2005JC003207.

van Gennip, S., Martin, A. P., Srokosz, M. A., Allen, J. T., Pidcock, R., Painter, S. C., et al. (2016). Plankton patchiness investigated using simultaneous nitrate and chlorophyll observations. *J. Geophys. Res. Oceans* 121, 4149–4156. doi: 10.1002/2016JC011789.

Vancoppenolle, M., Meiners, K. M., Michel, C., Bopp, L., Brabant, F., Carnat, G., et al. (2013). Role of sea ice in global biogeochemical cycles: emerging views and challenges. *Quaternary Science Reviews* 79, 207–230. doi: 10.1016/j.quascirev.2013.04.011.

Vantrepotte, V., and Mélin, F. (2009). Temporal variability of 10-year global SeaWiFS time-series of phytoplankton chlorophyll a concentration. *ICES Journal of Marine Science* 66, 1547–1556. doi: 10.1093/icesjms/fsp107.

Vantrepotte, V., and Mélin, F. (2011). Inter-annual variations in the SeaWiFS global chlorophyll a concentration (1997–2007). *Deep Sea Research Part I: Oceanographic Research Papers* 58, 429–441. doi: 10.1016/j.dsr.2011.02.003.

Vargas, M., Brown, C. W., and Sapiano, M. R. P. (2009). Phenology of marine phytoplankton from satellite ocean color measurements. *Geophys. Res. Lett.* 36, L01608. doi: 10.1029/2008GL036006.

Venables, H. J., Pollard, R. T., and Popova, E. E. (2007). Physical conditions controlling the development of a regular phytoplankton bloom north of the Crozet Plateau, Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* 54, 1949–1965. doi: 10.1016/j.dsr2.2007.06.014.

Venables, H., and Moore, C. M. (2010). Phytoplankton and light limitation in the Southern Ocean: Learning from high-nutrient, high-chlorophyll areas. *J. Geophys. Res.* 115, C02015. doi: 10.1029/2009JC005361.

Vernet, M., Geibert, W., Hoppema, M., Brown, P. J., Haas, C., Hellmer, H. H., et al. (2019). The Weddell Gyre, Southern Ocean: Present Knowledge and Future Challenges. *Rev. Geophys.* 57, 623–708. doi: 10.1029/2018RG000604.

Viviant, M., Trites, A. W., Rosen, D. A. S., Monestiez, P., and Guinet, C. (2010). Prey capture attempts can be detected in Steller sea lions and other marine predators using accelerometers. *Polar Biol* 33, 713–719. doi: 10.1007/s00300-009-0750-y.

Volk, T., and Hoffert, M. I. (1985). “Ocean Carbon Pumps: Analysis of Relative Strengths and Efficiencies in Ocean-Driven Atmospheric CO2 Changes,” in *Geophysical Monograph Series*, eds. E. T. Sundquist and W. S. Broecker (Washington, D. C.: American Geophysical Union), 99–110. doi: 10.1029/GM032p0099.

Volzke, S., McMahon, C. R., Hindell, M. A., Burton, H. R., and Wotherspoon, S. J. (2021). Climate influences on female survival in a declining population of southern elephant seals (*Mirounga leonina*). *Ecol Evol* 11, 11333–11344. doi: 10.1002/ece3.7919.

Von Schuckmann, K., Palmer, M. D., Trenberth, K. E., Cazenave, A., Chambers, D., Champollion, N., et al. (2016). An imperative to monitor Earth’s energy imbalance. *Nature Clim Change* 6, 138–144. doi: 10.1038/nclimate2876.

Watanabe, Y. Y., Ito, K., Kokubun, N., and Takahashi, A. (2020). Foraging behavior links sea ice to breeding success in Antarctic penguins. *Sci. Adv.* 6, eaba4828. doi: 10.1126/sciadv.aba4828.

Watanabe, Y. Y., and Papastamatiou, Y. P. (2023). Biologging and Biotelemetry: Tools for Understanding the Lives and Environments of Marine Animals. *Annu. Rev. Anim. Biosci.* 11, 247–267. doi: 10.1146/annurev-animal-050322-073657.

Wege, M., Salas, L., and LaRue, M. (2021). Ice matters: Life history strategies of two Antarctic seals dictate climate change eventualities in the Weddell Sea. *Glob Change Biol* 27, 6252–6262. doi: 10.1111/gcb.15828.

Williams, R. G., Ceppi, P., Roussenov, V., Katavouta, A., and Meijers, A. J. S. (2023). The role of the Southern Ocean in the global climate response to carbon emissions. *Phil. Trans. R. Soc. A.* 381, 20220062. doi: 10.1098/rsta.2022.0062.

Wright, S., Jeffrey, S., Mantoura, R., Llewellyn, C., Bjornland, T., Repeta, D., et al. (1991). Improved HPLC method for the analysis of chlorophylls and carotenoids from marine phytoplankton. *Mar. Ecol. Prog. Ser.* 77, 183–196. doi: 10.3354/meps077183.

Wunsch, C. (1998). The Work Done by the Wind on the Oceanic General Circulation. *J. Phys. Oceanogr.* 28, 2332–2340. doi: 10.1175/1520-0485(1998)028<2332:TWDBTW>2.0.CO;2.

Xi, H., Losa, S. N., Mangin, A., Soppa, M. A., Garnesson, P., Demaria, J., et al. (2020). Global retrieval of phytoplankton functional types based on empirical orthogonal functions using CMEMS GlobColour merged products and further extension to OLCI data. *Remote Sensing of Environment* 240, 111704. doi: 10.1016/j.rse.2020.111704.

Xing, X., Briggs, N., Boss, E., and Claustre, H. (2018). Improved correction for non-photochemical quenching of in situ chlorophyll fluorescence based on a synchronous irradiance profile. *Opt. Express* 26, 24734. doi: 10.1364/OE.26.024734.

Xing, X., Claustre, H., Blain, S., D'Ortenzio, F., Antoine, D., Ras, J., et al. (2012). Quenching correction for in vivo chlorophyll fluorescence acquired by autonomous platforms: A case study with instrumented elephant seals in the Kerguelen region (Southern Ocean). *Limnol. Oceanogr. Methods* 10, 483–495. doi: 10.4319/lom.2012.10.483.

Xing, X., Claustre, H., Uitz, J., Mignot, A., Poteau, A., and Wang, H. (2014). Seasonal variations of bio-optical properties and their interrelationships observed by Bio-Argo floats in the subpolar North Atlantic. *J. Geophys. Res. Oceans* 119, 7372–7388. doi: 10.1002/2014JC010189.

Xing, X., Morel, A., Claustre, H., Antoine, D., D'Ortenzio, F., Poteau, A., et al. (2011). Combined processing and mutual interpretation of radiometry and fluorimetry from autonomous profiling Bio-Argo floats: Chlorophyll *a* retrieval. *J. Geophys. Res.* 116, C06020. doi: 10.1029/2010JC006899.

Table of Acronyms

Table 1: Acronyms, definitions and units

Symbol	Explicit description	unit
ACC	Antarctic Circumpolar Current	
AES	Areas of Ecological Significance	
AODN	Australian Ocean Data Network (IMOS)	
ATL	Atlantic sector of the Southern Ocean	
BCP	Biological Carbon Pump	
BGC	Biogeochemical	
CCAMLR	Commission for the Conservation of Antarctic Marine Living Resources	
Chla	Chlorophyll a	
[Chla]	Chlorophyll a concentration	$mg.m^{-3}$
$[Chla^{Fluo}]$	[Chla] derived from fluorescence measurement, satellite-corrected	$mg.m^{-3}$
$[Chla^{LFM}]$	LFM prediction of $[Chla^{Fluo}]$	$mg.m^{-3}$
$[Chla^{Sat}]$	[Chla] derived from satellite measurement	$mg.m^{-3}$
CHL	[Chla] derived from satellite measurement	$mg.m^{-3}$
CHL_{MA}	Multi-annual component of CHL time series	$mg.m^{-3}$
CHL_{Seas}	Seasonal component of CHL time series	$mg.m^{-3}$
CHL_{SS}	Sub-seasonal component of CHL time series	$mg.m^{-3}$
$[CHL]_{spring-summer}$	Spring-summer climatological value of CHL	$mg.m^{-3}$
CROZEX	Crozet Natural Iron Bloom and Export Experiment	
CTD	Conductivity, Temperature, Depth	
$dark^{Fluo}$	Fluorescence dark signal	$mg.m^{-3}$
$dark^{KL}$	Vertical diffuse attenuation coefficient for the dark signal of L	m^{-1}
Σ_{CHL}	Integral over time of CHL annual cycle	$mg.m^{-3}.year$
EOV	Essential Ocean Variables	
ESA	European Space Agency	
$[F^{Fluo}]$	Smoothed, dark- and NPQ-corrected Fluo data	$mg.m^{-3}$
[F]	LFM-calibrated signal of $[F^{Fluo}]$	$mg.m^{-3}$
$[F^{LFM}]$	LFM prediction of $[F^{luo}]$	$mg.m^{-3}$
FDA	Functional Data Analysis	
Fluo	Chla fluorescence (raw)	$mg.m^{-3}$
Fluo*	Unsmoothed, dark- and NPQ-corrected Fluo data	$mg.m^{-3}$
FPC	Functional Principal Component	
FPCA	Functional Principal Component Analysis	
FPC*i*	i^{th} axis of FPCA	
$FPCi_{STD}$	Standard deviation of FPC*i* score	
fmicro	Fraction of microphytoplankton	%
fnano	Fraction of nanophytoplankton	%
fpico	Fraction of picophytoplankton	%
$f_{E\text{-}drop}$	Spatial frequency of SES tag spectrum energy-drop	m^{-1}
f_{mean}	Mean spatial frequency of SES tag light profiles	m^{-1}
Γ_e^{LFM}	LFM-derived Fluo calibration coefficient for tag *e*	

Γ^{Sat}	Satellite-derived Fluo calibration coefficient	
HNLC	High Nitrate - Low Chlorophyll	
HPLC	High Pressure Liquid Chromatography	
IMOS	Integrated Marine Observing System	
IND	Indian sector of the Southern Ocean	
K_L	Vertical diffuse attenuation coefficient for L	m^{-1}
KEOPS	Kerguelen Ocean and Plateau Study	
L	Light	µmol quanta.m^{-2}.s^{-1}
λ	Smoothing parameter of functional fit	
LFM	Linear Functional Model	
MEOP	Mammals Exploring the Oceans Pole to Pole	
[micro]	Concentration of microphytoplankton	mg.m^{-3}
MLD	Mixed Layer Depth	
MERIS	Medium Resolution Imaging Spectrometer	
MODIS	Moderate Resolution Imaging Spectroradiometer	
MPA	Marine Protected Area	
[nano]	Concentration of nanophytoplankton	mg.m^{-3}
NASA	National Aeronautics and Space Administration	
$NCHL_{Seas}$	Normalized seasonal component of CHL time series	dimensionless
$NCHL_{Seas + SS}$	Normalized sum of seasonal plus sub-seasonal component of CHL time series	dimensionless
NCEI	National Centers for Environmental Information (NOAA)	
NOAA	National Oceanic and Atmospheric Administration	
NPQ	Non-Photochemical Quenching	
OBPG	Ocean Biology Processing Group (NASA)	
OC	Ocean Colour	
OC-CCI	Ocean Color Climate Change Initiative (ESA)	
OISST	Optimal Interpolation Sea Surface Temperature (NOAA)	
OLCI	Ocean and Land Colour Imager	
OLS	Ordinary Least Squares	
PAC	Pacific sector of the Southern Ocean	
PCA	Prey Capture Attempt	
PF	Polar Front	
PFT	Phytoplankton Functional Type	
[pico]	Concentration of picophytoplankton	mg.m^{-3}
PSC	Phytoplankton Size Class	
PSL	Physical Sciences Laboratory (NOAA)	
$RMSE^{Fluo}$	Root-mean-square error of the functional fit applied to Fluo data	mg.m^{-2}
$RMSE^{LFM}$	Root-mean-square error of the LFM	mg.m^{-2}
S	Salinity	
SACCF	Southern ACC Front	
SAF	Subantarctic Front	
SAPIGH	Service d'Analyses de Pigments par HPLC (LOV)	
SeaWIFS	Sea-viewing Wide Field-of-view Sensor	
SES	Southern Elephant Seal	
SNO-MEMO	Système National d’Observation Mammifères Echantillonneurs du Milieu Océanique	

SO	Southern Ocean	
SOSCEx	Southern Ocean Seasonal Cycle Experiment	
SOTS	Southern Ocean Time Series (IMOS)	
SRDL	Satellite Relayed Data Logger	
SST	Sea Surface Temperature	ºC
STF	Subtropical Front	
SWOT	Surface Water and Ocean Topography	
T	Temperature	ºC
$<V>$	Water-column integrated content of variable *V*, calculated from $z=Z_{inf}$ to the surface (z=0)	$mg.m^{-2}$
$<V>_{\%ML}$	Percentage of $<V>$ contained within the mixed layer	%
$V(z)$	Value of variable *V* at depth *z*	$mg.m^{-3}$
V_{max}	Maximum value of variable *V* between $z=Z_{inf}$ and the surface (z=0)	$mg.m^{-3}$
V_{surf}	Surface value of variable *V*, computed as the mean between $z=Z_{pd}$ and the surface (z=0)	$mg.m^{-3}$
VAR_{Seas}	Relative contribution of CHL_{Seas} to total CHL variance	%
VIIRS	Visible Infrared Imaging Radiometer Suite	
WCSS	Within-cluster sum of squares	
Z_{eu}	Euphotic Depth	m
Z_{max}^{Fluo}	Depth of the maximum value of $[Chla^{Fluo}]$	m
Z_{inf}	Deeper bound of the LFM predictions	m
Z_{max}^{LFM}	Depth of the maximum value of $[Chla^{LFM}]$	m
Z_{MLD}	Mixed Layer Depth	m
Z_{NPQ}	“NPQ-layer” depth	m
Z_{pd}	Penetration Depth	m
Z_{sup}	Shallower bound of the LFM predictions	m

Index of Figures