REMOTE SENSING OF THE PHYTOPLANKTON COMMUNITY STRUCTURE IN COASTAL WATERS: A CASE STUDY IN THE SOUTH BRAZIL BIGHT

Andréa de Lima Oliveira

Doctorate Thesis of the Graduate Course in Remote Sensing, guided by Drs. Milton Kampel, and Natália Rudorff, approved in May 31, 2022.

URL of the original document:
<http://urlib.net/8JMKD3MGP3W34T/472R2QS>

INPE
São José dos Campos
2022

REMOTE SENSING OF THE PHYTOPLANKTON COMMUNITY STRUCTURE IN COASTAL WATERS: A CASE STUDY IN THE SOUTH BRAZIL BIGHT

Andréa de Lima Oliveira

Doctorate Thesis of the Graduate Course in Remote Sensing, guided by Drs. Milton Kampel, and Natália Rudorff, approved in May 31, 2022.

URL of the original document:
<http://urlib.net/8JMKD3MGP3W34T/472R2QS>

INPE

São José dos Campos

2022

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS
Serviço de Pós-Graduação - SEPGR

DEFESA FINAL DE TESE DE ANDRÉA DE LIMA OLIVEIRA
BANCA Nº 127/2022, REG 142921/2018

No dia 31 de maio de 2022, as 09h, por teleconferência, o(a) aluno(a) mencionado(a) acima defendeu seu trabalho final (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. O(A) aluno(a) foi APROVADO(A) pela Banca Examinadora, por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Doutora em Sensoriamento Remoto. O trabalho precisa da incorporação das correções sugeridas pela Banca Examinadora e revisão final pelo(s) orientador(es).

Título: "Remote Sensing of the Phytoplankton Community Structure in Coastal Waters: A Case Study in The South Brazil Bight"

Membros da banca:

Dra. Evlyn Márcia Leão de Moraes Novo – Presidente – INPE
Dr. Milton Kamlpel – Orientador – INPE
Dra. Natália Rudorff – Orientadora – INPE
Dra. Mariana Altenburg Soppa – Membro Externo – AWI – Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Division Climate Sciences, Physical Oceanography of the Polar Seas
Dr. Emmanuel Claude Devred – Membro Externo – BIO – Bedford Institute of Oceanography, Fisheries and Oceans Canada
Dr. Robert Brewin – Membro Externo – UoE – University of Exeter, Centre for Geography and Environmental Science, College of Life and Environmental Sciences

Declarações de aprovação dos membros estrangeiros Dr. Emmanuel Claude Devred e Dr. Robert Brewin anexas ao processo.

Documento assinado eletronicamente por **Milton Kampel**, **Pesquisador**, em 10/06/2022, às 09:47 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Natalia Rudorff Oliveira**, **Tecnologista**, em 10/06/2022, às 10:02 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Evlyn Marcia Leão de Moraes Novo**, **Pesquisador**, em 10/06/2022, às 11:34 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **mariana altenburg soppa (E)**, **Usuário Externo**, em 14/06/2022, às 04:25 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site http://sei.mctic.gov.br/verifica.html, informando o código verificador **9857165** e o código CRC **27C57269**.

Referência: Processo nº 01340.003513/2022-06

SEI nº 9857165

ACKNOWLEDGEMENTS

privilege to meet there.

I am thankful to the researchers who promptly and kindly answered my queries, especially, Áurea Ciotti, Andy Stock, Priscila Lange, and Robert Brewin.

I am thankful to the anonymous reviewers for the suggestions in the fourth chapter of this thesis, published in the Continental Shelf Research.

I am thankful to the members of thesis assessment committee for the valuable suggestions and questions.

I am thankful to my family and friends for supporting me in this long journey and keeping me motivated to face the challenges and uncertainties ahead. Especially, Rosa, Aman, Rafa, João, and Anderson.

ABSTRACT

Phytoplankton are responsible for most of the primary production in the oceans. Their dynamics are essential for ecological modelling and climate change studies. In this study, we analyzed distinct perspectives of remote sensing (RS) estimation of phytoplankton structure in optically-complex waters: (i) in a global perspective, we analyzed the challenges of estimating phytoplankton size classes (PSCs) from satellite, compared their performance in open-ocean and coastal waters, and discussed the elements that might be playing important roles; (ii) in a regional perspective, we used a coastal station in the South Brazil Bight (SBB), Southwestern Atlantic, as a case study to investigate how the phytoplankton pigments combined to other techniques (optical-microscopy and phytoplankton absorption coefficient) could be used to improve the estimation of phytoplankton assemblages in this region; (iii) and also, evaluated the application of remote sensing algorithms using abundance and spectral-based approaches applied to *in situ* and satellite-retrieval of PSCs to analyze their responses to temporal and spatial changes. Coastal waters presented a distinct pigments structure compared to the open ocean, however, their variability was not higher than open-ocean waters. The satellite-retrievals of chlorophyll-a concentrations (Chl$a$) presented higher errors for coastal waters, with mean absolute error of 55% and 72% for open-ocean and coastal, respectively. The bias indicated an overestimation for coastal waters 7% and an underestimation 9% for open-ocean waters. The PSC model performance, on the other hand, varied with the size class considered, with picophytoplankton presenting the worst metrics. The SBB region presents complex meteo-oceanographic dynamics, with local upwellings, oceanic fronts, meanders and eddies. The phytoplankton community in the ANTARES-Ubatuba station is characterized by mixed assemblages composed mainly of diatoms, nanoflagellates, and picophytoplankton in mostly oligomesotrophic conditions (0.1<Chl$a$<1 mg·m^{-3}). However, this scenario can shift to a high phytoplankton biomass condition (Chl$a$> 5 mg·m^{-3}) in some events. The locally tuned PSC model parameters were close to the values estimated in previous studies. The spectral-based approach is indicated to studies in a changing ocean because it is based on the spectral changes, instead of in previously set biomass relationships. The seasonality of the satellite-retrieved PSCs in the SBB showed an overall pattern to be out-of-phase with sea surface temperature (SST), similarly to Chl$a$. Our finds highlight the suitability of SST for future improvements in the regional tuning of the PSC model (correlation coefficient for Chl$a$ and SST ρ=-0.35). To the best of our knowledge, this is the first study to evaluate these types of RS PSC models in the coastal area of SBB. Further improvements are needed but this study shows promising results, helping to understand future changes in phytoplankton assemblages in this region.

Keywords: Phytoplankton Size Classes. time series. remote sensing. satellite. MODIS.

SENSORIAMENTO REMOTO DA ESTRUTURA DA COMUNIDADE FITOPLANCTÔNICA EM ÁGUAS COSTEIRAS: UM ESTUDO DE CASO NA BACIA DE SANTOS

RESUMO

O fitoplâncton é responsável pela maior parte da produção primária nos oceanos. Entender sua dinâmica é essencial para a modelagem ecológica e estudos de mudanças climáticas. Neste estudo, analisamos diferentes perspectivas de estimativa da estrutura da comunidade fitoplanctônica por sensoriamento remoto (SR) em águas opticamente complexas: (i) em uma perspectiva global, analisamos os desafios de estimar as classes de tamanho do fitoplâncton (do inglês *Phytoplankton Size Classes* - PSCs) a partir de dados de satélite, comparando seu desempenho em águas oceânicas e costeiras, e discutindo os elementos que podem estar desempenhando papéis importantes nessa relação; (ii) em uma perspectiva regional, utilizamos uma estação costeira na Bacia de Santos, Atlântico Sudoeste, como estudo de caso para investigar como os pigmentos fitoplanctônicos combinados a outras técnicas de microscopia óptica e coeficiente de absorção do fitoplâncton, poderiam ser utilizados para melhorar a estimativa de comunidades fitoplanctônicas nesta região; (iii) além disso, avaliamos a aplicação de algoritmos de sensoriamento remoto usando abordagens baseadas na abundância e em dados espectrais aplicadas à estimativa *in situ* e por satélite de PSCs para analisar suas respostas a mudanças temporais e espaciais. As águas costeiras apresentaram uma estrutura de pigmentos distinta, mas variabilidade similar às águas oceânicas. As estimativas das concentrações de clorofila-a (Chl$a$) por satélite apresentaram maiores erros para águas costeiras, com erro absoluto médio de 50% e 72% para mar aberto e costeiro, respectivamente. O viés indicou um excesso de estimativa para águas costeiras 7% e uma subestimação 9% para águas de mar aberto. O desempenho do modelo PSC, por outro lado, variou com a classe de tamanho do fitoplâncton considerada. A região da Bacia de Santos apresenta dinâmicas meteo-oceanográficas complexas, com ressurgências locais, frentes oceânicas, meandros e vórtices. A comunidade fitoplanctônica da estação fixa ANTARES-Ubatuba é caracterizada por agrupamentos mistos compostos principalmente por diatomáceas, nanoflagelados e picofitoplâncton em condições predominantemente oligomesotróficas (0,1<Chl$a$<1 mg·m^{-3}). No entanto, este cenário pode mudar para uma condição de alta biomassa fitoplanctônica (Chl$a$> 5 mg·m^{-3}) em alguns eventos. Os parâmetros do modelo PSC ajustados localmente foram próximos dos valores estimados em estudos anteriores reportados na literatura. A abordagem espectral é indicada para estudos em um oceano em mudança porque é baseada nas respostas espectrais, ao invés de se basear nas relações de biomassa pré-definidas. A sazonalidade das PSCs obtidas por satélite na Bacia de Santos mostrou em geral, um padrão fora de fase com variações na temperatura da superfície do mar (TSM), semelhante a Chl$a$. Nossos resultados indicam a relevância de se considerar variações da TSM em esforços futuros de aperfeiçoamento no ajuste regional de modelos para a estimativa de PSCs (coeficiente de correlação para Chl$a$ e TSM ρ=-0,35). Até onde sabemos, este é o primeiro estudo a avaliar esses tipos de modelos SR para estimativa de PSC na área costeira da Bacia de Santos. Ainda que mais esforços devam ser desenvolvidos para aperfeiçoar os modelos de estimativa de PSCs, este estudo apresentou resultados

promissores que podem ser úteis para ajudar no entendimento sobre mudanças em curso e futuras nas comunidades fitoplanctônicas nesta região.

Palavras-chave: classes de tamanho do fitoplâncton. série temporal. sensoriamento remoto. satélite. MODIS.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AMT	–	Atlantic Meridional Transect
BC	–	Brazil Current
CFUF	–	Cape Frio Upwelling Front
CW	–	Coastal Water
DIC	–	Dissolved Inorganic Carbon
DOC	–	Dissolved Organic Carbon
DP	–	Diagnostic Pigment
ESA	–	European Space Agency
GIOP-DC	–	Generalized Inherent Optical Properties framework with Default Configuration
HPLC	–	High performance liquid chromatography
MAD	–	Median absolute difference
MAPE	–	Mean absolute percentage error
MLR	–	Multiple linear regression
MPD	–	Median percentage difference
MERIS	–	Medium Resolution Imaging Spectrometer
MODIS	–	MODerate resolution Imaging Spectroradiometer
NASA	–	National Aeronautics and Space Administration
NOMAD	–	NASA bio-Optical Marine Algorithm Data set
OLCI	–	Ocean and Land Colour Instrument
PACE	–	Plankton, Aerosol, Cloud, ocean Ecosystem mission
PCA	–	Principal Component Analysis
PFT	–	Phytoplankton Functional Type
PIC	–	Particulate Inorganic Carbon
POC	–	Particulate Organic Carbon
PPW	–	Plata Plume Water
PRISMA	–	Prototype Research Instruments and Space Mission technology Advancement
PSC	–	Phytoplankton Size Class
RMSE	–	Root-mean square Error
SACW	–	South Atlantic Central Water
SBB	–	South Brazil Bight
SeaBASS	–	SeaWiFS Bio-optical Archive and Storage System
SeaWiFS	–	Sea-viewing Wide Field-of-view Sensor
SFF	–	Size-fractionated filtration
SSF	–	Subtropical Shelf Front
TW	–	Tropical Water

LIST OF SYMBOLS

λ	–	Wavelength
ρ	–	Pearson's correlation coefficient
$a(\lambda)$	–	Absorption coefficient (m^{-1})
$a_p(\lambda)$	–	Particulate absorption coefficient (m^{-1})
$a_{ph}(\lambda)$	–	Phytoplankton absorption coefficient (m^{-1})
$a_{sol}(\lambda)$	–	Absorption coefficient in the solution (λ)(m^{-1})
$a_{ph}^*(\lambda)$	–	Specific phytoplankton absorption coefficient (m^2(mg C)$^{-1}$)
$a_{cdom}(\lambda)$	–	Coloured dissolved organic matter absorption coefficient (m^{-1})
$a_d(\lambda)$	–	Detritus (non-algal) absorption coefficient (m^{-1})
$a_{m,n}(\lambda)$	–	Absorption coefficient of micro and nanophytoplankton (m^{-1})
$a_{pico}(\lambda)$	–	Absorption coefficient of picophytoplankton (m^{-1})
$a_m(\lambda)$	–	Absorption coefficient of microphytoplankton (m^{-1})
$a_n(\lambda)$	–	Absorption coefficient of nanophytoplankton (m^{-1})
$b_{bp}(\lambda)$	–	Backscattering of the particulate coefficient (m^{-1})
Chl$a$	–	Chlorophyll-a concentration (mg·m^{-3})
C_p	–	Picophytoplankton concentration of chlorophyll-a (mg·m^{-3})
$C_{m,n}$	–	Micro and nanophytoplankton concentration of chlorophyll-a (mg·m^{-3})
$C_{p,n}$	–	Pico and nanophytoplankton concentration of chlorophyll-a (mg·m^{-3})
C_n	–	Nanophytoplankton concentration of chlorophyll-a (mg·m^{-3})
C_{max}^p	–	Maximum asymptotic chlorophyll of picophytoplankton (mg·m^{-3})
$C_{max}^{p,n}$	–	Maximum asymptotic chlorophyll of pico and nanophytoplankton (mg·m^{-3})
$E_d(\lambda)$	–	Downwelling irradiance at a specific wavelength (λ) (W m^2)
F_p	–	Picophytoplankton fraction of chlorophyll-a (unitless)
$F_{m,n}$	–	Micro and nanophytoplankton fraction of chlorophyll-a (unitless)
$L_u(\lambda)$	–	Water upwelling radiance at a specific wavelength (λ) (W m^2 sr^{-1}),
L	–	Optical path length (m)
$OD(\lambda)$	–	Optical density at a specific wavelength (λ) (per unit beam cross-section))
$R_{rs}(\lambda)$	–	Remote sensing reflectance at a specific wavelength (λ) (sr^{-1})
$r_{rs}(-0,\lambda)$	–	Remote sensing reflectance in the subsurface at a specific wavelength (λ) (sr^{-1})
Q_a^*	–	Package effect index
S_p	–	Initial slope of picophytoplankton (unitless)
$S_{p,n}$	–	Initial slope of picophytoplankton and nanophytoplankton (unitless)
S_f	–	Size index (CIOTTI et al., 2002) (unitless)
SI	–	Size index (BRICAUD et al., 2004) (unitless)
TChl$a$	–	Total chlorophyll-a for the HPLC analysis (mg·m^{-3})

CONTENTS

1 INTRODUCTION

Marine phytoplankton are autotrophic organisms able to use sunlight and carbon dioxide dissolved in the seawater to produce organic matter. They are the foundation of most marine ecosystems' food web and play a key role in the carbon biogeochemical cycle. Considering climate change projections, an increase in sea surface temperature is occurring (FOX-KEMPER et al., 2021), which will likely change phytoplankton diversity, resulting in loss of ecological resilience with consequences to productivity and functioning of the marine environment (HENSON et al., 2021), and a higher frequency of harmful algal bloom events in coastal waters is expected (PORTNER et al., 2019). Understanding the phytoplankton assemblages' succession and dynamics can provide stakeholders with information to make better decisions in dealing with these changes.

Due to their relevance, there are many techniques used to identify and quantify phytoplankton assemblages *in situ*: one is the optical microscopy, in which a specialized technician identifies the taxonomic groups and counts the cells. This procedure is very time-consuming and is still considered the most reliable method for micro and nanophytoplankton cells, but is not recommended for cells smaller than $2mu$m (picophytoplankton) (EDLER; ELBRÄCHTER, 2010). Methods less time-consuming are also available, such as flow-cytometry and High Performance Liquid Chromatography (HPLC) which quantify the phytoplankton pigments; however, each one of them has drawbacks (IOCCG, 2014). For instance, these methods are not able to identify detailed taxonomy and flow-cytometry can miss cells that are not in the operational range of the instrument.

The chemotaxonomy applied to HPLC data which retrieves phytoplankton groups or size classes by the pigments' relative concentrations is one of the most used, due to its practicality in sampling, analysing and also covering all the phytoplankton size ranges. It is also important in the remote sensing context, due to its close relationship to the spectral absorption of the phytoplankton (BIDIGARE et al., 1990; BRICAUD et al., 2004). Chlorophyll-a is the main phytoplankton pigment, with high absorption in the blue and red spectrum of sunlight (BIDIGARE et al., 1990). However, other pigments also have an important contribution to light absorption, with their maximum absorption peak in slightly different wavelengths (BIDIGARE et al., 1990; BRICAUD et al., 2004). Thus, different pigment composition in the phytoplankton community results in a different absorption spectrum.

Monitoring and characterizing the phytoplankton assemblages on a global or even

regional scale is unfeasible considering solely *in situ* approaches, for this reason, models to derive phytoplankton assemblages using satellite data have been studied in the last decades (NAIR et al., 2008; IOCCG, 2014; RUDORFF; KAMPEL, 2012; MOUW et al., 2017; BRACHER et al., 2017). The usually available dataset validation input for these models is the diagnostic pigment analysis (DPA) obtained from data derived from HPLC (VIDUSSI et al., 2001; UITZ et al., 2006), giving the phytoplankton size classes, which is a way to classify the phytoplankton based on the usual size of each major taxonomic group (BRACHER et al., 2017). This size classification is usually also related to zooplankton preying preference and their role in the carbon cycle, e.g., diatoms are usually part of the microphytoplankton size class and are more likely to export carbon to the depth (JEFFREY et al., 2011) (this will be further described in the next Chapter 2).

There are three main approaches to access PSCs from remote sensing (RS) models: (i) the abundance-based approach, which receives as input the chlorophyll-a concentration (Chl$a$); (ii) the spectral-based approach, which receives spectral data as input, such as remote sensing reflectance, light absorption or backscattering coefficients; and (iii) the so-called ecological-based approach which adds to Chl$a$ and spectral data environmental variables, such as sea surface temperature, wind speed and stress (see IOCCG (2014), Bracher et al. (2017) and Mouw et al. (2017) for an overview of these approaches).

Most of the RS models were designed for global ocean data (BRACHER et al., 2017), except the ecological-based approaches (RAITSOS et al., 2008) which had regional development. However, local tuning to more complex waters such as continental shelf and coastal waters have also been successfully applied (see Liu et al. (2021) and Turner et al. (2021)). Regional models used to retrieve phytoplankton size classes and their role in the carbon cycle are receiving more attention due to the need to understand long-term regional variability (BRACHER et al., 2017; BREWIN et al., 2021). Coastal waters are also relevant for human activities, such as leisure activities and as food sources (i.e., fishing and marine aquaculture). Nonetheless, these waters are also highly sensitive to the changing environmental conditions, driven by temperature changes (FOX-KEMPER et al., 2021), which could intensify, for instance, the increase in algal blooms events (PORTNER et al., 2019), reduction in the fishing stocks or depletion of ideal conditions for the cultivated species.

The South Brazil Bight (SBB), in the Southwestern Atlantic, is a dynamic region with multi-scale processes affecting its biology (BRANDINI et al., 2018). This area is

used for oil and gas exploration, cargo transport (with important Ports in Paranaguá, Santos, Rio de Janeiro, among others), fishing, and leisure purposes. Due to discoveries of oil and gas reservoirs in the north portion of the South Brazil Bight in the 2000s, large infrastructure related to this industry were installed and the existing ones expanded in the following years (TEIXEIRA, 2013), inducing a population growth and change in the land cover and use (IWAMA et al., 2017). The population in the North Coast of São Paulo increased more than 25% from 2000 to 2010, and by projected values reached 329 thousands inhabitants in 2020, representing 45% growth from 2000 (SEADE, 2022). The population increment was not followed by a proportional improvement on the sanitation infrastructure. Thus environmental problems are intensified by this large population living by the sea or within 100 km from the coast (MMA, 2008). In addition, the geographical and geological characteristics of the region makes the population more exposed to landslides and floods (IWAMA et al., 2017).

The SBB also maintain a rich community of marine animals, both resident and migratory (MMA, 2007). Fishing is an important source of food and socioeconomic activity for the local population, and some fishing methods are considered a cutural identity for the traditional population (DIEGUES et al., 1999). The region is also considered naturally propitious to aquaculture due to its enclosed bays and usually protected waters, which is confirmed by the local production of mussels, scallops, and fish (FAGUNDES et al., 2004; BUENO et al., 2011). Tuning a regional phytoplankton size class RS model to the region would help to understand the phytoplankton community structure and the effects of the environmental dynamics on the phytoplankton assemblages. Considering this, the following section presents the general and specific objectives of this study.

1.1 Objectives

1.1.1 General objective

Study the application of remote sensing PSC models to coastal and continental shelves waters, based on phytoplankton pigment variability. And, at the local scale, evaluate the performance of locally-tuned PSC models to a coastal site in the South Brazil Bight, Southwestern Atlantic, for environmental studies of the phytoplankton community structure.

1.1.2   Specific objectives

  a) Analyse differences between oceanic and coastal waters based on phyto-plankton diagnostic pigments;

  b) Evaluate the fitting performance of an abundance PSC model, remote sensing-based to oceanic and coastal waters;

  c) Analyse the phytoplankton community structure and pigment composition of a coastal site in the South Brazil Bight;

  d) Analyse the optical properties and indexes of the phytoplankton assemblages;

  e) Evaluate the application of two remote sensing PSC models: one using the abundance-based approach, and another with the spectral-based approach, to determine the PSCs at a coastal site in the South Brazil Bight, using global and local tuning;

  f) Analyse the temporal variability of the PSCs at the coastal site in the South Brazil Bight, using a MODerate Resolution Imaging Spectrometer - Aqua (MODISA) remotely-sensed time series from 2002-2020.

This document structure is based on independent but complementing chapters, each chapter is a manuscript published or (to be) submitted to a scientific journal, except for the Introduction (Chapter 1), the Theoretical background (Chapter 2), and Final remarks (Chapter 6). The specific objectives will be met by the following chapters: a and b, Chapter 3; c and d by Chapter 4; and e and f by Chapter 5.

2 THEORETICAL BACKGROUND

2.1 Phytoplankton size classes: definition and concept

Photosynthesis is the process that transforms carbon dioxide and water into oxygen and organic matter. This process is the foundation of terrestrial and aquatic/marine food webs and one of the key processes in the carbon biogeochemical cycle (ROY et al., 2011; NAIR et al., 2008). On land, terrestrial plants are responsible for this production, but in the aquatic ecosystems, the tiny cells of different species of phytoplankton are the main producers.

Differently from forests, visually and tangibly diverse, phytoplankton are often seen as an uniform component of the marine primary producers, but they are composed of a diverse community (JEFFREY et al., 2011). However, their diversity is not easily assessed. Phytoplankton taxonomy identification requires some sort of *in situ* sampling, i.e., sampling the cells to identify the groups, by visual/optical identification (UTERMÖHL, 1958; OLSON; SOSIK, 2007), genomics sequencing (especially for picophytoplankton, which due to its size ($<2\mu$m) are difficult to identify by visual features) (IOCCG, 2014; RIBEIRO et al., 2018) or chemotaxonomy (ROY et al., 2011).

This identification is important due to the ecological role and characteristics of each marine phytoplankton taxonomic group. Diatoms, dinoflagellates, coccolithophorids, and cyanobacteria, all of these groups according to their physiologic characteristics, size, and composition will play different roles in the biogeochemical cycles (LEQUÉRÉ et al., 2005) and sustain the marine ecosystem's primary production. Not to mention the risk imposed by harmful algal bloom species, which should be regularly monitored to avoid marine animals and human intoxication (BABIN et al., 2008).

The phytoplankton functional types (PFTs) classification is a way to reduce biological complexity by aggregating taxonomic groups of phytoplankton based, mainly, on their biogeochemical roles (LEQUÉRÉ et al., 2005). Classifying phytoplankton by their main ecological trait was initially proposed by ecological studies (see the review by Nair et al. (2008)). This classification received an especial relevance by biogeochemical and climate modellers (such as LeQuéré et al. (2005) and Henson et al. (2021)), where they could be aggregated, simplifying the model, or isolated, making the model more complex, which is useful for diversity studies (see, for instance, Masuda et al. (2017)).

These types can be further combined in phytoplankton size classes (PSCs), as the

classification proposed by Sieburth et al. (1978). Even though this classification does not contemplate all the functional types usually used, there are still many ecological and biogeochemical processes related to cell size (NAIR et al., 2008) (see Table 2.1).

Table 2.1 - Phytoplankton Size Classes (PSCs) and Phytoplankton Functional Types (PFTs).

Size Class Diagnostic Pigment	PFT name	Description
Pico (<2 μm) Zeaxanthin Chlorophyll-b	Pico-autotrophs	Pico-eukaryotes and non N_2-fixing photosynthetic bacteria, such as, *Synechococcus* and *Prochlorococcus*. They have a high surface-to-volume ratio; are able to succeed in low nutrient and high light conditions; are broadly distributed in the oceans and have an important role in primary production but are insignificant in exporting matter to the depth.
	N_2-fixer	*Tricodesmium* [1] and N_2-fixing unicellular prokaryotes. They are able to use N_2 from atmosphere, which is advantageous in warm and nutrient-poor waters, but not efficient otherwise.
Nano (2-20 μm) 19'-hexanoyloxyfucoxanthin 19'-butanoyloxyfucoxanthin Alloxanthin	Calcifiers	Coccolithophorids. They have an important role in the marine carbonate flux, but high CO_2 concentration (low pH) reduces their calcification rate (RIEBESELL et al., 2000).
	DMS-producers	Phaeocystis and small (<20 μm) autotrophic flagellates produce dimethylsulfoniopropionate (DMSP) and convert it into DMS. These organisms affect the atmospheric sulfur cycle.
	Mixed-Phytoplankton (Chrysophyceae)	Chrysophyceae. Their size range vary from 2-200 μm, thus they are present in the nano and microphytoplankton, here the chrysophycea will be in the nano and dinoflagellates in the microphytoplankton.
Micro (20-200 μm) Fucoxanthin Peridinin	Silicifier	Diatoms. Due to their size, they contribute to carbon export far more effectively than smaller plankton through direct sinking of single cells, key-grazing pathways, and through mass sedimentation events at the end of the spring blooms when nutrients are depleted.
	Mixed-Phytoplankton (autotrophic dinoflagellates)	Dinoflagellates. They represent phytoplankton of heterogeneous size and taxonomic composition for which no distinct biogeochemical role is defined. This PFT constitutes the background biomass of phytoplankton, which do not bloom in the open ocean, have low seasonality, and no direct impact on the cycles of S, Si, or $CaCO_3$.

SOURCE: Adapted from LeQuéré et al. (2005) and Nair et al. (2008). Diagnostic pigments from Uitz et al. (2006).

The PSC classification is useful from an ecological and pragmatical perspective. In large cells, the pigments are "packaged" into chloroplasts (organelles responsible for photosynthesis in eukaryotic cells), and due to the self-shading, the absorption per unit of pigment concentration is less efficient than if they were homogeneously distributed into the cell cytoplasm (BRICAUD et al., 1995; BRICAUD et al., 2004; PRIEUR, 1981). Thus, when the specific absorption (absorption per pigment concentration) is calculated, small cells present higher values than large cells. This phenomenon, for the best of our knowledge, was first named as "package effect" by Kirk (1975),

---

[1] *Tricodesmium* usually form colonies, which would exclude them from picophytoplankton. However, due to the diagnostic pigment it is associated, we opted to maintain them on the picophytoplankton size-range.

but has been known at least since Duysens (1956).

Regarding the pigments, chlorophyll-a is the main photosynthetic pigment but is not the only one. Different pigments are related to different taxonomic groups (JEF-FREY, 1997; JEFFREY et al., 2011). The endosymbiosis hypothesis describes these differences as a result of a sequence of endosymbiosis events (JEFFREY et al., 2011), that is well established for cyanobacteria and chloroplasts, but is further extended to other groups of cells, considering a secondary and even tertiary endosymbiosis (DEL-WICHE, 1999). This hypothesis helps to explain why some groups share pigments (e.g. zeaxanthin which occurs in all cells) and others are exclusive to a group (e.g. peridinin occurs only in dinoflagellates) (see Jeffrey et al. (2011) and the references within).

Even though most of the pigments are not exclusively related to a specific taxonomic group, they are used as indicators of the presence and abundance of these groups, usually aggregated into PSCs (VIDUSSI et al., 2001; UITZ et al., 2006). These pigments and the phytoplankton size classes are included in the Table 2.1. The phytoplankton pigments will have an effect in the light absorption spectrum, and in the package effect described previously, which is used as an indicator of the groups and the PSCs present in the water.

In a broader perspective, considering the carbon cycle, marine phytoplankton are only a fraction of the particulate organic carbon (POC) pool, also composed of zoo-plankton, non-autotrophic bacteria and detritus (CEOS, 2014; MIDDELBURG, 2019; BREWIN et al., 2021). The other pools are dissolved inorganic carbon (DIC), dissolved organic carbon (DOC) and particulate inorganic carbon (PIC) (see Figure 2.1). As discussed by Brewin et al. (2021) and CEOS (2014), some of these pools are not fea-sibly retrieved by available technology of satellite sensors but are estimated through modelling. For this reason, the authors highlighted the need of improving the accu-racy of the satellite-retrievable pools, along with *in situ* autonomous detection and quantification strategies.

Figure 2.1 - Ocean carbon cycle and its sampling techniques.
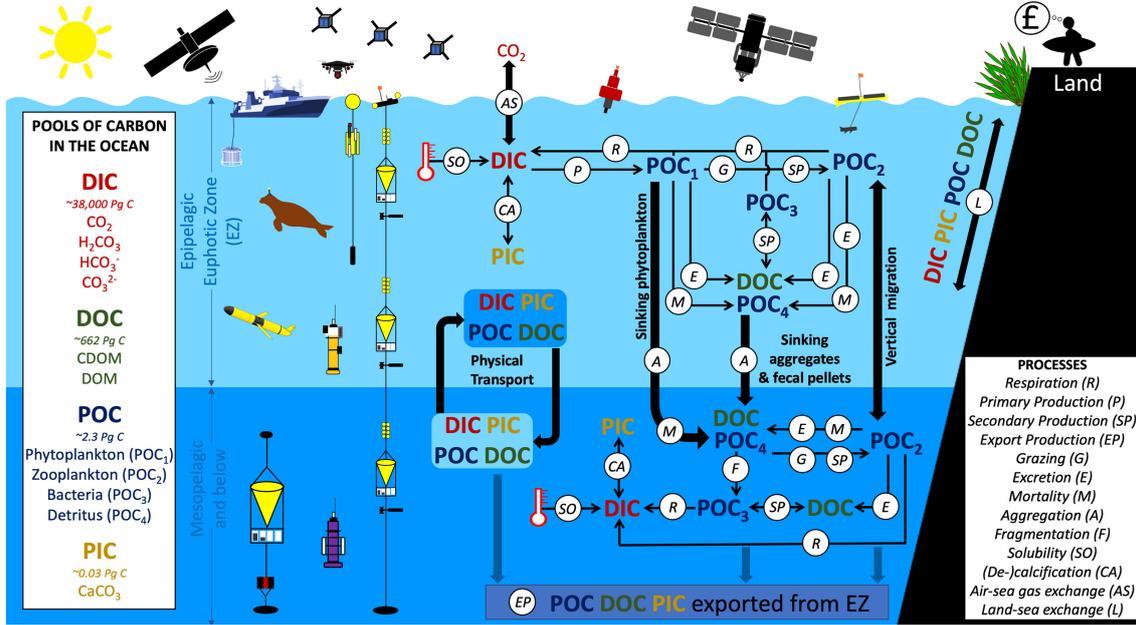
Pools, fluxes and processes that form the ocean biological carbon pump (OBCP), and current methods used to monitor them. Bold black text and thick black arrows represent the key export pathways and interactions with other domains (land and atmosphere). Global stocks of the different carbon pools in the ocean are given in the box on the left; the four major kinds of pools – DIC, DOC, POC and PIC – are given in different colours.

SOURCE: Brewin et al. (2021).

2.2 Obtaining phytoplankton size classes from satellite data

The phytoplankton bio-optics is modulated by their pigments and cells structure (i.e., size, shape, layers). The light scattered and absorbed by the cells suspended in the water will result in a colour, which when using a sensor, both *in situ* or on satellite platforms to measure top of atmosphere radiance (in this context, a radiant flux reflected by the water surface after cross the atmosphere) can be converted to the surface remote sensing reflectance following Gordon et al. (1988) (Equation 2.1) and atmosphere correction:

$$r_{rs}(-0, \lambda) = \frac{L_u(\lambda)}{E_d(\lambda)} = G(\lambda) \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \tag{2.1}$$

Where $r_{rs}(-0, \lambda)$ is the remote sensing reflectance in the subsurface (sr⁻¹), $L_u(\lambda)$ is the water upwelling radiance (W m² sr⁻¹), $E_d(\lambda)$ is the downwelling irradiance (W m²), $G(\lambda)$ is accounting for illumination conditions, sea surface properties, and the shape of the marine volume scattering function (GORDON et al., 1988), $b_b(\lambda)$ is the backscattering coefficient, and $a(\lambda)$ the absorption coefficient.

However, phytoplankton are not the only optically active constituents in ocean water. Water itself is optically active (scattering in the blue and absorbing in the red and infra-red), non-algal suspended particulate matter (detritus, inorganic matter), and coloured dissolved organic matter, all will affect the water-leaving radiance (MOBLEY, 2004).

The inherent optical properties (IOPs) can be expanded as a sum of their components (IOCCG, 2006). Thus, $a(\lambda)$ is represented in the Equation 2.2, where $a_w(\lambda)$ is the water absorption coefficient, a_{cdom} is the absorption due to coloured dissolved organic matter (CDOM), also known as yellow substances, Gelbstoff (yellow substance in German) and *gilvin* from Latin to pale yellow, suggested by Kirk (1976), a_{ph} is the absorption coefficient of phytoplankton, and a_d is the absorption due to non-algal or detritus particulate matter (organic and inorganic).

$$a(\lambda) = a_w(\lambda) + a_{cdom}(\lambda) + a_{ph}(\lambda) + a_d(\lambda) \qquad (2.2)$$

And $b_b(\lambda)$ is represented in the Equation 2.3, as the sum of b_{bw} backscattering of the water and b_{bp} backscattering of the particulate matter. CDOM is considered mostly absorbing, thus its backscattering is considered negligible (KIRK, 2011).

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda) \qquad (2.3)$$

These equations and the usual spectral behaviour of these constituents are used in semi-analytical approaches to retrieving IOPs from remote sensing reflectance. For satellite platforms, it is also required an atmosphere correction, which is basically removing the atmosphere and ocean-atmosphere interface signal (gases, aerosols, water, dust, glint, and whitecap) from the reflectance (IOCCG, 2010). This is not an easy task, especially for coastal waters as discussed by Hu et al. (2000), but this is not the focus of this study.

Morel and Prieur (1977) suggested a water classification that has been widely used.

By that time Chl*a* was widely sampled and one of the main biological variables of interest for ocean-colour studies. Thus, the authors separated two classes of water, the Case-1 waters are those in which all the active optical constituents co-vary with phytoplankton (MOREL, 1988), ideally composed only by phytoplankton and their pigment's byproducts such as Dissolved Organic Matter (DOM) and organic detritus (MOREL; PRIEUR, 1977). And Case-2 waters, would be ideally only composed of inorganic particles with no pigments (MOREL; PRIEUR, 1977).

This definition was later updated, making the definition of the Case-2 waters broader than the previous one. Thus, Case-2 waters are those waters where CDOM and inorganic particles make an important or dominant contribution to optical properties (MOREL, 1988), varying independently from phytoplankton. This classification was important for empirical algorithms to retrieve Chl*a*, because it was possible to relate the concentration of the pigments with changes in the remote sensing reflectance curves, when applied to Case-1 waters (MOREL, 1988).

For optically complex waters (Case-2), regional approaches and semi-analytical algorithms used to derive the IOPs are suggested (SATHYENDRANATH, 2000). However, even with these models, the satellite-retrieval of IOPs is limited, due to the ambiguity in the inversion as discussed by Defoin-Platel and Chami (2007). It means that different values of IOPs could result in the same remote sensing reflectance. The authors used synthesized data with some constraints to simulate realistic coastal-waters IOPs, and found that this ambiguity is higher under some conditions. For instance, highly absorbing waters (high CDOM) will result in low reflectance in the blue, which will increase ambiguity in the absorption coefficient estimations. Or, in another example, highly scattering waters will produce high reflectance and increase ambiguity in the backscattering estimations. The strategies suggested by the authors to handle these ambiguities were to invest in inverse models focused on local scales and seasons, the use of ancillary data, and the use of neighbours pixels to help reduce the ambiguity of the results.

The strategies for satellite-retrieval of IOPs and Chl*a* are important in this study because these are the main inputs to estimate the PSCs. There are different approaches to retrieving PSCs from remote sensing data. They are usually classified in three classes (IOCCG, 2014; BRACHER et al., 2017):

a) Abundance-based approaches: rely on remotely-sensed Chl*a* or other proxy of abundance based on (phytoplankon absorption coefficient, for instance),

10

they usually retrieve the dominant PSC or multiple PSCs. It is easy to apply because it is based on relatively simple equations. Some of these algorithms assume that there is a maximum asymptotic Chl$a$ for small cells above which Chl$a$ is attributed to large cells (CHISHOLM, 1992). The drawback of this approach is that it can not predict a changing relationship between PSCs and Chl$a$ in a changing ocean (BREWIN et al., 2011b; BRACHER et al., 2017).

b) Spectral-based approaches: there are different algorithms for this approach. They differ in the inputs, receiving remote sensing reflectance (or normalized radiance) (ALVAIN et al., 2005; ALVAIN et al., 2008; LI et al., 2013), absorption (CIOTTI et al., 2002; CIOTTI; BRICAUD, 2006; DEVRED et al., 2006), and backscattering (KOSTADINOV et al., 2009; KOSTADINOV et al., 2016). And having as output, size classes fraction (DEVRED et al., 2006), size index (CIOTTI et al., 2002), multiple taxa (ALVAIN et al., 2005), single taxa (WESTBERRY et al., 2005). Differently from the abundance-based approaches they are not based on an assumption of static relation between Chl$a$ and cells sizes, presumably being able to capture changes in these relations.

c) Ecological approaches: receive as input spectral data and Chl$a$ in addition to environmental variables, such as sea surface temperature (SST), wind speed, photosynthetically active radiation (PAR), mixed layer depth (MLD), location and season (RAITSOS et al., 2008; PALACZ et al., 2013). The idea is that by training the algorithm with these ecologically relevant data it will be able to give more accurate results for phytoplankton groups.

Specifically focusing on the retrieval of phytoplankton groups, Werdell et al. (2014) investigated the ambiguities in the inverse IOP models. They also investigated which conditions make them perform poorly and how they affected the performance in retrieving two phytoplankton groups (diatoms and the dinoflagellate, *Noctiluca miliaris*). The authors found out that a hyperspectral sensor reduced the root mean square error (RMSE) of the modelled IOPs, and consequently, the estimation of the abundance of the studied groups. They also found that the depth of the bloom (maximum Chl$a$ depth) highly influenced the bias observed in the results.

Intercomparison between the various approaches and algorithms are complicated by the differences on their results, and input data for the estimations, as mentioned previously. Despite this challenge, Brewin et al. (2011b) were able to adapt the

output results of the techniques so all of them retrieved the dominant PSC from satellite remote sensing. For the *in situ* comparison, the authors used HPLC (from different parts of the world) and phytoplankton cell counts (in the North Atlantic) databases, in contrast to most of the models, which usually only use HPLC for validation.

For the intercomparison, the authors adopted two criteria: (1) probability of detection, based on the scoring technique, i.e., giving different scores for correct classification, near-correct classification and incorrect classification; and (2) misclassification matrices, i.e., classifying the errors in two types: error of omission, when a satellite prediction fails to recognize a phytoplankton size class identified *in situ*, and error of commission when a satellite prediction incorrectly identifies a pixel as a different phytoplankton size class. Their results indicated that all the approaches (spectral, ecological and abundance-based approaches) presented a similar accuracy. However, each model had its performances varying according to the size class considered, the input satellite data sources and *in situ* validation data types. The authors highlighted that improving *in situ* observations would benefit ongoing research in this field.

2.2.1 Ocean colour space missions

Satellite-based ocean-colour remote sensing is dependent on the radiometers, sensors able to measure the water-leaving radiance. Technical features of the sensor data collection evolved since the first ocean-colour sensor was launched: (1) temporal resolution, how often they cover the same area, (2) spatial resolution, the spatial size of the pixel sampled by the sensor, (3) spectral resolution, number of electromagnetic wavelengths intervals which the sensor is able to sample, (4) radiometric resolution, number of grey-scale values, i.e., digital numbers the sensor can measure, and (5) signal-to-noise ratio.

The pioneer ocean-colour sensor was the Coastal Zone Color Scanner (CZCS) mission, launched in 1978 by the National Aeronautics and Space Administration (NASA), had only six bands (four of which were primarily designed for ocean colour) (NASA, 2019), and was used as a proof-of-concept for the remote sensing application for ocean biology (IOCCG, 2014), and it was first used to estimate chlorophyll-a concentration. The following missions were able to apply the learnt lessons from the CZCS to improve the technical aspects of the radiometers. Simultaneously, the understanding of ocean bio-optics improved, allowing the development of new approaches to ocean-colour retrievals (SATHYENDRANATH, 2000).

Different nations were able to launch their spectrally resolved ocean-colour radiometers into space since the CZCS was launched. Groom et al. (2019) reviewed the past, ongoing and future ocean-colour missions, their requirements and uses. The authors emphasized long-term data archive by different missions, such as the Aqua MODerate resolution Imaging Spectroradiometer onboard Aqua satellite (MODISA), operating since 2002, with 9 bands in the visible and centred in wavelengths especially projected to attend ocean-colour applications. Also, future missions, with hyperspectral sensors, which, as indicated by the authors, are the future in ocean-colour satellite sensors. The NASA Plankton Aerosol Cloud ocean Ecosystem mission (PACE), is an example and is expected to be launched in 2023/2024 (https://pace.oceansciences.org/mission.htm). The hyperspectral sensor promises to improve the phytoplankton functional types studies, by giving signals in wavelengths that were not provided by the previous sensors.

However, the hyperspectral data are not available for time series/climate change studies yet. Time-series studies require long-term data. Thus, initiatives similar to the European Space Agency Climate Change Initiative (ESA-CCI), which merged the data archive from different sensors (i.e., Sea-viewing Wide Field-of-View Sensor (SeaWiFS), MODIS, Medium Resolution Imaging Spectrometer (MERIS), Visible Infrared Imaging Radiometer Suite (VIIRS), Ocean and Land Colour Instrument (OLCI)) to produce a long and comparable database are mandatory for climate change studies. The challenges to such a merged database pass through standardizing the algorithms to atmosphere correction and retrieving the geophysical parameters, complicated by the differences between the sensors (SATHYENDRANATH et al., 2017).

Considering these challenges, and reducing their effect, a merged database is already available as mentioned above by ESA-CCI (see Sathyendranath et al. (2012) and Brewin et al. (2015b)). According to Groom et al. (2019), by 2029 the ocean-colour record "will start to be of sufficient duration to discriminate climate change impacts from natural variability, at least in some regions". Thus, past, current and future missions will form a continuous dataset of ocean-colour observations, aiding researchers to identify the changes in earth's climate and in the biology of marine ecosystems.

3 THE CHALLENGES OF FITTING PHYTOPLANKTON SIZE CLASS MODELS IN OPTICALLY-COMPLEX WATERS

3.1 Introduction

Coastal zones are, in general, intensively used for fishing, aquaculture, leisure, and transport of cargo. They are more vulnerable in sea level rise scenarios, with an increase in erosion and floods (PORTNER et al., 2019). Despite the small surface cover, they are composed of highly productive ecosystems and their role in the carbon biogeochemical cycle has been increasingly investigated (BORGES et al., 2005; DAI et al., 2022). The ecosystem's diversity and dynamics are a challenge for remote sensing studies due to intricate spatial distribution. The optical active constituents in these waters makes it even more complex.

In the initial phase of bio-optical model development, the models were usually focused on Case-1 waters, defined as waters in which all the optically-relevant constituents co-vary with phytoplankton (MOREL; PRIEUR, 1977), whereas Case-2 waters were defined as optically more complex where the optical constituents can vary independently of phytoplankton. Case-1 waters are usually open-ocean waters, even though the waters in particular locations can transition between Case-1 and Case-2 over time (LEE et al., 2006). More recently, several studies have broadened this classification, by applying the concept of optical water types (OWTs) which uses the reflectance spectra to classify the optical classes (MÉLIN; VANTREPOTTE, 2015; VANTREPOTTE et al., 2012). This approach emphasizes the diversity of OWTs in the coastal waters and how it can be used for improving the performance of semi-analytical algorithms for optically complex waters (VANTREPOTTE et al., 2012).

Semi-analytical algorithms for retrieval of absorption coefficients and particulate backscattering and empirical algorithms for retrieval of Chl*a* — a common input variable in models used to estimate phytoplankton size classes (PSCs) — from remote-sensing reflectance from either *in situ* or satellite data are known to perform better in Case-1 waters (DEFOIN-PLATEL; CHAMI, 2007). Most of the PSC models were also developed for application in open-ocean waters (BREWIN et al., 2010; DEVRED et al., 2011; SATHYENDRANATH et al., 2001; UITZ et al., 2006), though some models have been successfully applied to more optically-complex waters, such as coastal and continental shelf waters (LIU et al., 2021; GITTINGS et al., 2019; SUN et al., 2019; SUN et al., 2017; TURNER et al., 2021). In the latter cases, the authors usually carried out the regional tuning of satellite-retrieval algorithms for Chl*a* and PSCs. Despite such efforts, satellite-based methods for the retrieval of PSCs in optically

complex waters remain a challenge.

Bracher et al. (2017) described the lack of regional capability of PSC models as a gap, and identified the weak signal of the phytoplankton absorption in high absorbing waters (high CDOM) and the interference of high scattering conditions by non-algal particles as the main challenge, requiring local tuning and validation, usually not feasible due to the lack of *in situ* data or high uncertainties in the retrieved data. However, the possibility that the phytoplankton communities themselves might be different between coastal and open-ocean waters, with associated changes in their optical properties, thereby contributing to the poor performance of open-ocean algorithms in coastal waters, has not received much attention (but see Catlett and Siegel (2018) and Babin et al. (2003)).

Despite these obstacles, interest has been growing in using satellite data to monitor regional ecosystems to manage ecological services and model ecosystems for forecasting the effects of climate change (SATHYENDRANATH et al., 2017). At the same time, the advent of new satellite sensors—such as the Italian recently launched (March 2019), PRISMA (Prototype Research Instruments and Space Mission technology Advancement), the German missions: DESIS (DLR Earth Sensing Spectrometer) launched in June 2018 and EnMAP Environmental Mapping and Analysis Program launched in April 2022, and the US mission: PACE (Plankton, Aerosol, Cloud, ocean Ecosystem) with the sensor OCI (Ocean Colour Instrument), with high spectral resolution (hyperspectral missions)—holds the promise of better algorithms to monitor more complex waters, pushing the community to develop improved algorithms that would exploit the capabilities of hyperspectral sensors.

Considering this challenge, in this study, we discuss the possibility of aggregating global coastal-water phytoplankton diagnostic pigments samples to fit PSC models. The driving questions of this study were: (a) Are the pigments composition of open-ocean and coastal waters statistically different? (b) Are open-ocean waters more similar as a group than coastal waters regarding the phytoplankton diagnostic pigments? (c) Does the empirical Chl*a* algorithm perform better for open-ocean waters than for coastal water? (d) How do PSC model's parameters differ when fitted to open-ocean and coastal waters data?

The specific objectives are the following: (1) to analyse the differences in pigment characteristics between open-ocean and coastal waters; (2) test if open-ocean waters are statistically more similar as a group than coastal waters; (3) test if the satellite-based Chl*a* estimations from Aqua MODerate Imaging Spectrometer (MODIS) have

lower uncertainties in open-ocean waters than in coastal waters, and (4) test the differences in PSC model's parameters for the open-ocean and coastal waters, (BREWIN et al., 2010) and evaluate their performances.

3.2 Materials and methods

We used HPLC pigments data from 1989 to 2019 extracted from the NASA SeaBASS archive (WERDELL et al., 2003) for the study. Samples collected more than 50 km from the land were classified as open-ocean, following Stock and Subramaniam (2020) and samples collected within 50 km from the land were classified as coastal waters. The 50 km threshold was used as a first approach, but we acknowledge that other partitioning strategy could also be tested, such as, bathymetry (e.g., the 200m depth used by Brewin et al. (2015) to remove coastal and continental shelf waters from the analysed dataset), or divisions based on ecological and oceanographic dynamics, for instance, the Longhurst provinces (LONGHURST, 1998; LONGHURST, 2007).

Samples collected from depths greater than 10 m from the surface, located inland, and with Chl$a$ greater than 1000 mg·m^{-3} were not considered in the analysis. TChl$a$ lower than 0.0001 mg ·m^{-3}, the detection limit of HPLC analysis were set as zero. In addition, when there were data from the same date and geographic location only those closest to the surface were retained, to avoid duplicates. After applying these criteria, the dataset consisted of a total of 10100 samples, with 4668 samples from coastal waters and 5432 from open-ocean waters (Figure 3.1).We used R programming language for the statistical analysis. For organizing and extracting the pigment data from SeaBASS and for match-ups we adapted the codes shared by Stock and Subramaniam (2020).

Aiken et al. (2009) quality control steps for HPLC data require all carotenoids and accessory pigments to perform the regression and differences with TChl$a$ which are used as quality assurance for the extracted pigments. This step was not performed in this study, because many of the datasets shared on SeaBASS did not provide all the pigments required. However, here we relied on the quality assurance provided by the data providers and the quality check described by Werdell et al. (2003).

Figure 3.1 - Map with the location of the HPLC SeaBASS data used in this study, subset in (A) coastal (less than 50km from land) and (B) open-ocean waters (more than 50km from land).
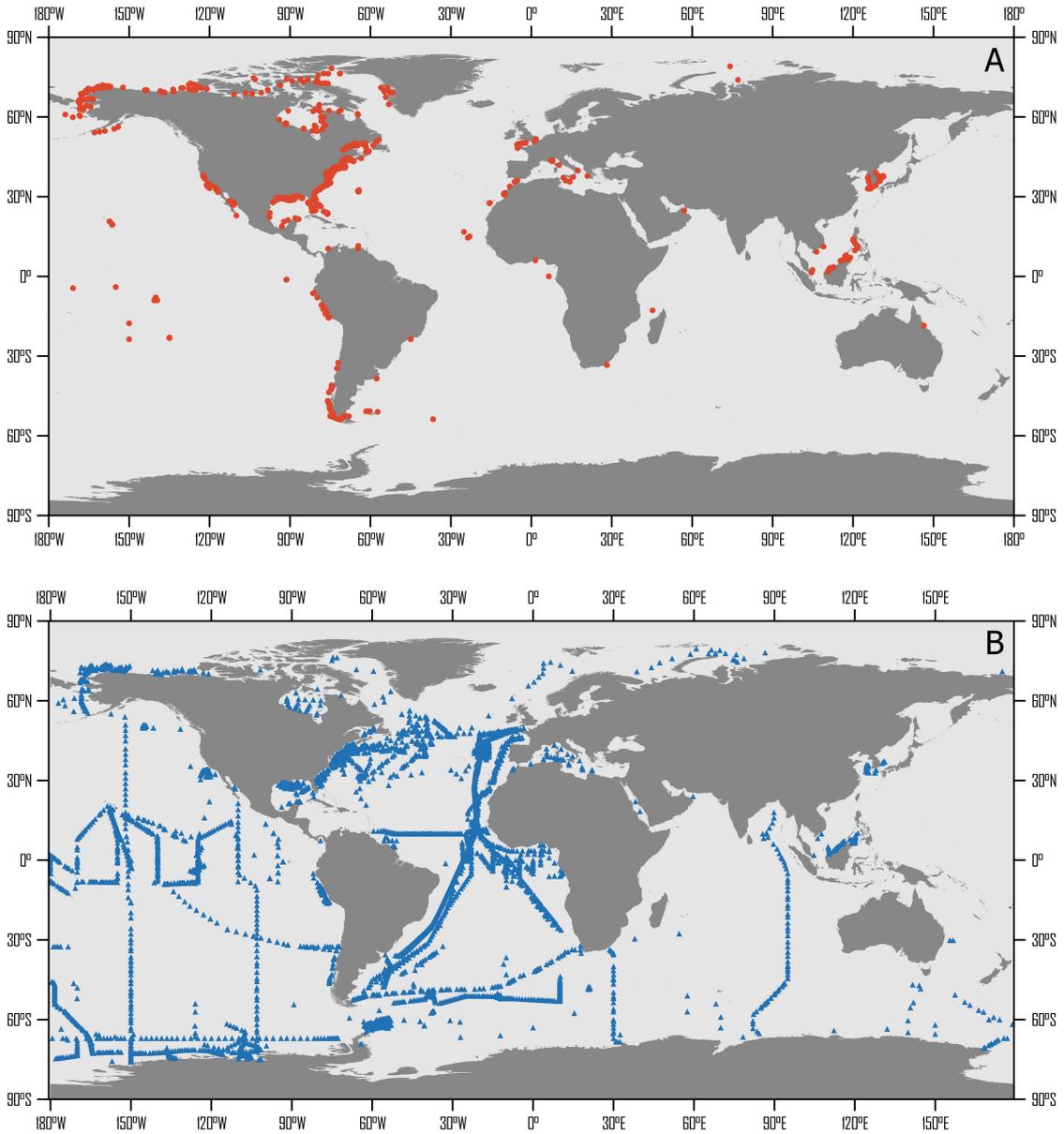
3.2.1 Diagnostic pigments

Considering that not all the pigments presented a normal distribution and that T-test requires a normal distribution to be applied, we applied both (T-test and

Wilcoxon test which can be applied to different distributions) to diagnostic pigments (DPs) normalized by total chlorophyll-a (TChl$a$) to compare coastal and open-ocean waters.

DPs are the pigments used in the equations defined by Vidussi et al. (2001) to estimate phytoplankton size classes, and are composed by the following seven pigments: fucoxanthin (Fuco), peridinin (Perid), zeaxanthin (Zea), total chlorophyll-a (TChl$a$) and total chlorophyll-b (TChl$b$), alloxanthin (Allo), 19'- hexonoyloxyfucoxanthin (Hex-fuco) and 19'- butanoyloxyfucoxanthin (But-fuco). TChl$a$ is the sum of monovinyl chlorophyll-a, divinyl-chlorophyll-a, chlorophyllide-a, chlorophyll-a allomers and epimers, this is important for the HPLC pigments quantification, for now on when we present the comparison between pigments, we use TChl$a$, but when we present the validation with satellite data, we use Chl$a$.

3.2.1.1 Cluster analysis and the elbow method

Cluster analysis (k-means) combined with the elbow method was applied to test if coastal waters presented more differences within clusters than those of the open-ocean waters. K-means was applied successively to the data set (diagnostic pigments normalized by TChl$a$) varying the number of clusters from 2 to 15 groups. Then the total distance within the clusters, represented by the total sum of squares of the clusters, were plotted against the number of clusters. A sharp reduction in the total sum of squares by increasing the number of clusters would indicate that the ideal number of clusters was achieved.

3.2.1.2 Principal component analysis

Principal component analyses are usually applied to reduce the dimensionality of a multivariate data set; however, it is also useful to provide an idea of the variables contributing the most to the data variability. In this study, it was applied to open-ocean and coastal-water subsets of diagnostic pigments and the divinyl chlorophyll-b concentrations. All pigments were normalized by TChl$a$, scaled, and centred previously to the analysis.

3.2.2 Match-ups

MODIS-Aqua was chosen for the match-ups because it provided the longest consistent source of ocean colour data from the same sensor, from 2002 to the present. However, merged databases from different ocean colour missions, such as ESA OC-CCI and Gloub-Colour which have applied statistical methods on the merging to

standardize the data are equally good candidates for this type of analysis and should be considered in future studies.

There are different match-ups strategies when considering PSC algorithms. The recommended approach is to use 3h as time window and 3 x 3 pixels as spatial window to MODIS in 1.1km nominal spatial resolution considering the *in situ* location (BAILEY; WERDELL, 2006). However, as discussed by Stock and Subramaniam (2020), to increase the number of match-ups, less strict criteria have been consistently applied. In this study, we used the same-calendar-date (keeping a 24 h window) NASA MODIS Level 3 Chl*a* with 4 km nominal spatial resolution and used a 3 x 3 pixel spatial window, obtaining the median value. Points which presented less than 3 valid pixels or where one or more pixels in the window presented 20% or more of difference from the median were excluded, following Stock and Subramaniam (2020).

For the coastal water the 4 km spatial resolution and the 24 h time window can be an additional source of uncertainty, due to the proximity with the land that could contaminate the pixel signal, the patchy nature of the distribution of the components on the sea surface and the short term variability more usual in coastal waters. This will be further discussed later on the text.

These data were compared with *in situ* Chl*a* for both subsets to test if there is a significant difference between the performances of remote sensing retrievals for coastal and open-ocean waters. NASA ocean colour default Chl*a* product is a merged algorithm composed of OC3M and CI (HU et al., 2012). The transition between CI and OC3M occurs at the range of Chl*a* between 0.15 and 0.2 mg· m^{-3}, when the CI is applied (https://oceancolor.gsfc.nasa.gov/atbd/chlor_a/), from this point we will refer to the default empirical chlorophyll-*a* algorithm of NASA as OCx.

For the open-ocean waters subset, to avoid non-open ocean conditions *in situ* and estimated Chl*a* higher than 3 mg·m^{-3} were excluded from the analysis, following Stock and Subramaniam (2020). This criterion was used only for the match-ups analysis, resulting in the removal of 9 match-ups, less than 1% of the match-ups. Thus we considered it did not substantially affect the statistical analysis considering the differences between coastal and open-ocean waters performance of Chl*a* empirical algorithm. After the application of this criteria a total of 948 match-ups were obtained, 401 for coastal and 547 for open-ocean waters (Figure 3.2).

Figure 3.2 - Map with the location of the match-ups obtained from MODISA 4 km nominal
spatial resolution for the same-calendar-date (A) coastal (less than 50km from
land) with 401 match-ups and (B) open-ocean waters (more than 50km from
land) with 547 match-ups.
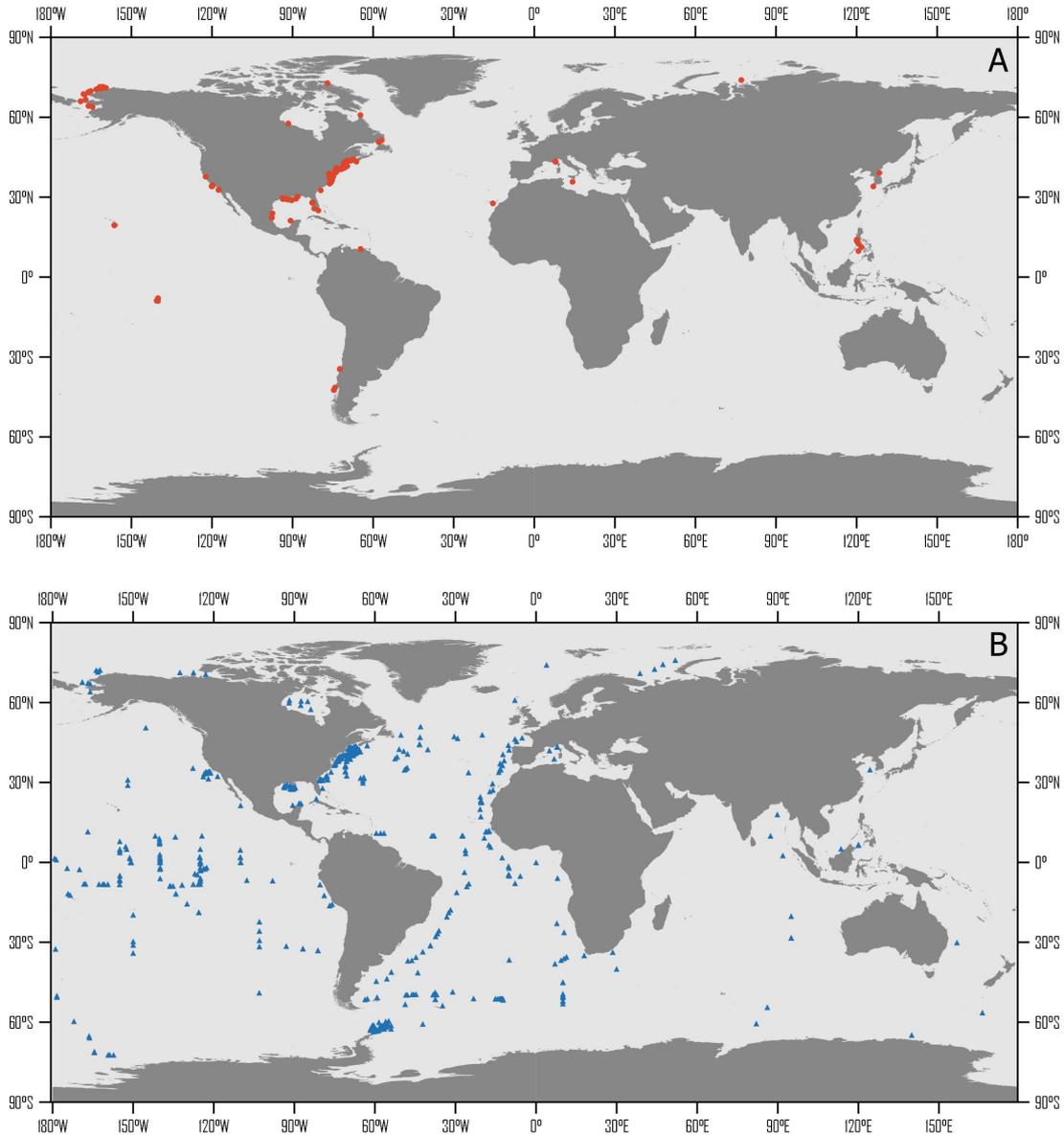
SOURCE: Author's production.

3.2.3 Statistical metrics

The statistical metrics applied to compare the match-ups and the *in situ* validation of the models were the determination coefficient (R^2), correlation coefficient (ρ), root-mean-square error (RMSE), median absolute difference (MAD), mean absolute error (MAE), and bias. These metrics were calculated using the following equations:

$$RMSE = 10^{\sqrt{\frac{\sum_{i=1}^{n}(\log_{10}(M_i)-\log 10(O_i))^2}{n}}} \tag{3.1}$$

$$MAD = 10^{median(|\log_{10}(M_i)-\log_{10}(O_i)|)} \tag{3.2}$$

$$MAE = 10^{\left(\frac{\sum_{i=1}^{n}|\log_{10}(M_i)-\log_{10}(O_i)|}{n}\right)} \tag{3.3}$$

$$bias = 10^{\left(\frac{\sum_{i=1}^{n}\log_{10}(M_i)-\log_{10}(O_i)}{n}\right)} \tag{3.4}$$

Where M_i and O_i are the modeled (satellite) and observed values, respectively, and n is the number of observations. When the metrics were applied to Chl*a* data the $\log_{10}$ transformation was applied, as presented in the Equations 5.5-5.8. When it was applied to percentages no log-transformation was applied.

Additionally, sea surface temperature (SST) from the Group for High Resolution Sea Surface Temperature (GHRSST), Level 4 multiscale ultrahigh resolution (MUR) product, with daily data with 0.01 degree spatial resolution (1.1km), for the match-ups dates were also downloaded (NASA, 2015) [1] and re-sampled in order to match the spatial resolution of the Chl*a* imagery. The ecology of phytoplankton and previous studies indicated SST as an important variable in phytoplankton physiology, growth, and abundance (GEIDER, 1987; GEIDER et al., 1997; STRAMSKI et al., 2002; BOUMAN et al., 2003; MARÃNÓN et al., 2014). Thus, this was used to discuss the possible relationship between the SST and the performance of the PSC model, which has already been reported by previous studies (BREWIN et al., 2017; TURNER et al., 2021; STOCK; SUBRAMANIAM, 2020).

---

[1] The GHRSST uses the SST from different sensors to produce a daily high spatial resolution products. For more details check ghrsst.org/ghrsst-data-services/products/.

3.2.4 Model parameterization

Brewin et al. (2010)'s model has been parameterized and tested for different regions with consistent results. To evaluate the feasibility of a generalized coastal-water PSC model, we parameterized the Brewin et al. (2010)'s model to the subsets of coastal and open-ocean waters. A validation dataset (8.6% and 10% of the SeaBASS data for coastal and open-ocean waters, respectively) was not included in the dataset used for the parameterization (the remaining data, around 90% of the SeaBASS initial dataset). We used the match-up dataset for the *in situ* and satellite validation of PSCs.

The satellite PSC outputs were compared to the *in situ* PSCs determined from the pigment dataset following Uitz et al. (2006). This method uses diagnostic pigments (DP) ratios (i.e., DP/$\sum$DP) to determine the fraction of each size class (micro, nano and picophytoplankton) for the Chl*a* (see Chapter 4 for more details). Following Uitz et al. (2006), we tuned the weights for each subset (open-ocean and coastal waters) using a multiple linear regression model and compared the statistics to the Uitz et al. (2006) results. The weights (W_i) were then used to estimate the C, using the product of the weight and the concentration of the respective pigment (P_i). Following Equation 3.5.

$$C \ = \ \sum_{i=1}^{7} W_i \cdot P_i \qquad (3.5)$$

Where C is the chlorophyll-*a* concentration estimated, W_i is the weight attributed to each pigment i, and P_i is the pigment concentration. The pigments in this sequence: (1) Fuco, (2) Perid, (3) Hex-Fuco, (4) But-fuco, (5) Allo, (6) TChl*b*, and (7) Zea.

Part of the Fuco (P_1) is due to the nanophytoplankton pool, since Fuco is also present in nanophytoplankton such as prymnesiophytes and chrysophytes. Thus, the correction described in Devred et al. (2011) was applied (see Equation 3.6-3.9). When the $P_{1,n}$ was equal to P_1, the value of $P_{1,n}$ was set as zero.

We also applied the correction described by Brewin et al. (2010) to account for Hex-fuco that is part of the pico-eukaryotes in ultra-oligotrophic conditions. We used the suggested threshold of Chl*a*<0.08 as described by Brewin et al. (2014a). Following the Equations 3.9 and 3.10.

$$P_{1,n} = 10^{(0.680 \cdot \log 10(P_3) + 0.680 \cdot \log 10(P_4))} \tag{3.6}$$

$$P_{1,m} = P_1 - P_{1,n} \tag{3.7}$$

$$F_m = \frac{P_{1,m} \cdot W_1 + P_2 \cdot W_2}{C} \tag{3.8}$$

$$F_n = \begin{cases} \frac{P_{1,n} \cdot W_1 + \sum_{i=3}^{5} W_i \cdot P_i}{C}, & \text{if } C > 0.08 \; mg \cdot m^{-3} \\ \frac{12.5 \cdot C \cdot W_3 \cdot P_3}{C} + \frac{P_{1,n} \cdot W_1 + \sum_{i=4}^{5} W_i \cdot P_i}{C}, & \text{if } C < 0.08 \; mg \cdot m^{-3} \end{cases} \tag{3.9}$$

and

$$F_p = \begin{cases} \frac{\sum_{i=6}^{7} W_i \cdot P_i}{C}, & \text{if } C > 0.08 \; mg \cdot m^{-3} \\ \frac{(-12.5 \cdot C + 1) \cdot W_3 \cdot P_3}{C} + \frac{\sum_{i=6}^{7} W_i \cdot P_i}{C}, & \text{if } C < 0.08 \; mg \cdot m^{-3} \end{cases} \tag{3.10}$$

The tuned weights of the DPs ratios were then used to estimate the PSCs fractions (following Equations 3.8-3.10), which were then multiplied by the Chl$a$ to estimate the chlorophyll for each size class, and used as input to parameterize Brewin et al. (2010)'s model. Thus, the chlorophyll-a concentrations of the picophytoplankton (C_p), nanophytoplankton (C_n) and microphytoplankton (C_m) and the total chlorophyll-a concentration ($Chla$) were fitted to the Equations 3.11 and 3.12 using a standard, non-linear least-squared fitting procedure (Levenberg-Marquardt). The fitting can be done with the chlorophyll concentrations in the logarithm transformed values or with the fractions, with no need to logarithmic transformation. Here both approaches were tested.

$$C_p = C_p^{max} \cdot [1 - exp(-S_p \cdot Chla)] \tag{3.11}$$

$$C_{n,p} = C_{max}^{n,p} \cdot [1 - exp(-S_{n,p} \cdot Chla)] \tag{3.12}$$

Where $C_{n,p}$ is the chlorophyll-a concentration of the pico and nanophytoplankton combined, C_p^{max} and $C_{max}^{n,p}$ are the chlorophyll-a asymptotic maximum associated with the picophytoplankton, and the combined nano and picophytoplankton. S_p and $S_{n,p}$ are the initial slopes.

These parameters were used to estimate the micro and nanophytoplankton fractions

using the following equations:

$$C_m = Chla - C_{n,p} \tag{3.13}$$

$$C_n = C_{n,p} - C_p \tag{3.14}$$

The bootstrap method (EFRON; TIBSHIRANI, 1994) was used to estimate these parameters, by re-sampling and fitting the model 1000 times (with replacement), thus the variability of the estimated parameters were computed. The R packages boot and minpack.lm were used for applying the bootstrapping method (CANTY; RIPLEY, 2021) with the non-linear least square Levenberg-Marquart fitting to the data (ELZHOV et al., 2016). The bias, standard error and the values in the confidence interval of 95% were also obtained.

3.2.5   Schematic diagram

The Figure 4.3 presents an overview of the methodology used in this chapter, including the databases used (SeaBASS, MODIS-Aqua, MUR) and the processing steps.

Figure 3.3 - Schematic diagram of the processing and analysis of this chapter.
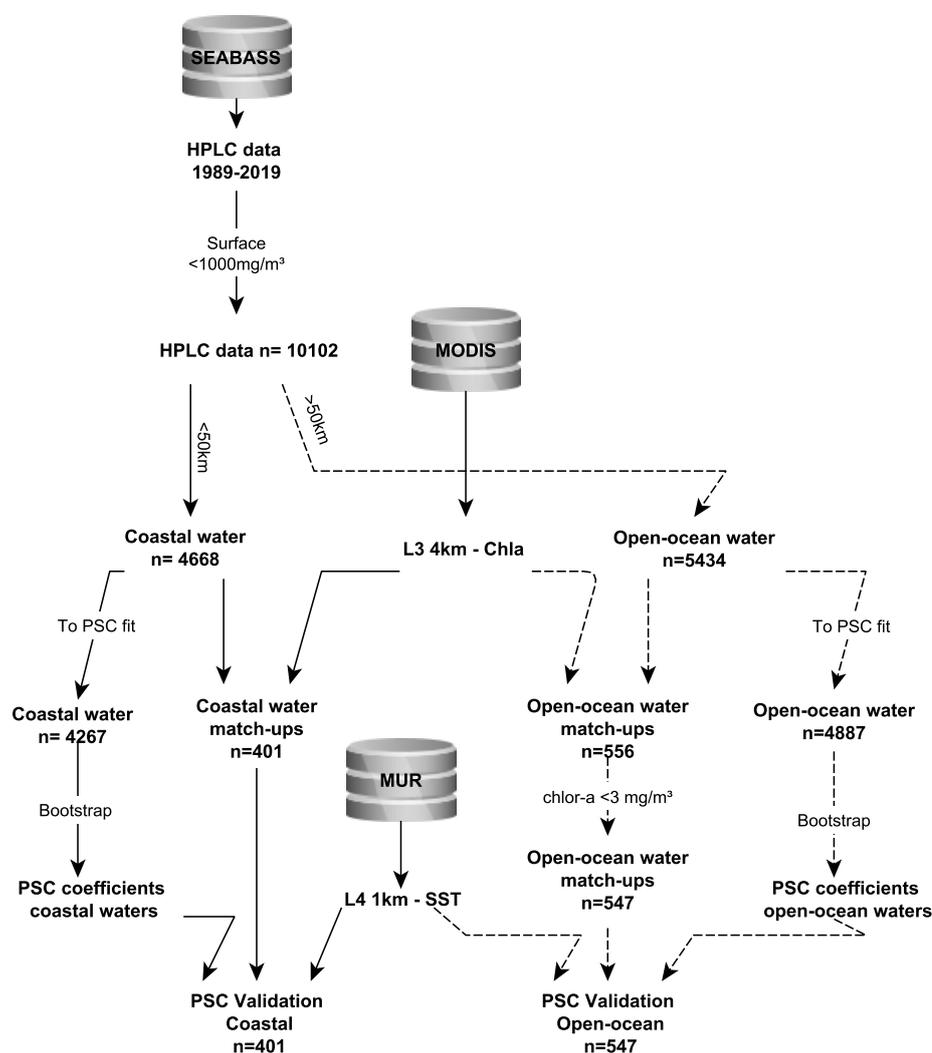
SOURCE: Author's production.

3.3 Results

The results of the statistical tests showed a significant difference (p-value<0.01) between coastal and open-ocean waters for all diagnostic pigments. Coastal waters presented higher concentrations of TChl*a* (Figure 3.4A) compared to open-ocean waters, with average of 14.96 ± 10.3 mg·m⁻³ and 0.55 ± 1.04 mg·m⁻³, for coastal and open-ocean waters, respectively. Fuco (Figure 3.4C) and Allo (Figure 3.4F)

normalized by TChl*a* were also higher in coastal waters; however, the histogram for these pigments showed more than one peak for open-ocean waters (multimodal and bimodal distribution, respectively, for Fuco and Allo) indicating potential subgroups. Zea, TChl*b*, But-fuco and Hexa-fuco presented higher contributions in open-ocean waters (Figure 3.4 B, E, F, G, and H, respectively). Zea presented a multimodal distribution for open-ocean waters (Figure 3.4B).

Table 3.1 - Diagnostic pigments concentrations and TChl*a* statistics in mg·m^{-3}, for coastal and open-ocean waters. Mean, minimum (min), maximum (max), standard deviation (Sd) and percentage of null concentrations (below the detection limit).

Pigments	Coast					Ocean				
	Mean	Min	Max	Sd	Zero(%)	Mean	Min	Max	Sd	Zero(%)
TChl*a*	4.74	0.02	330.3	10.3	0	0.59	0.0052	17.23	1.04	0
TChl*b*	0.2	0	5.8	0.32	6	0.04	0	1.13	0.07	15
Fuco	1.14	0	44.81	2.15	1	0.16	0	8.54	0.46	6
Zea	0.2	0	7.33	0.51	5	0.05	0	0.92	0.05	3
Perid	0.52	0	134.71	3.73	9	0.02	0	1.56	0.05	22
But-fuco	0.02	0	0.62	0.04	23	0.03	0	0.63	0.04	5
Hex-fuco	0.09	0	7.89	0.18	13	0.08	0	2.64	0.12	1
Allo	0.17	0	26.98	0.64	7	0.01	0	1.25	0.05	31

Figure 3.4 - Histogram of phytoplankton pigments considering coastal and open-ocean waters, pigments concentrations were normalized by total TChl$a$, thus they are dimensionless (except for TChl$a$ histogram).
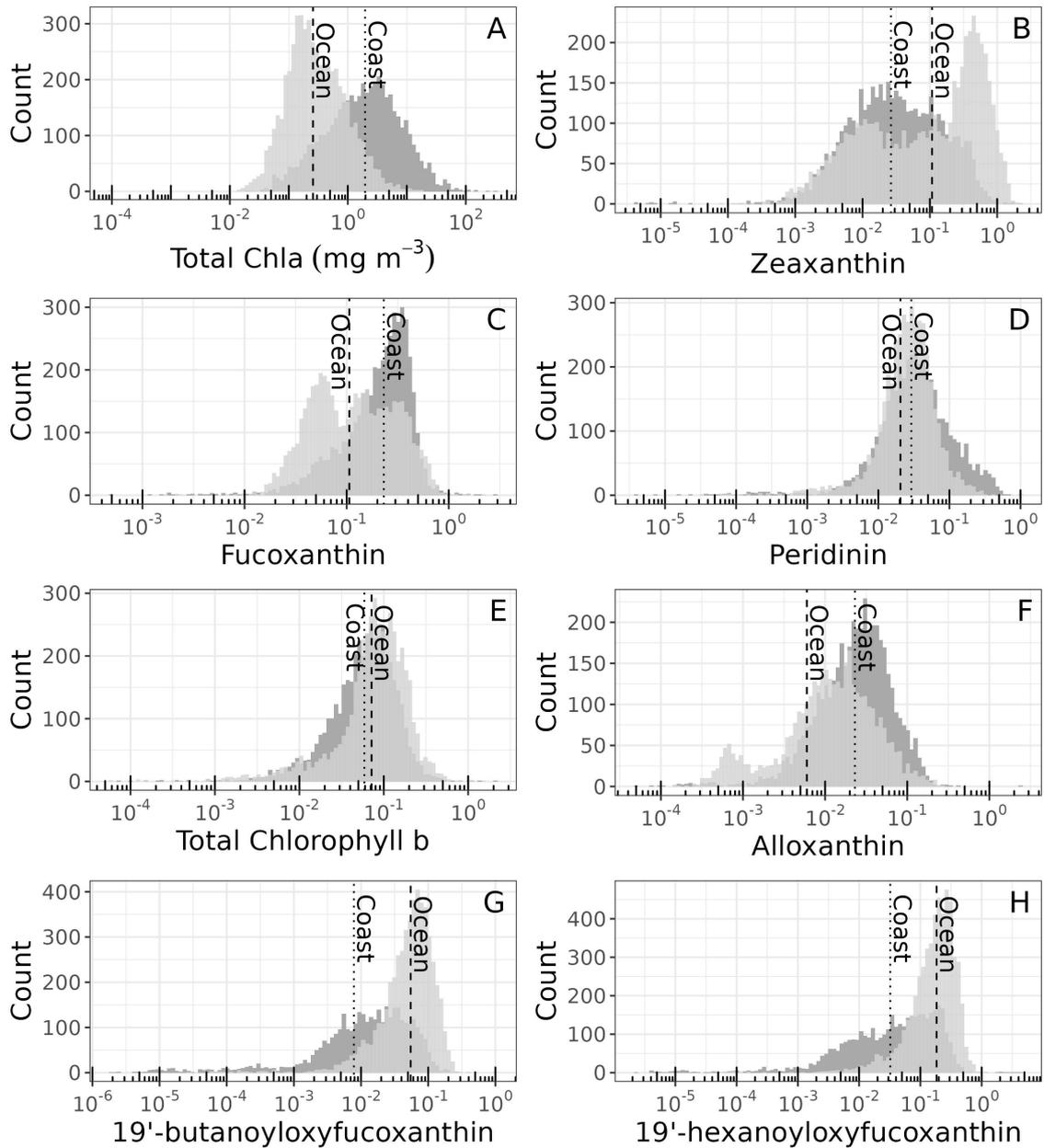
Light grey histograms represent open-ocean waters and dark grey coastal waters, vertical dashed black lines indicate the median TChl$a$ and pigments ratios for open-ocean waters, and vertical dotted black lines the median for coastal waters.

SOURCE: Author's production.

The k-means analysis and the elbow method indicated that the subset of open-ocean waters presented slightly higher distances within clusters than coastal waters (Figure 3.5). The maximum total sum of squares for coastal waters was 32669 and for open-ocean waters 38031, and 70707 when considering both subsets (dimensionless since they are representing the normalized concentrations of the pigments).

In addition, we applied the k-means clustering to the entire dataset considering two groups and then compared the groups separated regarding the distance from land (coastal and open-ocean waters). For the majority of the samples (75%) the geography clustering matched the pigments clustering (k-means). However, 18% of the coastal waters and 32% of the open-ocean waters presented different clustering results, i.e. were identified by the distance from the coast as one group and when considering the pigments k-means were clustered with the other group.

Figure 3.5 - Differences within groups and number of clusters for open-ocean and coastal waters subsets and both.
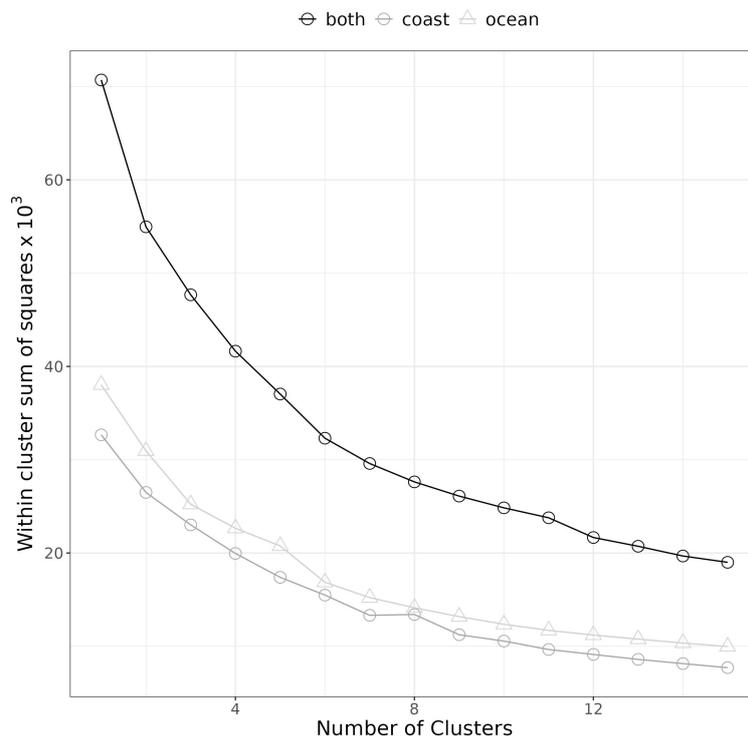
SOURCE: Author's production.

3.3.1   Diagnostic pigments data distribution

The two first principal components explained 49% of the variability of the diagnostic pigments ratios for open-ocean waters and 41% for coastal waters. Different patterns were observed for coastal and open-ocean waters. When considering both groups (Figure 3.6A) it is possible to distinguish the most parts of the coastal and open-ocean waters, but also latitude differences, with low latitudes more influenced by Zea and high latitudes by Fuco. In the ocean subset it is possible to clearly distinguish high and low latitudes (Figure 3.6B), while in the coastal water this pattern is not clear (Figure 3.6C), most part of the samples seems to be aggregated.

Some pigments ratios were aggregated in the PCA, which means that at least for the first and second principal components they vary in a similar pattern. For open-ocean waters these pigments are: (i) But-fuco, (ii) TChl*b*, and (iii) DVChl*b*. For the coastal waters the aggregations were observed for (i) Allo and Perid; (ii) Hex-fuco and But-fuco, and (iii) Zea and DvChl*b*.

Figure 3.6 - Principal components analysis results.
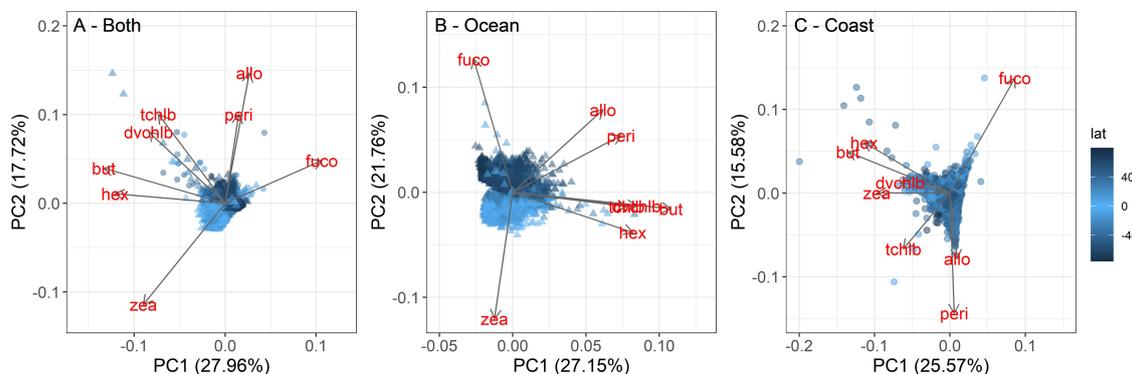
The colour gradient represents the latitudes, dark blue is the high latitudes (towards the poles) and light blue is the low latitudes (towards the equator). Triangle represents open-ocean water and circle coastal-water samples. The red labels in the plotting area are the pigments ratios variability axes for each set of data in the PCA, the longer the axes the higher the variability.

SOURCE: Author's production.

3.3.2 Match-ups performance

The match-ups results considering coastal and open-ocean waters presented an overall good agreement for both subsets, with R^2 0.67 of and 0.70 and RMSE of 1.83 and 2.15, for open-ocean and coastal waters, respectively (Figure 3.7 and Table 3.2). However, when the differences between the satellite estimations and the *in situ* measurements are analysed, the open-ocean waters presented lower errors than coastal waters— 40% open-ocean and 43% coastal, considering the median absolute difference. The same is observed in the mean absolute error, which indicated 55% and 72% error for open-ocean and coastal (both parameters are interpreted as the difference from 1). This indicates that coastal waters have more extreme values, i.e., conditions that made the Chl*a* estimations fail or overestimate, such as turbidity, dissolved organic matters and/or bottom reflectance in optically shallow waters—which were not removed in this overall analysis. The bias indicated an average overestimation of 7% for coastal waters and an underestimation of 9% for open-ocean waters (a bias of 1.07 indicates a overestimation of x1.07 i.e., 7% from the observed variable).

Figure 3.7 - Chlorophyll-a concentration match-ups for OCx of MODIS and HPLC from
SeaBASS.
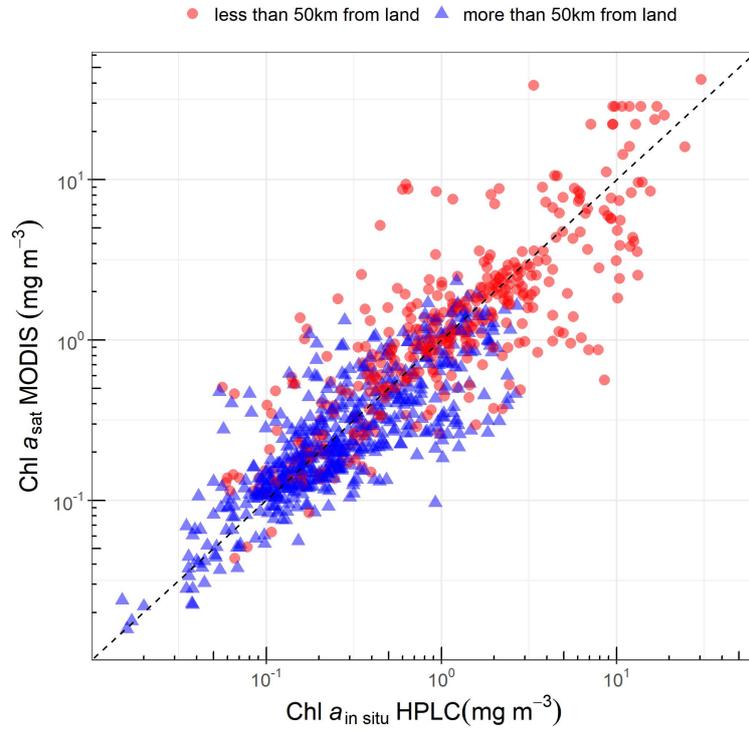
The dashed black line is the 1:1 line. Red circles were sampled less than 50km from land
and blue triangles were sampled more than 50km from land.

SOURCE: Author's production.

Table 3.2 - Match-ups statistics for coastal and open-ocean waters, with the coefficient of
determination (R^2), RMSE, MAD, MAE, Bias and number of match-ups (n).

	R^2	RMSE	MAD	MAE	Bias	n
Coast	0.70	2.15	1.43	1.72	1.07	401
Ocean	0.66	1.83	1.40	1.55	0.91	547
Both	0.76	1.97	1.41	1.62	0.97	948

3.3.3 PSC model parameters and performance

Before parameterizing the satellite remote sensing PSC model (BREWIN et al., 2010), we applied a parameterization to the *in situ* PSC model based on DP ratios (UITZ et al., 2006), as previously explained. The DP ratio weights have been tuned by different studies for global and regional datasets. We estimated the weights for open-ocean and coastal-water datasets, and compared their values with the values obtained by other studies (Table 3.3). The weights estimated for coastal waters were higher than the weights for open-ocean waters, except for the TChl*b* ratio. The pigment But-fuco ratio had negative values, to avoid this, the multiple linear regression (MLR) model was constrained to positive values, resulting in a null value for this pigment. The same occurred for both datasets (open-ocean and coastal) and this was also observed in the weights estimated by Chase et al. (2020). However, other studies opted for keeping negative values, such as Turner et al. (2021) for Allo.

Table 3.3 - Diagnostic pigments ratio weights for open-ocean, coastal and both types of waters compared to other studies. The multiple linear regression was performed constraining the weights to positive values. For brevity, in the table Turner et al. (2021) (T21), Chase et al. (2020) (C20), Brewin et al. (2014a) (B14a).(*) MLR estimated in this study.

DPs	Uitz	Ocean*	Coast*	Both*	T21	C20	B14a
Fuco	1.41	1.77	1.95	1.94	2.20	2.62	1.72
Perid	1.41	1.17	1.99	1.99	1.08	1.32	1.27
Hex-fuco	1.27	0.93	1.09	0.91	0.86	0.87	0.68
But-fuco	0.35	0	0	0	3.63	0	1.42
Allo	0.6	3.70	4.53	4.53	-0.10	2.64	4.96
TChl*b*	1.01	1.62	1.38	1.39	1.21	0.94	0.81
Zea	0.86	0.58	1.82	1.82	0.99	1.52	1.28

The TChl*a* estimated using the DP ratios and weights following Uitz et al. (2006) was used to compare with the measured TChl*a* of the *in situ* dataset and repa-

rameterization, to check if the new tuning improved the statistics and reduced the errors (Table 3.4). All the coefficients of determination were higher than 0.9, but the median absolute difference and mean absolute error were lower when adopting the weights estimated with the multiple linear regression (MLR) reparameterization of this study. This reduction in the differences was higher for the coastal-water dataset, in which when adopting Uitz et al. (2006) weights presented a median absolute difference of 42%, and when adopting the tuned weights it was reduced to 9%, the same was observed in the bias which presented a an understimation of 31% with Uitz's weights and an overestimation of 4% with the MLR weights. Considering this reduction in the differences, the tuned weights were adopted to determine the *in situ* PSCs for the present work.

Table 3.4 - Statistics comparing the DPs weights comparing the chlorophyll-*a* estimated by the weights and the chlorophyll-*a* concentration measured *in situ*. Considering (UITZ et al., 2006) weights and multiple linear regression fitting (MLR).

	R^2	MAD	MAE	Bias
Uitz	0.97	1.33	1.35	0.77
Uitz open-ocean	0.94	1.23	1.27	0.83
Uitz coastal	0.97	1.42	1.47	0.69
Both MLR	0.96	1.14	1.22	1.10
Open-ocean MLR	0.96	1.2	1.24	0.84
Coastal MLR	0.97	1.09	1.15	1.04

The bootstrapping results for Brewin et al. (2010)'s model re-parameterization to open-ocean and coastal waters are presented in Table 3.5. The results using the fitting applied to fractions presented values of D_i closer to 1 than when applying the fit to C_i in the logarithm transform version (data not shown), thus we opted to keep the fraction fit parameters. The maximum asymptotic values were higher for the coastal-water dataset with C_p^{max} of 0.371 mg·m⁻³ and $C_{max}^{p,n}$ of 1.397 mg·m⁻³. While the initial slopes were higher for open-ocean waters with S_p of 7.210 and $S_{p,n}$ of 1.291. All the estimated parameter were significant with p-value<0.001. Comparing the

parameters estimated in this study with the parameters from other studies (TURNER et al., 2021; LAMONT et al., 2018a; BREWIN et al., 2010; BREWIN et al., 2014b) its is possible to observe that almost all of them were closer to the parameters estimated for open-ocean waters, with low asymptotic maximum chlorophyll and high initial slopes (Table 3.6). Except for the C_p^{max} estimated with the size-fractionated filtration by Brewin et al. (2014b), which presented a C_p^{max} of 0.730 mg·m⁻³. The D_i values were higher than 0.5 for open-ocean and coastal waters, but the values for coastal water were lower, which as also observed in the results presented by Turner et al. (2021).

Table 3.5 - Bootstrapping statistics (1000 resampling with replacement) for the estimated PSC model parameters. The values in parenthesis are the 2.5% and 97.5% confidence intervals on the bootstrap parameter distribution. The bootstrap standard error (SE) of the estimations is also presented. All the parameters presented p-value<0.001.

Par	Ocean		Coast	
	Value	SE	Value	SE
C_p^{max}	0.107 (0.1021 to 0.1122)	0.003	0.371 (0.3260 to 0.4107)	0.0216
S_p	7.210 (6.710 to 7.683)	0.248	1.397 (1.157 to 1.642)	0.1236
$C_{p,n}^{max}$	0.687 (0.6409 to 0.7290)	0.0224	2.840 (2.427 to 3.232)	0.205
$S_{p,n}$	1.291 (1.201 to 1.385)	0.0470	0.216 (0.1787 to 0.2521)	0.0187

Table 3.6 - Comparison of parameters estimated for PSC model. SFF stands for size-fractionated filtration, the method used to estimate the size-fractionated Chl$a$. All the others used the diagnostic pigment analysis approach, with HPLC (high performance liquid chromatography). The D_i factor is the product of S_i and C_i^{max}, where i is p or p, n.(*) Parameters from all data fit, see Brewin et al. (2014b) supporting information ts01.

Ref	C_p^{max}	S_p	D_p	$C_{p,n}^{max}$	$S_{p,n}$	$D_{p,n}$
Ocean	0.107	7.210	0.772	0.687	1.291	0.886
Coast	0.371	1.397	0.518	2.840	0.216	0.613
Brewin et al. (2010)	0.107	6.801	0.727	1.057	0.851	0.899
Brewin et al. (2014b)*(SFF)	0.730	1.047	0.763	2.611	0.364	0.951
Brewin et al. (2014b)*	0.155	4.615	0.712	1.204	0.774	0.930
Turner et al. (2021)	0.15	3.6	0.54	0.81	0.97	0.786
Lamont et al. (2018a)	0.11	8.94	0.983	0.72	1.36	0.979

Using the validation subset of the dataset, i.e., the values that coincide with the MODIS match-ups, the reparameterized Brewin et al. (2010)'s model was applied to *in situ* TChl$a$ and satellite estimated Chl$a$ (OCx) and the results compared to the size fractions estimated by the diagnostic pigment analysis of the reparameterized Uitz et al. (2006) model. The Chl$a$ of each PSC presented better results for nano and microphytoplankton with R^2 >0.6 for *in situ* validation and >0.40 for satellite validation (considering open-ocean and coastal waters) (Table 3.7) than the results obtained for picophytoplankton.

The results for picophytoplankton chlorophyll concentrations were poor, with ρ<0.6, open-ocean waters presented the lowest correlation coefficient (0.29) for satellite validation, the bias indicated an overestimation of 16% for C_{pico} of open-ocean for *in situ* and satellite validation, respectively. It seems to be affected by the maximum limit established. It is possible to see in the scatter plot the dots being flatted in this upper limit for *in situ* (see Figures 3.8 A) and satellite estimations (3.9A). The mean absolute error was similar for picophytoplankton for open-ocean and coastal waters (around 90%). With the bias close to 1. However, this was not observed in the mean

absolute error of the microphytoplankton with values of 1.91 and 1.75, indicating an average error of 91% and 75%, for open-ocean and coastal water, respectively. The bias indicated an average overestimation of 39% and 29%, for open-ocean and coastal water. The scatter plots of the validation are also presented (Figure 3.8 and Figure 3.9).

For the fractional contributions, the algorithm performance were relative poor for all size classes (considering the correlation coefficient), especially for the fractional nanophytoplankton, which presented negative values for the coastal water dataset. Considering the MAE and the bias, it presented errors of, 12 to 22%, and bias ranging from -3 to 1% (notice that for fractions no log-transformation were used, thus the results are presented as fraction, unitless). The performances of the satellite-retrievals were poorer than *in situ* validation. Coastal and open-ocean performances for fractions were better for open-ocean waters with higher correlation coefficients and lower errors (Table 3.7 and plots not shown).

The performances reported in this study for *in situ* validation were very similar to the reported by other authors when applying the Brewin et al. (2010) model (see Table 3.8). The correlation coefficient reported were usually higher than 0.9 for C_m, presenting lower values for C_p, with values ranging from 0.40 (TURNER et al., 2021) to 0.64 (BREWIN et al., 2015), which were very similar to the values reported by this study. The RMSE varied from 0.24 to 0.34 (BREWIN et al., 2015) considering all the size classes, presenting a lower range than the reported in this study, which varied from 0.23 to 0.47.

It is possible to observe that low SST are related to overestimation of picophyto-plankton size classes (Figures 3.8 A and D; Figures 3.9 A and D). On the other hand, high sea surface temperature is related to the overestimation of microphytoplankton size classes, for open-ocean and coastal waters. It is illustrated by the residuals of the fractions versus SST (see Figure 3.10). The picophytoplankton is overestimated for temperatures lower than 10°C (Figures 3.10 A and D), which is also observed for pico and nanophytoplankton (Figures 3.10 A and D). The opposite pattern is observed in microphytoplankton (Figures 3.10 C and F).

Table 3.7 - Validation statistics of the PSC model. The Chl$a$ metrics were calculated in the $\log_{10}$-space. The fractions metrics were in linear space. RMSE is root-mean-squared error, MAE mean absolute error and bias.

Open-ocean waters

Variable	*in situ*				Satellite			
	ρ	RMSE	MAE	bias	ρ	RMSE	MAE	bias
C_p	0.32	2.95	1.90	1.16	0.29	2.98	2.0	1.16
C_n	0.90	1.70	1.47	1.12	0.70	2.40	1.89	1.04
C_m	0.90	2.26	1.91	1.39	0.78	2.96	2.34	1.2
F_p	0.69	0.17	0.13	-0.03	0.33	0.19	0.14	-0.01
F_n	0.32	0.15	0.12	0	0.13	0.16	0.13	0.01
F_m	0.57	0.2	0.15	0.02	0.56	0.2	0.15	0

Coastal waters

Variable	*in situ*				Satellite			
	ρ	RMSE	MAE	bias	ρ	RMSE	MAE	bias
C_p	0.60	2.44	1.92	1.18	0.48	2.69	2.08	1.25
C_n	0.80	2.614	2.02	1.13	0.64	3.32	2.45	1.22
C_m	0.94	2.21	1.75	1.29	0.81	3.38	2.53	1.38
F_p	0.61	0.16	0.12	0.01	0.47	0.18	0.14	0
F_n	-0.16	0.19	0.15	-0.02	-0.23	0.19	0.16	-0.02
F_m	0.52	0.23	0.2	0.01	0.35	0.26	0.22	0.01

Table 3.8 - Statistic metrics comparison for the validation of PSC. Considering this study and others. (*) To compare the RMSE from this study to others, here we did not applied the conversion back from log-space with the base 10.

Parameter	ρ	RMSE	N	Distribution	Reference
C_m	0.90	0.35*	547	Global	This study (Open-ocean)
C_n	0.90	0.23*	547	Global	This study (Open-ocean)
C_p	0.32	0.47*	547	Global	This study (Open-ocean)
C_m	0.94	0.34*	401	Global	This study (Coastal)
C_n	0.80	0.42*	401	Global	This study (Coastal)
C_p	0.60	0.39*	401	Global	This study (Coastal)
C_m	0.91	0.34	5841	Global	Brewin et al. (2015)
C_n	0.93	0.24	5841	Global	Brewin et al. (2015)
C_p	0.64	0.26	5841	Global	Brewin et al. (2015)
C_m	0.86		418	NW Atlantic Ocean	Turner et al. (2021)
C_n	0.67		418	NW Atlantic Ocean	Turner et al. (2021)
C_p	0.40		418	NW Atlantic Ocean	Turner et al. (2021)
C_m	0.98	0.30	374	Southern Africa	Lamont et al. (2018b)
C_n	0.86	0.33	374	Southern Africa	Lamont et al. (2018b)
C_p	0.45	0.30	374	Southern Africa	Lamont et al. (2018b)

Figure 3.8 - PSC validation for the *in situ* dataset, for open-ocean and coastal waters.
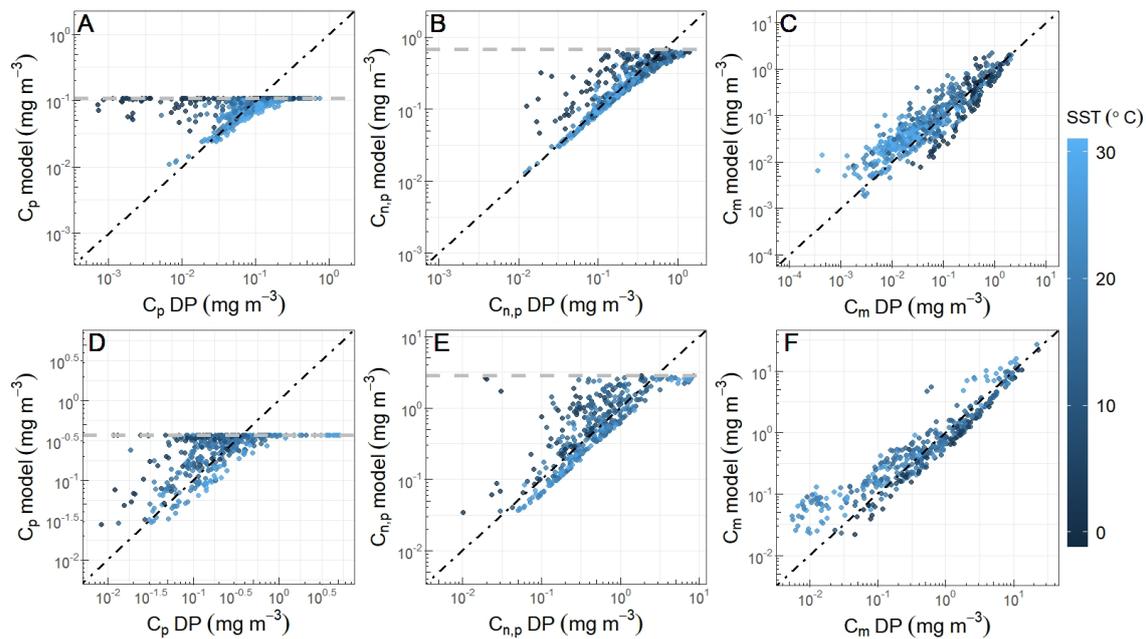
A, B and C are the plots for open-ocean waters, and D, E and F are the plots for coastal waters. The dashed-dotted black line is the 1:1 line and the dashed grey line is the model threshold for picophytoplankton (A and D) and pico and nanophytoplankton (B and E). The colour gradient scale represents the sea surface temperature in Celsius degrees.

SOURCE: Author's production.

Figure 3.9 - PSC validation for the satellite match-ups, for open-ocean and coastal waters.
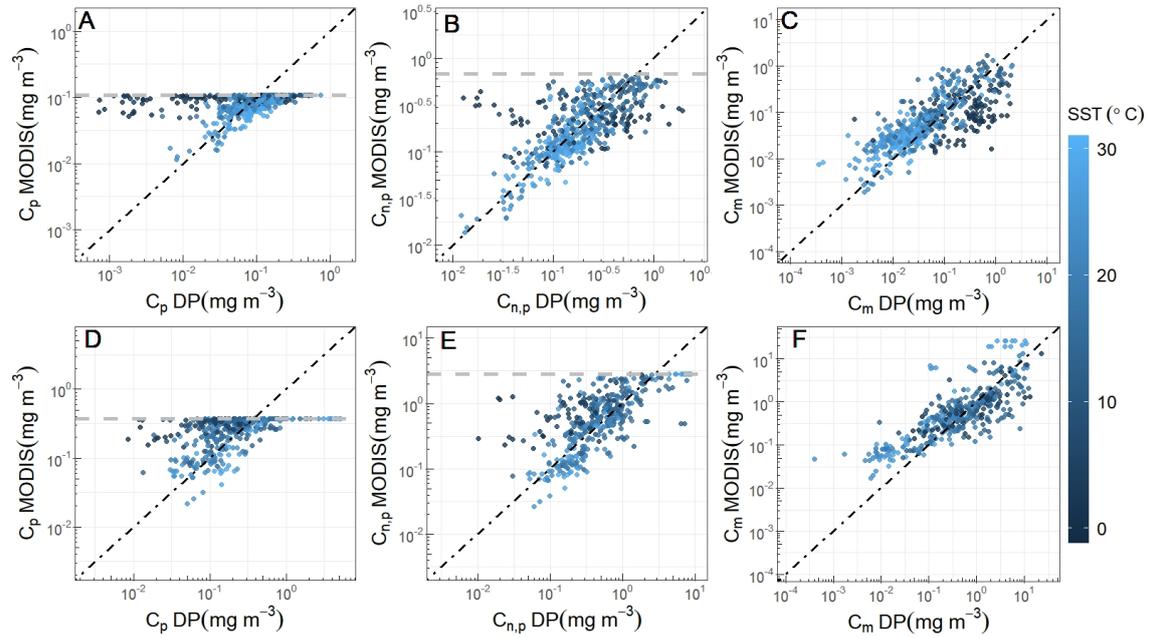
A, B and C are the plots for open-ocean waters, and D, E and F are the plots for coastal waters. The dashed-dotted black line is the 1:1 line and the dashed grey line is the model threshold for picophytoplankton (A and D) and pico and nanophytoplankton (B and E). The colour gradient scale represents the sea surface temperature in Celsius degrees.

SOURCE: Author's production.

Figure 3.10 - Residuals for the fractions of *in situ* dataset, for open-ocean and coastal waters.
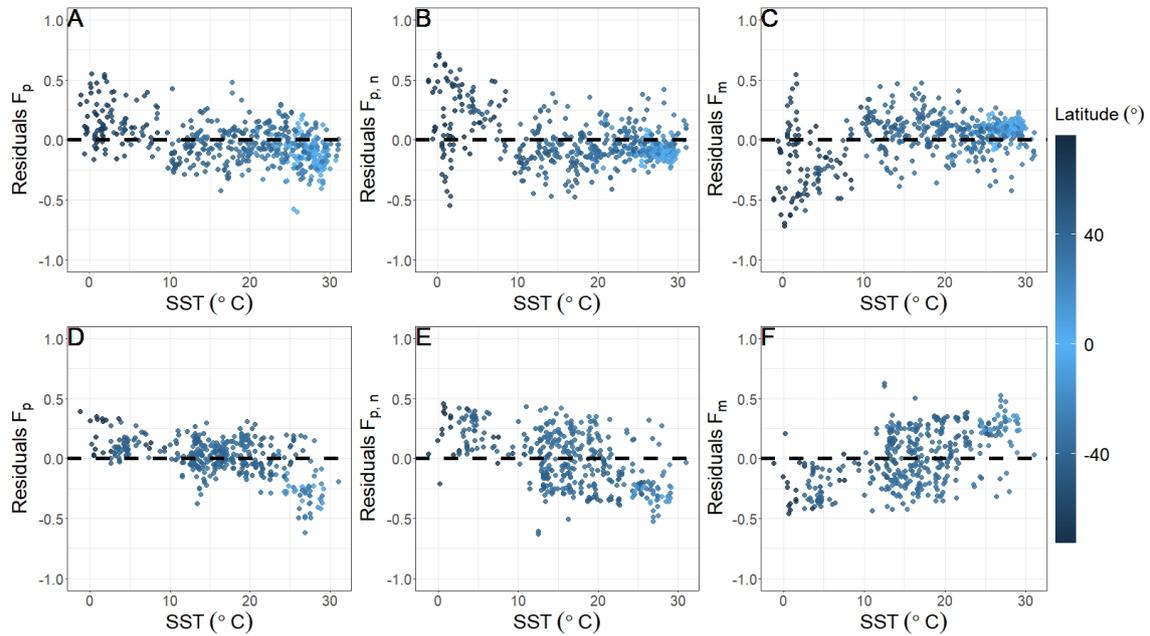
A, B and C are the plots for open-ocean waters, and D, E and F are the plots for coastal waters. The dashed black line is the zero line for picophytoplankton (A and D) and pico and nanophytoplankton (B and E). The colour gradient scale represents the latitudes in degrees.

SOURCE: Author's production.

3.3.4 HPLC from SeaBASS

The pattern observed in the PCA for the open-ocean data, with aggregations of high and low latitudes, which was not observed for coastal waters, may actually be related to the distribution of the data sampled (Figure 3.11 Table 3.9). Most of the data sampled in coastal waters are at latitudes lower than 45°N and 45°S (93%), while for open-ocean waters it is more evenly distributed, even though low latitudes still have more samples (67%), especially for spring and winter (for each hemisphere, austral seasons for the southern hemisphere and boreal for the northern hemisphere), even in the open-ocean samples. The seasons preferentially sampled are summer, fall, and spring, following this order, with the samples taken during winter being usually at low latitudes.

Considering open-ocean waters, high latitudes are usually more productive than low latitudes, with higher average concentrations of TChl*a*, 0.917 mg·m^{-3} and 0.422

mg·m^{-3}, respectively for high and low latitudes. High latitudes also present higher concentrations of Fuco, 0.321 mg·m^{-3} compared with 0.088 mg·m^{-3} at low latitudes (Table 3.9). Regarding coastal waters, samples from low latitudes presented higher average TChl$a$, 4.993 mg·m^{-3} compared with 1.422 mg·m^{-3} at high latitudes. Indeed, all the pigments presented higher concentrations at low-latitude coastal waters. Considering the percentage of samples with some pigments that were below detection limits, open-ocean waters sampled at low latitudes presented high percentages of non-detectable concentrations of Allo (40%) and Perid (26%).

Figure 3.11 - SEABASS HPLC samples in open-ocean (more than 50km from land) and coastal water (less than 50km from land), in high and low latitudes (45° threshold).
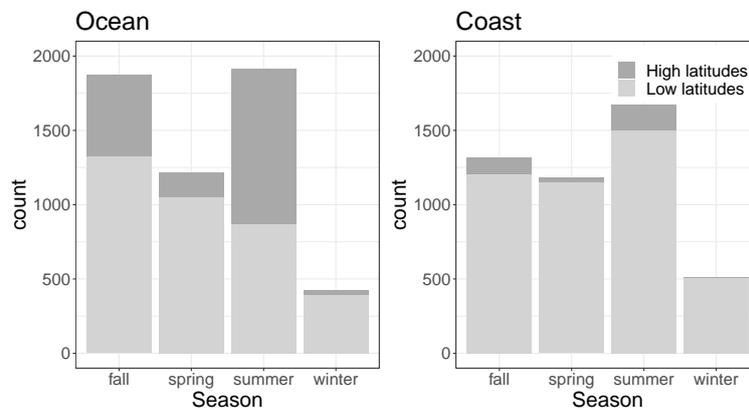
SOURCE: Author's production.

Table 3.9 - Diagnostic pigments concentrations and TChl*a* statistics for open-ocean and coastal waters at high and low latitudes.

	Ocean									
Pigments	**High latitudes (N=1801)**					**Low latitudes (N=3631)**				
	Mean	Min	Max	Sd	Zero(%)	Mean	Min	Max	Sd	Zero(%)
TChl*a*	0.917	0.019	16.685	1.42	0	0.422	0.005	17.227	0.729	0
TChl*b*	0.044	0	1.116	0.081	16	0.040	0	1.129	0.063	14
Fuco	0.321	0	8.536	0.669	0	0.088	0	5.219	0.289	8
Zea	0.01	0	0.227	0.02	7	0.064	0	0.919	0.054	0.2
Perid	0.027	0	1.56	0.075	15	0.013	0	1.079	0.039	26
But-fuco	0.037	0	0.571	0.049	2	0.021	0	0.634	0.029	7
Hex-fuco	0.108	0	1.252	0.119	1	0.073	0	2.642	0.116	1
Allo	0.024	0	1.248	0.079	12	0.01	0	0.591	0.029	40

	Coast									
Pigments	**High latitudes (N=328)**					**Low latitudes (N=4354)**				
	Mean	Min	Max	Sd	Zero(%)	Mean	Min	Max	Sd	Zero(%)
TChl*a*	1.422	0.039	24.498	2.601	0	4.993	0.021	330.304	10.618	0
TChl*b*	0.083	0	0.516	0.078	2.4	0.205	0	5.802	0.333	6
Fuco	0.562	0.001	11.337	1.303	0	1.185	0	44.809	2.195	0.5
Zea	0.012	0	0.207	0.022	18	0.217	0	7.33	0.53	4
Perid	0.034	0	0.513	0.059	12	0.558	0	134.71	3.857	9
But-fuco	0.013	0	0.257	0.029	18	0.023	0	0.621	0.039	23
Hex-fuco	0.039	0	0.581	0.073	19.5	0.095	0	7.886	0.189	12
Allo	0.036	0	0.734	0.08	6	0.175	0	26.98	0.659	7

3.4 Discussion

3.4.1 PSC models

The PSC model and the validation analysis presented in this study for open-ocean and coastal waters subsets performed, in general, worse than other studies (see Brewin et al. (2010), Turner et al. (2021), Lamont et al. (2018a)). Despite the overall poor results observed, we could notice some interesting features which can be related to latitudinal and seasonal differences. These features are the overestimation of picophytoplankton and nanophytoplankton and the underestimation of microphytoplankton in low SST, observed for open-ocean and coastal waters.

Other studies have discussed latitudinal and seasonal differences in the phytoplankton pigments. Kramer and Siegel (2019), for instance, studied the clustering and variability of phytoplankton pigments, considering a global data set (which also included coastal water samples) and local data sets (time-series data from coastal locations

only). The authors found a latitudinal pattern in the EOF (empirical orthogonal function), with the mode 1 positively related to the microphytoplankton pigments (Fuco, Perid) positive at high latitudes, and negatively related to picophytoplankton pigments (Zea, DvChl*b*), negative at low latitudes— a similar pattern observed in our study. The authors also reported more phytoplankton groups retrieved for local data sets when analysed isolated, than the number of groups retrieved by the global data set.

Babin et al. (2003) observed that for some coastal waters the spectrum of $a_{ph}(\lambda)$ as a function of chlorophyll-*a* described in Bricaud et al. (1995) presented significant variability, the authors attributed this variability mainly to high concentrations of phaeopigments, which play a minor role in open-ocean waters. These examples illustrate the issue discussed in this study, since differences in pigments composition (not only in the diagnostic pigments) may affect well-established relationships for open-ocean conditions, causing them to fail in some coastal-water conditions.

These relationships are not constant and may also fail in a changing environment, as discussed by Sathyendranath et al. (2017). Thus, even for open-ocean waters, it is important to keep track of these relationships, using up-to-date data to check and validate them. Because, as the authors put it, in the context of climate change "*the past may not be a reliable guide to the future*".

The main point though is that the open-ocean waters present this general ecological pattern of diatoms (Fuco see Jeffrey et al. (2011)) in high latitude spring and cyanobacteria (Zea see Jeffrey et al. (2011)) in oligotrophic tropical waters, which is usually not so clearly observed in coastal waters due to more complex and diverse phytoplankton assemblages, here illustrated by the ratios of diagnostic pigments and TChl*a*. Admittedly, diatoms are also found in coastal waters, but at least in some regions, their abundance tends to be episodically determined, for example, by the intensity of local upwelling. Thus, for these waters, the size classes may not be as structured as in the open ocean, which makes it difficult to infer them using chlorophyll-based algorithms designed for the open ocean. That is why regional validation and tuning of phytoplankton algorithms are required, and why it would not be reasonable to just adopt the global ocean model algorithms for regional applications.

Regional tuning of PSC algorithms is a two-step process: First, in equation Equation 3.5 (UITZ et al., 2006), the pigment-specific weights have to be adjusted, and then, the regionally-tuned equation has to be used to partition the total chlorophyll-a into

the three size classes (see Equations 3.8-3.10), before the parameters of the size-class model are fitted. In our study, the re-assignment of weights to the diagnostic pigments was found to be particularly important in coastal waters and contributed to reducing the uncertainties in the algorithm performance. We note, however, that Chase et al. (2020) and Brewin et al. (2014a) did not observe significant differences in the PSC fractions when using different weights for the western North Atlantic Ocean, and Atlantic Ocean, respectively.

The dataset used for the PSC model parameterization seems to play an essential role in the performance. In this study, we used a multi-year (1989-2019) and global dataset for the parameterization. The application of this model in different regions and for different time-frame results in different parameters and performances (BREWIN et al., 2010; TURNER et al., 2021; LAMONT et al., 2018a). Brewin et al. (2015), for instance, also compared a multi-year dataset, from 1992 to 2012 and obtained relatively better statistics for PSCs, with coefficient of correlation >0.9 considering C_m and $C_{p,n}$ and >0.6 for C_p. The authors removed stations collected on water with depths lower than 30m. They analysed 5841 HPLC samples, which 91% were located in waters with depth >200m and the remaining 9% located in continental shelves, and applied a more strict quality control in the HPLC data.

It is important to highlight that environmental conditions such as sea surface temperature (SST) may be used to improve their performances. As observed in the validation figures (Figures 3.8 and 3.9), contributed to some of the spread in the fits, and was related to overestimation in microphytoplankton (at high temperatures) and picophytoplankton (at low temperatures). Sea surface temperature, as shown by other authors (BREWIN et al., 2017; TURNER et al., 2021; STOCK; SUBRAMANIAM, 2020) has proven to improve the PSC models' performance, thus it is highly recommended to be included as a tuning variable, as in Brewin et al. (2017). Approaches used in ecological models, in which environmental conditions are allowed to modulate PSC retrievals, as in Stock and Subramaniam (2020), are also worthy of further exploration. Our results indicated that temperature influence on PSC models is also valid for coastal waters, and not only for open-ocean waters.

3.5 Final considerations

In this paper, we assessed the differences in pigment composition in coastal waters relative to open-ocean waters, and why it is a challenge to fit phytoplankton size class models in coastal areas. According to our findings, the variability in pigment composition in coastal waters was not higher than that in the open ocean. However,

the pigment structure in coastal waters does appear to be distinct from that in the open ocean, and this difference might explain, at least partially, why it is not straightforward to apply open-ocean algorithms in coastal waters. The OCx (default empirical Chl$a$ algorithm) presented higher errors for coastal waters. On the other hand, the PSC model performances varied according to the size-fraction considered. *In situ* validation presented, in general, better performance than satellite validation for coastal and open-ocean waters. The parameters of the model were dependent on the re-parameterization dataset.

Considering these findings, the following issues must be addressed to fitting regional PSC models for open-ocean and coastal waters:

1. Are the assumptions in the PSC models regarding the pigment structure maintained in the region?

2. Are the satellite retrievals for chlorophyll-a concentration (for example, OCx) reliable for the region?

In addition, recent studies have found that adding other variables, such as sea surface temperature, wind speed, and sea height anomalies as inputs to PSC models can also improve their performances (BREWIN et al., 2017; TURNER et al., 2021; STOCK; SUBRAMANIAM, 2020), which is relevant for regional studies in an ecological-based perspective (IOCCG, 2014). This was suggested for temperature in our results. Moreover, efforts to increase *in situ* data collection are still required to advance the retrieval of satellite-derived PSCs. It would be particularly advantageous if multiple methods (e.g. HPLC, flow cytometry, optical microscopy, size-fractionated filtration, among others) could be applied to the same samples, given the limitations of individual methods for measuring size classes and taxonomic groups. The insights that we can obtain when the same samples are subjected to multiple methods for measuring size classes and taxonomic composition are investigated in the next chapter (Chapter 4).

4 PHYTOPLANKTON ASSEMBLAGES AND OPTICAL PROPERTIES IN A COASTAL REGION OF THE SOUTH BRAZIL BIGHT

4.1 Introduction [1]

The marine phytoplankton comprise diverse photosynthetic organisms that sustain the oceanic animal population (JEFFREY et al., 2011)— and, therefore, the fishery production. They are also responsible for removing almost a third of the carbon dioxide released into the atmosphere by human activities (GRUBER et al., 2019; SABINE et al., 2004), and play essential roles in biogeochemical cycles of carbon (all phytoplankton taxonomic groups), silicate (diatoms), calcium (coccolithophores), and nitrogen (cyanobacteria) (LEQUÉRÉ et al., 2005). Due to their global importance, studying marine phytoplankton abundance and taxonomic diversity is a rich research field, with different methodological approaches, e.g., optical microscopy, flowcytometry, quantitative cell imagining, genomic sequencing, and phytoplankton diagnostic pigment analysis (DPA) in use.

Diagnostic pigment analysis is a technique based on the composition of pigments within phytoplankton taxonomic groups, and use different statistical methods to identify these groups (JEFFREY et al., 2011; KRAMER; SIEGEL, 2019; MACKEY et al., 1996) from data on major phytoplankton pigments. The drawback of this approach is that some pigments can occur in more than one group (e.g., fucoxanthin, is a biomarker of diatoms but occurs in other groups as well), whereas others may occur exclusively within a specific group, but do not occur in all the species of the group (e.g., peridinin, dinoflagellates) (JEFFREY et al., 2011). In addition, interpreting the results may be complicated by the adaptive nature of the pigments, such that pigment composition may vary in different ecological and environmental conditions (IRIGOIEN et al., 2004; SCHLÜTER et al., 2000; ZAPATA et al., 2004). Nevertheless, phytoplankton community structure can be retrieved with reasonable confidence when using quality-controlled pigment data and appropriate statistical approaches (CATLETT; SIEGEL, 2018; KRAMER et al., 2018; KRAMER; SIEGEL, 2019).

There are some studies in which different approaches to measuring phytoplankton size classes were compared. Brewin et al. (2014a) compared the size-fractionated filtration with DPA. They observed that DPA overestimated nano-sized fraction

---

and underestimated pico-sized fraction compared with SFF in mainly oligotrophic waters and drew attention to the importance of using more than one single method to determine phytoplankton size structure *in situ*. Chase et al. (2020) compared flowcytometry and imaging (flowcytoBot) with DPA in mixed trophic conditions (0.18-5.14 mg·m^{-3}) and observed that DPA overestimated micro-sized and pico-sized fractions compared with flowcytometry and underestimated nano-sized fraction. The authors then applied this information to the derivation of the model parameters to improve the phytoplankton size fraction estimation.

The South Brazil Bight (SBB) has great importance for the fishery, oil and gas industries in the South-Western region of the South Atlantic Ocean (CASTELLO et al., 2009; MATSUURA, 1996; VASCONCELLOS; GASALLA, 2001). Previous studies in the region have reported on the phytoplankton community structure, size classes, and dynamics using different approaches such as optical microscopy (**??**BRANDINI et al., 2014; GAETA et al., 1999; MOSER et al., 2016); chemotaxonomy (LIMA et al., 2019); flowcytometry (RIBEIRO et al., 2016), and size-fractionated chlorophyll-*a* concentration (SUSINI et al., 1995). Although each of these studies has focused on a different part of the SBB, they have all highlighted the importance of the water masses to describe the distribution of the phytoplankton assemblages. Nevertheless, these studies spanned relatively short periods (usually less than a year), and therefore could not investigate phytoplankton community dynamics at the annual time scale.

Information on the complex oceanographic dynamics, which drive phytoplankton primary production in the SBB, is essential for understanding the variability in the phytoplankton community structure in the region (BRANDINI et al., 2018). Three water masses interact in the SBB: Tropical Water (TW, T>20°C, S>36psu), transported southward by the Brazil Current (BC) in the surface layer, mainly oligotrophic; the nutrient-rich South Atlantic Central Water (SACW, T: 6-20°C, S: 34.6-36psu), transported beneath the TW (> 200 m); and the Coastal Water (CW, with variable temperature and lower salinity) closer to the shore, influenced by the mixing of local river discharges and shelf waters (CASTRO; MIRANDA, 1998). The Plata Plume Water (PPW), as described by Möller et al. (2008), plays a seasonal role in the region during the austral winter. The La Plata River plume dispersion in the S-SE Brazilian continental shelf during the austral winter is associated with the northward displacement of the westerlies and the associated Ekman transport, which drives low-salinity and nutrient-enriched waters against the coast, enhancing local productivity (PIOLA et al., 2000; PIOLA et al., 2008). Piola et al. (2000) refer to this seasonal incursion across the SBB as the Subtropical Shelf Front (SSF)

associated with the transport of a mixture of the PPW and Sub Antarctic Water.

Brandini et al. (2014) investigated the temporal cross-shelf dynamics of the phytoplankton community structure in the south region of the SBB from November 2005 to June 2006 and suggested that the variability of phytoplankton assemblages is driven by inputs other than the continental runoff into these waters. According to these authors, phytoplankton diversity and biomass increases in winter, mostly induced by the PPW advection and in summer, when it is driven by the intrusion of the nutrient-rich SACW into the mid-shelf. Influx of river outflow into the vicinities of the sampling site are known to have a low volume (107 m^3·s^{-1} is the average river discharge into the north coast of São Paulo) (COMITÊ DE BACIAS HIDROGRÁFICAS DO LITORAL NORTE (CBHLN), 2017). It has therefore been suggested that the PPW is one of the main sources of nutrient-rich, cold, and low-salinity waters into the region in autumn-winter (PIOLA et al., 2008; CIOTTI et al., 2018), associated with the SSF (BRANDINI et al., 2018). Other mesoscale and sub-mesoscale features, such as coastal eddies, could also induce mixing or stratification of the water column and potentially change phytoplankton assemblages (e.g., CARVALHO et al., 2019, LÉVY et al., 2014, SIEGEL et al., 2011). Wind-driven mixed-layer deepening by local and mesoscale winds are also known to play an important role in the dynamics of the phytoplankton assemblage in the region (GAETA et al., 1999).

In the present study, we characterize the phytoplankton assemblages and the environmental drivers of their succession at a coastal time-series station located in the inner portion of the SBB, off Ubatuba coast (23.60°S-44.96°W): the Antares-Ubatuba station. The phytoplankton community structure and optical properties were analysed using microscopy data, DPA, and the phytoplankton spectral absorption. The phytoplankton pigment composition obtained from HPLC is analysed applying hierarchical cluster analysis to indicate the main taxonomic groups (JEFFREY et al., 2011; KRAMER et al., 2020; KRAMER; SIEGEL, 2019). PSC are also determined using DPA (e.g., UITZ et al. 2006) and compared with the microscopy dataset. A locally-tuned version of the DPA is also tested for comparison. Relations between the DPA taxonomic groups and PSCs with the specific phytoplankton absorption coefficient ($a_{ph}^*(\lambda)$), package effect index (Q_a^*) (BRICAUD et al., 2004), and size index derived from the absorption spectra (S_f) (CIOTTI et al., 2002) are analysed to characterize the optical properties of the phytoplankton assemblages.

4.2 Materials and methods

4.2.1 Study area

The study area is located in the region described by Castro and Miranda (1998) as the South Brazil Bight (23–28.5°S) limited to the north by Cape Frio (23°S) and to the south by Cape Santa Marta (28.5°S). The shelf is divided into three zones: inner-, mid- and outer-shelf, regarding its hydrodynamic differences (CASTRO et al., 2006; CASTRO; MIRANDA, 1998). The time-series station (from here on referred to as Antares-Ubatuba) is indicated in the inner-shelf of the SBB, on the isobath of 40-43 m, depending on tide height (Figure 4.1).

Figure 4.1 - Location of the South Brazil Bight and the Antares-Ubatuba time-series station (23.60°S-44.96°W) on the north coast of São Paulo, Ubatuba, Brazil, SW-S Atlantic (isobaths in meters).
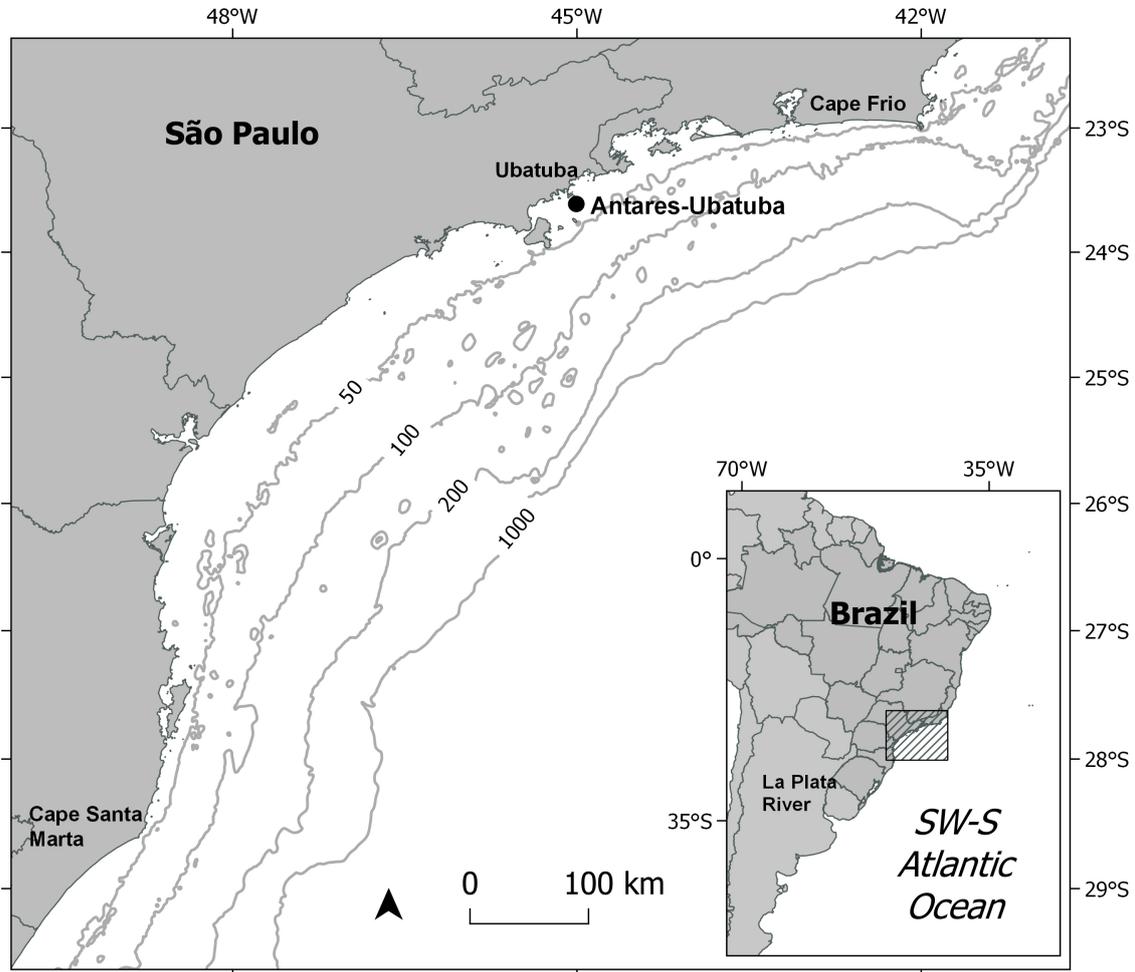
SOURCE: Author's production.

As previously mentioned, meteo-oceanographic processes with different temporal and spatial scales influence the study region's water quality. Local and mesoscale wind field (speed, direction, and duration) is a key parameter related to upwellings, mixing of the water column, which affects the phytoplankton assemblages (??GAETA et al., 1999). In particular, the processes that most affect the water quality of the SBB are: (i) the passage of atmospheric cold fronts that occur throughout the year, which are more frequent and severe during the austral winter and spring (CASTRO et al., 2015), changing the prevailing wind direction from Northeast to Southeast,

and mixing the water column; (ii) the intrusion of the SACW on the continental shelf mainly during the austral spring and summer, driven by continuous north-easterly winds (CASTRO et al., 2006), especially at the Cape Frio and Cape Santa Marta upwelling cells, which in addition, may form the Cape Frio Upwelling Front (CFUC) as described by Cerda and Castro (2014) and could reach the study area (see also Brandini et al. (2018)) (see Figure 4.2A-B); (iii) mesoscale eddies observed at the meandering onshore front of the Brazil Current that flows southwards following the shelf-break, which also favours the intrusion of SACW onto the shelf (CAMPOS et al., 2000; SILVEIRA et al., 2000); and (iv) the northward advection of the PPW, described as the SSF (BRANDINI et al., 2018; PIOLA et al., 2000)(Figure 4.2C-D).

Figure 4.2 - MODIS/Aqua 8 days composition of sea surface temperature (A and C) and chlorophyll-a concentration (B and D) showing the oceanic fronts that may influence the study area (dashed lines in black): Cape Frio Upwelling Front CFUF (A and B) observed from October 7th to 14th, 2012, and Shelf Sub-tropical Front SSF (C and D) observed from June 9th to 16th, 2016.
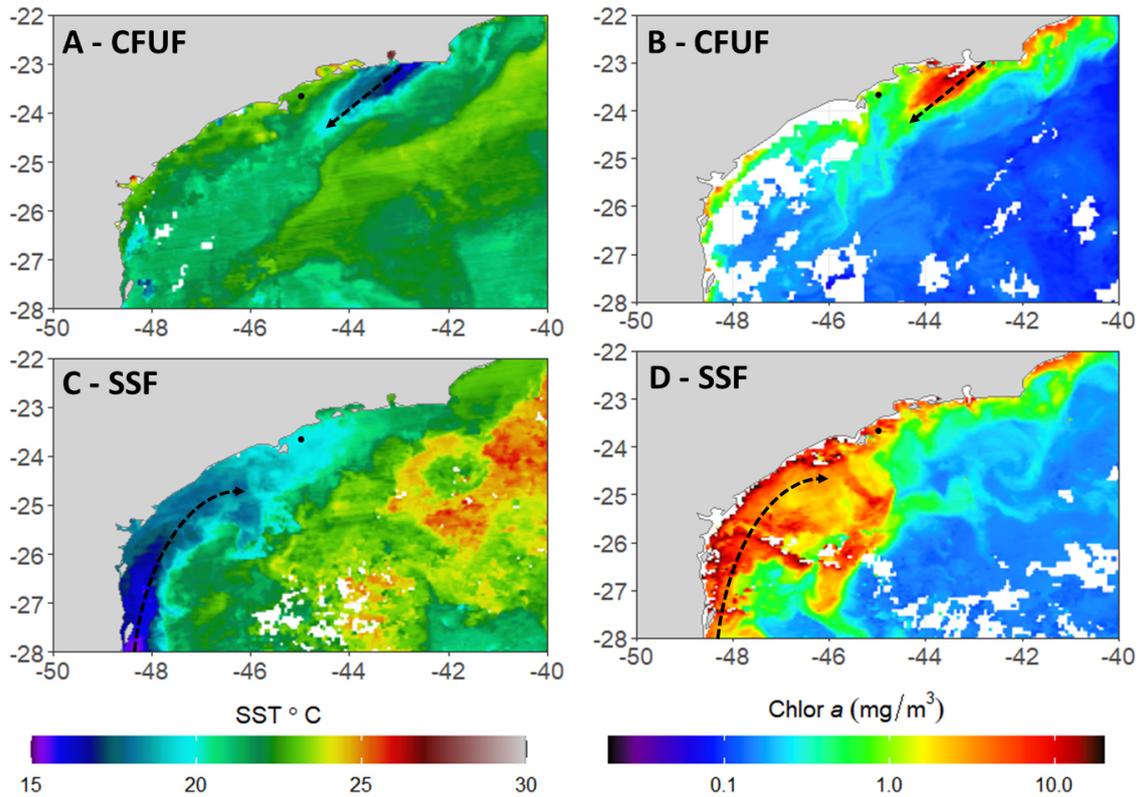
SOURCE: Author's production.

4.2.2 Measurements

4.2.2.1 Phytoplankton pigments and taxonomy

The phytoplankton diagnostic pigment concentrations were determined by High Performance Liquid Chromatography. The samples were collected during 41 campaigns between July 2012 to May 2017. Surface water samples (1 m) were collected with a Niskin bottle, and 1 L of sample was filtered through glass fiber filters (0.7 μm) GF/F (Whatman®). Filters were stored and transported to the laboratory in liquid nitrogen and then maintained in an ultra-freezer at -80°C. All the samples (stored in dry ice) were sent for analysis at the NASA Goddard Space Flight Center, following quality control protocols and using the HPLC method described in VanHeukelem and Hooker (2011).

For the qualitative and quantitative analyses of phytoplankton taxonomy, surface seawater samples were collected and fixed with formaldehyde neutralized with hexamethylenetetramine (final concentration 0.4%). An aliquot of 50 mL of the sample was placed in sedimentation chambers (72 h decantation) and observed following the method described in Utermöhl (1958), under an inverted microscope (Zeiss Axiovert 35). Pico- and nanoflagellates were counted in transects (for at least 400 organisms) and separated according to their size. Nano (other than flagellates) and microplankton were counted in the entire chamber and identified up to the lowest possible taxon (genera and species). The optical microscopy analysis present limitations in the quantification of pico-sized cells (SHERRARD et al., 2006), so its quantification was most likely underestimated, but we still decided to maintain these data for a rough comparison with the HPLC DPA.

The phytoplankton biovolumes were calculated from the number of cells using geometric shapes that best suited each phytoplankton group commonly occurring in the study region, as analysed by optical microscopy. Since measuring the geometrical dimensions of the cells is very time consuming, only 13 samples (out of 41) had their biovolumes estimated. The mathematical equations used here were taken from Sun and Liu (2003), and the cell linear dimensions (i.e., height, width, depth, and more depending on the geometric shape) were taken from microscopy observations . The biovolumes were corrected to account for shrinkage caused by sample fixation, following Montagnes et al. (1994). After calculating the biovolume, the carbon biomass was estimated using the equations described in Menden-Deuer and Lessard (2000) (see Table 4.1) and then multiplied by the number of cells per litre of the respective taxonomic groups. The final biomass was converted to mgC·m^{-3}.

Table 4.1 - Equations applied to estimate the biomass in pgC per cell from the biovolume of each major phytoplankton group identified by microscopy, and the reference group, according to Menden-Deuer and Lessard (2000). V is the biovolume (μ m ^{3}).

Major groups from this study	Reference group	Equation for biomass estimation (*pg C per cell*)
Flagellates and coccolithophorids (picoplankton and nanoplankton)	Mixed protist phytoplankton $<3000\mu m^3$	$0.261 \cdot V^{0.86}$
Diatoms $<3000\mu m^3$	Diatoms $<3000\mu m^3$	$0.288 \cdot V^{0.811}$
Diatoms $>3000\mu m^3$	Diatoms $>3000\mu m^3$	$0.116 \cdot V^{0.881}$
Dinoflagellates (nanoplankton and microplankton)	Dinoflagellates	$0.76 \cdot V^{0.819}$

4.2.2.2 Phytoplankton absorption coefficients

For the particulate absorption coefficient, surface water samples were collected with a Niskin bottle, and 1 L of sample was filtered through glass fibre GF/F filters (0.7 μm) (Whatman®The filters were folded and packed in an aluminium foil envelope and immediately stored in liquid nitrogen until the analysis. The particulate (a_p) and detritus (a_d) (after depigmentation with 10% sodium hypochlorite NaClO, 0.1% of active Cl (TASSAN; FERRARI, 1995)) absorption coefficients were measured using an integrated sphere mouthed in a dual-beam spectrophotometer (Shimadzu UV-2450) (MITCHELL et al., 2002). Phytoplankton absorption was obtained by subtracting a_d from a_p. In the beginning of the time-series sampling, the method used for the particulate absorption was the quantitative filter pad technique with the transmittance method (T) (MITCHELL et al., 2002), since the Antares-Ubatuba station, even though coastal, is not much subjected to riverine or resuspension sediment load. In January 2016, however, the Transmittance-Reflectance (T-R) Method (TASSAN; FERRARI, 1995; TASSAN; FERRARI, 2002) was also implemented to compare and evaluate improvements. The T method overestimated the absorption signal, especially at the shorter wavelengths ($R^2 = 0.98$, RPD = 17%, RMSE= 0.0078 m^{-1} at 440 nm). Hence, the T-R method was selected as the "new" standard for the station. An empirical spectral adjustment was applied to the T method dataset to match the T-R measurements and obtain a consistent time-series minimizing the effects of methodological uncertainties.

The spectrally-resolved specific absorption coefficient was calculated by dividing the phytoplankton absorption coefficient at each wavelength (λ) by the corresponding total chlorophyll-a concentration (TChl$a$) (i.e. the sum of monovinyl chlorophyll a, divinyl-chlorophyll a, chlorophyllide a and chlorophyll a allomers and epimers).

4.2.2.3 Sea water properties

Temperature and salinity were measured using a calibrated Conductivity-Temperature-Depth (CTD) profiler (SeaBird and Falmouth NXIC). Dissolved oxygen concentration (DO_2) and inorganic nutrients, i.e., ammonium (NH_{4+}), nitrate (NO_{3-}), nitrite (NO_{2-}) and total inorganic phosphorous (H_2PO_{4-}), were also measured and used to determine the trophic index (TRIX) according to Vollenweider et al. (1998) and Giovanardi and Vollenweider (2004) for coastal waters (Equation 4.1):

$$TRIX = \frac{[\log_{10}(PO_4 \cdot\ TN \cdot\ TChla \cdot\ D\%O_2) + 1.5]}{1.2} \qquad (4.1)$$

where, TChl$a$ is the total chlorophyll-a concentration (mg·m^{-3}); $D\%O_2$ is the percent deviation of the oxygen concentration from saturation conditions, TN is the dissolved inorganic nitrogen given by $N-NO_{3-} + N-NO_{2-} + N-NH_{4+}$(mg·m^{-3}); and H_2PO_4 is the reactive inorganic phosphorus calculated as $P-H_2PO_{4-}$(mg·m^{-3}).

DO_2 was measured using the Winkler technique described in Strickland and Parsons (1968). Nutrient concentrations were determined according to their specific methods i.e., Ammonium (NH_4) according to Aminot and Marcel (1983), Nitrate (NO_3) + Nitrite (NO_2) according to Armstrong et al. (1967), Phosphate (PO_4) according to Murphy and Riley (1962) and Silicate ($Si(OH)_4$) according to HANSEN, H. P.; KOROLEFF (1983). Until 2014, they were manually determined on discrete samples; thereafter, concentrations were obtained by continuous flow analysis using an Auto-Analyzer (Seal AA3) (the automatic and the manual approaches present differences lower than 0.5% for the analyzed nutrients). The trophic status (Equation 4.1) was classified according to the TRIX value(PRIMPAS; KARYDIS, 2011), i.e., ultraoligotrophic (<1.6), oligotrophic (1.6-2.8), mesotrophic (2.8-4.0), eutrophic (4.0-5.3), and dystrophic (>5.3).

4.2.3 Data analysis

4.2.3.1 Diagnostic pigments analysis

To identify general phytoplankton taxonomic groups from the concentrations of the diagnostic pigments, a hierarchical cluster analysis was adopted, following Latasa and Bidigare (1998), Catlett and Siegel (2018), Kramer and Siegel (2019). We applied the following steps: (1) calculate a Pearson correlation matrix for the pigments normalized by TChl$a$ (such that the correlation index varies from -1 to 1); (2) estimate the correlation distances as $1 - \rho$, where ρ is the Pearson correlation; (3)

estimate the cluster linkages using the Ward minimum variance method (squared distance) (MURTAGH; LEGENDRE, 2014); (4) evaluate the cluster validity by estimating the correlation between the distance matrix (calculated in the step 2) and the cophenetic distances (i.e. distances between pigments in the dendrogram) for the hierarchical clustering result. The analysis was carried out using R programming language.

To determine the fraction of TChl*a* associated with each taxonomic group identified in the cluster analysis, we used the Uitz et al. (2006) model. We also applied a multiple linear regression model to estimate the coefficients of each diagnostic pigment. The results obtained in the regional approach were compared to the global approach (UITZ et al., 2006). A similar approach proposed by Chase et al. (2020) was adopted for the locally-tuned model and CHEMTAX, to avoid assigning each phytoplankton type into a single size class. Instead, here we used microscopic analysis to determine the proportion of each phytoplankton type that fell into nano and micro size classes. If the fractions in both size classes were non-negligible, then that type was split into two size classes accordingly.

The size index proposed by Bricaud et al. (2004) was calculated for each campaign, using the proportions of each phytoplankton size class, and reducing them to single number representative of the mean size of the sample, see Equation 4.2. It was applied to compare the computed optical properties (described in the next Section 4.2.3.2).

$$SI = \frac{(50 \cdot microplankton\% + 5 \cdot nanoplankton\% + 1 \cdot picoplankton\%)}{100} \qquad (4.2)$$

Where micro, nano, and picoplankton in the equation are the percentages of each phytoplankton size class estimated by DPA. This size index is given in micrometer (μm).

For comparison, we also determined the relative contribution of phytoplankton groups to the overall chlorophyll-a biomass using CHEMTAX v1.95 chemical taxonomy software (MACKEY et al., 1996) with class-specific accessory pigments and chlorophyll-*a* ratios. CHEMTAX uses factor analysis and the steepest-descent algorithm to best fit the data to an initial matrix of pigment ratios, i.e., the ratio between the respective accessory pigments and chlorophyll-*a*. The initial pigment ratios of major algal classes used here were obtained from an average output matrix determined from regional data collected over the SBB, compiled from Lima et al.

(2019), wherein the chemotaxonomic groups were identified according to Jeffrey et al. (2011).

In the case of CHEMTAX, to be consistent with the ratios used by the authors (LIMA et al., 2019) the chlorophyll-*a* was calculated as the sum of Monovinyl Chlorophyll-*a* and Divinyl Chlorophyll-*a*, and is hence different from the TChl*a* used in the hierarchical cluster analysis. In addition, some modifications were necessary to adapt the pigment dataset analysed in this work to the input pigments used in Lima et al. (2019). In Lima et al. (2019) the $\beta\epsilon$-carotenoid and the $\beta\beta$-carotenoid were resolved and considered separately. However, in the present study these pigments were summed and considered as $\beta\beta$-carotenoid. This latter pigment was found in most of the phytoplankton groups identified by the CHEMTAX ratios in Lima et al. (2019).

The locally-tuned phytoplankton groups and size fractions obtained from the diagnostic pigment analysis were compared with the taxonomic groups and carbon biomass estimated from the microscopic analysis—converting number of cells per litre to biovolume (μm^3) and then to biomass in (mg C· m^{-3}).

4.2.3.2   Computed phytoplankton optical properties

The measured phytoplankton absorption coefficient and TChl*a* were used to compute the specific phytoplankton absorption coefficient (i.e., the absorption coefficient normalized by TChl*a*, $a_{ph}^*(\lambda)$), where λ is the wavelength. The specific absorption coefficient was computed for all wavelengths λ from 400 to 700 nm.

A package effect index ($Q_a^*(440)$) at wavelength λ=440 nm and a size index (S_f) were then computed from the normalized absorption spectra. The package effect index ($Q_a^*(\lambda)$) is defined as the ratio of the absorption coefficient of the phytoplankton pigments ($a_{ph}(\lambda)$) within the cells by the absorption coefficient of the same pigments if they were dispersed in a solution ($a_{sol}(\lambda)$) (BRICAUD et al., 2004; MOREL; BRICAUD, 1981; SATHYENDRANATH et al., 1987) (Equation 4.3):

$$Q_a^*(\lambda) = \frac{a_{ph}(\lambda)}{a_{sol}(\lambda)} \tag{4.3}$$

The package effect index decreases from a maximum of 1, when no package effect is observed, to smaller numbers (>0), as the package effect increases. The value of $a_{sol}(\lambda)$ is calculated using the concentration of each phytoplankton pigment in

the sample and the specific absorption coefficient of each pigment when in solution ($a^*_{sol,i}$) (Equation 4.4). We also applied the missing term correction (Equation 4.5), empirically obtained by Bricaud et al. (2004), which is especially important for oligotrophic waters. Bricaud et al. (2004) observed that the measured $a_{ph}(\lambda)$ was sometimes higher than the spectra reconstructed from pigments (i.e., $a_{sol}(\lambda)$). They argued that measurement uncertainties could not explain the discrepancy, as they would have introduced a random noise instead of systematic bias. The authors adopted the hypothesis of Nelson et al. (1993) that the observed difference was due to pigments that were not measured by the HPLC (carotenoids or phycobiliproteins) or to other light-absorbing compounds (cytochromes, flavins, quinones) extracted in the measurement of the $a_{ph}(\lambda)$. Hence, a "missing term", $a_{miss}(\lambda)$, was introduced to account for the bias. For λ=440 nm,we have:

$$a_{sol}(440) = \sum C_i a^*_{sol,i}(440) + a_{miss}(440) \qquad (4.4)$$

with

$$a_{miss}(440) = 0.0525 \cdot TChla^{0.855} \qquad (4.5)$$

The index of package effect (Q^*_a) was only calculated for 440 nm.

The size index (S_f) was also calculated to verify the effect of the package effect on the phytoplankton absorption spectra and the phytoplankton size groups obtained from the diagnostic pigment analysis. The size index was obtained following Ciotti et al. (2002) and Ciotti and Bricaud (2006), using the absorption coefficient by phytoplankton normalized by the mean spectral absorption from 400 to 700nm ($a_{<ph>}(\lambda)$), and a non-linear least squared regression fit to obtain S_f (Equation 4.6):

$$a_{<ph>}(\lambda) = [S_f \cdot \bar{a}_{pico}(\lambda)] + [(1 - S_f) \cdot \bar{a}_{micro}(\lambda)] \qquad (4.6)$$

where $\bar{a}_{micro}(\lambda)$ and $\bar{a}_{pico}(\lambda)$ are the endmember vectors provided by Ciotti and Bricaud (2006) for the largest (micro) and smallest (pico) cells.

4.2.4 Schematic diagram

Schematic diagram of the measurements and analysis.

Figure 4.3 - Schematic diagram of measurements and analysis of the Chapter 4.
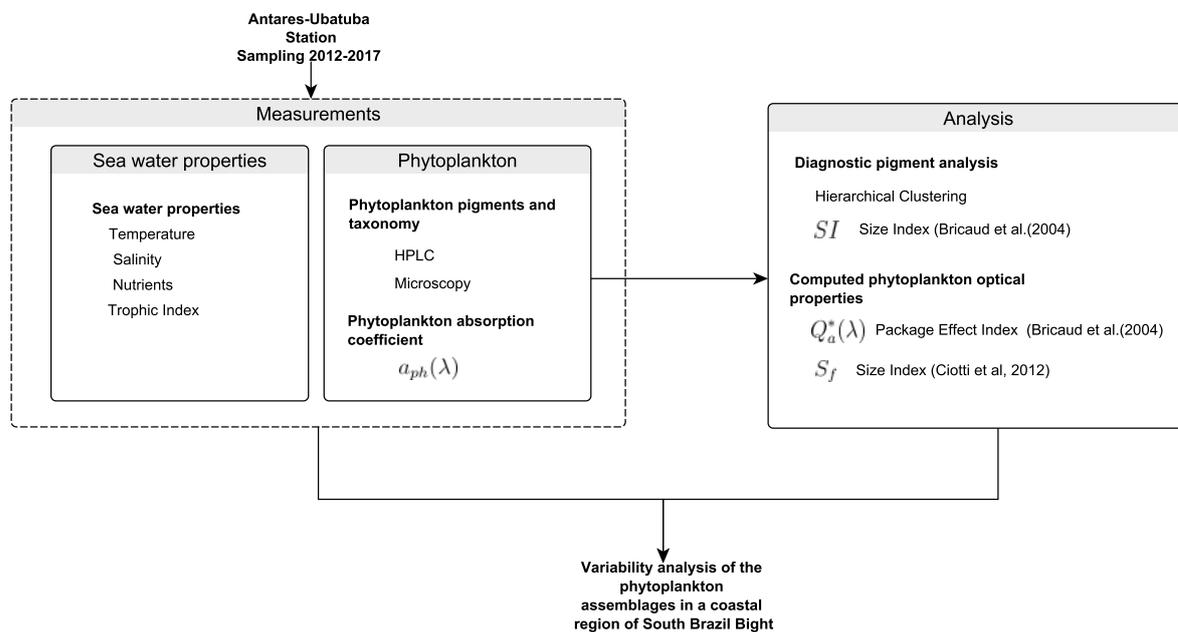
SOURCE: Author's production.

4.3   Results

4.3.1   Phytoplankton pigments and taxonomy

Typically, the study site may be characterized as oligo-mesotrophic with average TChl$a$ concentration lower than 1 mg·m^{-3}, including algal bloom events. If excluding the blooms, the average TChl$a$ is 0.5 mg·m^{-3}. When TChl$a$ was higher than 5mg·m^{-3}, i.e., ten times higher than mean conditions, it was considered as a bloom event.

Although TChl$a$ has a relatively low variability (mean 0.8 ±1.6 mg·m^{-3}, and median 0.4 mg·m^{-3}), some exceptionally high values were registered within the sampling period: one on November 13th, 2012 (10 mg·m^{-3}) and another on February 25th, 2014 (6.22 mg·m^{-3}), indicating algal bloom events (Figure 4.4). Microscopy data on November 13th, 2012 revealed a multispecies bloom, composed mainly of *Mesodinium major*, a mixotrophic ciliate known to present symbiosis with cryptophytes (JOHNSON et al., 2016) and a variety of diatoms, especially those belonging to the genus *Pseudo-nitzschia*. In the second event (February 25th, 2014), the mi-

croscopy data indicated a dinoflagellate bloom composed mainly of *Dinophysis spp*, a mixotrophic dinoflagellate. Simultaneously with this event, an extensive *Mesodinium* bloom occurred near the sampling point, which lasted for months (CETESB, 2018). Alloxanthin (Allo) concentrations during these events were higher compared with other diagnostic pigments. In November 2012, Allo (1.3 mg·m^{-3}) and fucoxanthin (Fuco) (1.6 mg·m^{-3}) concentrations were proportionally higher, whereas zeaxanthin (Zea) concentration was lower than usual (0.5 mg·m^{-3}). These results are in contrast with the event of February 2014, which presented high concentrations only of Allo (0.9 mg·m^{-3}) and a very low concentration of Fuco (0.05 mg·m^{-3}).

Figure 4.4 - Barplot of the relative proportion of eight phytoplankton diagnostic pigments from the Antares-Ubatuba time-series station (23.60°S-44.96°W).
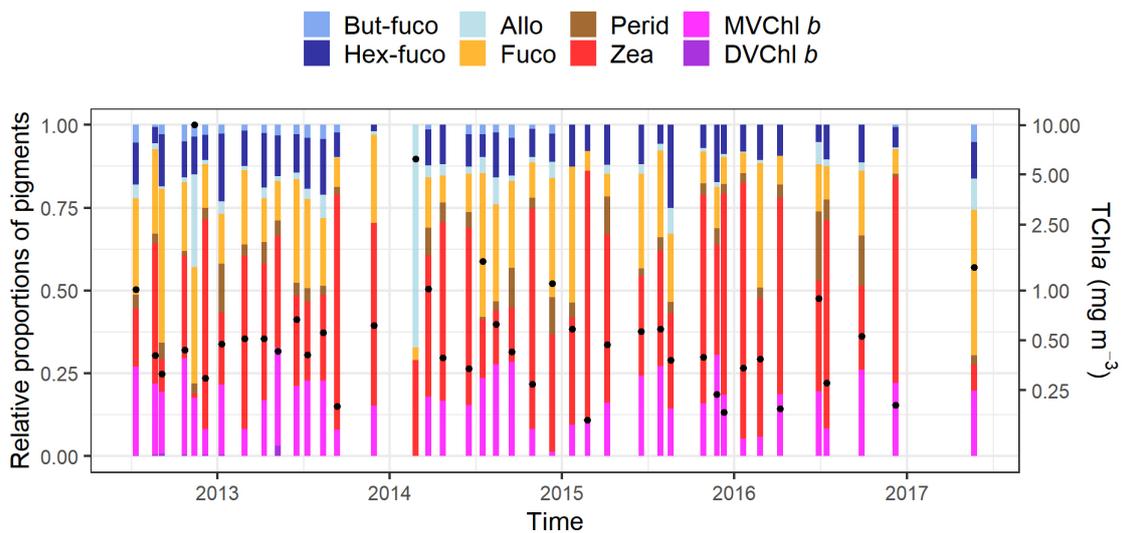
The black dots represent TChl*a*.
SOURCE: Author's production.

Non-bloom conditions were characterized usually by higher proportions of Zea (mean 0.122±0.058 mg·m^{-3}) followed by Fuco (mean, 0.077±0.082 mg·m^{-3}) and total chlorophyll-*b* (TChl*b* , monovinyl chlorophyll-*b* + divinyl chlorophyll-*b*) (mean 0.060±0.045 mg·m^{-3}). Peridinin (Perid), Allo, 19'-Butanoyloxyfucoxanthin (But-fuco) and 19'-Hexanoyloxyfucoxanthin (Hex-fuco) were also present in lower concentrations all year round, throughout the sampling period.

In order to test if there were statistically significant differences between the seasons aggregated as spring/summer and autumn/winter, a t-test was performed (excluding the bloom events) for pigments that presented a normal distribution (verified by a Shapiro-Wilk normality test on the log-transformed pigments concentrations), and a non-parametric Wilcoxon test for those pigments that did not present a normal distribution. Some pigments had significantly higher concentrations during the austral autumn and winter, i.e., TChl*b* , Hex-fuco, and Allo (see Table 4.2). Zea was the only pigment that showed a higher concentration in the spring-summer, but the differences were not statistically significant.

Table 4.2 - Summary of diagnostic pigments and TChl*a* (austral spring/summer and autumn/winter) for 2012-2017. Two algal bloom events that occurred in spring and summer were not included here (see text). (Mean ± standard deviation; n = number of observations).* indicates the significant difference between the seasons (p-value<0.05).

Pigments concentration (mg·m^{-3})	Abbreviation	Spring /Summer (n=16)	Autumn /Winter (n=23)
Total chlorophyll-*a* (monovinyl chlorophyll-*a* + divinyl-chlorophyll-*a* + chlorophyllide-*a* + chlorophyll-*a* allomers and epimers)	TChl*a*	0.421±0.231	0.590±0.350
Total chlorophyll-*b* (monovinyl chlorophyll-*b* + divinyl chlorophyll *b*)	TChl*b*	**0.037±0.026***	**0.076±0.050**
Total chlorophyll-*c* (chl *c1* + chl *c2* (chl *c1c2*) + chl *c3*)	TChl*c*	0.042±0.028	0.068±0.057
19'-Butanoyloxyfucoxanthin	But-fuco	0.004±0.006	0.008±0.011
19'-Hexanoyloxyfucoxanthin	Hex-fuco	**0.028±0.016***	**0.040±0.021**
Alloxanthin	Allo	**0.005±0.008***	**0.015±0.018**
Diadinoxanthin	Diadino	0.016±0.017	0.020±0.015
Diatoxanthin	Diato	0.001±0.002	0.001±0.001
Fucoxanthin	Fuco	0.061±0.060	0.089±0.095
Peridinin	Perid	0.016±0.022	0.020±0.026
Zeaxanthin	Zea	0.136±0.062	0.112±0.055
Divinyl-chlorophyll *a*	DVChl*a*	0.006±0.009	0.005±0.008
Chlorophyll *c1+c2*	Chl*c12*	0.029±0.020	0.048±0.041
Chlorophyll *c3*	Chl*c3*	0.013±0.010	0.020±0.019
Photoprotective Carotenoids	PPC	0.187±0.088	0.179±0.081
Photosynthesis Carotenoids	PSC	0.108±0.092	0.158±0.123

The phytoplankton cell counts obtained from microscopy analysis showed a dominance of nano-flagellates (64%±10%, Figure 4.5A) and preserved the highest proportion when converted to bio-volume and biomass (40%±18%, Figure 4.5C). This agreement was not observed for the other taxonomic groups (Figure 4.5). Picoplank-

ton (cyanobacteria and heterotrophic bacteria) was the second most abundant group when considering cell counts (29%±10%) but presented a lower contribution when converted to biomass (4%±2%). The diatoms and dinoflagellates (in both micro and nanoplankton size classes) showed a higher contribution when converted to biovolume and biomass by virtue of their large sizes. Diatoms were mostly represented in the micro-size (18%±25% of the total phytoplankton biomass) and presented a minor fraction of nano-size (5%±11%). Dinoflagellates had an approximately even distribution between micro (12%±7%) and nano (15%±11%) size classes.

Figure 4.5 - Average proportions of phytoplankton cells belonging to each taxon from microscopic number of cells (A), its estimated biovolume (B) and its estimated biomass (C), all based on microscopic analysis (N = 13).
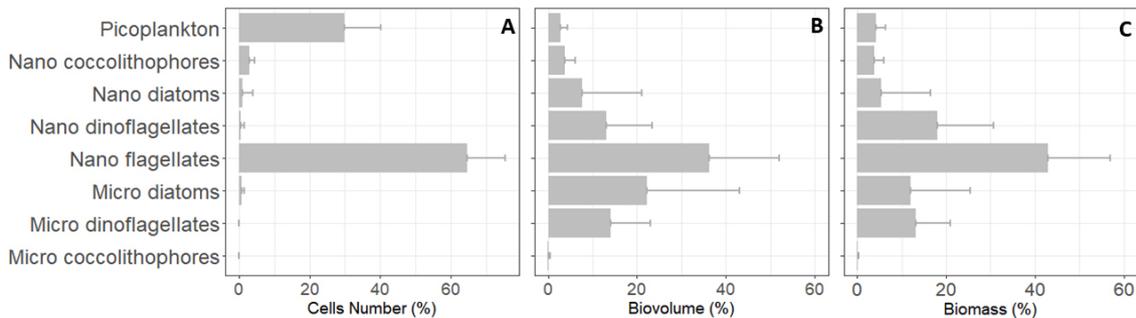
SOURCE: Author's production.

We also compared the total carbon biomass determined by microscopy analysis with TChl$a$, using log-transformed linear regression. The agreement was good and significant for the 13 samples analysed (R^2 = 0.57, p < 0.005, see Figure 4.6A). The Carbon-to-TChl$a$ ratio presented an average of 35±26, ranging from 11 to 102 (Figure 4.6B). The values found in the literature range from 15 to 176 (SATHYENDRANATH et al., 2009).

Figure 4.6 - (A) Scatterplot of TCha versus carbon biomass estimated from biovolume derived from microscopy. (B) Carbon to TChl$a$ ratio for each of the campaigns for which microscopy data were available (N=13).
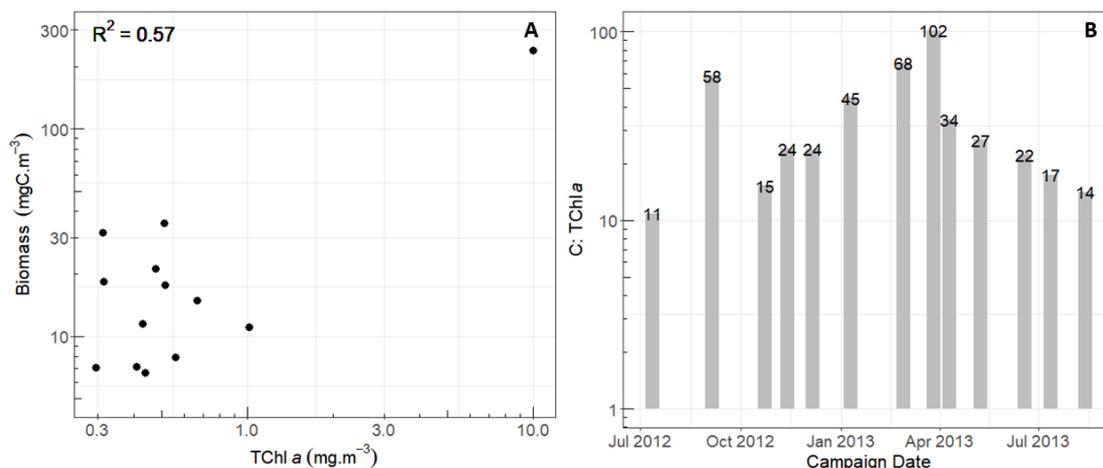
SOURCE: Author's production.

4.3.2 Diagnostic pigments analysis

The correlation matrix of the diagnostic pigments (DPs) was calculated for both the concentrations of pigments normalized to TChl$a$ and the absolute pigment concentrations. For the correlation matrix, we considered 18 pigments: total chlorophyll-a (TChl$a$), 19'-butanoyloxyfucoxanthin (But-fuco), 19'-hexanoyloxyfucoxanthin (Hex-fuco), alloxanthin (Allo), diadinoxanthin (Diadino), diatoxanthin (Diato), fucoxanthin (Fuco), peridinin (Perid), zeaxanthin (Zea), divinyl chlorophyll-a (DVChl$a$), monovinyl chlorophyll-b (MVChl$b$), divinyl chlorophyll-b (DVChl$b$), chlorophylls $c1$ and $c2$ (Chl$c1$+$c2$), chlorophyll $c3$ (Chl$c3$), lutein (Lut), neoxanthin (Neo), violaxanthin (Viola) and prasinoxanthin (Pras).

The normalized correlation matrix was then used in the hierarchical cluster analysis. DVChl$b$ was not included in the cluster analysis because it presented no detectable concentration in more than 70% of the campaigns. The cophenetic correlation coefficient between the distance matrix (between observations) and the dendrogrammatic distance (between the model points), calculated using the Ward method, was 0.56, suggesting a positive correlation and a significant result (p-value << 0.01).

The dendrogram from the hierarchical cluster analysis presented six clusters of phy-

toplankton taxonomic groups, as described in Jeffrey et al. (2011). The classified groups corresponded to diatoms, dinoflagellates, cryptophytes, cyanobacteria, haptophytes, and green algae (Figure 4.7). For the size classes, based on microscopic analysis we grouped diatoms, dinoflagellates, cryptophytes, and haptophytes into one size class of combined micro and nanophytoplankton, and cyanobacteria and green algae were assigned to pico-phytoplankton, as the dendrogram indicated an ecological grouping (co-occurrence) of these assemblages.

The ratio Allo and Chl$c12$ to TChl$a$ increased with TChl$a$ ($\rho = 0.88$ for Allo and $\rho = 0.43$ for Chl$c12$), suggesting an association of cryptophytes with higher TChl$a$ conditions (the bloom events). In contrast, the ratios of Zea and MVChl$b$ to TChl$a$ decreased with TChl$a$ ($\rho = -0.38$ for Zea and $\rho = -0.28$ for MVChl$b$), suggesting that cyanobacteria and green algae assemblages were more characteristic of oligo-mesotrophic conditions. The ratios of the other diagnostic pigments to TChl$a$ did not show any clear patterns associated with TChl$a$, probably indicating that they were associated with mixed assemblages.

Figure 4.7 - Hierarchical cluster analysis of phytoplankton pigments normalized by TChl$a$.
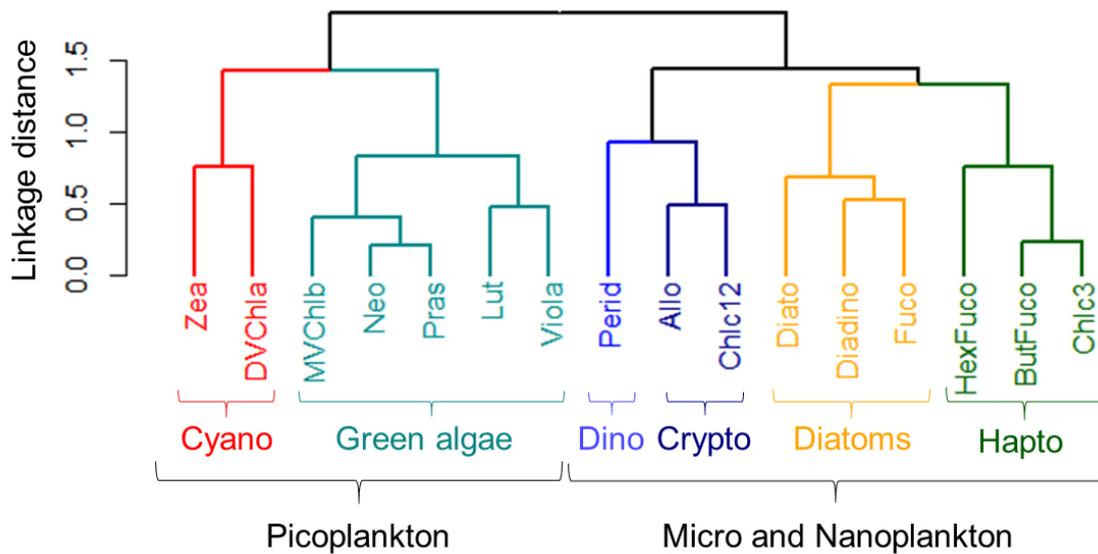
Six taxonomic phytoplankton groups related to each cluster, according to our interpretation, are suggested as follows: diatom (orange), dinoflagellate (blue), cryptophyte (dark blue), cyanobacteria (red), haptophytes (dark green), and green algae (green).

SOURCE: Author's production.

To estimate each taxonomic group's contribution to TChl$a$ from the hierarchical cluster analysis for each campaign, we applied a local tuning of the Uitz et al. (2006) global phytoplankton size fraction model. After applying a multiple linear regression model to the selected DPs, we observed that when considering But-fuco by itself in the model, we obtained a negative weight for this pigment. We also tried constraining the model to only positive coefficients, but in this case, the But-fuco was zero, and the Hex-fuco was 2.36, and the comparison of the haptophytes DPA derived biomass with that determined by microscopy analysis was poorer. Thereafter, we considered the Hex-fuco combined with But-fuco in this dataset, because these pigments were both appeared to be representative of haptophytes in the hierarchical cluster analysis. This step improved the correspondence with the microscopic analysis. The algal bloom events were excluded from this analysis because when they were included, the results changed substantially: the Allo coefficient was very much overestimated (> 5, compared to 0.69 without bloom events). The equations and weights of the locally and globally (UITZ et al., 2006) tuned DPs are presented in Table 4.3.

One way to evaluate the model performance is to examine whether the sum of the weighed DPs ($\sum$DP) matched the TChl$a$. This test showed that the locally tuned model had 11% relative percentage difference (RMSE=0.07 mg·m^{-3}). In contrast, the globally tuned model had 30% difference (RMSE=0.20 mg·m^{-3}), even though the coefficients of determination were equal to each other (R^2=0.96). The higher the DP weight, the higher its concentration relative to TChl$a$. In our case, all weights were somewhat higher than the globally-tuned model (Table 4.3).

Table 4.3 - Phytoplankton groups and the related Diagnostic Pigment (DP) weights are shown for the global regression model (UITZ et al., 2006) and the locally-tuned model from this study (excluding bloom events). Equations considering the locally tuned weights are also shown. AIC refers to the Akaike Information Criterion. (*) indicates p-values < 0.05.

Phytoplankton Groups	DPs	Uitz et al.(2006) Global	Local tuned AIC = -94.7 R^2=0.96	Equations locally-tuned
Diatoms	Fuco	1.41	**2.29***	$f_{diat} = 2.29 \cdot \dfrac{Fuco}{\sum DP}$
Cyanobacteria	Zea	0.86	**0.92***	$f_{cyano} = 0.92 \cdot \dfrac{Zea}{\sum DP}$
Dinoflagellate	Perid	1.41	**1.90***	$f_{dino} = 1.90 \cdot \dfrac{Perid}{\sum DP}$
Cryptophytes	Allo	0.6	0.69	$f_{crypto} = 0.69 \cdot \dfrac{Allo}{\sum DP}$
Haptophytes	Hex-fuco	1.27	**1.34***	$f_{hapto} = 1.34 \cdot \dfrac{(Hex-fuco+But-fuco)}{\sum DP}$
	But-fuco	0.35	**1.34***	
Green algae	TChl$b$	1.01	1.58*	$f_{green} = 1.58 \cdot \dfrac{TChlb}{\sum DP}$

Using the locally-tuned DP model (Table 4.3), we estimated each taxonomic group's contribution to TChl$a$ from the hierarchical cluster analysis for each campaign (Figure 4.8). Diatoms were the most abundant group in 51% of the campaigns (N = 21). These campaigns were generally characterized by high TChl$a$ (mean 1.09±2.06 mg·m^{-3}) and low sea surface temperature (22.6±2.8°C), as shown in Figure 4.8. Even though diatoms were the dominant group in these cases, usually all six phytoplankton groups were also present, characterizing a mixed assemblage, except for two cases: November 2013, when dinoflagellates were absent, and January 2015, when cryptophytes were absent.

Cyanobacteria were the second most abundant group, i.e., they had more TChl$a$ associated with them than any other taxonomic group in 36% of the campaigns (N = 15). These campaigns were associated with low TChl$a$ (mean 0.35±0.21 mg·m^{-3}) and high surface temperature (24.5±2.4°C) (Figure 4.8). Although cyanobacteria presented the highest contribution in these cases, all the other groups were usually present, indicating a mixed assemblage. Only in two of those campaigns, one or two other groups were absent, i.e., cryptophytes or/and dinoflagellates, and these were the cases with the lowest TChl$a$ (0.16 mg·m^{-3}) of the entire series.

Figure 4.8 - Temporal variation in the TChl*a* with the stations colour-coded according to the most abundant phytoplankton group estimated by the locally-tuned DPA; dashed line is the *in situ* sea surface temperature.
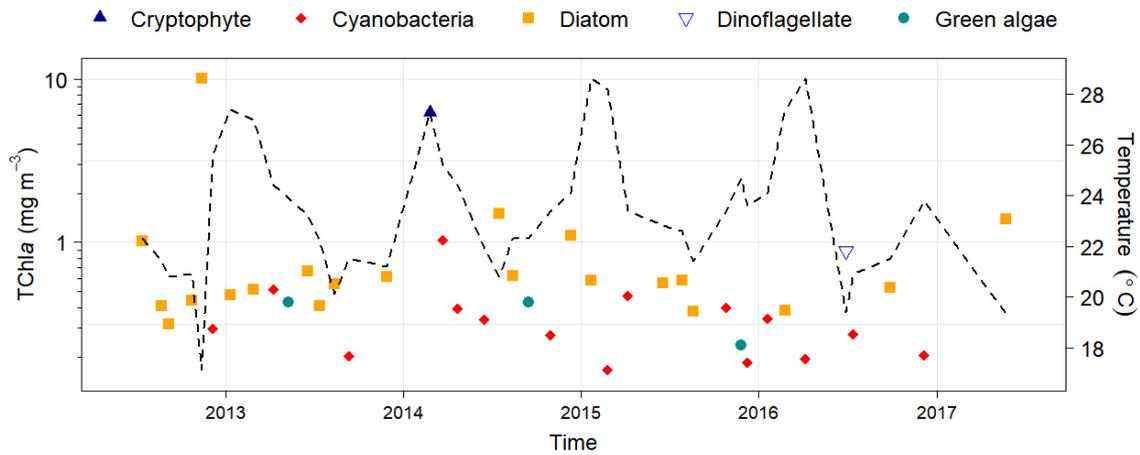
Note that the surface temperature includes stations where HPLC data were not available.

SOURCE: Author's production.

According to the locally-tuned model (DPA), green algae were the most abundant group at only three campaigns. Even on these occasions, when the TChl*a* associated with green algae was higher than that associated with any of the other taxonomic groups, the TChl*a* associated with green algae was not much higher than those associated with cyanobacteria and diatoms. The close association between cyanobacteria and green algae explains why green algae were grouped with cyanobacteria in the cluster analysis (Figure 4.7). The mean TChl*a* for these campaigns was generally lower than the overall average (0.36±0.11 mg·m^{-3}).

In June 2016, dinoflagellates showed the highest proportion of associated TChl*a* than any of the other taxonomic groups. The TChl*a* was higher than average (0.89 mg·m^{-3}), and the surface temperature lower than average (20°C). In the February 2014 campaign, cryptophytes were dominant (57%), showing the second-highest TChl*a* (6.22 mg·m^{-3}). For this sample, three phytoplankton groups were absent (dinoflagellates, green algae, and haptophytes), whereas diatoms were present in low proportion (10%), and cyanobacteria represented 32% according to the locally tuned model based on pigments concentrations.

The comparison of the global Uitz model for phytoplankton size classes and phytoplankton groups with the locally-tuned model and the CHEMTAX output matrix required an adaptation of the results, as illustrated in Table 4. The CHEMTAX analysis was based on pigment ratios obtained from a previous study, but in the same SBB region, using a dataset collected over a much broader spatial domain (LIMA et al., 2019). In that study, the authors showed significant positive correlations between prasinoxanthin (exclusive to prasinophytes) and concentrations of Lut and Neo (pigments present in chlorophytes and all types of prasinophytes). Thus, in their output pigment ratio matrix, most parts of these pigments were attributed to prasinophytes. However, the hierarchical cluster analysis of the present study indicated that it was not possible to resolve different groups of green algae in the Antares-Ubatuba dataset, and they were kept in a single group (combined chlorophytes and prasinophytes).

Microscopic analysis indicated that a non-negligible proportion of dinoflagellates and diatoms were part of the nano-sized fraction (around 50% of dinoflagellates and 30% of diatoms). For this reason, a similar approach from Chase et al. (2020) was adopted for the locally-tuned Uitz model, in which Perid (dinoflagellates) and Fuco (diatoms), which are usually entirely attributed to micro-sized cells, were fractionated into micro and nano size classes on the basis of the microscopic data: Perid was evenly divided between nano- and microplankton, and Fuco, had 30% attributed to nanoplankton and 70% to microplankton (Table 4.4).

Table 4.4 - Size classes and taxonomic groups identified using Uitz et al. (2006) model, locally-tuned, and the CHEMTAX method.

Uitz et al. (2006) Global	Locally-tuned	CHEMTAX
Microplankton	Diatoms (70%)	Diatoms (70%)
	Dinoflagellates (50%)	Dinoflagellates (50%)
Nanoplankton	Diatoms (30%)	Diatoms (30%)
	Dinoflagellates (50%)	Dinoflagellates (50%)
	Cryptophytes	Cryptophytes
	Haptophytes	Haptophytes
Picoplankton	Green algae	Prasinophytes
		Chlorophytes
	Cyanobacteria	Prochlorococcus
		Synechococcus

Overall, the three approaches agreed well for the microplankton ($R^2 > 0.84$ and RMSE < 0.19) and reasonably well for the picoplankton fractions ($R^2 > 0.76$ and RMSE < 0.09). High differences were observed for the nanoplankton fraction, especially when comparing the regional CHEMTAX with the global Uitz model ($R^2 = 0.47$ and RMSE $= 0.18$) (Table 4.5. The locally-tuned model and the regional CHEMTAX presented good agreement for all size classes, with a lower RMSE (0.08-0.09). This result is perhaps not surprising since both the locally-tuned model and the CHEMTAX method were run with the local observations as input. But it highlights the value of having local observations to improve the quality of regional algorithms.

Table 4.5 - Statistical parameters comparing the three approaches (UITZ et al., 2006) Global model, locally-tuned model, and CHEMTAX method) used to estimate the class sizes fractions (R^2 = coefficient of determination; RMSE = Root Mean Square Error).

	Locally-tuned *vs.* Uitz Global		Locally-tuned *vs.* CHEMTAX		Uitz Global *vs.* CHEMTAX	
	R^2	RMSE	R^2	RMSE	R^2	RMSE
Microplankton	0.90	0.10	0.85	0.09	0.84	0.19
Nanoplankton	0.77	0.11	0.73	0.08	0.47	0.18
Picoplankton	0.92	0.05	0.81	0.09	0.76	0.09

In order to compare the results obtained by the locally-tuned model and phytoplankton carbon biomass proportions determined by microscopy analysis (by estimating the biovolume and then estimating the carbon biomass in mgC·m^{-3} and then proportions for each taxonomic group), we compared the average proportions of the pigment biomass (TChl*a*) (see Figure 4.9A-B). In this case, nano- and micro-sized diatoms presented similar proportions, but with DPA proportions a little higher than the microscopy ones. Nano and micro-sizes dinoflagellates were underestimated by the DPA, which could be in part related to the variability in the C:Chl*a* ratio and partly to the absence of Peridinin in some dinoflagellate species (as will be further discussed). Nanoplankton biomass, composed of nanoflagellates and coccolithophorids was underestimated by the locally-tuned model, whereas picoplankton was highly underestimated by the microscopy analyses. This was expected as optical microscopy analysis is difficult for identifying and counting pico-sized cells and cannot be considered a reliable estimate for this group. Although it is difficult to draw any robust conclusion from this type of comparison, since each of the methods has its own limitations and drawbacks (IOCCG, 2014), it is nevertheless considered useful as it at least indicates if the estimates in the phytoplankton community are consistent across the different methods used in the analysis. Furthermore, carbon-based estimates of phytoplankton community structure and those based on chlorophyll are expected to differ from each other, as the C:Chl*a* ratio changes with photoacclimation as well as with community structure. Comparison of carbon-based and

chlorophyll-based community structure is therefore useful for understanding variability in the C:Chl$a$ ratio in natural communities.

Figure 4.9 - Average proportions of the major phytoplankton groups according to carbon biomass estimated from (A) microscopy and (B) TChla fractions estimated by the locally-tuned DP model. The nanoplankton group biomass by microscopy is an aggregation of nano-sized flagellate and coccolithophorids. In the DP locally-tuned model, TChla biomass aggregated the haptophytes and cryptophytes, whereas cyanobacteria and green algae account for the picoplankton fraction.
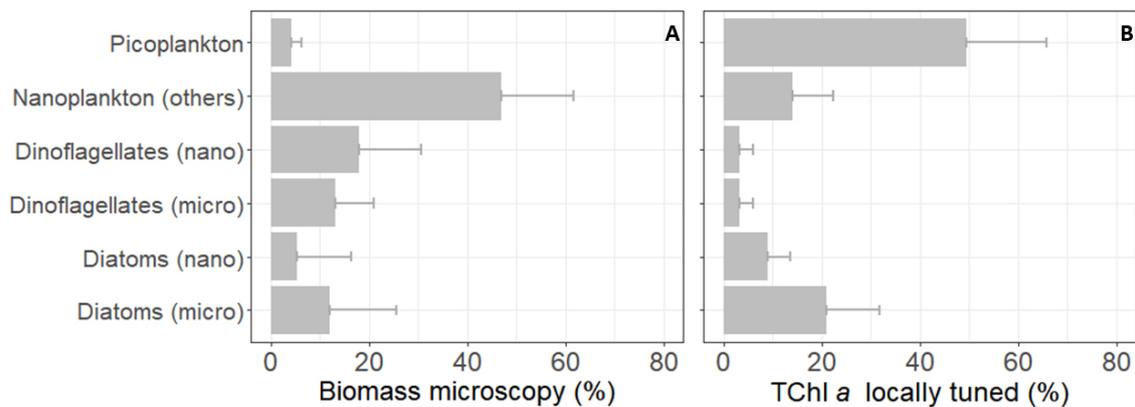
The grey line over the bars are the standard deviation.
SOURCE: Author's production.

4.3.3 Computed phytoplankton optical properties

In general, campaigns with a higher proportion of diatoms presented $a_{ph}^*(\lambda)$ values close to or below the average spectrum, with mean $a_{ph}^*(440)$ of 0.06 ± 0.02 m^2(mg Chl$a$)$^{-1}$ (see Figure 4.10). Two campaigns with the most flattened spectra (November 2012 and May 2017), with values of $a_{ph}^*(440)$ of 0.015 and 0.023 m^2(mg Chl$a$)$^{-1}$, presented high proportions of diatoms (52% and 58%, respectively) and a low proportion of cyanobacteria (<5%). Both samples were associated with the two lowest sea surface temperatures of the study period (17 °C and 19.3°C, respectively), indicating upwelling events. Nitrate concentrations were 0.71 μM/L and 3.72 μM/L, respectively for November 2012 and May 2017, above the average for the region (0.44 $\pm$ 0.87 μM/L). Silicate concentration was above average (5.31 $\pm$ 2.34 μM/L) only for May 2017 (8.18 μM/L), but was close to average in November 2017 (4.41 μM/L).

The first case occurred in the austral spring (November 2012) and was associated with the influence of the CFUF, promoting the multispecific bloom event with a variety of diatoms and *Mesodinium major* detected in our time-series. In the second case (May 2017, austral autumn) a local upwelling cell was observed near the study site.

Figure 4.10 - The specific absorption coefficients ($a^*_{ph}(\lambda)$) for different taxonomic groups with a higher relative contribution of a particular group, even though it is not necessarily the dominant group. (A) diatoms, (B) cyanobacteria, and (C) dinoflagellates, cryptophytes, or green algae.
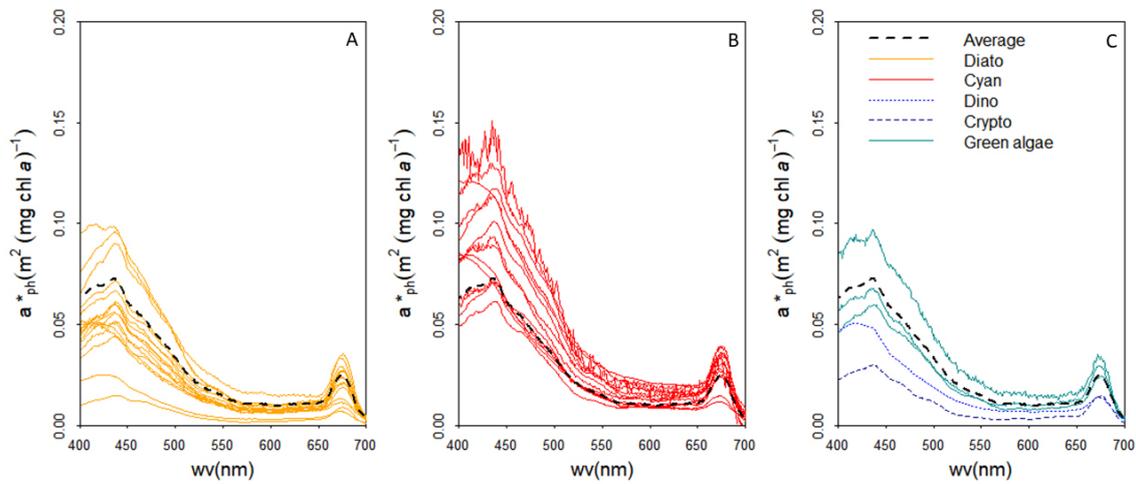
The dashed black curves are the average curve considering all the samples.

SOURCE: Author's production.

The campaigns with high proportions of cyanobacteria presented a^*_{ph} values above or remarkably close to the average spectrum (Figure 4.10B), with mean $a^*_{ph}(440)$ of 0.09 ± 0.02 m²(mg Chl*a*)$^{-1}$. The proportion of cyanobacteria varied from 28 to 64% for these campaigns, meaning that although this group was the one with the highest relative contribution, it was not necessarily the dominant group (i.e., in some cases, the other groups'contributions were similar).

The samples with the highest relative contributions of cryptophytes, dinoflagellates, or green algae to TChla presented $a^*_{ph}(\lambda)$ values mostly below or close to the average (Figure 4.10C). The flattened spectrum corresponds to a bloom event with a high contribution of *Dinophysis spp.* collected in February 2014, with $a^*_{ph}(440)$ of 0.03

m^2(mg Chl$a$)$^{-1}$. As previously mentioned, *Dinophysis spp.* is known to show symbiosis with smaller phytoplankton species, such as cryptophytes. The campaigns with the high relative contribution of dinoflagellates (June 2016) and green algae (May 2013, Sep 2014, and Nov 2015) presented $a_{ph}^*(\lambda)$ values closer to the average (Figure 4.10C).

The samples that presented the lowest values of the packaging effect index at 440 nm (Q_a^*) (BRICAUD et al., 2004), which should indicate the presence of larger cells (and hence higher flattening of the absorption spectra), were those with a high proportion of diatoms, dinoflagellates, and cryptophytes (Figure 4.11 A). The campaigns with a high contribution of diatoms (>30%) and low contribution of cyanobacteria (<30%) presented a mean $Q_a^*(440)$ of 0.54 $\pm$ 0.19 (N = 14). On the other hand, the campaigns with high cyanobacteria (>30%) and low diatom (<30%) contributions presented a mean $Q_a^*(440)$ of 0.74 $\pm$ 0.19 (n = 13). The mean $Q_a^*(440)$ for all stations was 0.62 $\pm$ 0.20. The $Q_a^*(440)$ values presented a significant relation with TChl$a$ (ρ = 0.76, p-value<0.01) and with cyanobacteria (ρ = 0.68, p-value<0.01) but a poor relation with diatoms (ρ = -0.39, p-value<0.05) and the micro (ρ = 0.36, p-value<0.05) and nano-sized fractions (i.e., including diatoms and dinoflagellates, along with haptophytes, and cryptophytes) (ρ = 0.58, p-value<0.01). $Q_a^*(440)$ was positively related only with cyanobacteria (ρ = 0.68, p-value<0.01).

Figure 4.11 - (A) Package effect index (BRICAUD et al., 2004) at 440 nm *vs.* TChl$a$. (B) Size index *vs.* TChl$a$.
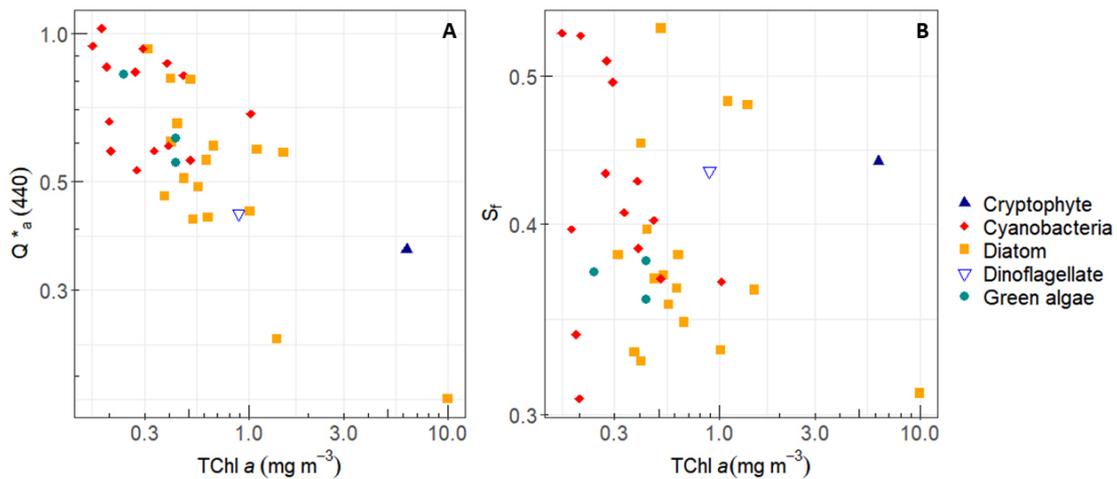
Colours and shapes differentiate the most abundant groups in each sample.

SOURCE: Author's production.

The size index (S_f) of Ciotti et al. (2002) varied from 0.31 to 0.54, with an average value of 0.40 ± 0.06 (Figure 4.11B). This relatively low variation, and average value around 0.4, denotes the prevalence of mixed assemblages in the Antares-Ubatuba site, mostly dominated by nano-size class. The samples with a high contribution of diatoms (>30%) and low cyanobacteria (<30%) usually presented lower S_f values (0.38 ± 0.05) due to the package effect of larger cells. Campaigns with a high contribution of cyanobacteria (>30%) and low contribution of diatoms (<30%) usually presented higher values of S_f (0.43 ± 0.07), indicating a high abundance of smaller cells. Two apparent exceptions were observed in September 2013 and April 2016, with an elevated proportion of cyanobacteria (53% and 43%, respectively) and low S_f values (0.31 and 0.34, respectively). The CHEMTAX results for the campaign of September 2013 indicated only 14% of *Trichodesmium* colonies, and zero percent for the April 2016 campaign. Instead, both campaigns were dominated by *Synechococcus* cyanobacteria, which does not explain the low S_f values. Values of S_f presented a poor correlation with TChl$a$ (ρ = -0.17, non-significant), as well as with the taxonomic groups and the associated aggregated size classes determined from pigment data.

A different perspective is obtained when the phytoplankton groups are aggregated into size classes (following Table 4.4) and an average size index (SI) is computed (an SI value of 1 represents dominance of picoplankton and 50 indicates dominance of microplankton) (BRICAUD et al., 2004) and compared with the $Q_a^*(440)$ and S_f values (Figure 4.12). An inverse relation is expected in these plots from theoretical considerations: i.e., when SI increases $Q_a^*(440)$ and S_f should decrease. This was observed for $Q_a^*(440)$ (ρ = -0.42, p-value<0.01), with exceptions for: (1) February 2014 (*Dynophysis sp.* bloom) which presented low $Q_a^*(440)(0.36)$ and low SI (8.47 μm), and (2) September 2012, which presented high $Q_a^*(440)$ (0.93) and high SI (33.43 μm) (Figure 4.12A). For the S_f and SI comparison, the relation was highly scattered but still presented the expected negative correlation but no significant (ρ = -0.21, non-significant) (Figure 4.12B).

Figure 4.12 - (A) Package effect $Q_a^*(440)$ (BRICAUD et al., 2004) at 440 nm *vs.* Size index (SI) (BRICAUD et al., 2004). (B) Size index (S_f) *vs.* Size index (SI) from Bricaud et al. (2004).
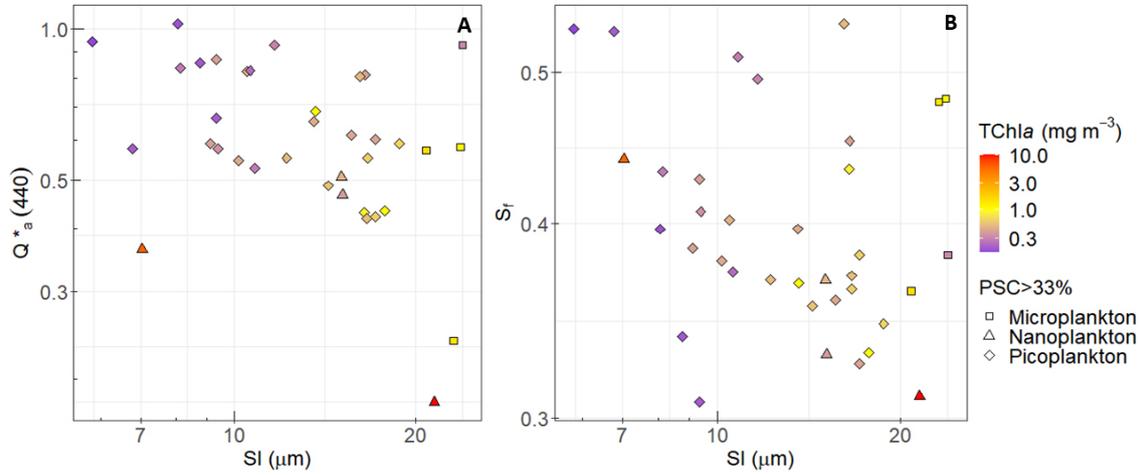
Shapes identify the most abundant size class in each sample. Colour scale is the TChla.

SOURCE: Author's production.

Comparing the value of the specific absorption coefficient of phytoplankton, the index of packaging effect and the size index as indicators of phytoplankton community structure, we see that $a_{ph}^*(\lambda)$ and $Q_a^*(440)$ showed a much better agreement with the phytoplankton groups determined by the DP analysis and with the TChl$a$ than the S_f index, which showed more dispersion. A probable explanation is that the S_f index is totally independent of the pigment data set, whereas both $a_{ph}^*(\lambda)$ and $Q_a^*(440)$ make use of the pigment data in their calculation. The low correlation in the S_f values, when examined as a function of TChl$a$ and the DPA PSCs, could be partly due to uncertainties in the absorption data set (in the dataset studied here the mean coefficient of variation for a_{ph} in 440 nm was 8.9% between replicates).

4.3.4 Phytoplankton groups and environmental conditions

The trophic index varied between 1.99 and 4.83 throughout the sampling period with, 15% of the months classified as oligotrophic (1.6–2.8 units), 76% as mesotrophic (2.8–4.0) and 9% as eutrophic conditions (4.0–5.3) (see Supplementary Material S4).

The average temperature and salinity during the study period were 23.5°C and 34 psu, respectively. Sea surface temperature was highly variable at the study site,

ranging from 17.1 °C to 29.7 °C. The Coastal Water, with relatively higher temperature and lower salinity, was the most prevalent water mass, being dominant at the study site during more than 60% of the campaigns, followed by the Plata Plume Water, with low temperature and salinity, and the South Atlantic Central Water, with low temperature and high salinity (Figure 4.13).

Figure 4.13 - Temperature-Salinity (TS) diagram. The sampling campaigns are colour-coded according to the most abundant taxonomic groups.
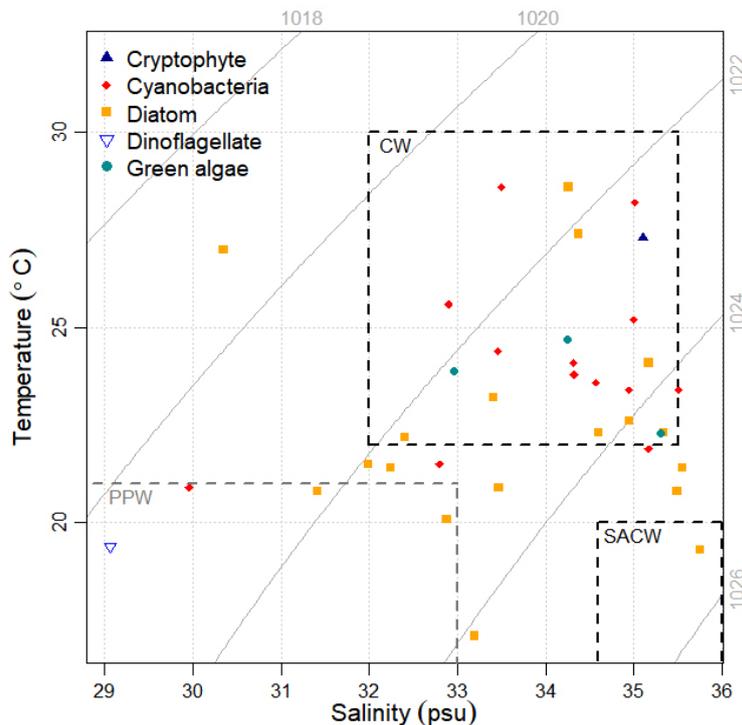
The dashed boxes represent the domain of specific water masses: Coastal Water (CW), South Atlantic Central Water (SACW), and Plata Plume Water (PPW).

SOURCE: Author's production.

The TChl$a$ concentration usually increases with lower temperatures (ρ = -0.25) (associated with the PPW or SACW intrusions), mostly under conditions favourable for diatom occurrence, such as mixed water column, with high nutrient concentrations. Well-mixed and nutrient- rich waters have been known to favour the predominance of diatoms (GLIBERT, 2016; MARGALEF, 1978).

The CW was characterized by mixed assemblages with the most abundant groups

in the samples formed either by cyanobacteria or diatoms, in most cases. The few samples that matched the PPW (N = 4) did not show a clear pattern regarding TChl$a$ (0.51 ± 0.28 mg·m^{-3}) or the dominance of a specific taxonomic group that would characterize the presence of this water mass. This nutrient-enriched water mass transported northward during the austral autumn-winter and forming the SSF (see Figure 4.2) is expected to be an important source of nutrients and new production for the SBB (BRANDINI et al., 2018). However, the Antares-Ubatuba site is located at its northern limit, and the new nutrients may be already exhausted. The only campaign under the influence of the PPW and SSF with relatively higher TChl$a$ (0.90 mg·m^{-3}) and dominance of dinoflagellates (indicated by locally-tuned DPA) occurred in June 2016. On this date, a public environmental report noted the presence of an algal bloom associated with *Noctiluca scintillans* and *Dinophysis acuminata* in the region, which was also related to cases of food poisoning from consumption of bivalves in the area (COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO (CETESB), 2017). Low turbulence of the water column is the foremost favourable condition for these organisms (GLIBERT, 2016; MARGALEF, 1978), whereas their nutritional strategies such as the practice of mixotrophy (LEE, 2008) are also relevant.

The SACW was observed to reach the surface layer only once in the entire series (May 2017), according to the TS diagram (Figure 4.13). In this case, the station was dominated by diatoms, as indicated both by the DPA (58%), and the TChl$a$ was higher than average (1.3 mg·m^{-3}). The SACW is an important source of new nutrients and promotes new production at the SBB, with coastward sub-surface intrusions and upwelling cells, which occur especially during the austral spring-summer (BRANDINI et al., 2018; METZLER et al., 1997). This event (May 2017) was characterized by the unusual formation of three upwelling cells off Rio de Janeiro, Santos bight, and São Sebastião Island (which is close to the Antares-Ubatuba station), starting 3 days before the campaign (Figure 4.14 top).

Figure 4.14 - Analysed sea surface temperature (NASA/JPL, 2015) for May 21th, 2017 (top) and November 21th, 2012 (bottom), upwelling events (identifiable from their light blue colour).
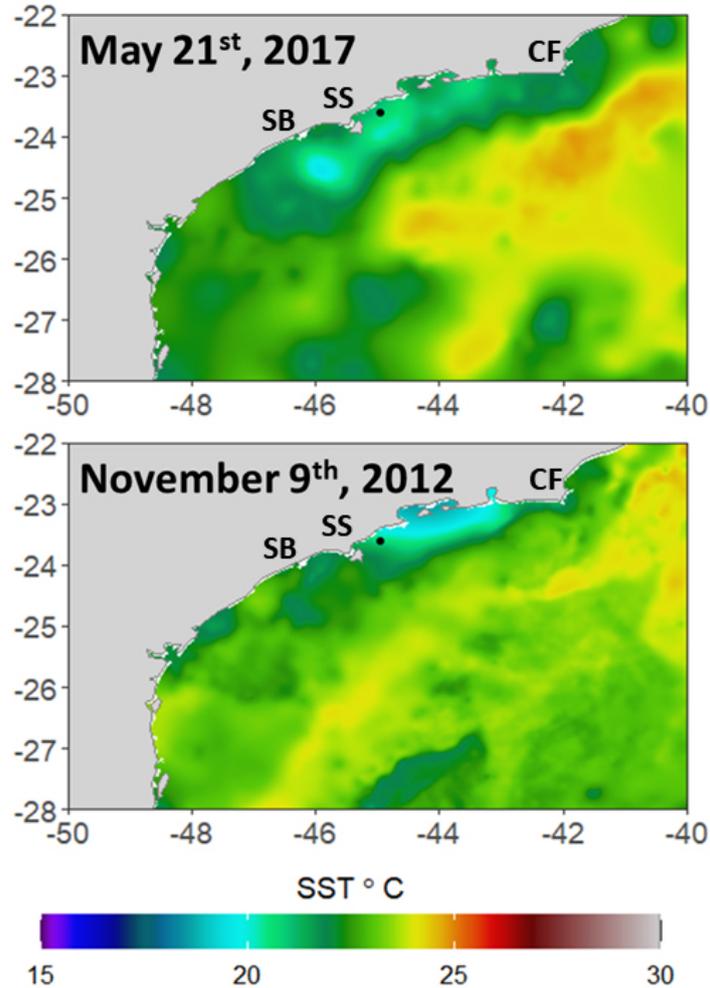
The black point in the map indicates the Antares-Ubatuba time-series station. The references in the images are CF for Cape Frio, SS for Sao Sebastiao and SB for Santos Bight.

SOURCE: Author's production.

Intrusions of the SACW at the Cape Frio upwelling cell may also remotely influence the Ubatuba site during strong upwelling events, forming the southward CFUF (BRANDINI et al., 2018), which is better detected using satellite SST and Chl$a$ maps rather than the TS diagram, since SACW mixes with CW at the study site. These intrusions are usually rich in nutrients leading to an increase in TChl$a$, and blooming of phytoplankton assemblages associated with new production, such as diatoms. For

instance, the nutrients concentrations were 0.92 μM/L for silicate and 0.34 μM/L for nitrate, in September 2012 (two months before the CFUF formation), increasing to 4.41 μM/L of silicate and 0.71 μM/L of nitrate in November 2012, when the CFUF reached the study site (Figure 4.14 bottom). These nutrient inputs are reflected in the TChl$a$ that increased from 0.40 to 10.0 mg·m^{-3} in this period.

However, nutrient depletion and/or high grazing rates can limit high phytoplankton biomass accumulation after the upwelling events, during the relaxation phase (CARBONEL; VALENTIN, 1999). In both major Cape Frio upwelling events recorded in our series, with the CFUF reaching the Ubatuba site, and with the highest TChl$a$ values observed, i.e., November 2012 (10 mg·m^{-3}) and February 2014 (6.22 mg·m^{-3}), grazing rates were also high, as indicated by the (Pheophorbide-a + Pheophytin-a)/TChl$a$ ratio (0.19 and 0.13, respectively). These high grazing rates are likely one of the reasons for mixotrophic phytoplankton to be dominant or co-dominant in these two events (*Mesodinium major* and *Dinophysis*, respectively).

In the first case (November 2012), the sea surface temperature measured at the Ubatuba site indicated the presence of the SACW (17 °C), even though the salinity was close to 33.5 psu. Satellite SST confirmed the influence of the CFUF, which started 4 days before the campaign (November 9_{th}) (Figure 13 bottom). The phytoplankton community was co-dominated by diatoms and *Mesodinium major* with symbiotic association with cryptophytes (as indicated by the DPA with 52% diatoms and 12% cryptophytes, and microscopy analyses with 60% nanoflagellates and 35% diatoms). Since the upwelling event had started just 4 days before the campaign, the nutrients were still abundant (as previously noted), sustaining high phytoplankton biomass accumulation with good physiological status of the population (low senescence rate of 0.02, calculated as Chlorophyllide-a/TChl$a$).

In the second case (February 2014), although the CFUF was present in the study region (as indicated by regional satellite SST (NASA, 2015) and Chl$a$ (MODIS-Aqua), not shown here), the sea surface temperature measured at the Antares-Ubatuba station was much higher (27°C), indicating the prevalence of CW at surface waters. The DPA indicated dominance of cryptophytes (57%), which in this case, and differently from the first event, was associated with the dominance of dinoflagellates (*Dinophysis*), as indicated by microscopy analysis, with symbiotic association. The proportion of diatoms, indicated by both DPA (12%) and microscopy was also much lower. In this case, the upwelling event started more than two months before the campaign (in December 2013). Hence, the sampling happened at the end of the

relaxation phase, and the nutrients were below the average for the region, 1.36 and 0.15 μM/L, for silicate and nitrate, respectively.

4.4 Discussion

The Antares-Ubatuba coastal time-series station was characterized by oligo-mesotrophic waters in most of the analysed period, with generally a relatively low variability of TChl$a$, except for episodic events that promote phytoplankton blooms, occurring mainly during the spring-summer season, and were associated with SACW upwelling in Cape Frio eventually reaching the study site. The two observed blooms were similar in their high concentrations of the cryptophyte biomarker pigment, Allo. However, the first bloom event was associated with the mixotrophic ciliate *Mesodinium major*, co-dominated with diatoms, as shown by microscopy analysis. According to microscopy, the second event was a dinoflagellate bloom, although not matched by a high concentration of peridinin, a dinoflagellate biomarker pigment. These discrepancies between the phytoplankton group indicated by DPA and the dominant groups indicated by microscopy reveal the existence of a symbiotic (klep-toplasty) relationship between the ciliate and cryptophytes (JOHNSON et al., 2016) in the first bloom event; and between *Dinophysis* and *Mesodinium major* in the second event (RIAL et al., 2013). In addition, fucoxanthin was present in much lower concentrations during the second event, which was likely a consequence of sampling in different phases of the bloom, with the second event corresponding to a relaxation phase (with nutrient depletion). The microscopy analysis revealed indeed a much lower abundance of diatoms, and the fucoxanthin present was also likely shared by the dominant dinoflagellates, as this pigment is present even in those species that do not present peridinin (JEFFREY et al., 2011). These findings highlight the importance of utilising complementary methods of HPLC pigment measurement and microscopy analysis to understand better the dynamics of phytoplankton assemblages and their physiological conditions.

All 17 pigments (except for DVChl$b$) analysed were present all year round, suggesting the coexistence of mixed assemblages, as previously reported for the Ubatuba site (**??**). Some pigments, especially diagnostic of nano- and microplankton (JEFFREY, 1997; UITZ et al., 2006), had higher mean concentrations during the autumn-winter (see Table 4.2). The autumn-winter season is more influenced by 'continuous' inputs of new nutrient sources into the euphotic layer by convection and wind-driven mixed-layer deepening, and the northward advection of the SSF with PPW (BRANDINI et al., 2018; PIOLA et al., 2000). The increase in phytoplankton biomass during the

austral winter is relatively well known and reported, especially in the southern SBB region (BRANDINI et al., 2014). However, in the north portion of the SBB the increase in the biomass due to the SSF was not clearly observed, which indicates nutrient depletion along the SSF.

Pigment information can be used as a taxonomic tool to describe the phytoplankton community and as a proxy for physiological responses under distinct environmental conditions, such as light availability and grazing pressure (JEFFREY, 1997). In almost the entire data set, the most dominant diagnostic pigment was Zea (see Figure 3). This pigment is an essential photoprotective carotenoid, usually related to the cyanobacteria group (HIRATA et al., 2011; UITZ et al., 2006), although it also occurs in other groups of phytoplankton (JEFFREY et al., 2011). High Zea to TChl$a$ ratios are expected under high-light conditions or during the senescent stage (declining concentrations of pigments) (LIMA et al., 2019). Therefore, it could indicate the physiological state of the cells and different light regimes depending on the season (e.g., summer and winter). In this study, Zea concentrations and the sum of Photoprotective Carotenoids (PPC) were not significantly different between spring-summer and autumn-winter periods (see Table 4.1), indicating that seasonal irradiance difference is not an important issue for this subtropical coastal region.

The relative content of chlorophyll-a degradation products can be used as a proxy for grazing pressure and the senescence of phytoplankton cells (JEFFREY, 1997). High proportions of pigments indicative of grazing (Pheophorbide-a +Pheophytin-a) were more frequently observed in our data set than that indicative of senescence (Chlorophyllide-a), suggesting that in the study area, phytoplankton are promptly consumed.

The phytoplankton groups identified in the microscopy analysis and indicated by the DPA using hierarchical cluster analysis have already been reported in the SBB region (BRANDINI et al., 2014; LIMA et al., 2019; MOSER et al., 2014), and more specifically, for the coast of Sao Paulo (**??**VILLAC et al., 2008). For the Ubatuba site, **??**) mentioned the dominance of nanoflagellates (Chlorophyceae, Prasinophyceae, Haptophyceae, Chrysophyceae, and Cryptophyceae), followed by diatoms (Bacillariophyceae), dinoflagellates (Dinophyceae), coccolithophorids (Haptophyceae), silicoflagellates (Chrysophyceae), filamentous cyanobacteria (Cyanophyceae) and euglenophytes (Euglenophyceae).

Various global approaches to obtain the phytoplankton groups through DPA are similar to each other, but there are small differences in strategies adopted to couple

proportions of pigments with particular taxonomic groups or size classes. These differences are especially important to differentiate picoeukaryotes from nano- and micro-eukaryotes in oligotrophic waters with TChl$a$ < 0.25 mg·m^{-3} (HIRATA et al., 2011) and ultra-oligotrophic waters with TChl$a$ < 0.04 mg·m^{-3} (BREWIN et al., 2010). In this study, we did not have such oligotrophic conditions, and the focus was on the phytoplankton taxonomic groups identified in the hierarchical cluster analysis. For this reason, we used the Uitz et al. (2006) approach.

However, it should be noticed that when aggregating the phytoplankton groups defined by the hierarchical cluster analysis (i.e., diatoms, dinoflagellates, cryptophytes, haptophytes, green algae, and cyanobacteria) into the PSCs (i.e., pico-, nano- and microplankton), some of these groups could be wrongly aggregated. Considering this, we tuned the PSC aggregation, in the case of dinoflagellates and diatoms, following our microscopy analysis, which indicated that around 50% of the dinoflagellates and 30% of diatoms were part of nanoplankton. The contribution of diatoms and dinoflagellates to the nanoplankton fraction has also been indicated by Chase et al. (2020). In their study, the authors evaluated the accuracy of DPA using imaging flow-cytometry data. Their results showed that the DPA overestimated microphytoplankton and picophytoplankton compared with flow cytometry data, and underestimated nanophytoplankton contribution to total carbon biomass. We also found a similar pattern in our analyses comparing the DPA to microscopy data, with overestimation of diatoms, mostly composed of micro-sized cells, and underestimation of the nano fraction. Hence, limitations of both methods, i.e., DPA and microscopy, associated in the former method with ecophysiological variability of the DP ratios, and in the latter method with over-representation of nano-sized groups, for instance, are still a challenge when it comes to improving size class assignments. This result is especially important for our study area, which presents a high proportion of nanophytoplankton, suggesting that additional adjustments should be made to improve the accuracy of results when applying the DPA approach to determine phytoplankton groups and PSCs. Despite our attempt to tune the distribution of dinoflagellate and diatoms in nano and microplankton, the Antares-Ubatuba microscopy dataset analysed here was limited to a few samples; therefore, further analysis with a more extensive dataset using techniques such as flow cytometry should be considered in future studies. We also have to recognise that, because of C:Chl$a$ ratios that might vary between different classes of phytoplankton, and with photo-acclimation, carbon-based and chlorophyll-based biomass may not always match each other, even when the methods are perfect.

Diatoms biomass estimated by cell metrics may be influenced by the volume of the vacuole, that may vary depending on the species, ranging from 22% to 70% of the total cell volume (SICKO-GOAD et al., 1984), leading to overestimation in the diatom biomass (see Figure 4.10). We were not able to estimate the volume of the vacuole by microscopy in the fixed cells, but this must be considered in future studies. Furthermore, the assumptions to convert biovolume to carbon mass may not apply for some diatom groups, as the allometric approach requires a simplification and "ignores inherent species-specific variability to provide average estimates and grounds for comparison" (MENDEN-DEUER; LESSARD, 2000).

The locally tuned DP model proposed in this study presented a strong agreement with the regional CHEMTAX results, which was more detailed regarding the phytoplankton groups. In addition, CHEMTAX analysis has already been used to characterize the phytoplankton community in the shelf break of the SBB, focusing, however, on the distribution of the cyanobacteria *Trichodesmium* (see Lima et al. (2019)).

With regard to the phytoplankton optical properties, the $a_{ph}^*(\lambda)$ presented a flattened curve for the stations with a high relative contribution of diatoms (usually micro- and nanoplankton) and sharped curves for stations with a high relative contribution of cyanobacteria (picoplankton), as expected (see Ciotti et al. (2002), Devred et al. (2006), Sathyendranath et al. (2001)). The results observed for the $Q_a^*(440)$ also reinforced this expected pattern (i.e., stations with a higher contribution of diatoms had lower values whereas those with a higher contribution of cyanobacteria had higher $Q_a^*(440)$ values) (BRICAUD et al., 2004). On the other hand, S_f results were not so straightforward, presenting a low agreement with TChl$a$ and the phytoplankton groups and size classes determined by the DPA.

Some samples presented low values of S_f, even though cyanobacteria were more abundant. It is not clear why S_f values were relatively low, considering that cyanobacteria are predominantly picoplankton size organisms, and we expected that the campaigns dominated by these organisms would present high S_f values (>0.5) (CIOTTI et al., 2002; CIOTTI; BRICAUD, 2006). A possible explanation for the relatively S_f lower values could have been the presence of *Trichodesmium*, a filamentous cyanobacterium ($\sim$1 mm length), which would change the packaging effect associated with cyanobacteria (DETONI; CIOTTI, 2020). However, *Trichodesmium* colonies were detected through CHEMTAX in only one of these campaigns, and they were not detected in the microscopy analysis. Another possibility is that the presence of

mixotrophic symbiotic species could be contributing to lower S_f values. Microscopy analysis showed the occasional presence of a species of diatom *Hemialus sp.* that is known to occur in a symbiotic relationship with the cyanobacteria *Richelia sp.* (requiring further analysis). We also note that the base vectors for calculation of S_f, as proposed by Ciotti et al. (2002) and Ciotti and Bricaud (2006), were not tuned for the region, which could have affected the performance of the method.

Even though S_f algorithms have already been applied in the SBB region (see Ciotti and Bricaud (2006)), it was evaluated on a spatial distribution basis, in the presence of high gradients in TChl*a* (from productive coastal waters to oligotrophic offshore conditions). Temporal variability at a coastal station, such as those explored in the present study, is more subtle, and therefore likely less suited for analysis by the S_f index. Another point that must be considered here is that S_f estimation was independent of the HPLC dataset, since we used the mean absorption spectra normalization (see Ciotti et al. (2002)), different from the $Q_a^*(440)$ and $a_{ph}^*(\lambda)$ which required the concentrations of the pigments (see Bricaud et al. (2004)) for their calculation. Hence, a_{ph}^* and Q_a^* were not entirely independent of the pigment data, probably compensating for measurement uncertainties in a_{ph}. Phytoplankton optical properties are the basis of some remote sensing algorithms to retrieve phytoplankton functional types and phytoplankton size classes (DEVRED et al., 2006; DEVRED et al., 2011; MOISAN et al., 2017; MOUW et al., 2017; NAIR et al., 2008; RUDORFF; KAMPEL, 2012; SATHYENDRANATH et al., 2001). For this reason, results and relations that diverge from those expected for remote sensing algorithms should be investigated to evaluate the particularities of the phytoplankton assemblages in the study area and to improve potential applications and identify sources of uncertainties. Abundance-based approaches, relating the PSC to the TChl*a* (BREWIN et al., 2010; IOCCG, 2014; UITZ et al., 2006) or approaches that combine abundance and optical properties (BREWIN et al., 2010), could perhaps be more suitable than approaches based solely on specific optical properties for this coastal region. Further investigations should be undertaken in the future to explore these possibilities.

Overall, our results showed a similar scenario as described by Moser et al. (2014), mainly characterized by the presence of a typical (sub)tropical phytoplankton community, i.e., dominated by stress-tolerant pico- and nanoplanktonic taxa, associated with oligotrophic waters of the Brazil Current (with TW on the surface). The Brazil Current is a western boundary current with intense mesoscale and sub-mesoscale activities that influence the inner-shelf of the study area (CALADO et al., 2006; SOUTELINO et al., 2013). Brazil Current meandering has been associated with SACW

upwelling (CAMPOS et al., 2000), advection of cyclonic and anticyclonic eddies (CALADO et al., 2006), and interactions with shelf fronts under the influence of river discharge (MÖLLER et al., 2008). These dynamic processes affect the availability of nutrients and water column stratification, causing alterations in the phytoplankton community, increasing the contribution of ruderal taxa (R-strategists), and reducing the stress-tolerant taxa (MOSER et al., 2014). When such processes occur, the phytoplankton community shifts to diatom-dominated assemblage; however, this scenario could quickly change to a poor biomass condition due to nutrient depletion or as grazing rates increases.

Ribeiro et al. (2016) characterized the phytoplankton community in the SBB using flow cytometry, focusing on the distribution of pico- and nanoplankton that are poorly studied in the region. The authors reported that the cyanobacteria *Prochlorococcus* were more abundant in oceanic, oligotrophic, surface waters (TW). Our CHEMTAX results showed that the relative contribution of *Synechococcus* was always higher than *Prochlorococcus*, indicating, as suggested by the TS diagram (Figure 4.13), that TW hardly reaches the inner-shelf region (see Cerda and Castro (2014)).

The occurrence of major algae bloom events in the study area was registered at the Ubatuba coastal time series station in 2012 and 2014, associated with the CFUF. Minor bloom events with potential occurrence of red tides were also reported by the São Paulo state water agency (CETESB, 2018; COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO (CETESB), 2017), such as in 2016, associated to the PPW SSF, and in 2017 with local upwelling. Although these events cannot be considered as frequent in the SBB, their early identification and alert are important due to their potential impacts on social and economic activities (CIOTTI et al., 2018). As already suggested, a long term monitoring strategy should be promoted in the SBB region (see Brandini et al. (2018), Ciotti et al. (2018)).

4.5 Conclusions and final considerations

In the present study, a time-series of HPLC pigments, microscopy analysis and bio-optical data acquired in a coastal region of the SBB was analysed to characterize the phytoplankton composition and the associated oceanographic conditions. The DPA with the hierarchical cluster analysis provided a set of phytoplankton groups that were applied to the locally-tuned Uitz et al. (2006) model, which yielded results that were consistent with other studies in the region and with the microscopy analysis. Changes in the phytoplankton assemblages were also evident in the optical propri-

eties such as $a^*_{ph}(\lambda)$, $Q^*_a(440)$ and S_f. Further analysis is required to understand the S_f variability in the study area.

From the dataset analysed here, two major algal bloom events were identified in 2012 and 2014 (TChl$a > 5$ mg·m^{-3}), and two minor blooms in 2016 and 2017, caused by different meteo-oceanographic processes, and dominated by different phytoplankton species. These events were also reported by the São Paulo state water agency (CETESB), but considering the low sampling frequency of the Antares-Ubatuba station (with monthly collections at best), other minor bloom events could have occurred in the study region without being noticed or reported in the literature. Moreover, we analysed a relatively short period (2012–2017) in which we had simultaneous HPLC, microscopy and optical data, and although we have captured the main processes that promote phytoplankton blooms in the SBB, other processes such as mesoscale features e.g., eddies and meanders of the Brazil Current, and the passage of strong atmospheric cold fronts, are also likely to influence the phytoplankton assemblages and succession. Considering the importance that some blooms can present, it is essential to maintain efforts to implement and to sustain continuous sampling in time-series stations. Long time-series are also required for climate-change studies (BREWIN et al., 2015a; GROOM et al., 2019; SATHYENDRANATH et al., 2017).

Based on the microscopy and the locally-tuned DPA, we can conclude that the phytoplankton community is characterized by mixed assemblages composed mainly of diatoms, nanoflagellates, and picoplankton (*Synechococcus* cyanobacteria and green algae) in mainly oligomesotrophic conditions, that are occasionally changed by the intrusion of oceanic fronts (CFUF and SSF) and local upwelling. When these events result in the intrusion of nutrient-rich waters, the assemblages change to higher biomass conditions, usually accompanied by diatom dominance. The CFUF and local upwelling potentially causes bloom events, more likely occurring in austral summer and spring. A summary diagram was included to illustrate these findings (Figure 4.15).

Figure 4.15 - Diagram with the oceanographic conditions and the potential consequences for the phytoplankton assemblages. Darker green in the diagram illustrates higher concentrations of chlorophyll-*a*, and the symbols representing the phytoplankton groups are illustrative, representing most likely groups found in the described conditions. Even though they are not exactly quantitative results, they illustrate expected abundances of these groups at the sea surface, in the study area.

Normal Conditions	Local Upwelling	Plata River Plume Water (Subtropical Shelf Front)	Cape Frio Upwelling (Cape Frio Front)
35 >Salinity >32 Temperature>22°C	36 >Salinity >34 Temperature<20°C	Salinity<33 Temperature <22°C Requires satellite SST to confirm front	36 >Salinity >34 Variable temperature Requires satellite SST to confirm front
• Oligo-mesotrophic conditions • High diversity • Most common groups are diatoms, green algae, and cyanobacteria.	• Increase in chlorophyll • Mostly diatoms • High diversity, specially in the beginning of the upwelling • **Potential algal bloom**	• May cause an increase in chlorophyll • Mostly diatom and dinoflagellates, but may also be associated to cyanobacteria depending on the depletion of nutrients • **Potential algal bloom**	• High chlorophyll • Low phytoplankton diversity (mostly diatoms) • Grazing activity and nutrient depletion causes shift in the dominant species from autotrophs to mixotrophic species (such as ciliates and dinoflagellates) • **Potential algal bloom**

Diagram elaborated by the authors with use of images from Tracey Saxby, Integration and Application Network (ian.umces.edu/media-library).

SOURCE: Author's production.

5 REMOTE SENSING ESTIMATION OF PHYTOPLANKTON SIZE CLASSES IN A SUBTROPICAL COASTAL SITE: AN EVALUATION

5.1 Introduction

Understanding the phytoplankton community variability is key to managing marine areas and anticipating the effects of climate change in marine ecosystems. In the last two decades, several approaches were designed to obtain remote sensing estimations of phytoplankton size classes (BREWIN et al., 2010; DEVRED et al., 2011; SATHYEN-DRANATH et al., 2001; UITZ et al., 2006). These models have been tested both globally and locally, including coastal and continental shelf waters (LIU et al., 2021; GITTINGS et al., 2019; SUN et al., 2019; TURNER et al., 2021).

The challenge in applying these models to coastal waters is that the models' assumptions may not suit well to all types of waters (as discussed in Chapter 3), due to the diverse phytoplankton assemblages (see KRAMER et al. 2020 and KRAMER; SIEGEL 2019). For instance, different sources of nutrients may drive changes in the phytoplankton communities by adding nutrients in the euphotic zone through upwelling which can potentially change the dominant group into the water and increase overall abundance (see SARKER et al. 2020) or by river inputs adding nutrients and potentially changing the light penetration in the water column which can indirectly select phytoplankton groups (see SARKER et al. 2020 and WILTSHIRE et al. 2015), and grazing pressure over the phytoplankton through a top-down pressure in the phytoplankton community (see WILTSHIRE et al. 2015), changing the abundance and dominance according to the herbivores preference. All those processes are potentially able to induce changes in the predefined relationships established between Chl*a* and the phytoplankton size classes derived from the DPA, which are the basis for most of the PSC models.

Among the available PSC models, the abundance-based approaches, such as the one described in Brewin et al. (2010), seem to be one of the most extensively applied and re-tuned to different regional seas (LIU et al., 2021; GITTINGS et al., 2019; SUN et al., 2019; TURNER et al., 2021; LAMONT et al., 2018b). The fact that they rely only on Chl*a* makes it easier to apply the model to satellite data since Chl*a* products are routinely estimated and available for download. This is not the case with spectral-based models, such as the one proposed by Devred et al. (2011), as it requires the phytoplankton absorption coefficient in different wavelengths. This turns the application of this model more time-consuming, especially considering optically-complex waters, where the semi-analytical models to retrieve phytoplankton absorption co-

efficients are susceptible to the interference of high coloured dissolved organic matter absorption and high scattering by non-algal particles (DEFOIN-PLATEL; CHAMI, 2007). Indeed, empirical algorithms for Chl*a* are also affected by the interference of other inherent optically active constituents, that is why a regional tuning for this algorithm is also indicated.

In this study, we will look forward to answering the following research questions: (1) Do these models respond well in this particular subtropical area? (2) Is the performance of the models maintained when applied to satellite data? That will be answered with the following specific objectives: (a) to parameterize and evaluate the results of two PSC models (the abundance-based approach proposed by Brewin et al. (2010) and the spectral-based approach by Devred et al. (2011)) to a subtropical coastal area in the South Brazil Bright (SBB), (b) to use the models to identify the dominant size class applying them to satellite data (MODISA imagery) with Generalized Inherent Optical Properties framework with Default Configuration (GIOP-DC) to Chl*a* and phytoplankton absorption coefficients, and (c) use the models to characterize the spatial and seasonal variability of Chl*a* and phytoplankton size structure in the SBB.

5.2 Methods

5.2.1 Study area

The study area and the sampling station were described in the Chapter 4, Section 4.2.1. The sampling period was from July 2006 to July 2019. From 2006 to 2015 it was monthly sampled, and from 2016 on it was less frequently sampled, varying from 3 to 6 months interval between campaigns.

5.2.2 Measurements

The phytoplankton absorption coefficients, phytoplankton diagnostic pigments, nutrients, and temperature were measured as described in the Chapter 4, Section 4.2.2.

5.2.2.1 Flurimetric method for chlorophyll-*a* concentration

HPLC analysis were not available for the entire time-series period. Thus, for the most part of the series the Chl*a* available was the obtained by the fluorometric method. Chl*a* was measured following the method described in Welschmeyer (1994) and Shoaf and Lium (1976). Samples were filtered in Whatman GF/F (0.7 μm), kept in liquid nitrogen until laboratory analysis, when the filters were then immersed in

5 ml of a solution of 90% acetone:DMSO (6:4 by volume) (SHOAF; LIUM, 1976) and kept at -4ºC for 24 hours for the pigment extraction, and finally measured in a calibrated Turner 10- AU-005 fluorometer (WELSCHMEYER, 1994). The total chlorophyll-*a* concentrations obtained by HPLC were used to correct the fluorometric Chl*a* (see Appendix A) to standardized the concentrations used in both PSC approaches.

5.2.2.2   Absorption coefficient of CDOM

For the coloured dissolved organic matter (CDOM) absorption coefficient the water samples were filtered in 0.2 µm polycarbonate membranes (Whatman nucleopore) following Mitchell et al. (2002). Samples were stored in pre-combusted glass bottles wrapped with aluminum foil and kept under refrigeration (4°C) until laboratory analysis. The filtered samples were put to room temperature before the analysis to avoid bias due the thermal difference between the samples and the reference water (Milli-Q water). Using a 10 cm quartz cell, the absorbance spectra were measured from 250 to 850 nm in a spectrophotometer (Shimadzu, UV 2450). The CDOM absorption coefficient was then calculated using the following equation:

$$a_{cdom}(\lambda) = \frac{2.303}{L} \cdot [(OD_{sample}(\lambda) - OD_{blank}(\lambda)) - OD_{offset}] \qquad (5.1)$$

Where L is the optical pathway length in meters, i.e., the quartz cell length (0.1m), $OD_{sample}(\lambda)$ is the optical density of the sample, $OD_{blank}(\lambda)$ is the the Milli-Q water optical density, measured in the spectrophotometer and the OD_{offset} is the mean signal in the offset interval at 683-687nm.

5.2.3   Model parameterization

Even though these models have already been applied to retrieve three size classes (pico, nano and microphytoplankton) (BREWIN et al., 2010; DEVRED et al., 2011), in the Chapter 4 we showed that for the available *in situ* data set of the ANTARES-Ubatuba station the distinction between micro and nanophytoplankton populations is complicated when using solely diagnostic pigments. Therefore, we decided to keep only two populations, i.e., picophytoplankton and a combined population of micro and nanophytoplankton.

5.2.3.1   Abundance-based approach (BREWIN et al., 2010)

The abundance-based approach was described previously in the Chapter 3, Section 3.2.4. The samples used for the parameterization were collected from 2012-2017 and

had the HPLC analysis. Initially, the number of samples was 41, but two of them were removed from the parameterization because they were identified as algal bloom events, thus the final number of samples for the parameterization was 39.

5.2.3.2 Spectral-based approach (DEVRED et al., 2011)

The re-parameterization of the Devred et al. (2011) model requires phytoplankton absorption coefficient and Chl$a$ (n=80). The phytoplankton absorption coefficient was fitted to the Chl$a$ to retrieve the specific phytoplankton absorption coefficient of two-components size classes population:

$$a_{ph}(\lambda) = C_{max}^p \cdot [a_{pico}(\lambda) - a_{m,n}(\lambda)] \cdot [1 - exp(-S_p \cdot Chla)] + a_{m,n}(\lambda) \cdot Chla \quad (5.2)$$

Where a is the phytoplankton absorption coefficient in a specific wavelength (λ), C_{max}^p and a_{pico} are the maximum asymptotic chlorophyll and the specific phytoplankton absorption coefficient for picophytoplankton, respectively, $a_{m,n}$ is the combined component of micro and nanophytoplankton and S_p is the initial slope of picophytoplankton.

Devred et al. (2011) have not specifically classified the populations in pico, nano and microphytoplankton. Instead, the authors classified them as small, medium and large cells. Here, we opted to use the pico, nano and microphytoplankton classification to avoid confusion with the other model (BREWIN et al., 2010). In addition, Devred et al. (2011) combined the small and medium cells, while considering the local characteristics of the study area, and we decided to combine the medium and large cells, as described previously. Thus, from now on, small cells will be referred as picophytoplankton, and medium and large cells as micro and nanophytoplankton.

After tuning the specific phytoplankton absorption coefficients for each size class, the values for a set of wavelengths were used to estimate the Chl$a$ for each size class, using the phytoplankton absorption coefficient as input. Devred et al. (2011) used five SeaWiFS wavelengths, i.e., 443, 490, 510, 555, 670 nm, but here, we opted to use ten MODIS wavelengths, i.e., 412, 443, 469, 488, 531, 547, 555, 645, 667, 678 nm. The equation system was solved as described in Devred et al. (2011), with a least-squares with equality and inequality constraints method, using the limSolve package in R (SOETAERT et al., 2009; VANDENMEERSCHE et al., 2009). The method does not require Chl$a$ as an input, thus it is possible to estimate the Chl$a$ for each

size fraction by solving the following equation system:

$$a_{ph}(\lambda_1) = a^*_{pico}(\lambda_1) \cdot C_p + a^*_{m,n}(\lambda_1) \cdot C_{m,n}$$
$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad (5.3)$$
$$a_{ph}(\lambda_{10}) = a^*_{pico}(\lambda_{10}) \cdot C_p + a^*_{m,n}(\lambda_{10}) \cdot C_{m,n}$$

Where $a_{ph}(\lambda_i)$ is the measured phytoplankton absorption coefficient at the i wavelength (i from 1 to 10), C_p is the chlorophyll-a concentrations for picophytoplankton, and $C_{m,n}$ for micro and nanophytoplankton. A lower constrain of 0.005 mg· m⁻³ was defined for the possible values of Chl$a$ concentrations for each size class, following Devred et al. (2011).

The data set with the diagnostic pigments was used as an independent set of data (i.e., not used in the parametrization) to validate and evaluate the spectral approach. It is important to notice that after estimating the parameters in the regional tuning of Devred's model it can also be applied as an abundance-based approach, by using the S_p estimated by solving the Equation 5.2. C^p_{max} as defined in Devred et al. (2011) is obtained by solving the Equation 5.4, below:

$$C^p_{max} \cdot S_p = 1 \qquad\qquad (5.4)$$

5.2.3.3   Inter-comparison (DPA and PSC models)

Regarding the diagnostic pigment analysis (DPA) and the PSC models, an inter-comparison was performed between the PSCs estimated by the DPA (local (Chapter 4) and global DPA (UITZ et al., 2006)) and the PSC models (local (this chapter) and global parameters (BREWIN et al., 2010; DEVRED et al., 2011)). The idea of this comparison is to check if the results of the re-parameterization of the PSC models are actually giving better results than the parameters estimated previously considering the global ocean.

Firstly, considering the locally-tuned DPA presented in Chapter 4 and the results from the re-parameterization of PSC models performed in this chapter; the results obtained for the chlorophyll fraction for each size class were compared. Secondly, the statistical errors were calculated for other combinations, the global parameters from PSC models and the fractions estimated using the locally-tuned DPA (4). Finally, the global parameters were used to estimate the PSC fractions, the global-DPA from Uitz et al. (2006) and global parameters from Brewin et al. (2010) and Devred et

al. (2011) were compared.

5.2.3.4 Residuals and correlation analysis

Following the method adopted by Brewin et al. (2019), after estimating the PSC fractions in the *in situ* data a correlation analysis was performed between the residual (the estimated by the model minus the measured) and the environmental variables available, such as sea surface temperature, nutrients and salinity. The idea of this analysis is the identify if any of these variables still have a correlation to the "unexplained part of the measured values". Brewin et al. (2019) used intervals of temperature in the samples to fit the parameters of the model to those bins and this way used temperature as an input in the model. However, in this study, due to the limited number of HPLC samples available, we will not be able to further apply these results to improve the local model.

5.2.4 Statistical metrics

The statistical metrics applied to compare the estimated values by the models and the measured values were the determination coefficient (R^2), correlation coefficient (ρ), root-mean-square error (RMSE), median absolute difference (MAD), mean absolute error (MAE), and bias. These metrics were calculated using the following equations:

$$RMSE = 10^{\sqrt{\frac{\sum_{i=1}^{n}(\log_{10}(M_i) - \log 10(O_i))^2}{n}}} \tag{5.5}$$

$$MAD = 10^{median(|\log_{10}(M_i) - \log_{10}(O_i)|)} \tag{5.6}$$

$$MAE = 10^{\left(\frac{\sum_{i=1}^{n} |\log_{10}(M_i) - \log_{10}(O_i)|}{n}\right)} \tag{5.7}$$

$$bias = 10^{\left(\frac{\sum_{i=1}^{n} \log_{10}(M_i) - \log_{10}(O_i)}{n}\right)} \tag{5.8}$$

Where M_i and O_i are the modeled and observed values, respectively, and n is the number of observations. When the metrics were applied to Chl*a* data the $\log_{10}$ transformation was applied, considering the known *log*10 distribution of Chl*a* and a_{ph} (CAMPBELL, 1995; WERDELL et al., 2013), as presented in the Equations 5.5-5.8.

5.2.5 Satellite data application

For satellite application NASA MODISA level 1A images were downloaded and processed using SEADAS v8.1 to level 3, with 1.1 km spatial resolution, applying Generalized Inherent Optical Properties framework with Default Configuration (GIOP-DC) described in Werdell et al. (2013) and obtaining the Chl$a$, $a_{ph}(\lambda)$, $a_{dg}(\lambda)$ and $b_{bp}(\lambda)$ for the following wavelengths 412, 443, 469, 488, 531, 547, 555, 645, 667 and 678 nm. The images were then aggregated per day (for validation exercise) and per month (for time-series analysis from 2002-2020). The Chl$a$ was also obtained by the default NASA algorithm, which is the empirical algorithm OC3M, merged with the Colour Index (CI) algorithm proposed by Hu et al. (2012) in the interval of Chl$a$ between 0.15 and 0.2 mg·m^{-3}. From now on we will refer to the default empirical chlorophyll-a algorithm as OCx. For the match-ups we considered the same-calendar-day satellite images, a window of 3 x 3 pixels, only the dates that had 5 or more valid pixels and coefficient of variation lower than 20% for Chl$a$ and $a_{ph}(\lambda)$. The validation exercise first considered the validation of the input variables, Chl$a$ (n=23) and $a_{ph}(\lambda)$ (n=21), and then the estimated values for the PSC fractions.

Additionally, reanalysis of sea surface temperature (SST) from the Group for High-Resolution Sea Surface Temperature (GHRSST), Level 4 multiscale ultrahigh-resolution (MUR) in an optimal interpolation approach on a global 0.01 degree grid (1km) (NASA, 2015) were also downloaded and averaged per month and for the study area to discuss possible relationships involving changes in sea surface temperatures and PSCs.

5.2.5.1 Sensitivity analysis with uncertainties for fractional Chl$a$ and a_{ph} inputs in the PSC models results

The uncertainties in the Chl$a$ and $a_{ph}(\lambda)$ derived from remote sensing reflectances introduces errors in the PSC results. A sensitivity analysis was performed to estimate the effect of these uncertainties, considering the uncertainty of 30% for satellite-derived Chl$a$ and $a_{ph}(\lambda)$. They were applied in the PSC models and then compared with the phytoplankton DPA results.

5.2.6 Schematic diagram

Schematic diagram of the measurements and analysis.

Figure 5.1 - Schematic diagram of measurements and analysis of the Chapter 5.
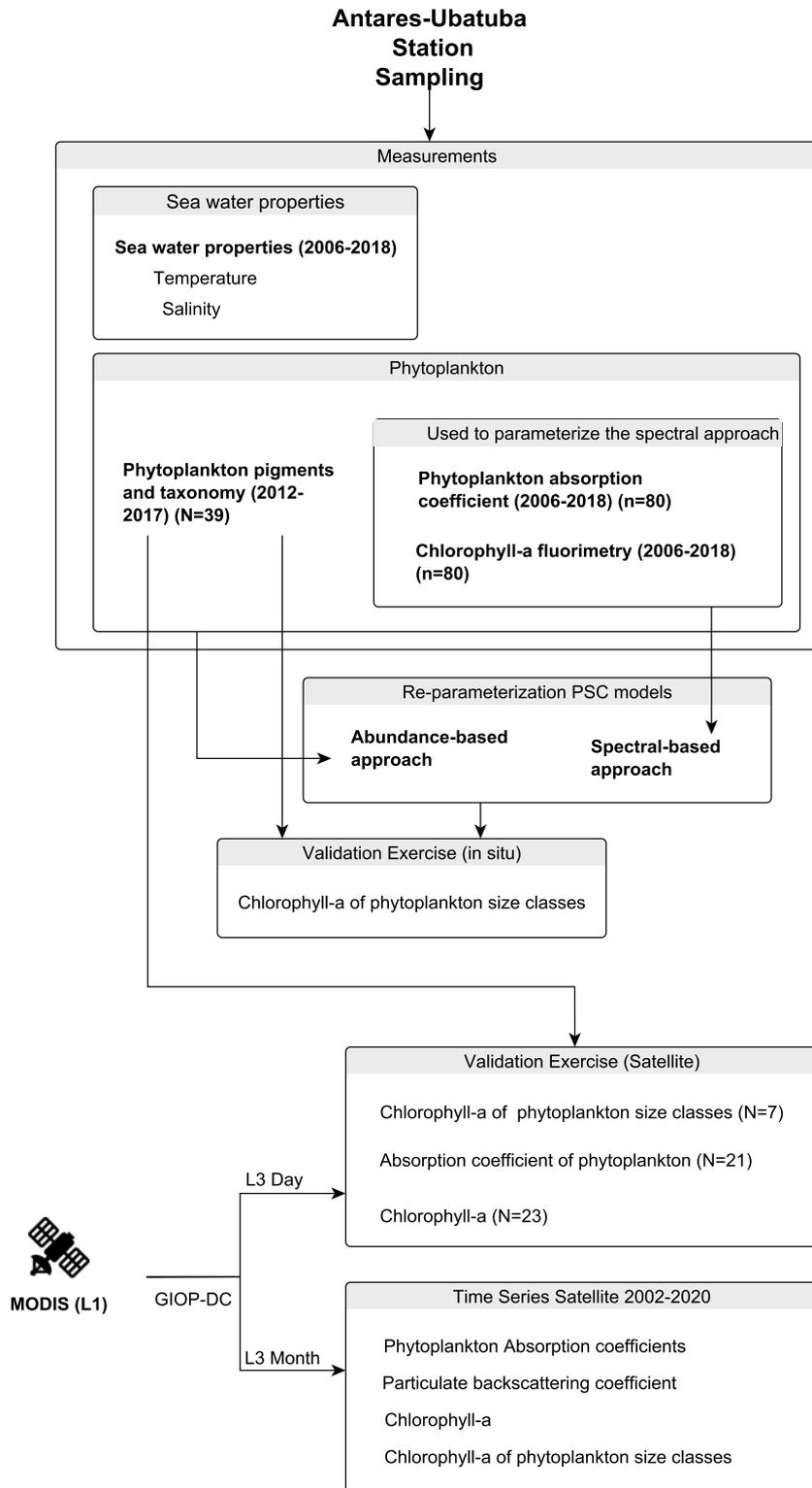
SOURCE: Author's production.

5.3 Results

5.3.1 Model parameterization

The abundance-based and the spectral-based approaches, as described in Brewin et al. (2010) and Devred et al. (2011) are interchangeable, i.e., after estimating one it is possible to estimate the other model's parameters. The parameters estimated in this study were compared to parameters reported for other regions (Table 5.1). We can observe that the initial slope for picophytoplankton (S_p) was relatively low for the abundance approach (1.72) in comparison to the spectral-based approach (4.81) and even comparing to other regions (Table 5.1). The S_p for other regions varied from 3.6 in the Northeast U.S. continental shelf (TURNER et al., 2021) to 8.94 in the Southern Africa region (LAMONT et al., 2018b). The values for the C_{max}^p were less variable, but the value was higher for the Brewin et al. (2010) approach (0.40) than for Devred et al. (2011) (0.21), and also higher compared to other studies which varied from 0.11 for the global ocean (BREWIN et al., 2010) and Southern Africa (LAMONT et al., 2018b) to 0.26 for China coastal waters (West Pacific) (HUAN et al., 2021).

The specific absorption coefficients for picophytoplankton at 443 nm ($a_{pico}^*(443)$) were 0.132 and 0.148, for abundance and spectral-based approach, respectively, values similar to the estimated by other studies, varying from 0.142 in a coastal water of China (HUAN et al., 2021) to 0.21m^{-1} in the Atlantic Ocean (DEVRED et al., 2011). The spectral curves of the specific absorption coefficient for phytoplankton obtained with the two approaches used here, indicated notable differences, more pronounced around 443 nm and 665 nm, especially for the picophytoplankton. Even though the specific absorption coefficients are from the same area, they were estimated with different subsets of the Antares-Ubatuba station dataset. The fit using the Devred et al. (2011) approach presented a curve with high peaks (Figure 5.2 A), while the Brewin et al. (2010) approach presented a flatter curve (Figure 5.2 B). Regarding the specific absorption coefficient for the micro and nanophytoplankton size classes ($a_{m,n}^*$), they were more similar to microphytoplankton values reported by other studies (0.022 to 0.031 m^{-1}) (see Table 5.1). In the present study, the $a_{m,n}^*$ values varied from 0.013 (abundance approach) to 0.028 m^{-1} (spectral approach).

Table 5.1 - Comparison of the model's parameters from this study and in other regions. The two first rows are the parameters estimated in this study, considering the Brewin et al. (2010) and Devred et al. (2011) as spectral and abundance models (since they are interchangeable as explained in the text). The area for the other studies are in parenthesis. Parameters is blank means them were not estimated in the study.

Model	S_p	C^p_{max}	$a^*_{pico}(443)$	$a^*_{m,n}(443)$	$a^*_n(443)$	$a^*_m(443)$
This study (Brewin 2010)	1.72	0.40	0.132 (0.102-0.163)	0.013 (0.009-0.016)		
This study (Devred 2011)	4.81	0.21	0.148 (0.126-0.170)	0.028 (0.015-0.041)		
Brewin et al. (2010) (Atlantic Ocean)	6.80	0.11				
Brewin et al. (2011a) (NOMAD)	5.12	0.15	0.155		0.033	0.022*
Devred et al. (2011) (NW Atlantic)			0.18		0.052	0.031*
Devred et al. (2011) (NOMAD)			0.21		0.064	0.022*
Huan et al. (2021) (West Pacific, coast)	3.85	0.26	0.142		0.0701	0.030
Lamont et al. (2018b) (Southern Africa, coastal)	8.94	0.11				
Turner et al. (2021) (Northeast U.S. continental shelf)	3.6	0.15				

Figure 5.2 - Specific absorption (a^*_{ph}) for picophytoplankton and micro and nanophytoplankton, considering Devred et al. (2011) (A) and Brewin et al. (2010) (B) models, with the confidence interval at 95%.
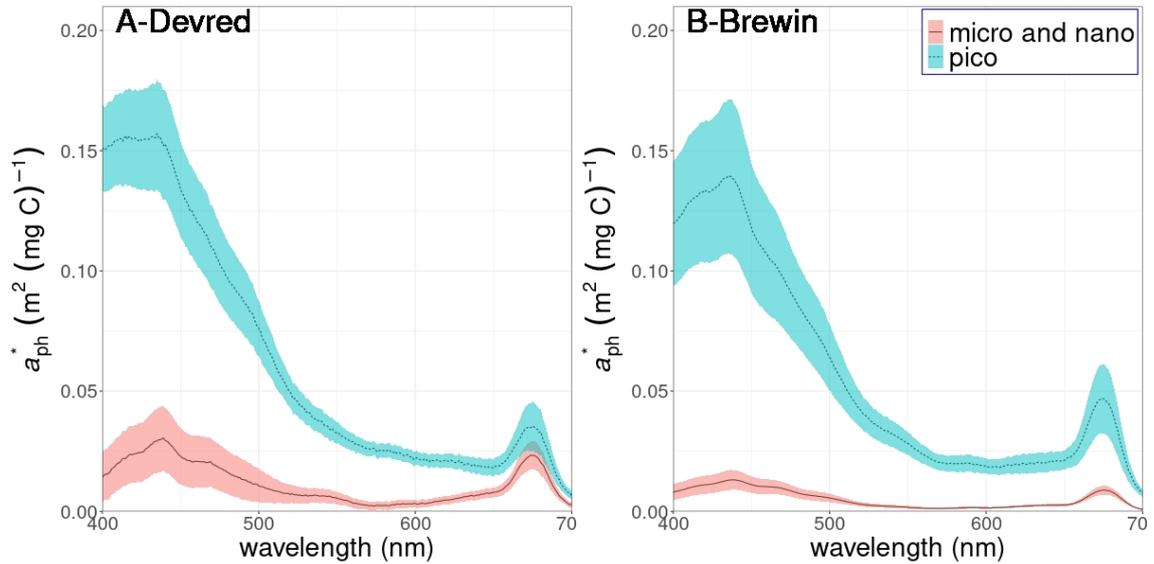
SOURCE: Author's production.

Regarding the inter-comparison for local and global parameters of PSC models in the Antares-Ubatuba station, our results indicated that using locally-tuned parameters reduced the uncertainties of the estimations. The local parameters bias varied from underestimating 27% (for spectral-based approach) to overestimating 12%

(abundance-based approach). Considering the global parameters the bias varied from underestimation of 54% to overestimation of 200% (spectral-based) (see Table 5.2). Interestingly, for the spectral-based approach, the micro and nanophytoplankton size class presented better correlation and determination coefficients for the global parameters, ρ=0.82 (global) and 0.63 (local). However, the picophytoplankton size class was extremely poor for the global parameters, ρ = -0.19 (global) and 0.56 (local).

Table 5.2 - Diagnostic pigments analysis and PSC compared considering the *in situ* data set. Mean absolute error (MAE), root-mean square error (RMSE) and bias for the chlorophyll-*a* concentrations were calculated in log10 and then converted back for the fractions it was estimated with log transformation. ([a]) Brewin et al. (2010) and Devred et al. (2011) original parameters were used to estimate the fractions and compared with the tuned estimations for this study data set; ([b]) Brewin et al. (2010) and Devred et al. (2011) original parameters were used to estimate the fractions and compared with the fractions estimated using Uitz et al. (2006).(*) indicates p-value<0.01.

	Variable	ρ	MAE	RMSE	Bias
This study (abundance)	C_p	0.76*	1.23	1.31	1.02
	$C_{m,n}$	0.94*	1.26	1.36	1.12
	$F_{m,n}$	0.63*	0.09	0.02	0.02
This study (spectral)	C_p	0.56*	1.49	1.61	0.73
	$C_{m,n}$	0.63*	2.2	3.08	0.77
	$F_{m,n}$	0.41*	0.2	0.04	0.04
Brewin et al. (2010)[a]	C_p	0.63*	2.16	2.3	0.47
	$C_{m,n}$	0.94*	1.63	1.76	1.62
	$F_{m,n}$	0.67*	0.26	0.26	0.26
Devred et al. (2011)(NW)[a]	C_p	-0.19	8.65	12.07	0.12
	$C_{m,n}$	0.82*	3.08	3.37	3.03
	$F_{m,n}$	0.51*	0.43	0.43	0.43
Brewin et al. (2010) *vs.* UITZ [b]	C_p	0.67*	2.18	2.34	0.46
	$C_{m,n}$	0.87*	1.81	2.16	1.8
	$F_{m,n}$	0.65*	0.29	0.28	0.28
Devred et al. (2011) (NW) *vs.* UITZ[b]	C_p	-0.24	8.84	12.49	0.11
	$C_{m,n}$	0.76*	3.39	3.95	3.33
	$F_{m,n}$	0.49*	0.45	0.48	0.45

Considering the performance of the PSC models, the abundance approach presented better statistics for the *in situ* validation for all the size-classes and fractions (Table 5.2, Figures 5.3 and 5.4). For the spectral approach, some stations fall in the lower

threshold of 0.005 mg· m^{-3} (Figure 5.3 B). Both approaches underestimated the pico-phytoplankton fraction in stations with algal blooms (Figure 5.3 A and C). However, for the spectral approach as the maximum asymptotic chlorophyll-a concentration and the slope are not directly used in the estimation it is possible to notice that some values were slightly higher than C_{max}^{p} threshold (0.21 mg·m^{-3}) (Figure 5.3 A), implying that for the spectral approach the boundaries are less strict than for the abundance approach.

The statistics for the proportional fractions of the PSCs were consistently poorer than the Chl$a$ of each size class. For the spectral-based approach they were very poor, which could be related the Chl$a$ estimation of the spectral-based approach. The spectral approach Chl$a$ is estimated from the phytoplankton absorption coefficients, for each size class. Comparing the measured Chl$a$ and the estimated Chl$a$ from the spectral-based approach, the statistics were reasonable, with R^2= 0.56, ρ= 0.75, RMSE= 0.274mg·m^{-3} and MDP = 15%, but is a potential source of uncertainties for the PSC estimations.

Considering that the abundance-based approach here was fitted and "validated" with the same dataset (n=39, Chl$a$ average=0.520 , min=0.165, max=1.497mg · m^{-3}), except for the two stations with algal blooms that were removed from the fitting dataset. The abundance approach has an advantage compared to the spectral-based approach in the *in situ* comparison, which was fitted with an independent dataset (n=80, with Chl$a$ average=0.480, min=0.131, max= 2.42mg · m^{-3}), t-test and non-parametric Wilcoxon text did not detect significant difference for the Chl$a$ of these two subsets of Antares-Ubatuba station dataset. However, this must be highlighted and the results considered with caution.

Figure 5.3 - Scatter-plots of the chlorophyll-*a* concentration for each size class modelled by the abundance and spectral-based approaches and *in situ* values.
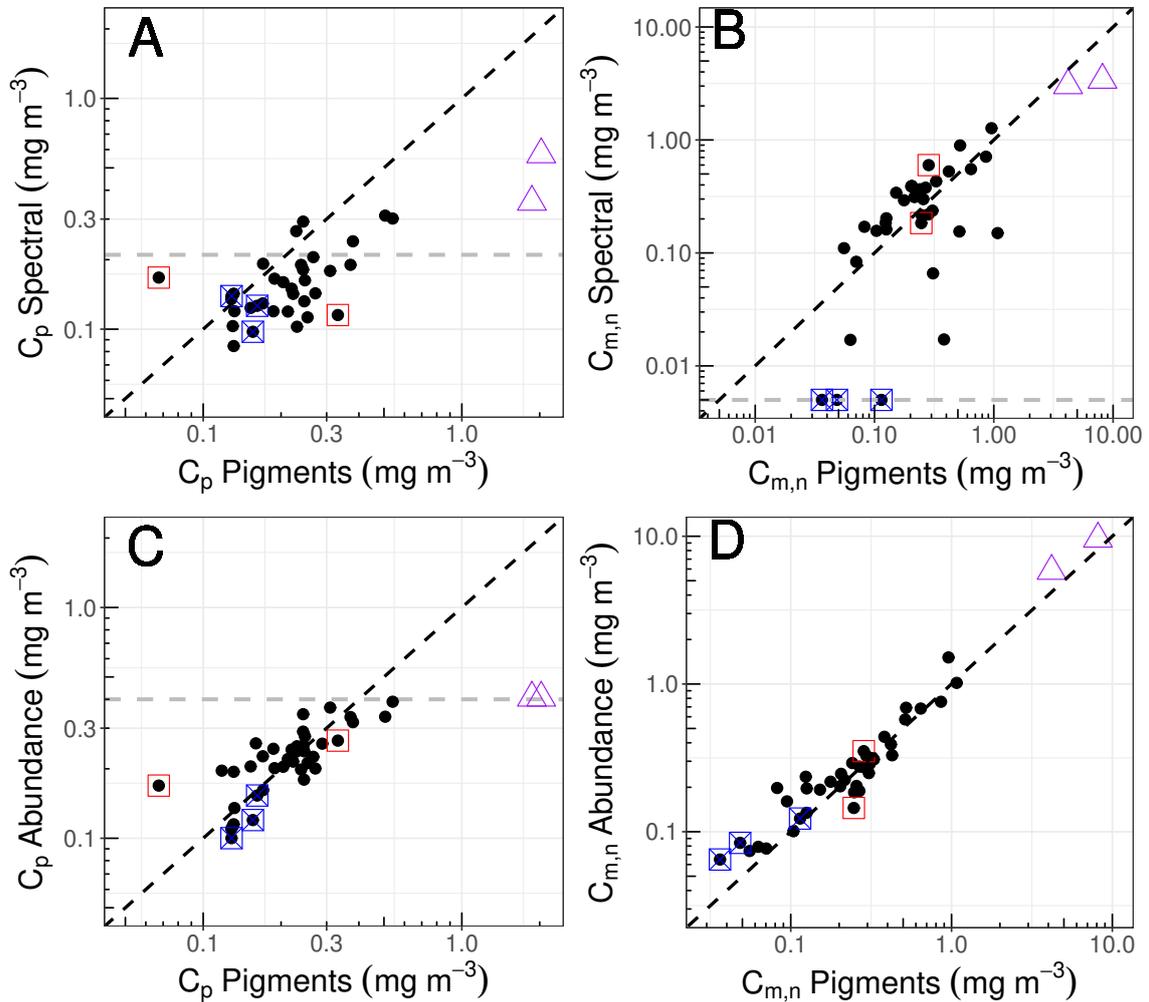
Bloom stations are represented as the purple triangles, the red squares are stations identified as outliers in the spectral-based model estimations for picophytoplankton (DEVRED et al., 2011), the blue squares with a 'x' are the stations in the threshold of the spectral-based model. The diagonal dashed black line is the 1:1 reference, the dashed grey horizontal lines are the threshold of the models, lower bound of 0.005 mg· m^{-3} in Figure B and upper bound in 0.4 mg· m^{-3} in Figure C.

SOURCE: Author's production.

Figure 5.4 - Plots of the PSC determined using the abundance-based and spectral-based models *vs.* the DPA (fractions of each size class varying from 0-1).
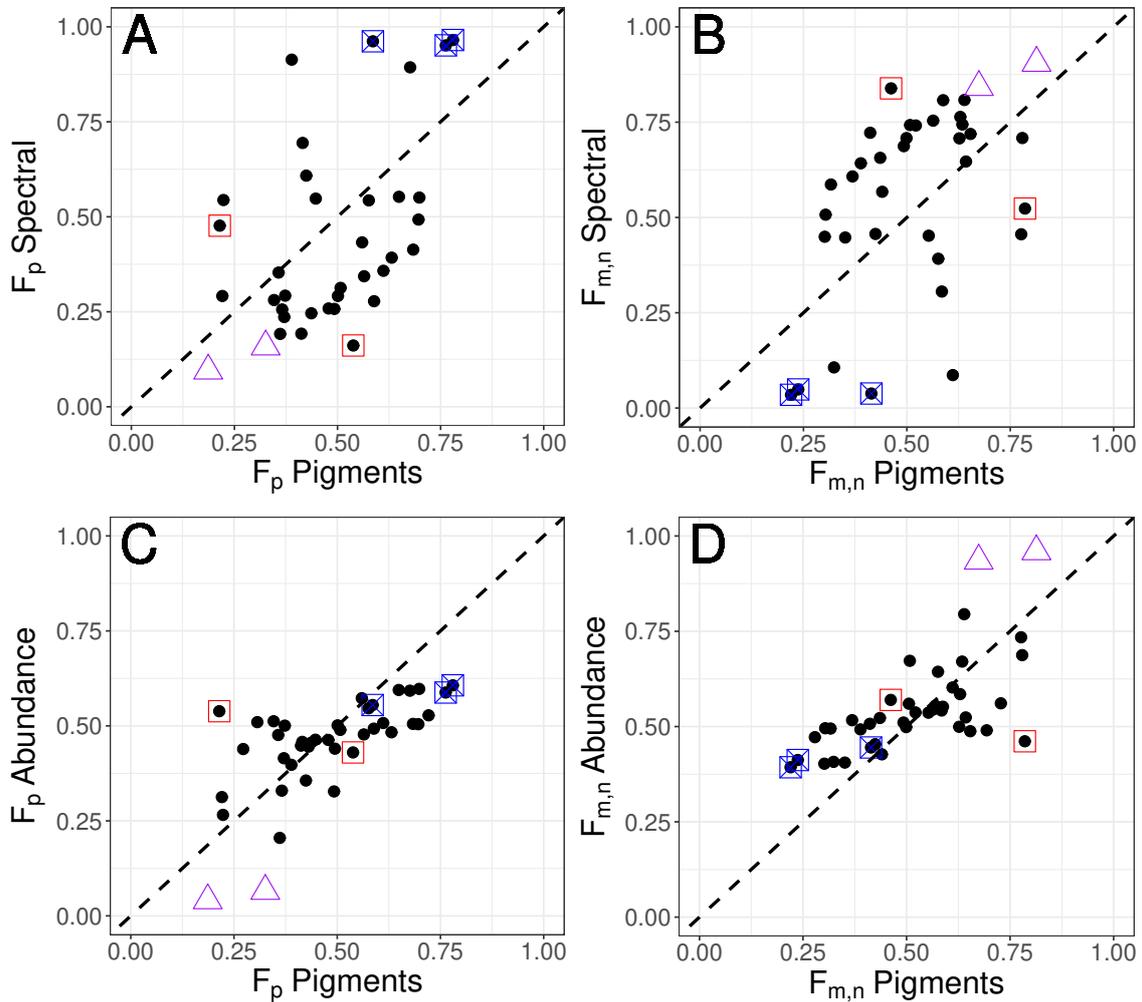
Bloom stations are represented as the purple triangles, the red squares are stations identified as outliers in the spectral-based model chlorophyll estimations for picophytoplankton, the blue squares with a 'x' are the stations in the threshold of spectral-based model. The diagonal dashed black line is the 1:1 reference.

SOURCE: Author's production.

For the correlation analysis, Chl*a* was negatively correlated with temperature. This was also observed by Brewin et al. (2017) when the authors reported a ρ of -0.67, which indicates that higher fractions of small cells are related to high temperatures. Here this relation was not so pronounced (ρ=-0.35) but was significant (p-value<0.05)(Table 5.3), indicating that temperature could be used to tune the mod-

els as performed by Brewin et al. (2017) and Brewin et al. (2019). For residual analysis, the micro and nanophytoplankton fraction was positively correlated to silicate concentrations for the abundance approach, which could be explained by the diatoms' dependence on silicate. However, the n is too small for any robust conclusion (n=10), and this pattern was not observed in the spectral approach. For the other parameters, correlation were not significant .

Table 5.3 - Residuals analysis and correlation analysis of Chl$a$ with environmental variables.(*) is used to highlight p-value <0.05.

Variable	$F_{m,n} spectral$			$F_{m,n} abundance$			$Chla$		
	ρ	p-value	n	ρ	p-value	n	ρ	p-value	n
Temperature	-0.254	0.146	34	0.007	0.965	39	**-0.35**	**0.029***	**39**
Salinity	0.317	0.088	30	0.228	0.195	34	0.17	0.336	34
$Chla - a$	0.178	0.313	34	0.024	0.884	39	-	-	-
$a_{cdom}(440nm)$	0.016	0.930	34	0.079	0.631	39	0.057	0.73	39
Silicate	0.525	0.119	10	**0.638**	**0.047***	**10**	0.238	0.508	10
Phosphate	0.558	0.093	10	-0.222	0.538	10	0.119	0.744	10

5.3.2 Satellite application

5.3.2.1 Validation exercise and sensitivity analysis

The application of PSC models to satellite data is highly dependent on the reliability of the retrievals of the input variables for the region. The number of match-ups available for our study area was relatively low, 23 for Chl$a$ and 21 for phytoplankton absorption coefficient, but it was still possible to perform a validation exercise, to draw some conclusions on the reliability of the results for the area. The Pearson's correlation coefficient for all the variables were higher than 0.5 with p-value <0.001 (Table 5.4).

Chl$a$ was clearly overestimated by OCx (not shown) and GIOP-DC satellite retrieval (Figure 5.5), the determination coefficient was 0.61 and the bias was 2.28, which indicates an overestimation of 128% . This could have higher impact in the abundance-approach estimations of PSCs. The overestimation of Chl$a$ by NASA empirical algorithms in the region have been reported by Giannini et al. (2013) using in-water remote sensing reflectance in the southern part of the Brazilian continental shelf (<25°S). For the OC3M algorithm, the authors reported a mean absolute

percentage difference of 58% and mean relative percentage difference of 35%, with a determination coefficient of 0.78, with 137 samples. However, the overestimation reported here was much higher.

The phytoplankton absorption coefficients presented better statistics for the wavelengths related to chlorophyll-a absorption peak at 412 nm (Table 5.4), with relatively low bias (underestimation of 3%). The station with the highest value, indicated with a red square in the Figure 5.6, was sampled on October 16th 2008, the *in situ* data showed a low sea surface temperature of 22.1°C and salinity of 34.82 psu, the *in situ* Chl$a$ was also above the average with 2.33 mg · m [-3] (Figure 5.6). The SST images for this date seem to present upwelling cells close to the sampling area, and began 5 days previous to the sampling date (data not shown).

Table 5.4 - Statistical metrics for satellite-derived values and *in situ* measured for Chl$a$ and $a_{ph}(\lambda)$. Satellite-derived Chl$a$ and $a_{ph}(\lambda)$ from MODIS GIOP-DC. Number of match-ups (n), Pearson's correlation coefficient (ρ), mean absolute error (MAE), root-mean square error (RMSE), median percentage difference (MPD), and bias. (*) indicates p-value<0.001 for the correlation.

Variable	n	ρ	MAE	RMSE	MPD	Bias
$Chla_{corr}$	23	0.78*	2.277	2.492	135.2%	2.277
$a_{ph}(412)$	21	0.70*	1.294	1.411	18%	0.968
$a_{ph}(443)$	21	0.71*	1.31	1.442	24.5%	1.186
$a_{ph}(469)$	21	0.69*	1.406	1.56	33.8%	1.344
$a_{ph}(488)$	21	0.67*	1.492	1.656	45.9%	1.445
$a_{ph}(531)$	21	0.65*	1.411	1.606	22.6%	1.226
$a_{ph}(547)$	21	0.62*	1.436	1.612	23.5%	1.078
$a_{ph}(555)$	21	0.60*	1.462	1.631	28.3%	1.04
$a_{ph}(645)$	21	0.65*	1.553	1.835	31%	1.526
$a_{ph}(667)$	21	0.70*	1.644	1.919	49%	1.632
$a_{ph}(678)$	21	0.70*	1.572	1.866	41.9%	1.543

Figure 5.5 - Scatterplot of Chl$a$ for MODIS GIOP estimations and *in situ* measurements.
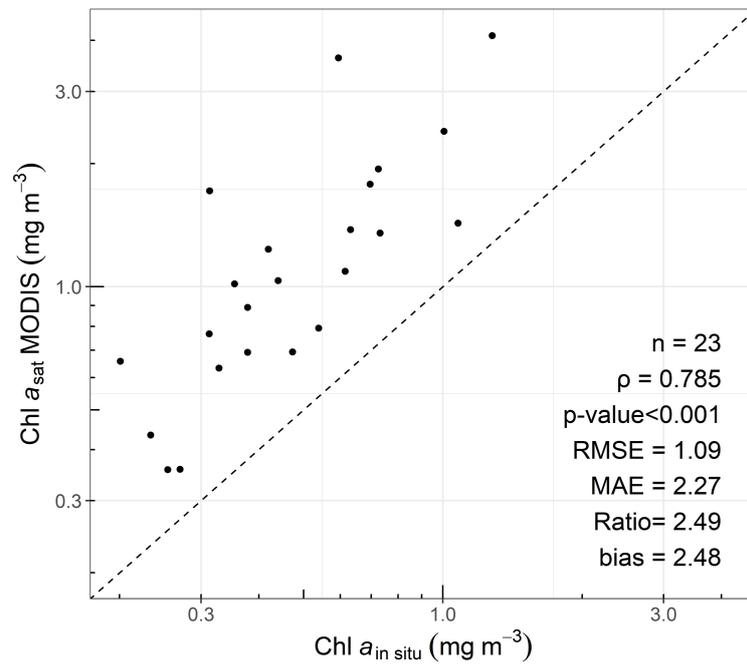
The *in situ* chlorophyll-a concentrations were obtained from HPLC and fluorometry. The fluorometry results were corrected to match HPLC TChl$a$ (see Appendix). The dashed black line is the 1:1 line.

SOURCE: Author's production.

Figure 5.6 - Scatterplots of the phytoplankton absorption coefficient with MODIS GIOP-DC estimations.
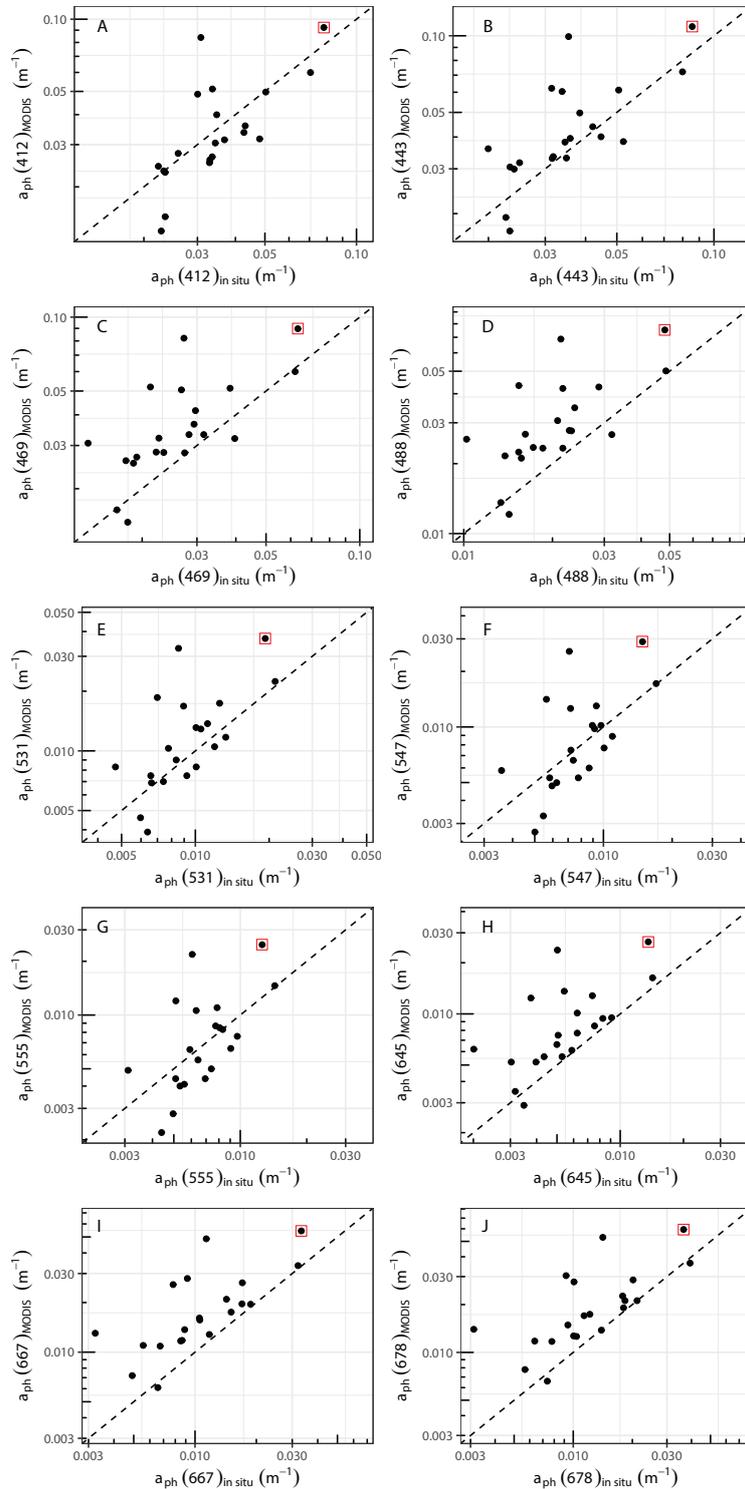
Dashed black line is the 1:1 line. The red square indicates the sample on October 14th of 2008, this seems to be related to a high biomass condition and overestimated for satellite estimations, mainly in longer wavelengths.

SOURCE: Author's production.

Considering the correlation between the Chl$a$ and the phytoplankton absorption, as described in Bricaud et al. (1998), the measured *in situ* presented reasonable results with a coefficient of determination of 0.69 (Figure 5.7), lower than the reported for Case-1 waters ($R^2>0.9$ see Bricaud et al. (1998)). For the satellite retrievals it is possible to observe that by using the GIOP retrieval this condition is maintained (Figure 5.7B), resulting in a high coefficient of determination ($R^2=0.97$). This is not observed when comparing the the OCx estimations for chlorophyll-a (not shown). The GIOP-DC uses the OCx and Bricaud et al. (1998) as an initial guess for the phytoplankton absorption coefficients $a_{ph}(\lambda)$, but after getting the final $a_{ph}(\lambda)$ estimates the new Chl$a$.

Figure 5.7 - Regression between phytoplankton absorption coefficient at 443nm and Chl$a$: (A) *in situ* measurements, (B) MODIS-Aqua estimations (GIOP-DC for Chl$a$ and phytoplankton absorption coefficient).
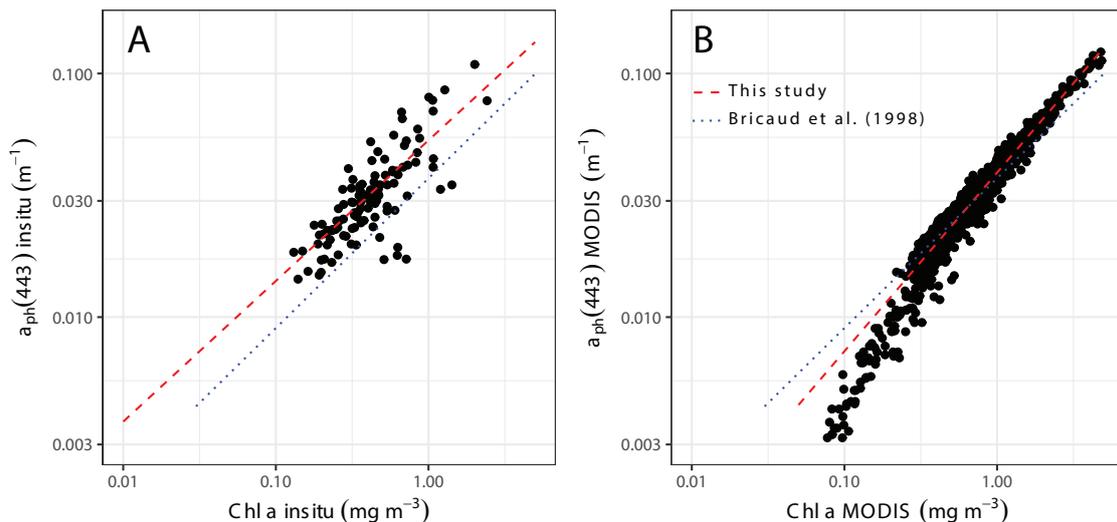
Dashed red line is the predicted following the function of $a_{ph}(\lambda) = A \cdot C^B$ with estimated parameters by fitting the equation, as described in Bricaud et al. (1998). The dotted blue line is the Bricaud et al. (1998) curve, with the parameters at 443 nm (A=0.0371 and B=0.6145).(A) the regression to *in situ* data at 443 nm had as parameters A=0.0532 and B=0.5774 ($n = 116$, $\rho = 0.83$, p-value<0.001) and (B) the regression to satellite retrievals to the time-series returned the parameters A= 0.0394 and B=0.7274 ($n = 1112$, $\rho = 0.98$, p-value<0.001).

SOURCE: Author's production.

Considering the reliability of the IOPs derived from the semi-analytic model (GIOP-

DC), except for a_{dg}, the a_{ph} and the b_{bp} seems to be closely related to Chl$a$. The correlation coefficients at 443 nm wavelength were -0.13, 0.60, and 0.98, for a_{dg}, a_{ph} and b_{bp} (in log_{10} space), respectively all with p-value<0.001, considering the median values and applying the correlation of Pearson to all the data available in the time-series, the Figures 5.8A and B illustrates this for part of the data. The *in situ* Pearson correlation coefficient for a_{cdom} was also negative -0.05, but not significant (p-value>0.01). The a_{ph} followed the function described in Bricaud et al. (1998) (Figure 5.7B), as described in the previous paragraph. For a_{dg} it is possible to spot retrievals which presented high Chl$a$ and low a_{dg} (Figure 5.8C), and also the opposite (i.e., low Chl$a$ and high a_{dg}) which may be related to the conditions observed in the region, where CDOM is likely to vary independently from Chl$a$ (CARVALHO et al., 2014), due to inputs from nearby coastal runoffs or from other sources, such as the La Plata River plume (PIOLA et al., 2008) during the austral winter. High CDOM concentrations could be related to the overestimated satellite retrievals of Chl$a$ and with the poor results in the validation exercise (Table 5.4).

Figure 5.8 - Reflectance and IOPs spectrum from MODIS estimated using GIOP-DC.
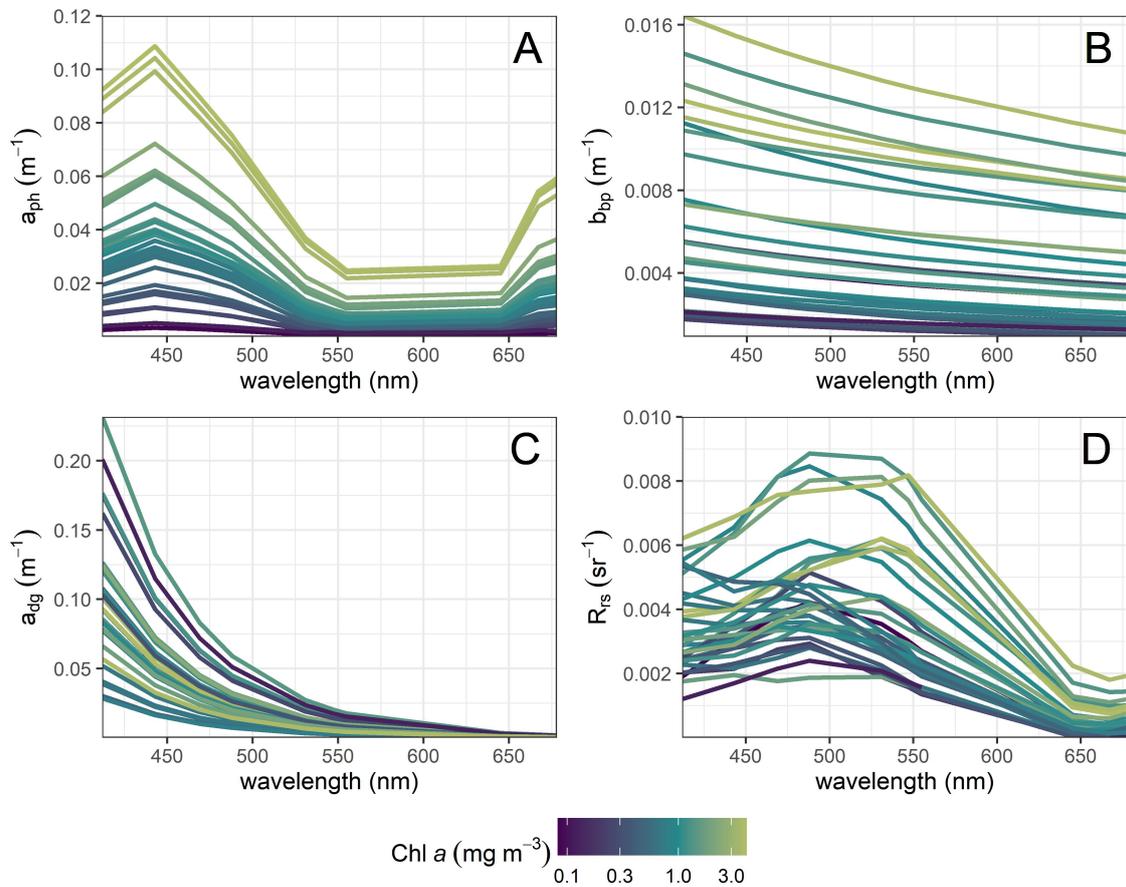
(A) phytoplankton absorption coefficient, (B) particulate backscattering coefficient, (C) non-algal particulate and coloured dissolved organic matter absorption coefficient, and (D) remote sensing reflectance.

SOURCE: Author's production.

The satellite-retrievals from spectral-based approach Chl*a* and the *in situ* Chl*a* were compared to check if the values retrieved were overestimated for satellite data, as the OCx and GIOP-DC. Larger difference could affect the comparison between the PSCs fractions. The Chl*a* estimated by the spectral-based approach had poor performance (R^2=0.39, ρ=0.62, RMSE=0.614 mg ·m^{-3}, MPD=42%), and also overestimated the Chl*a* compared to the *in situ* values. However, the overestimation was not so pronounced as the GIOP-DC, which presented a MPD of 135%.

The match-ups considering the *in situ* phytoplankton pigments analysis presented a low number (n=7), but still, some conclusions can be drawn from these data.

The Pearson's correlation coefficients were not significant (p-value>0.05) and did not present strong correlation (ρ<0.5) for both models and size classes (Table 5.5). The errors and the bias indicated that for the spectral based approach the picophytoplankton presented a underestimation of around 20%, and an overestimation of 170% for the nano and microphytoplankton. This was also observed for the abundance based, presenting an overestiamtion of all size classes, of 38% for pico and 250% for nano and microphytoplankton, which is related to the overestimation observed in the GIOP-DC estimation of Chl$a$.

Taking into account this overestimation of the Chl$a$ observed for the GIOP estimations. We tested the effect on the PSC estimations of an empirical correction, presented in the Equation 5.9. This correction improved the statistic of the size classes, presenting a reasonable reduction of the bias, with underestimations of 2% and 1.5%, respectively for picophytoplankton and nano and microphytoplankton (Table 5.5).

$$Chl_{OCx\ <corrected>} = 10^{(-0.3458\ +\ 0.6069\ \cdot\ log_{10}(Chl_{OCx}))} \tag{5.9}$$

Figure 5.9 - Validation exercise for the chlorophyll-*a* fractions for the PSCs models, including the error bars resulted from the sensitivity analysis. The *in situ* values were derived from DPA analysis.
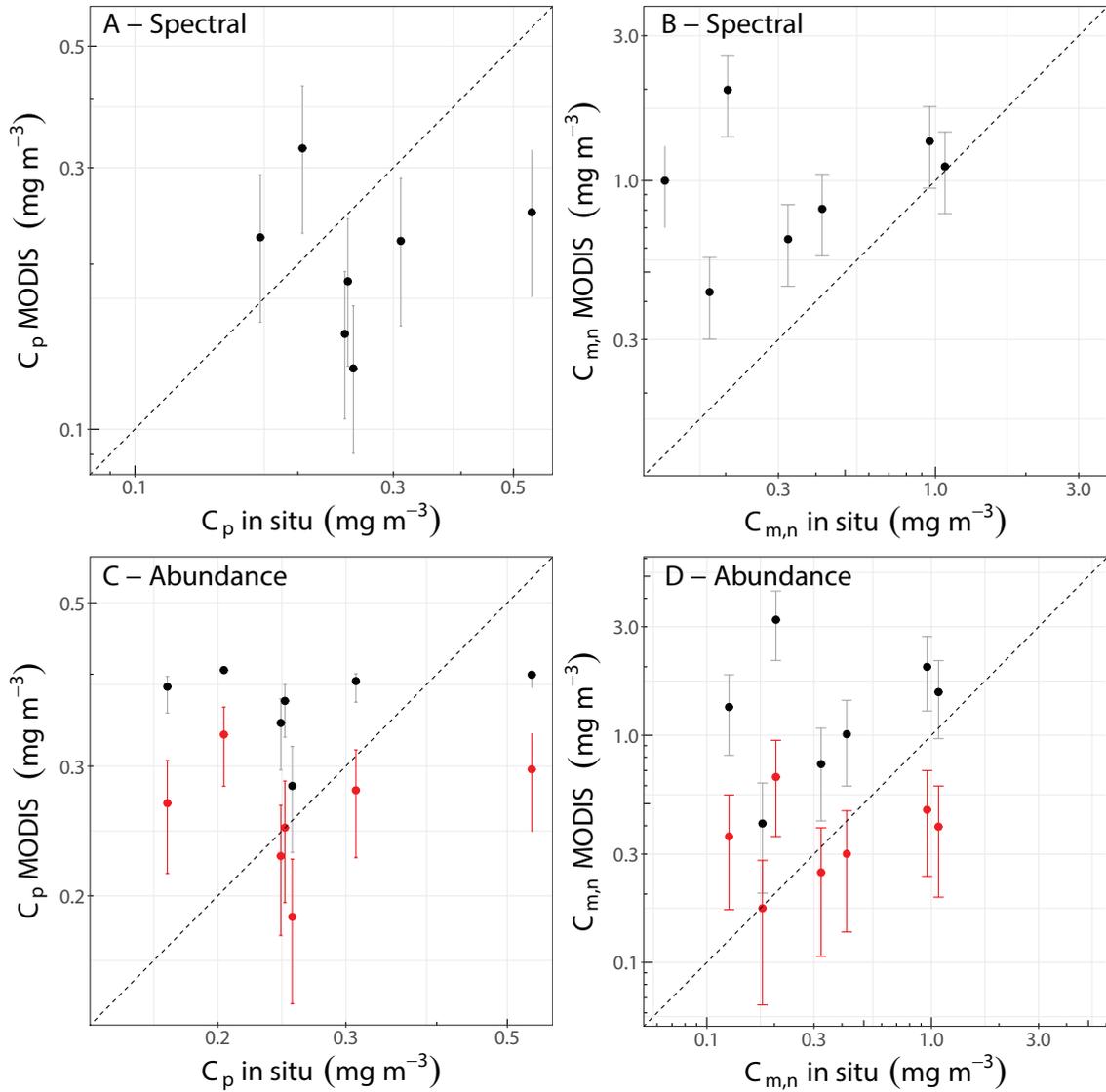
The error bars were estimated with the $\pm\,30\%$ of the Chl*a* and phytoplankton absorption coefficient used as input. A and B represent the estimations for the spectral-based approach $(a_{ph}(\lambda))$ and C and D represent the estimation for the abundance-based approach (Chl*a*) the black dots are the estimated values for the Chl*a* estimated by GIOP and the red dots are the values corrected for the overestimation observed in the GIOP.

SOURCE: Author's production.

Table 5.5 - Statistics of the validation exercise for the PSC chlorophyll-a, considering MODIS estimated PSCs chlorophyll-a and PSC estimated by DPA. The correlation coefficient and the p-value were calculated in the $log_{10} - space$. The statistics for the corrected Chl$a$ for the abundance-based approach were also included.

Variable	n	ρ	RMSE	MAE	Bias
C_p $spectral$	7	0.06	1.657	1.603	0.771
$C_{m,n}$ $spectral$	7	0.248	3.566	2.703	2.703
C_p $abundance$	7	0.146	1.605	1.51	1.385
$C_{m,n}$ $abundance$	7	0.268	4.549	3.506	3.506
C_p $abundance$ (corrected)	7	0.144	1.432	1.334	0.979
$C_{m,n}$ $abundance$ (corrected)	7	0.255	2.161	1.914	0.985

5.3.2.2 Satellite time series analysis

Analysing these results from a time-series perspective (using monthly averages), it is possible to observe that the spectral-based approach presented higher micro and nanophytoplankton fractions, ranging from 70% to 90%. The abundance approach presented a broader range from 32% to 86%, varying in a similar pattern with Chl$a$ (Figure 5.10A and B). The estimations for the abundance approach with the corrected Chl$a$ (Figure 5.10A, solid dark purple line) presented a more restrictive range from 34% to 63% but presented lower fractions of the micro and nanophytoplankton. It seems that the micro and nanophytoplankton size fractions were out-of-phase with SST. Cross-correlation analysis between the SST and PSCs fractions showed a negative correlation with lag=0, and a positive correlation within lag of 5-8 months. A Pearsons's correlation test presented negative coefficients for the spectral and abundance based models with significance results (p-value<0.01), ρ = -0.57 and -0.25, respectively for spectral and abundance-based approaches. It is worth mentioning that both models performed poorly for the satellite-retrieval of the fractions of the PSC (data not shown), but we opted to keep the fractions in the time series and the monthly average images because they are easier to interpret in a single image than the Chl$a$ for each size class, and the variability would be similar to the Chl$a$ of the PSCs.

Chl$a$ presented high peaks during the winter months, for some years, e.g., 2015 to 2020, but in some years this was not observed, such as for 2012 to 2014, in which it seems to present higher Chl$a$ throughout the year (Figure 5.10B). The correlation test and the lag cross-correlation presented negative correlation between

Chl*a* and SST (ρ=-0.57, p-value<0.01). A seasonality pattern was not clear for the phytoplankton absorption coefficient at 443. There were peaks throughout the years but they were not consistent with the seasons. Some abrupt changes were noticed at the end of 2013, middle of 2005, 2009, 2016 and 2017 which could be related to algal bloom events (Figure 5.10C). The correlation test showed weak and not-significant correlation at 443nm, but negative and significant at 645nm (ρ=-0.17, p-value<0.05). The high peaks at the phytoplankton absorption coefficient could also be related to the overestimation driven by the increase in CDOM absorption.

CDOM and detritus absorption coefficient presented a clear seasonal pattern, with peaks usually during the winter-spring seasons (July to October) (Figure 5.10C), which could be related to cold water coming from the southern region (PIOLA et al., 2008), supported by the sea surface temperature time series. The cross-correlation and correlation test presented significant negative correlation for SST and $a_{dg}(443)$ (ρ=-0.55, p-value<0.01), the cross-correlation presented positive correlation with a lag of 6 months. This was also observed for the particulate backscattering coefficient at 443nm (ρ=-0.52, p-value<0.01), also presenting positive correlation with a lag of 6 months (Figure 5.10E). In this study, due to the lack of *in situ* backscattering coefficient, it was not possible to validate the GIOP-DC results.

Figure 5.10 - Time-series of monthly averages of satellite estimations micro and nanophytoplankton fractions, chlorophyll-a concentration, the inherent optical properties derived from GIOP-DC, and sea surface temperature, plotted with all the other variables in dashed purple line.
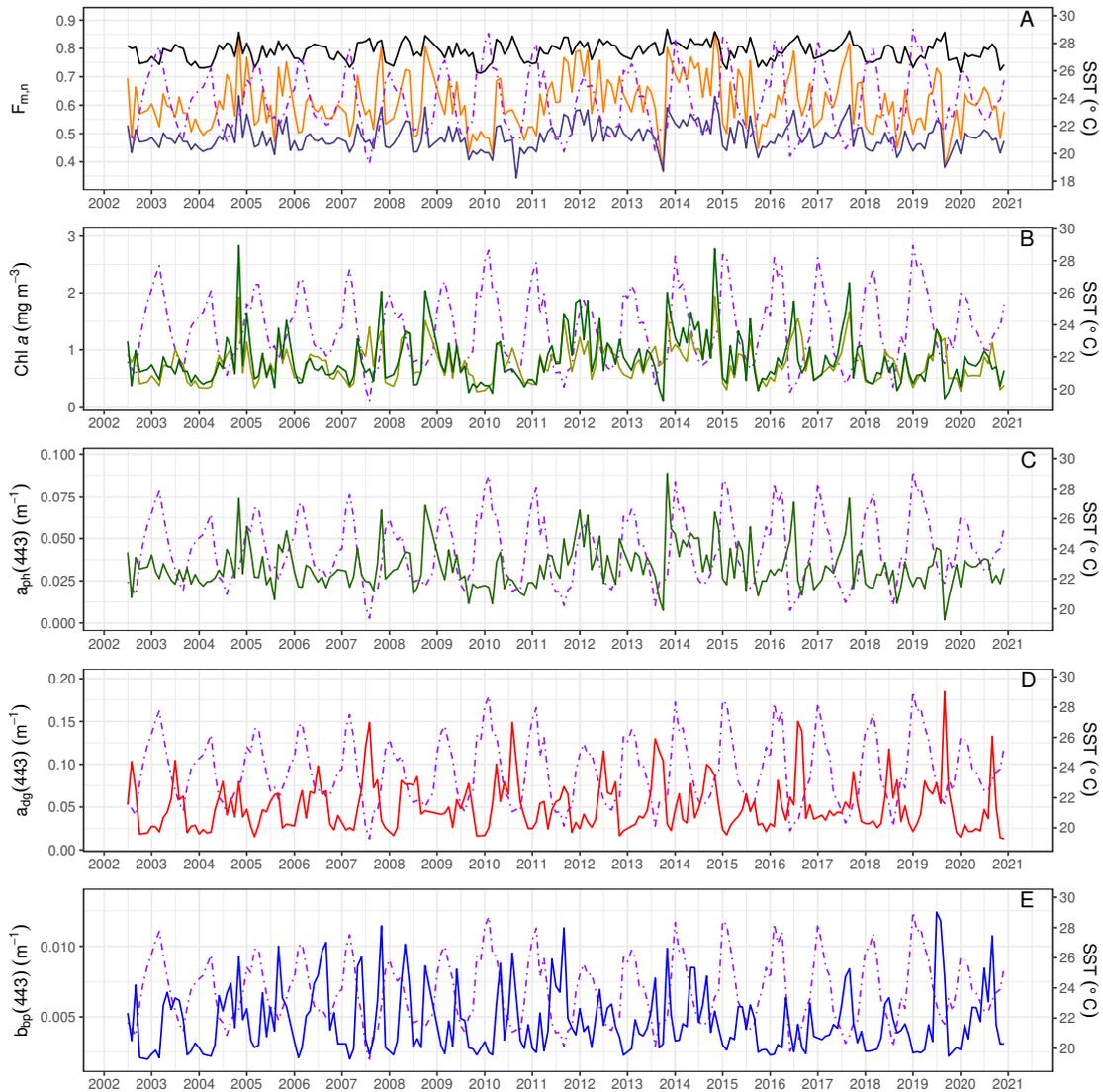
Sea surface temperature (SST) is plotted as the secondary y-axis in all the plots in dashed purple line. (A) Phytoplankton fractions of micro and nano-size, the solid orange line for the abundance-based approach (BREWIN et al., 2010), the solid dark purple line for the abundance-based approach with the corrected Chl$a$ as input, and the solid black line for spectral-based approach (DEVRED et al., 2011), (B) Chlorophyll-a OCx in light-green and from GIOP-DC in dark green (no-correction applied), (C) Phytoplankton absorption coefficient at 443nm, (D) Non-algal particulate and coloured dissolved organic matter absorption coefficient, and (E) Particulate backscattering coefficient.

SOURCE: Author's production.

Monthly average maps for the micro and nanophytoplankton fractions for the South Brazil Bight show that areas closer to the coast present higher fractions of micro and nanophytoplankton. For the Antares-Ubatuba station this fraction was usually higher than 50% (for the abundance approach) and higher than 70% for the spectral approach. From June to August it is possible to notice an expansion of micro and nanophytoplankton fractions further on the continental shelf, with a gradual retraction of its domain to a minimum area during January (Figure 5.11). A similar pattern is observed in the spectral approach, but for this approach, the domain of micro and nanophytoplankton seems to extend further to the offshore, which could indicate a problem in the fitting of the model to these areas, especially in the autumn and winter months (May to September) (Figure 5.12).

For sea surface temperature the seasonal variability is noticeable with high temperatures during the austral summer (December to March), and low temperatures during the austral winter (June to September). Some upwelling events could be masked by the monthly time averaging. High temperatures dominate the area during the austral summer months (December to March), when it is possible to spot the cold water near the coast of Cape Frio, in the north area of the map, indicating the upwelling domain (5.13 images from January and February). In the austral autumn and winter months (May to September) the cold water starts expanding the domain in the continental shelf coming from the South.

Figure 5.11 - Micro and nanophytoplankton fraction from the abundance approach (BREWIN et al., 2010) (non-correction applied to Chl*a* from GIOP-DC), monthly averages for 2002-2020.
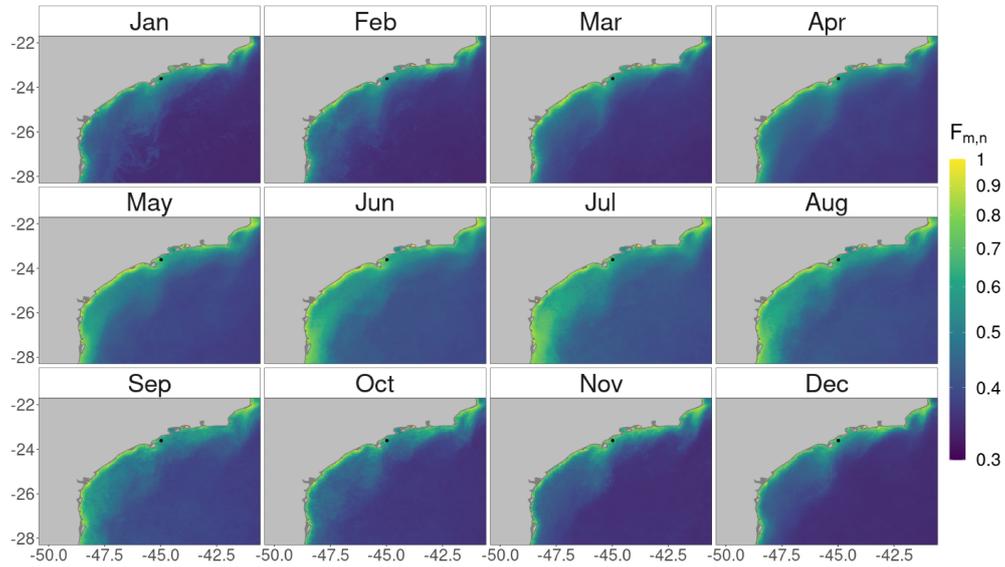
SOURCE: Author's production.

Figure 5.12 - Micro and nanophytoplankton fraction, monthly averages the spectral approach (DEVRED et al., 2011) for 2002-2020.
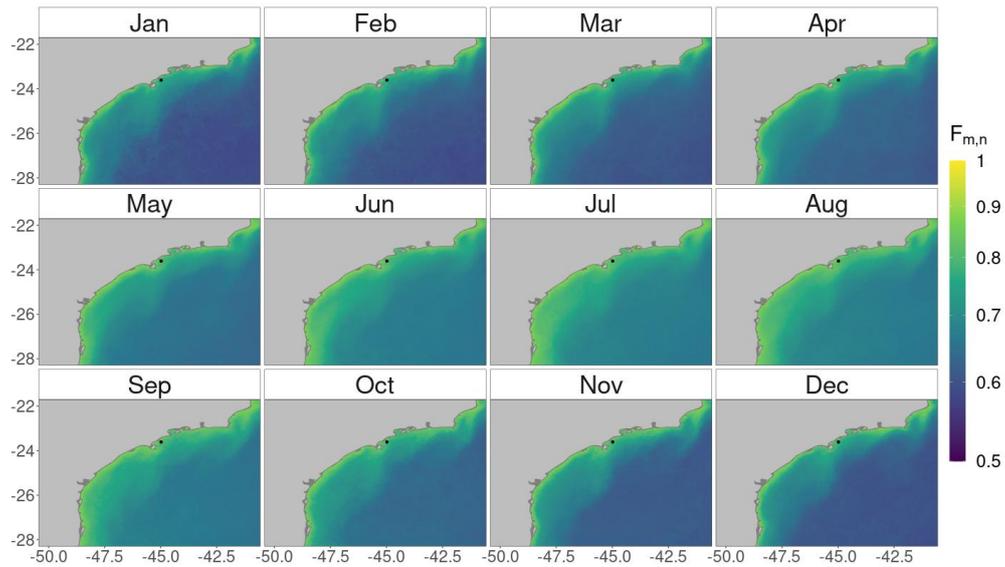
SOURCE: Author's production.

Figure 5.13 - Sea surface temperature averages for 2002-2020.
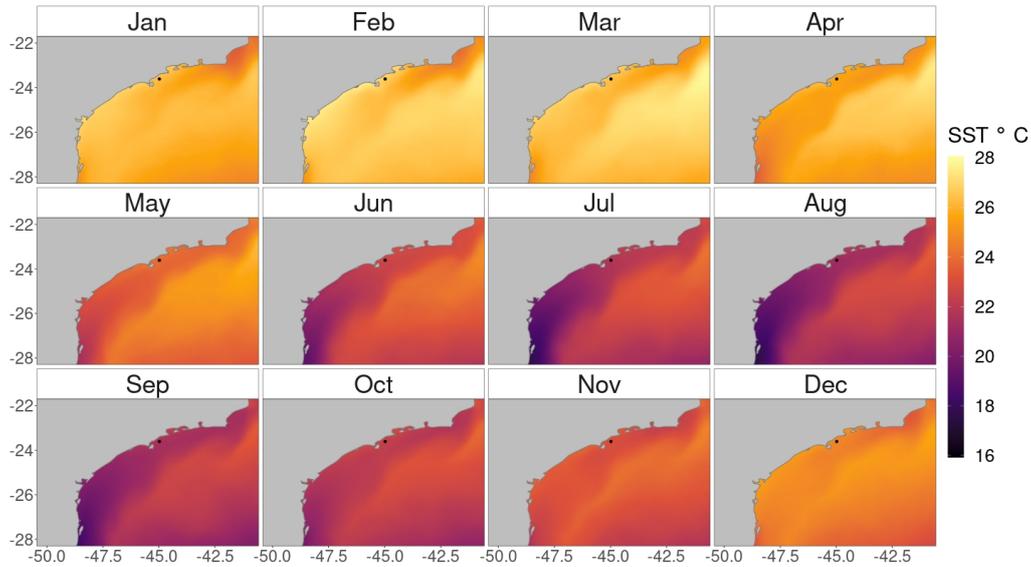
SOURCE: Author's production.

5.4 Discussion

5.4.1 PSC Models performances and satellite application

Few studies focused on the evaluation of PSC models in coastal and productive shelf waters. Liu et al. (2018) reported that the picophytoplankton fraction presented a poor performance for all the models tested in the northwest Atlantic. The authors suggested that it could be related to the under-representation of pico-dominated samples in productive waters. However, this is not necessarily the case in the Antares-Ubatuba region, which does have an important fraction of picophytoplankton, as discussed in the previous Chapter 4, but a low variability in the fractions sampled in the time series could be a drawback for the model fitting.

Regarding the spectral-based approach, the selection of an appropriate semi-analytical model used to retrieve the phytoplankton absorption, as discussed by Brewin et al. (2010) could have an impact on the results of the PSC model. The GIOP-DC, selected in this study, uses the Bricaud et al. (1998) model to estimate an initial phytoplankton absorption from Chl*a*. According to the discussion presented in Brewin et al. (2010), estimations of the phytoplankton absorption which make assumptions on the shape of the phytoplankton absorption spectrum would not be

119

appropriate. However, using the Quasi-Analytic Algorithm (QAA) model, as suggested by the authors, which allows more flexible assumptions the phytoplankton absorption spectrum (LEE et al., 2002), resulting in a failure of the estimations in the wavelengths around 665 nm. Here, we opted to use the GIOP-DC.

Liu et al. (2018) found that the Devred et al. (2011) model applied as an abundance-based approach performed better when compared to other models considering the accuracy of all the size classes retrieved. However, it could be related to low accuracy in the $a_{ph}(\lambda)$ satellite retrievals. The authors also discussed the uncertainties of satellite retrieved $a_{ph}(\lambda)$ to the model's estimations. More specifically, they found an underestimation of $a_{ph}(\lambda)$ in the three wavelengths analysed, i.e., 443, 490 and 555 nm. These models would benefit from a regionally tuned semi-analytical model to retrieve IOPs and also a locally tuned Chl$a$ model to reduce the effect of overestimation of chlorophyll common in optically complex waters (DEFOIN-PLATEL; CHAMI, 2007). The use of regional tuned models to satellite-derived Chl$a$ was successfully adopted in other regions, improving their results (LIU et al., 2021; GITTINGS et al., 2019; SUN et al., 2019).

Our results presented a better performance for the abundance-based approach for the *in situ* "validation". However, this result must be interpreted with caution, as previously warned in the results section, because the validation of the abundance approach was performed with the same data set of the fitting. Thus, this apparently better performance could be related to the non-independence of the dataset used in the fitting and in the "validation". This was not the case for the spectral-based approach, validated with a different subset of the Antares-Ubatuba station dataset. The application of a correction to the input Chl$a$ in the satellite validation exercise reduced the uncertainties in the PSC estimations. This is a good indication that the use of a local algorithm for Chl$a$ could improve the PSC estimations.

The drawback of using the phytoplankton absorption coefficient is that it is also carrying uncertainties into the model since its estimation depends on the performance of semi-analytical models to retrieve bio-optical parameters from satellite reflectance, which is still a challenge, especially for coastal waters. However, the estimation relying only on Chl$a$ may not be able to capture more subtle changes in the phytoplankton communities and it is more sensitive to satellite-retrieved Chl$a$ overestimation, as indicated by the sensitivity analysis performed in this study. The abundance approach is more sensitive to Chl$a$ uncertainties than the spectral-based approach to uncertainties in phytoplankton absorption coefficient, and also

this variability is more consistent through different size classes. In contrast, the abundance-based approach presented a higher difference in the estimation for the micro and nanophytoplankton fraction. However, the sensitivity analysis adopted here for the spectral-based approach used a fixed percentage for all wavelengths, arguably a more appropriate approach for the analysis of spectral data is applying a wavelength-dependent uncertainty.

Abundance-based models are limited by the assumption that the relation between abundance (Chl$a$) and the phytoplankton size classes is "constant" or stable, and does not account for changes in this relation. On the other hand, the spectral models do not rely on this dependence and would be able to reflect changes in the phytoplankton communities in a climate change scenario. Despite the performance slightly worse than the abundance-based approach for *in situ* validation in this study, the spectral-based approach could be potentially better using hyperspectral data to derive $a_{ph}(\lambda)$ and the PSCs as shown by Zhang et al. (2021), when evaluating the use of hyperspectral data to retrieval phytoplankton pigments from $a_{ph}(\lambda)$. The spectral-based approach is still more time-consuming and computationally costing than the abundance-based approach, which is not a problem for modern computers. Applying a sensitivity analysis to find the ideal number of wavelengths to its application could reduce the computational cost and improve quality of the results, especially for its application in hyperspectral sensors, such as the PRecursore IperSpettrale della Missione Applicativa (PRISMA) and Ocean color Instrument (OCI) from Plankton, Aerosol, cloud and Ecosystem (PACE) mission.

5.4.2 Spatial-temporal seasonality

The pattern observed for the monthly average fraction of micro and nanophytoplankton using the abundance-based approach, presented a clear increase in the proportion of this size class near the coast, during the austral autumn and winter months (Figure 5.11). This pattern is likely enhanced in some years, which could be captured by the multi-annual monthly average image but not so clearly captured in the time series. However, this result must be interpreted with caution, as this model was re-tuned considering only one coastal station, the model may not respond well out of the range of the area where it was adjusted to. The same caution is valid for the spectral-based approach.

Some interesting patterns were observed considering the time series. Micro and nanophytoplankton fraction seems to be out-of-phase with SST (Figure 5.10A). Even though it is not so clear as the $a_{dg}(\lambda)$, it is possible to see an overall pattern.

It could indicate that SST could be a good variable to improve PSCs estimations in the region, which was also indicated by the residuals analysis (Table 5.3). However, other variables should also be investigated, such as mixed layer depth and wind stress, especially due to the complex dynamics of the region, with the occurrence of upwellings, meanders and eddies (CASTRO, 2014).

Considering the large difference in the fractions estimated by the abundance and the spectral-based approaches, the abundance-based approach presented fractions that are more close to the fractions estimated in the Chapter 4, with higher proportions of nano and microphytoplankton but in some conditions the picophytoplankton domain. The spectral-based approach, presented proportions much higher for the micro and nanophytoplankton than that indicated by the Chapter 4 (Figure 4.12, the picophytoplankton represented more than 33% in most many of the *in situ* samples). Thus, the variability may be similar, but the values of the proportions must be further studied.

The particle backscattering and the CDOM and detritus absorption coefficients are also out-of-phase with the sea surface temperature (Figure 5.10 D and E). This is particularly interesting considering the possible sources of particulate matter and CDOM in the region. Gonçalves-Araujo et al. (2019) studied the absorption coefficients of CDOM and particulate matter in different waters masses in the SBB, from near the coast to offshore areas, and showed a difference in the optically active components in the Coastal Water (CW) and the Tropical Water (TW) (as discussed in the Chapter 4). The CW was dominated by particulate matter and CDOM, but the authors observed a good correlation between CDOM and Chl*a*, indicating an autochthonous production of CDOM. However, their results were based on the data from a unique cruise during the austral summer, from 25th to 30th of January, 2013. It means they were not able to observe the high-CDOM and low-temperature conditions, present during the austral winter in the region (GONÇALVES-ARAUJO et al., 2019). This must be further investigated to understand the dynamics of those components in the area, with *in situ* validation for those components. Indeed, there is a project underway in the SBB region ("Chemical and Biological Characterization of the Pelagic System of the Santos Basin"), whose data may be used in complementary analyses, in due course.

5.5 Conclusion and final remarks

The evaluation of PSC models for this subtropical area showed that despite the less computational costing of applying an abundance-based model and the relatively

better statistics of the retrievals than the spectral-based model, its uncertainties are more variable, and it is constrained by the conditions in which the fitting data set was sampled. The *in situ* validation presented better performance than the satellite retrieval validation exercise, indicating that uncertainties in the satellite-derived input variables may propagate to higher levels of uncertainties for the derived phytoplankton size fractions and associated chlorophyll concentrations. Moreover, the satellite-derived Chl*a* used for the abundance-based PSC models, is more subject to uncertainties in coastal waters, with external sources of CDOM and non-algal particulate matter. Thus, if the goal is to evaluate changes in the phytoplankton assemblage, it could be recommended to invest in the spectral-based approach, which with the perspective of availability of hyperspectral data from new ocean color remote sensing missions (e.g., PRISMA and PACE) would allow selecting the most suitable wavelengths to estimate the PSCs, reducing the computational-cost of this approach.

The PSC fractions were different for both methods indicating that differences in the applied model will certainly change the estimations (and the conclusions) of the studies. The seasonality of the PSCs in the SBB showed an overall pattern to be out of phase with the SST seasonality (i.e., when the SST is low the micro and nano fraction is high), similar to Chl*a*, which is also observed in the structure of the micro and nanophytoplankton dominance in the continental shelf extension. This finding highlight the suitability of using SST for future improvements in the regional tuning of the PSC model. On the other hand, the use of one coastal station to characterize the seasonality of the entire SBB seems to be inadequate for the extension of the area, thus further studies to validate and improve PSC models in the SBB are required.

6 FINAL REMARKS

> "Only with good *sea truth* data can we have confidence in the climate products generated using the algorithms."
>
> ———————————————
> Sathyendranath et al. (2017)

There are many aspects that could be enhanced to retrieve phytoplankton size classes from satellite imagery for optically-complex waters. In the remote sensing realm, we can list: (i) atmosphere correction, (ii) semi-analytical and empirical ocean colour models, (iii) phytoplankton size class models, (iv) sensor resolution (spatial, temporal and spectral), and (v) sensors able to measure different physical properties of the water/atmosphere interface. In the *in situ* realm, considering the data used to validate and adjust the PSC models, the most important aspects are: (i) technique(s) used to quantify and identify the PSC, (ii) seasonality of the PSCs in the area, (iii) physical aspects that are potentially driving the PSC distribution which could be used to improve the PSC models, and (iv) optical aspects of the water being studied.

Most of these aspects were not the focus of this study, instead, we were able to focus on only two PSC models. But our findings highlighted the importance of using different techniques for the *in situ* quantification of PSCs, and the aspects of each PSC model considered. Coastal waters have unique pigment compositions which characterize the phytoplankton assemblages and thus require regional tuning for PSCs obtained from DPA and RS models. Regarding these last, the abundance-based approach can derive reasonable estimates if the input variable has low uncertainty (Chl$a$), which is not always the case for optically-complex waters. Moreover, once tuned, this approach does not allow changes in the relation between Chl$a$ and PSCs, which can occur in a changing environment. The spectral based approach, on the other hand, can provide PSC estimates in more optically-complex waters, provided with the phytoplankton absorption coefficient, and is more flexible in the Chl$a$ - PSCs relations.However, this approach is also influenced by spectral uncertainties which can lead to higher uncertainties in the retrieved parameters. SST can be an important parameter determining this relation, both globally and locally, at the Ubatuba coastal site, and could be further used to improve the PSC estimates by RS models, which need to be further tested.

The ANTARES-Ubatuba coastal station is part of a complex system with different meteo-oceanographic processes governing the phytoplankton assemblages and size classes i.e., with the influence of coastal upwelling during spring-summer and wind-induced mixing and coastal currents transporting rich waters from continental sources in autumn-winter. This work contributes to the indication of challenges and requirements of PSC models applied to coastal waters and for the understanding of phytoplankton assemblages and PSCs at a coastal site at the South Brazil Bight. To the best of our knowledge, this is the first work to evaluate these types of space-borne PSC models in the coastal waters of the South Brazil Bight, using a relatively long time series.

Future perspectives of PSC investigations should invest in autonomous sampling platforms such as Biogeochemical Argo (BGC-Argo) Floats (REMBAUVILLE et al., 2017) and in algorithms able to fully explore the potential of the sensors with hyperspectral resolution. Ocean and atmosphere dedicated hyperspectral satellite missions, such as PACE are expected to play an important role in the improvement of PSC estimates from space, especially at optically-complex waters, taking advantage of improved atmospheric correction and retrieval of Chl$a$ and IOPs.

Our results are a first step in the investigation of PSC algorithms in the ANTARES-Ubatuba station. As mentioned earlier, many other directions could have been taken in this investigation, such as focusing on the atmospheric correction performance and its influence on the models' performances, or on the selection of the best semi-analytical models for the SBB waters, or on the evaluation of all available PSC and PFT models. However, all these possible approaches would face similar challenges, i.e., the quantity, frequency, types, and quality of *in situ* data available and the scarcity of match-ups for validation. For this reason, as highlighted by Sathyendranath et al. (2017), it is extremely important to maintain and build on the *in situ* datasets and improve the sampling methods for evaluating the performance of the satellite sensors and products. Thus, the effort to keep long-term *in situ* sampling, is essential for climate change products and satellite sensors validation, which will improve the ability to assess the ocean-colour products and built reliable information for stakeholders.

REFERENCES

AIKEN, J.; PRADHAN, Y.; BARLOW, R.; LAVENDER, S.; POULTON, A.; HOLLIGAN, P.; HARDMAN-MOUNTFORD, N. Phytoplankton pigments and functional types in the Atlantic Ocean: A decadal assessment, 1995-2005. **Deep-Sea Research Part II: Topical Studies in Oceanography**, v. 56, n. 15, p. 899–917, 2009. ISSN 09670645. 17

ALVAIN, S.; MOULIN, C.; DANDONNEAU, Y.; BREON, F. Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. **Deep-Sea Research I**, v. 52, p. 1989–2004, 2005. 11

ALVAIN, S.; MOULIN, C.; DANDONNEAU, Y.; LOISEL, H. Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view. **Global Biogeochemical Cycles**, v. 22, n. 3, p. 1–15, 2008. ISSN 08866236. 11

AMINOT, A.; MARCEL, C. **Manuel des analyses chimiques en milieu marin**. Brest: CNEXO, 1983. 376 p. ISBN 2902721102 9782902721108. 57

ARMSTRONG, F. A. J.; STEARNS, C. R.; STRICKLAND, J. D. H. The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment. **Deep Sea Research and Oceanographic Abstracts**, v. 14, n. 3, p. 381–389, 1967. ISSN 0011-7471. Available from: <https://www.sciencedirect.com/science/article/pii/0011747167900824>. 57

BABIN, M.; MOREL, A.; FOURNIER-SICRE, V.; FELL, F.; STRAMSKI, D. Light scattering properties of marine particles in coastal and open ocean waters as related to the particle mass concentration. **Limnology and OceanographyLimnology and Oceanography**, v. 48, n. 2, p. 843–859, 2003. 16, 45

BABIN, M.; ROESLER, C. S.; CULLEN, J. J. **Real-time coastal observing systems for marine ecosystem dynamics and harmful algal blooms : theory, instrumentation and modelling**. [S.l.: s.n.], 2008. 807 p. ISBN 9789231040429ͦ9231040421. 5

BAILEY, S. W.; WERDELL, P. J. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. **Remote Sensing of Environment**, v. 102, n. 1-2, p. 12–23, 2006. ISSN 00344257. 20

BIDIGARE, R. R.; ONDRUSEK, M. E.; MORROW, J. H.; KIEFER, D. A. In-vivo absorption properties of algal pigments. **Ocean Optics X**, v. 1302, p. 290, 1990. ISSN 0277786X. 1

BORGES, A. V.; DELILLE, B.; FRANKIGNOULLE, M. Budgeting sinks and sources of CO2 in the coastal ocean: diversity of ecosystem counts. **Geophysical Research Letters**, v. 32, n. 14, p. 1–4, 2005. ISSN 00948276. 15

BOUMAN, H. A.; PLATT, T.; SATHYENDRANATH, S.; LI, W. K. W.; STUART, V.; FUENTES-YACO, C.; MAASS, H.; HORNE, E. P. W.; ULLOA, O.; LUTZ, V.; KYEWALYANGA, M. Temperature as indicator of optical properties and community structure of marine phytoplankton : implications for remote sensing. **Marine Ecology Progress Series**, v. 258, p. 19–30, 2003. 22

BRACHER, A.; BOUMAN, H. A.; BREWIN, R. J. W.; BRICAUD, A.; BROTAS, V.; CIOTTI, A. M.; CLEMENTSON, L.; DEVRED, E. Obtaining phytoplankton diversity from ocean color : a scientific roadmap for future development. **Frontiers in Marine Science**, v. 4, p. 1–15, 2017. 2, 10, 11, 16

BRANDINI, F. P.; Nogueira Jr, M.; SIMIÃO, M.; CODINA, J. C. U.; NOERNBERG, M. A. Deep chlorophyll maximum and plankton community response to oceanic bottom intrusions on the continental shelf in the South Brazilian Bight. **Continental Shelf Research**, v. 89, p. 61–75, 2014. ISSN 0278-4343. Available from: <http://dx.doi.org/10.1016/j.csr.2013.08.002>. 50, 51, 83

BRANDINI, F. P.; TURA, P. M.; SANTOS, P. P. Ecosystem responses to biogeochemical fronts in the South Brazil Bight. **Progress in Oceanography**, v. 164, p. 52–62, 2018. ISSN 00796611. 2, 50, 51, 54, 79, 80, 82, 87

BRASIL. MINISTÉRIO DO MEIO AMBIENTE (MMA). **Atlas de sensibilidade ambiental ao óleo da Bacia Marítima de Santos**. Brasília: MMA, 2007. 124 p. 3

BREWIN, R. J.; CIAVATTA, S.; SATHYENDRANATH, S.; JACKSON, T.; TILSTONE, G.; CURRAN, K.; AIRS, R. L.; CUMMINGS, D.; BROTAS, V.; ORGANELLI, E.; DALL'OLMO, G.; RAITSOS, D. E. Uncertainty in ocean-color estimates of chlorophyll for phytoplankton groups. **Frontiers in Marine Science**, v. 4, 2017. ISSN 22967745. 22, 46, 47, 104, 105

BREWIN, R. J.; DEVRED, E.; SATHYENDRANATH, S.; LAVENDER, S. J.; HARDMAN-MOUNTFORD, N. J. Model of phytoplankton absorption based on

three size classes. **Applied Optics**, v. 50, n. 22, p. 4535–4549, 2011. ISSN
15394522. 100

BREWIN, R. J.; HARDMAN-MOUNTFORD, N. J.; LAVENDER, S. J.;
RAITSOS, D. E.; HIRATA, T.; UITZ, J.; DEVRED, E.; BRICAUD, A.; CIOTTI,
A.; GENTILI, B. An intercomparison of bio-optical techniques for detecting
dominant phytoplankton size class from satellite remote sensing. **Remote
Sensing of Environment**, v. 115, n. 2, p. 325–339, 2011. ISSN 00344257.
Available from: <http://dx.doi.org/10.1016/j.rse.2010.09.004>. 11

BREWIN, R. J.; LAVENDER, S. J.; HARDMAN-MOUNTFORD, N. J.;
HIRATA, T. A spectral response approach for detecting dominant phytoplankton
size class from satellite remote sensing. **Acta Oceanologica Sinica**, v. 29, n. 2, p.
14–32, 2010. ISSN 0253505X. 119

BREWIN, R. J.; MORÁN, X. A. G.; RAITSOS, D. E.; GITTINGS, J. A.;
CALLEJA, M. L.; VIEGAS, M.; ANSARI, M. I.; AL-OTAIBI, N.;
HUETE-STAUFFER, T. M.; HOTEIT, I. Factors regulating the relationship
between total and size-fractionated chlorophyll-a in coastal waters of the Red Sea.
Frontiers in Microbiology, v. 10, p. 1–16, 2019. ISSN 1664302X. 96, 105

BREWIN, R. J.; RAITSOS, D. E.; DALL'OLMO, G.; ZAROKANELLOS, N.;
JACKSON, T.; RACAULT, M. F.; BOSS, E. S.; SATHYENDRANATH, S.;
JONES, B. H.; HOTEIT, I. Regional ocean-colour chlorophyll algorithms for the
Red Sea. **Remote Sensing of Environment**, v. 165, p. 64–85, 2015. ISSN
00344257. Available from: <http://dx.doi.org/10.1016/j.rse.2015.04.024>.
88

BREWIN, R. J.; SATHYENDRANATH, S.; MÜLLER, D.; BROCKMANN, C.;
DESCHAMPS, P. Y.; DEVRED, E.; DOERFFER, R.; FOMFERRA, N.; FRANZ,
B.; GRANT, M.; GROOM, S.; HORSEMAN, A.; HU, C.; KRASEMANN, H.;
LEE, Z. P.; MARITORENA, S.; MÉLIN, F.; PETERS, M.; PLATT, T.;
REGNER, P.; SMYTH, T.; STEINMETZ, F.; SWINTON, J.; WERDELL, J.;
WHITE, G. N. The Ocean Colour Climate Change Initiative: III. A round-robin
comparison on in-water bio-optical algorithms. **Remote Sensing of
Environment**, Elsevier Inc., v. 162, p. 271–294, 2015. ISSN 00344257. Available
from: <http://dx.doi.org/10.1016/j.rse.2013.09.016>. 13

BREWIN, R. J.; SATHYENDRANATH, S.; PLATT, T.; BOUMAN, H.;
CIAVATTA, S.; DALL'OLMO, G.; DINGLE, J.; GROOM, S.; JÖNSSON, B.;

KOSTADINOV, T. S.; KULK, G.; LAINE, M.; MARTÍNEZ-VICENTE, V.; PSARRA, S.; RAITSOS, D. E.; RICHARDSON, K.; RIO, M.-H.; ROUSSEAUX, C.; SALISBURY, J.; SHUTLER, J. D.; WALKER, P. Sensing the ocean biological carbon pump from space: a review of capabilities, concepts, research gaps and future developments. **Earth-Science Reviews**, v. 217, p. 103604, 2021. ISSN 00128252. Available from: <https://doi.org/10.1016/j.earscirev.2021.103604Received>. 2, 7, 8

BREWIN, R. J. W.; SATHYENDRANATH, S.; HIRATA, T.; LAVENDER, S. J.; BARCIELA, R. M.; HARDMAN-MOUNTFORD, N. J. A three-component model of phytoplankton size class for the Atlantic Ocean. **Ecological Modelling**, v. 221, n. 11, p. 1472–1483, 2010. ISSN 0304-3800. Available from: <http://dx.doi.org/10.1016/j.ecolmodel.2010.02.014>. xiii, xvi, xvii, xxiv, 15, 17, 23, 24, 33, 34, 35, 36, 37, 44, 46, 84, 86, 91, 92, 93, 94, 95, 99, 100, 101, 116, 118

BREWIN, R. J. W.; SATHYENDRANATH, S.; JACKSON, T.; BARLOW, R.; BROTAS, V.; AIRS, R.; LAMONT, T. Influence of light in the mixed-layer on the parameters of a three-component model of phytoplankton size class. **Remote Sensing of Environment**, Elsevier Inc., v. 168, p. 437–450, 2015. ISSN 0034-4257. Available from: <http://dx.doi.org/10.1016/j.rse.2015.07.004>. 17, 37, 39, 46

BREWIN, R. J. W.; SATHYENDRANATH, S.; LANGE, P. K.; TILSTONE, G. Comparison of two methods to derive the size-structure of natural populations of phytoplankton. **Deep-Sea Research Part I**, v. 85, p. 72–79, 2014. ISSN 0967-0637. Available from: <http://dx.doi.org/10.1016/j.dsr.2013.11.007>. xv, 23, 33, 46, 49

BREWIN, R. J. W.; SATHYENDRANATH, S.; TILSTONE, G.; LANGE, P. K.; PLATT, T. A multicomponent model of phytoplankton size structure. **Journal of Geophysical Research: Oceans**, n. 119, p. 3478–3496, 2014. ISSN 21699275. xv, 35, 36

BRICAUD, A.; BABIN, M.; MOREL, A.; CLAUSTRE, H. Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: analysis and parameterization. **Journal of Geophysical Research**, v. 100, n. C7, p. 321–332, 1995. ISSN 0148-0227. Available from: <http://doi.wiley.com/10.1029/95JC00463>. 6, 45

BRICAUD, A.; CLAUSTRE, H.; RAS, J.; OUBELKHEIR, K. Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations. **Journal of Geophysical Research: Oceans**, v. 109, n. 11, p. 1–12, 2004. ISSN 21699291. xii, xxi, 1, 6, 51, 58, 59, 60, 75, 76, 77, 85, 86

BRICAUD, A.; MOREL, A.; BABIN, M.; ALLALI, K.; CLAUSTRE, H. Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters : analysis and implications for bio-optical models. **Journal of Geophysical Research**, v. 103, n. C13, p. 31033–31044, 1998. 109, 110, 119

BUENO, R. S.; MARQUES, H. L. d. A.; ROMA, R. P. C. R. Crescimento e sobrevivência da vieira Nodipecten nodosus (Linnaeus, 1758), (Mollusca: Pectinidae) em diferentes estruturas de cultivo na Praia Grande do Bonete, Ubatuba, Estado de São Paulo. **Biotemas**, v. 23, n. 1, 2011. ISSN 2175-7925. 3

CALADO, L.; GANGOPADHYAY, A.; SILVEIRA, I. C. da. A parametric model for the Brazil Current meanders and eddies off southeastern Brazil. **Geophysical Research Letters**, v. 33, n. 12, p. 1–5, 2006. ISSN 00948276. 86, 87

CAMPBELL, J. W. The lognormal distribution as a model for bio-optical variability in the sea. **Journal of Geophysical Research**, v. 100, n. C7, 1995. ISSN 01480227. 96

CAMPOS, E. J.; VELHOTE, D.; SILVEIRA, I. C. Shelf break upwelling driven by Brazil current cyclonic meanders. **Geophysical Research Letters**, v. 27, n. 6, p. 751–754, 2000. ISSN 00948276. 54, 87

CANTY, A.; RIPLEY, B. D. **boot: Bootstrap R (S-Plus) Functions**. 2021. 25

CARBONEL, C. A.; VALENTIN, J. L. Numerical modelling of phytoplankton bloom in the upwelling ecosystem of Cabo Frio (Brazil). **Ecological Modelling**, v. 116, n. 2-3, p. 135–148, 1999. ISSN 03043800. 81

CARVALHO, A. d. C. d. O.; MENDES, C. R. B.; KERR, R.; AZEVEDO, J. L. L. de; GALDINO, F.; TAVANO, V. M. The impact of mesoscale eddies on the phytoplankton community in the South Atlantic Ocean: HPLC-CHEMTAX approach. **Marine Environmental Research**, v. 144, n. 2019, p. 154–165, 2019. ISSN 18790291. Available from: <https://doi.org/10.1016/j.marenvres.2018.12.003>. 51

CARVALHO, M.; CIOTTI, A. M.; GIANESELLA, S. M. F.; CORRÊA, F. M. P. S.; PERINOTTO, R. R. C. Bio-optical properties of the inner continental shelf

off santos estuarine system, southeastern Brazil, and their implications for ocean color algorithm performance. **Brazilian Journal of Oceanography**, v. 62, n. 2, p. 71–87, 2014. ISSN 1982436X. 110

CASTELLO, J. P.; SUNYÉ, P. S.; HAIMOVICI, M.; HELLEBRANDT, D. Fisheries in southern Brazil: a comparison of their management and sustainability. **Journal of Applied Ichthyology**, v. 25, n. 3, p. 287–293, 2009. ISSN 01758659. 50

CASTRO, B.; BRANDINI, F.; PIRES-VANIN, A.; MIRANDA, L. Multidisciplinary oceanographic processes on the Western Atlantic continental shelf between 4°N and 34°S. In: ROBINSON, A. R.; BRINK, K. H. (Ed.). **The sea**. Cambriged: Harvard University Press, 2006. v. 14, p. 259–293. ISBN 0471115452. 52, 54

CASTRO, B. M. Summer/winter stratification variability in the central part of the South Brazil Bight. **Continental Shelf Research**, v. 89, p. 15–23, 2014. ISSN 02784343. Available from: <http://dx.doi.org/10.1016/j.csr.2013.12.002>. 122

CASTRO, B. M.; MIRANDA, L. B. Physical oceanography of the western Atlantic Continental Shelf located between 4º N and 34º S. In: ROBINSON, A. R.; BRINK, K. H. (Ed.). **The sea**. [S.l.]: John Wiley Sons, 1998. p. 209– 251 p. 50, 52

CASTRO, B. M. de; PEREIRA, A. F.; CAROLI, A. de; FOLONI-NETO, H.; PASCHOAL, G. C. A.; SILVEIRA, I. C. A.; AMOR, C. C. Correntes e massas de água na plataforma continental. In: MARTINS, R. P.; GROSSMANN-MATHESON, G. S. (Ed.). **Meteorologia e oceanografia**. Elsevier, 2015. p. 191–254. ISBN 9788535262087. Available from: <http://linkinghub.elsevier.com/retrieve/pii/B978853526208750012X>. 53

CATLETT, D.; SIEGEL, D. A. Phytoplankton pigment communities can be modeled using unique relationships with spectral absorption signatures in a dynamic coastal environment. **Journal of Geophysical Research: Oceans**, v. 123, n. 1, p. 246–264, 2018. ISSN 21699291. 16, 49, 57

CERDA, C.; CASTRO, B. M. Hydrographic climatology of South Brazil Bight shelf waters between Sao Sebastiao (24°S) and Cabo Sao Tome (22°S). **Continental Shelf Research**, v. 89, p. 5–14, 2014. ISSN 02784343. Available from: <http://dx.doi.org/10.1016/j.csr.2013.11.003>. 54, 87

CHASE, A. P.; KRAMER, S. J.; HAËNTJENS, N.; BOSS, E. S.; KARP-BOSS, L.; EDMONDSON, M.; GRAFF, J. R. Evaluation of diagnostic pigments to estimate phytoplankton size classes. **Limnology and Oceanography: Methods**, 2020. xv, 33, 46, 50, 58, 70, 84

CHISHOLM, S. Phytoplankton size. In: FALKOWSKI, P. G.; WOODHEAD, A. D.; VIVIRITO, K. (Ed.). **Primary productivity and biogeochemical cycles in the sea**. New York, NY: [s.n.], 1992. p. 213–237. 11

CIOTTI, A. M.; BRICAUD, A. Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. **Limnology and Oceanography: Methods**, v. 4, n. 7, p. 237–253, 2006. ISSN 15415856. 11, 60, 85, 86

CIOTTI, A. M.; FERREIRA, A.; GIANNINI, M. F. Seasonal and event-driven changes in the phytoplankton communities in the Araçá Bay and adjacent waters. **Ocean and Coastal Management**, v. 164, p. 14–31, 2018. ISSN 09645691. Available from: <https://doi.org/10.1016/j.ocecoaman.2018.03.024>. 51, 87

CIOTTI, A. M.; LEWIS, M. R.; CULLEN, J. J. Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. **Limnology and Oceanography**, v. 47, n. 2, p. 404–417, 2002. xxi, 11, 51, 60, 76, 85, 86

COMITÊ DE BACIAS HIDROGRÁFICAS DO LITORAL NORTE (CBHLN). **Plano de bacias hidrográficas do Litoral Norte**. [S.l.], 2017. 225 p. Available from: <http://www.ambiente.sp.gov.br/crhi/2010/LN.zip>. 51

COMMITTEE ON EARTH OBSERVATION SATELLITES (CEOS). **Ceos strategy for carbon observations from space**. [S.l.], 2014. 7

COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO (CETESB). **Relatório de qualidade das águas costeiras no estado de São Paulo 2016**. São Paulo, 2017. 178 p. Available from: <http://praias.cetesb.sp.gov.br/publicacoes-relatorios/>. 79, 87

_____. **Relatório de qualidade das águas costeiras no estado de São Paulo 2017**. São Paulo, 2018. 287 p. Available from: <http://cetesb.sp.gov.br/aguas-costeiras/publicacoes-e-relatorios/>. 62, 87

DAI, M.; SU, J.; ZHAO, Y.; HOFMANN, E. E.; CAO, Z.; CAI, W.-J.; GAN, J.; LACROIX, F.; LARUELLE, G. G.; MENG, F.; MÜLLER, J. D.; REGNIER, P. A.; WANG, G.; WANG, Z. Carbon fluxes in the coastal ocean: synthesis, boundary processes, and future trends. **Annual Review of Earth and Planetary Sciences**, v. 50, n. 1, p. 593–626, 2022. ISSN 0084-6597. 15

DEFOIN-PLATEL, M.; CHAMI, M. How ambiguous is the inverse problem of ocean color in coastal waters? **Journal of Geophysical Research: Oceans**, v. 112, n. 3, p. 1–16, 2007. ISSN 21699291. 10, 15, 92, 120

DELWICHE, C. F. Tracing the thread of plastid diversity through the tapestry of life. **The American Naturalist**, v. 154, n. S4, p. S164–S177, 1999. ISSN 0003-0147. 7

DETONI, A. M. S.; CIOTTI, A. M. Trichome abundance, chlorophyll content and the spectral coefficient for light absorption of Trichodesmium slicks observed in the Southwestern Atlantic. **Journal of Plankton Research**, v. 42, n. 2, p. 135–139, 2020. ISSN 0142-7873. 85

DEVRED, E.; SATHYENDRANATH, S.; STUART, V.; MAASS, H.; ULLOA, O.; PLATT, T. A two-component model of phytoplankton absorption in the open ocean: theory and applications. **Journal of Geophysical Research: Oceans**, v. 111, n. 3, p. 1–11, 2006. ISSN 21699291. 11, 85, 86

DEVRED, E.; SATHYENDRANATH, S.; STUART, V.; PLATT, T. A three component classification of phytoplankton absorption spectra : application to ocean color data. **Remote Sensing of Environment**, v. 115, n. 9, p. 2255–2266, 2011. ISSN 0034-4257. Available from: <http://dx.doi.org/10.1016/j.rse.2011.04.025>. xiii, xiv, xvi, xvii, xxiv, 15, 23, 86, 91, 92, 93, 94, 95, 96, 99, 100, 101, 103, 116, 118, 120

DIEGUES, A. C.; ARRUDA, R. S. V.; SILVA, V. C. F. da; FIGOLS, F. A. B.; ANDRADE, D. **Os Saberes Tradicionais e a Biodiversidade no Brasil**. São Paulo: USP, 1999. 189 p. 3

DUYSENS, L. The flattening of the absorption spectrum of suspensions as compared to that of solutions. **Biochimica et Biophysica Acta**, v. 19, p. 1–12, 1956. 7

EDLER, L.; ELBRÄCHTER, M. The Utermöhl method for quantitative phytoplankton analysis Lars. In: KARLSON, B.; CUSACK, C.; BRESNAN, E.

(Ed.). **Microscopic and molecular methods for quantitative phytoplankton analysis**. Paris: [s.n.], 2010. p. 110. Available from: <http://hab.ioc-unesco.org/index.php?option=com_oe&task= viewDocumentRecord&docID=5440>. 1

EFRON, B.; TIBSHIRANI, R. **An Introduction to the bootstrap**. [S.l.: s.n.], 1994. ISBN 9780412042317. 25

ELZHOV, T. V.; MULLEN, K. M.; SPIESS, A.-N.; BOLKER, B. **minpack.lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds**. 2016. Available from: <https://cran.r-project.org/package=minpack.lm>. 25

FAGUNDES, L.; GELLI, V. C.; OTANI, M. N.; VICENTE, M. C. M.; FREDO, C. E. Perfil sócio-econômico dos mitilicultores do litoral paulista. **Informações Econômicas, SP**, v. 34, n. 5, p. 47–59, 2004. 3

FOX-KEMPER, B.; HEWITT, H.; XIAO, C.; AðALGEIRSDÓTTIR, G.; DRIJFHOUT, S.; EDWARDS, T.; GOLLEDGE, N.; HEMER, M.; KOPP, R.; KRINNER, G.; MIX, A.; NOTZ, D.; NOWICIKI, S.; NURHATI, I.; RUIZ, J.; SALLÉE, J.; SLANGEN, A.; YU, Y. Ocean, cryosphere and sea level change. In: MASSON-DELMOTTE, V.; ZHAI, P.; PIRANI, A.; CONNORS, S.; PÉAN, C.; BERGER, S.; CAUD, N.; CHEN, Y.; GOLDFARB, L.; GOMIS, M.; HUANG, M.; LEITZELL, K.; LONNOY, E.; MATTHEWS, J.; MAYCOCK, T.; WATERFIELD, T.; YELEKÇI, O.; YU, R.; ZHOU, B. (Ed.). **Climate change 2021: the physical science basis. contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change science basis**. Cambridge: Cambridge University Press, 2021. v. 2018, p. 1–257. 1, 2

Fundação sistema estadual de análise de dados (SEADE). **População do Estado de São Paulo por municípios (2000-2021)**. 2022. Available from: <https://repositorio.seade.gov.br/dataset/ 08103e8c-95f6-4627-856b-a92071b97207/resource/ 4a05d680-ab13-428a-abef-3c6f69a6f800/download/populacao_municipal. csv>. 3

GAETA, S. A.; RIBEIRO, S. M. S.; METZLER, P. M.; FRANCOS, M. S.; ABE, D. S. Environmental forcing on phytoplankton biomass and primary productivity of the coastal ecosystem in Ubatuba region, southern Brazil. **Revista Brasileira de Oceanografia**, v. 47, n. 1, p. 11–27, 1999. 50, 51, 53

GEIDER, R. J. Light and temperature dependence of the carbon to chlorophyll a ratio in microalgae and cyanobacteria: implications for physiology and growth of phytoplankton. **New Phytologist**, v. 106, n. 1, p. 1–34, 1987. ISSN 14698137. 22

GEIDER, R. J.; MACINTYRE, H. L.; KANA, T. M. Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a:carbon ratio to light, nutrient-limitation and temperature. **Marine Ecology Progress Series**, v. 148, n. 1-3, p. 187–200, 1997. ISSN 01718630. 22

GIANNINI, M. F. C.; GARCIA, C. A. E.; TAVANO, V. M.; CIOTTI, Á. M. Effects of low-salinity and high-turbidity waters on empirical ocean colour algorithms: an example for Southwestern Atlantic waters. **Continental Shelf Research**, v. 59, p. 84–96, 2013. ISSN 02784343. Available from: <http://dx.doi.org/10.1016/j.csr.2013.04.013>. 105

GIOVANARDI, F.; VOLLENWEIDER, R. A. Trophic conditions of marine coastal waters: experience in applying the Trophic Index TRIX to two areas of the Adriatic and Tyrrhenian seas. **Journal of Limnology**, v. 63, n. 2, p. 199–218, 2004. ISSN 11295767. 57

GITTINGS, J. A.; BREWIN, R. J.; RAITSOS, D. E.; KHEIREDDINE, M.; OUHSSAIN, M.; JONES, B. H.; HOTEIT, I. Remotely sensing phytoplankton size structure in the Red Sea. **Remote Sensing of Environment**, v. 234, p. 111387, 2019. ISSN 00344257. Available from: <https://doi.org/10.1016/j.rse.2019.111387>. 15, 91, 120

GLIBERT, P. M. Margalef revisited: a new phytoplankton mandala incorporating twelve dimensions, including nutritional physiology. **Harmful Algae**, v. 55, p. 25–30, 2016. ISSN 15689883. Available from: <http://dx.doi.org/10.1016/j.hal.2016.01.008>. 78, 79

GONÇALVES-ARAUJO, R.; RÖTTGERS, R.; HARAGUCHI, L.; BRANDINI, F. P. Hydrography-driven variability of optically active constituents of water in the south brazilian bight: biogeochemical implications. **Frontiers in Marine Science**, v. 6, p. 1–11, 2019. ISSN 22967745. 122

GORDON, R.; BROWN, O. B.; EVANS, H.; BROWN, W.; SMITH, C.; BAKER, K.; CLARK, D. K. A semianalytic radiance model of ocean color. **Journal of Geo**, v. 93, n. D9, p. 10,909–10,92, 1988. 8, 9

GROOM, S. B.; SATHYENDRANATH, S.; BAN, Y.; BERNARD, S.; BREWIN, B.; BROTAS, V.; BROCKMANN, C.; CHAUHAN, P.; CHOI, J. K.; CHUPRIN, A.; CIAVATTA, S.; CIPOLLINI, P.; DONLON, C.; FRANZ, B. A.; HE, X.; HIRATA, T.; JACKSON, T.; KAMPEL, M.; KRASEMANN, H.; LAVENDER, S. J.; PARDO-MARTINEZ, S.; MELIN, F.; PLATT, T.; SANTOLERI, R.; SKAKALA, J.; SCHAEFFER, B.; SMITH, M.; STEINMETZ, F.; VALENTE, A.; WANG, M. Satellite ocean colour: current status and future perspective. **Frontiers in Marine Science**, v. 6, 2019. ISSN 22967745. 13, 88

GRUBER, N.; CLEMENT, D.; CARTER, B. R.; FEELY, R. A.; HEUVEN, S. van; HOPPEMA, M.; ISHII, M.; KEY, R. M.; KOZYR, A.; LAUVSET, S. K.; MONACO, C. L.; MATHIS, J. T.; MURATA, A.; OLSEN, A.; PEREZ, F. F.; SABINE, C. L.; TANHUA, T.; WANNINKHOF, R. The oceanic sink for anthropogenic CO 2 from 1994 to 2007. **Science**, v. 363, n. 6432, p. 1193–1199, 2019. ISSN 10959203. 49

HANSEN, H. P.; KOROLEFF, F. Determination of nutrients. In: GRASSHOFF, K.; KREMLING, K.; EHRHARDT, M. (Ed.). **Methods of seawater analysis**. Weinheim, Germany: Wiley-VCH Verlag GmbH, 1983. Available from: <http://doi.wiley.com/10.1002/9783527613984.ch10>. 57

HENSON, S. A.; CAEL, B. B.; ALLEN, S. R.; DUTKIEWICZ, S. Future phytoplankton diversity in a changing climate. **Nature Communications**, v. 12, n. 1, p. 1–8, 2021. ISSN 20411723. Available from: <http://dx.doi.org/10.1038/s41467-021-25699-w>. 1, 5

HIRATA, T.; HARDMAN-MOUNTFORD, N. J.; BREWIN, R. J.; AIKEN, J.; BARLOW, R.; SUZUKI, K.; ISADA, T.; HOWELL, E.; HASHIOKA, T.; NOGUCHI-AITA, M.; YAMANAKA, Y. Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. **Biogeosciences**, v. 8, n. 2, p. 311–327, 2011. ISSN 17264170. 83, 84

HU, C.; CARDER, K. L.; MULLER-KARGER, F. E. Atmospheric correction of SeaWiFS imagery over turbid coastal waters: a practical method. **Remote Sensing of Environment**, v. 74, n. 2, p. 195–206, 2000. ISSN 0034-4257. Available from: <https://www.sciencedirect.com/science/article/pii/S0034425700000808>. 9

HU, C.; LEE, Z.; FRANZ, B. Chlorophyll a algorithms for oligotrophic oceans : a novel approach based on three-band reflectance difference. **Journal of Geophysical Research**, v. 117, p. 1–25, 2012. 20, 97

HUAN, Y.; SUN, D.; WANG, S.; ZHANG, H.; QIU, Z.; BILAL, M.; HE, Y. Remote sensing estimation of phytoplankton absorption associated with size classes in coastal waters. **Ecological Indicators**, v. 121, p. 107198, 2021. ISSN 1470160X. Available from: <https://doi.org/10.1016/j.ecolind.2020.107198>. 99, 100

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). Summary for policymakers. In: PORTNER, H.-O.; ROBERTS, D.; MASSON-DELMOTTE, V.; ZHAI, P.; TIGNOR, M.; POLOCZANSKA, E.; MINTENBECK, K.; ALEGRIA, A.; NICOLAI, M.; OKEM, A.; PETZOLD, J.; RAMA, B.; WEYER, N. (Ed.). **IPCC special report on the ocean and cryosphere in a changing climate**. [S.l.]: In press, 2019. 1, 2, 15

SATHYENDRANATH, S. (Ed.). **Remote sensing of ocean colour in coastal, and other optically-complex, waters edited**. Dartmouth, Canada: IOCCG, 2000. 145 p. ISSN 1098-6030. ISBN 3. Available from: <http://www.vliz.be/imis/imis.php?module=ref&refid=134621&request=147862>. 10, 12

INTERNATIONAL OCEAN-COLOUR COORDINATING GROUP (IOCCG). **Remote sensing of inherent optical properties : fundamentals, tests of algorithms, and applications**. 2006. 126 p. ISSN 00222836. ISBN 9781896246567. Available from: <ioccg.org/reports/report5.pdf>. 9

INTERNATIONAL OCEAN-COLOUR COORDINATING GROUP (IOCCG). **Atmospheric correction for remotely sensed ocean colour products 2010**. [S.l.], 2010. ISBN 9781896246611. 9

INTERNATIONAL OCEAN-COLOUR COORDINATING GROUP (IOCCG). **Phytoplankton functional types from Space**. [S.l.], 2014. 156 p. Available from: <http://epic.awi.de/36000/>. 1, 2, 5, 10, 12, 47, 86

IRIGOIEN, X.; HULSMAN, J.; HARRIS, R. P. Global biodiversity patterns of marine phytoplankton and zooplankton. **Nature**, v. 429, n. 6994, p. 863–867, 2004. ISSN 00280836. 49

IWAMA, A. Y.; BATISTELLA, M.; FERREIRA, L. d. C.; TEIXEIRA, L. R. Dinâmica de cobertura e uso da terra e implicações sobre as áreas de riscos geodinâmicos e conservação no Litoral Norte de São Paulo. In: FERREIRA, L. d. C.; SCHMIDT, L.; BUENDIA, M. P.; CALVIMONTES, J.; VIGLIO, J. E. (Ed.). **Clima de tensão: ação humana, biodiversidade e mudanças**

climáticas. Campinas: Editora da Unicamp, 2017. p. 55–84. ISBN 9788526813700. 3

JEFFREY, S. W. Application of pigment methods to oceanography. In: JEFFREY, S. W.; MANTOURA, R.; WRIGHT, S. (Ed.). **Phytoplankton pigments in oceanography: guidelines to modern methods**. Paris: UNESCO, 1997. p. 661. 7, 82, 83

JEFFREY, S. W.; WRIGHT, S. W.; ZAPATA, M. Microalgal classes and their signature pigments. In: ROY, S.; LLEWELLYN, C. A.; EGELAND, E. S.; JOHNSEN, G. (Ed.). **Phytoplankton pigments: characterization, chemotaxonomy, and application in oceanography**. Cambridge: United Kingdom: Cambridge University Press, 2011. p. 3–77. ISBN 9780511732263. 2, 5, 7, 45, 49, 51, 59, 66, 82, 83

JOHNSON, M. D.; BEAUDOIN, D. J.; LAZA-MARTINEZ, A.; DYHRMAN, S. T.; FENSIN, E.; LIN, S.; MERCULIEF, A.; NAGAI, S.; POMPEU, M.; SETÄLÄ, O.; STOECKER, D. K. The genetic diversity of Mesodinium and associated cryptophytes. **Frontiers in Microbiology**, v. 7, p. 2017, 2016. ISSN 1664302X. 61, 82

KIRK, J. T. A theoretical analysis of the contribution of algal cells to the attenuation of light within natural waters. **New Phytologist**, v. 77, n. 2, p. 341–358, 1975. ISSN 14698137. 6

_____. Yellow substance (gelbstoff) and its contribution to the attenuation of photosynthetically active radiation in some inland and coastal south-eastern Australian waters. **Marine and Freshwater Research**, v. 27, n. 1, p. 61–71, 1976. Available from: <https://doi.org/10.1071/MF9760061>. 9

_____. **Light and photosynthesis in aquatic ecosystems**. 3. ed. New York: Cambridge University Press, 2011. 665 p. ISBN 9780521151757. 9

KOSTADINOV, T. S.; MILUTINOVI, S.; MARINOV, I.; CABRÉ, A. Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. **Ocean Science**, v. 12, n. 2, p. 561–575, 2016. ISSN 18120792. 11

KOSTADINOV, T. S.; SIEGEL, D. A.; MARITORENA, S. Retrieval of the particle size distribution from satellite ocean color observations. **Journal of Geophysical Research: Oceans**, 2009. ISSN 21699291. 11

KRAMER, S. J.; ROESLER, C. S.; SOSIK, H. M. Bio-optical discrimination of diatoms from other phytoplankton in the surface ocean: evaluation and refinement of a model for the Northwest Atlantic. **Remote Sensing of Environment**, v. 217, p. 126–143, 2018. ISSN 00344257. Available from: <https://doi.org/10.1016/j.rse.2018.08.010>. 49

KRAMER, S. J.; SIEGEL, D. A. How can phytoplankton pigments be best used to characterize surface ocean phytoplankton groups for ocean color remote sensing algorithms? **Journal of Geophysical Research: Oceans**, p. 7557–7574, 2019. ISSN 21699291. 44, 49, 51, 57, 91

KRAMER, S. J.; SIEGEL, D. A.; GRAFF, J. R. Phytoplankton community composition determined from co-variability among phytoplankton pigments from the NAAMES field campaign. **Frontiers in Marine Science**, v. 7, p. 1–15, 2020. ISSN 22967745. 51, 91

LAMONT, T.; BARLOW, R. G.; BREWIN, R. J. Variations in remotely-sensed phytoplankton size structure of a cyclonic eddy in the southwest Indian Ocean. **Remote Sensing**, v. 10, n. 7, 2018. ISSN 20724292. 35, 36, 44, 46

LAMONT, T.; BREWIN, R. J. W.; BARLOW, R. G. Seasonal variation in remotely-sensed phytoplankton size structure around southern Africa. **Remote Sensing of Environment**, v. 204, p. 617–631, 2018. ISSN 0034-4257. Available from: <https://doi.org/10.1016/j.rse.2017.09.038>. 39, 91, 99, 100

LATASA, M.; BIDIGARE, R. R. A comparison of phytoplankton populations of the Arabian Sea during the Spring Intermonsoon and Southwest Monsoon of 1995 as described by HPLC-analyzed pigments. **Deep Sea Research Part II: Topical Studies in Oceanography**, v. 45, n. 10, p. 2133–2170, 1998. ISSN 0967-0645. Available from: <http://www.sciencedirect.com/science/article/pii/S0967064598000666>. 57

LEE, D. I.; CHO, H.-S.; JEONG, S.-B. Distribution characteristics of marine litter on the sea bed of the East China Sea and the South Sea of Korea. **Estuarine, Coastal and Shelf Science**, v. 70, n. 1-2, p. 187–194, 2006. ISSN 02727714. Available from: <http://dx.doi.org/10.1016/j.ecss.2006.06.003>. 15

LEE, R. E. **Phycology**. 4. ed. Cambridge: [s.n.], 2008. 547 p. ISBN 9780521864084. 79

LEE, Z.; CARDER, K. L.; ARNONE, R. A. Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. **Applied Optics**, v. 41, n. 27, p. 5755, 2002. ISSN 0003-6935. 120

LEQUÉRÉ, C.; HARRISON, S. P.; PRENTICE, I. C.; BUITENHUIS, E. T.; AUMONT, O.; BOPP, L.; CLAUSTRE, H.; CUNHA, L. C. D.; GEIDER, R.; GIRAUD, X.; KLAAS, C.; KOHFELD, K. E.; LEGENDRE, L.; MANIZZA, M.; PLATT, T.; RIVKIN, R. B.; SATHYENDRANATH, S.; UITZ, J.; WATSON, A. J.; WOLF-GLADROW, D. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models ´. **Global Change Biology**, v. 11, p. 2016–2040, 2005. 5, 6, 49

LÉVY, M.; JAHN, O.; DUTKIEWICZ, S.; FOLLOWS, M. J. Phytoplankton diversity and community structure affected by oceanic dispersal and mesoscale turbulence. **Limnology and Oceanography: Fluids and Environments**, v. 4, n. 1, p. 67–84, 2014. 51

LI, Z.; LI, L.; SONG, K.; CASSAR, N. Estimation of phytoplankton size fractions based on spectral features of remote sensing ocean color data. **Journal of Geophysical Research: Oceans**, v. 118, n. 3, p. 1445–1458, 2013. ISSN 21699291. 11

LIMA, C. R.; MENDES, C. R. B.; TAVANO, V. M.; DETONI, A. M. S.; SECCHI, E. R. Chemotaxonomy-based mapping of phytoplankton communities in the subtropical Southwestern Atlantic Ocean, with emphasis on the marine cyanobacterium Trichodesmium. **Progress in Oceanography**, v. 172, p. 77–88, 2019. ISSN 00796611. Available from: <https://doi.org/10.1016/j.pocean.2019.01.008>. 50, 59, 70, 83, 85

LIU, B.; D'SA, E. J.; MAITI, K.; RIVERA-MONROY, V. H.; XUE, Z. Biogeographical trends in phytoplankton community size structure using adaptive sentinel 3-OLCI chlorophyll a and spectral empirical orthogonal functions in the estuarine-shelf waters of the northern Gulf of Mexico. **Remote Sensing of Environment**, v. 252, p. 112154, 2021. ISSN 00344257. Available from: <https://doi.org/10.1016/j.rse.2020.112154>. 2, 15, 91, 120

LIU, X.; DEVRED, E.; JOHNSON, C. Remote sensing of phytoplankton size class in northwest Atlantic from 1998 to 2016: Bio-optical algorithms comparison and application. **Remote Sensing**, v. 10, n. 7, p. 1–25, 2018. ISSN 20724292. 119, 120

LONGHURST, A. R. **Ecological geography of the sea**. San Diego: Academic Press, 1998. 398 p. ISBN 0-12-455559-4. 17

_____. **Ecological Geography of the Sea**. 2nd. ed. Elsevier, 2007. ISBN 9780124555211. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780124555211X50001>. 17

MACKEY, M. D.; MACKEY, D. J.; HIGGINS, H. W.; WRIGHT, S. W. CHEMTAX- a program for estimating class abundances from chemical markers : application to HPLC measurements of phytoplankton. **Marine Ecology Progress Series**, v. 144, p. 265–283, 1996. 49, 58

MARÃNÓN, E.; CERMEÑO, P.; HUETE-ORTEGA, M.; LÓPEZ-SANDOVAL, D. C.; MOURIÑO-CARBALLIDO, B.; RODRÍGUEZ-RAMOS, T. Resource supply overrides temperature as a controlling factor of marine phytoplankton growth. **PLoS ONE**, v. 9, n. 6, p. 20–23, 2014. ISSN 19326203. 22

MARGALEF, R. Life-forms of phytoplankton as survival alternatives in an unstable environment. **Oceanologica Acta**, v. 1, n. 4, p. 493–509, 1978. ISSN 00320935. Available from: <https://archimer.ifremer.fr/doc/00123/23403/>. 78, 79

MASUDA, Y.; YAMANAKA, Y.; HIRATA, T.; NAKANO, H. Competition and community assemblage dynamics within a phytoplankton functional group: simulation using an eddy-resolving model to disentangle deterministic and random effects. **Ecological Modelling**, v. 343, p. 1–14, 2017. ISSN 03043800. Available from: <http://dx.doi.org/10.1016/j.ecolmodel.2016.10.015>. 5

MATSUURA, Y. A probable cause of recruitment failure of the Brazilian sardine Sardinella aurita population during the 1974/75 spawning season. **South African Journal of Marine Science**, v. 17, p. 29–35, 1996. ISSN 02577615. 50

MÉLIN, F.; VANTREPOTTE, V. How optically diverse is the coastal ocean? **Remote Sensing of Environment**, v. 160, p. 235–251, 2015. ISSN 00344257. Available from: <http://dx.doi.org/10.1016/j.rse.2015.01.023>. 15

MENDEN-DEUER, S.; LESSARD, E. J. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. **Limnology and Oceanography**, v. 45, n. 3, p. 569–579, 2000. ISSN 00223018. xvi, 55, 56, 85

METZLER, P. M.; GLIBERT, P. M.; GAETA, S. A.; LUDLAM, J. M. New and regenerated production in the South Atlantic off Brazil. **Deep-Sea Research Part I: Oceanographic Research Papers**, v. 44, n. 3, p. 363–384, 1997. ISSN 09670637. 79

MIDDELBURG, J. J. **Marine carbon biogeochemistry: a primer for earth system scientists**. Springer Open, 2019. ISBN 978-3-030-10821-2. Available from: <http://link.springer.com/10.1007/978-3-030-10822-9>. 7

Ministério do Meio Ambiente (MMA). **Macrodiagnóstico da zona costeira e marinha do Brasil**. Brasília, 2008. 242 p. 3

MITCHELL, B. G.; KAHRU, M.; WIELAND, J.; STRAMSKA, M. **Determination of spectral absorption coefficients of particles, dissolved material and phytoplankton for ocean discrete water samples**. [S.l.], 2002. 56, 93

MOBLEY, C. D. **Light and water radiative transfer in natural waters**. [S.l.: s.n.], 2004. 1–4 p. 9

MOISAN, T. A.; RUFTY, K. M.; MOISAN, J. R.; LINKSWILER, M. A. Satellite observations of phytoplankton functional type spatial distributions, phenology, diversity, and ecotones. **Frontiers in Marine Science**, v. 4, p. 1–24, 2017. 86

MÖLLER, O. O.; PIOLA, A. R.; FREITAS, A. C.; CAMPOS, E. J. The effects of river discharge and seasonal winds on the shelf off southeastern South America. **Continental Shelf Research**, v. 28, n. 13, p. 1607–1624, 2008. ISSN 02784343. 50, 87

MONTAGNES, D. J. S.; BERGES, J. A.; HARRISON, P. J.; TAYLOR, F. J. R. Estimating carbon, nitrogen, protein, and chlorophyll a from volume in marine phytoplankton. **Liminology and Oceanography**, v. 35, n. 5, p. 1044–106,, 1994. Available from: <https://doi.org/10.4319/lo.1994.39.5.1044>. 55

MOREL, A. Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). **Journal of Geophysical Research**, v. 93, n. C9, p. 10749, 1988. ISSN 0148-0227. 10

MOREL, A.; BRICAUD, A. Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. **Deep Sea Research Part A, Oceanographic Research Papers**, v. 28, n. 11, p. 1375–1393, 1981. ISSN 01980149. 59

MOREL, A.; PRIEUR, L. Analysis of variations in ocean color. **Limnology and Oceanography**, v. 22, n. 4, p. 709–722, 1977. ISSN 19395590. 9, 10, 15

MOSER, G. A.; CASTRO, N. O.; TAKANOHASHI, R. A.; FERNANDES, A. M.; POLLERY, R. C.; TENENBAUM, D. R.; VARELA-GUERRA, J.; BARRERA-ALBA, J. J.; CIOTTI, A. M. The influence of surface low-salinity waters and cold subsurface water masses on picoplankton and ultraplankton distribution in the continental shelf off Rio de Janeiro, SE Brazil. **Continental Shelf Research**, v. 120, p. 82–95, 2016. ISSN 18736955. Available from: <http://dx.doi.org/10.1016/j.csr.2016.02.017>. 50

MOSER, G. A.; TAKANOHASHI, R. A.; BRAZ, M. d. C.; LIMA, D. T. de; KIRSTEN, F. V.; GUERRA, J. V.; FERNANDES, A. M.; POLLERY, R. C. G. Phytoplankton spatial distribution on the Continental Shelf off Rio de Janeiro, from Paraíba do Sul River to Cabo Frio. **Hydrobiologia**, v. 728, n. 1, p. 1–21, 2014. ISSN 15735117. 83, 86, 87

MOUW, C. B.; HARDMAN-MOUNTFORD, N. J.; ALVAIN, S.; BRACHER, A.; BREWIN, R. J. W.; BRICAUD, A.; CIOTTI, A. M.; DEVRED, E. A Consumer 's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. **Frontiers in Marine Science**, v. 4, n. 41, p. 1–19, 2017. 2, 86

MURPHY, J.; RILEY, J. P. A modified single solution method for the determination of phosphate in natural waters. **Analytica Chimica Acta**, v. 27, p. 31–36, 1962. ISSN 0003-2670. Available from: <https://www.sciencedirect.com/science/article/pii/S0003267000884445>. 57

MURTAGH, F.; LEGENDRE, P. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? **Journal of Classificatio**, n. 31, p. 274–295, 2014. 58

NAIR, A.; SATHYENDRANATH, S.; PLATT, T.; MORALES, J.; STUART, V.; FORGET, M. H.; DEVRED, E.; BOUMAN, H. Remote sensing of phytoplankton functional types. **Remote Sensing of Environment**, 2008. ISSN 00344257. 2, 5, 6, 86

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA). **GHRSST level 4 MUR global foundation sea surface temperature analysis (v4.1)**. NASA PO.DAAC, 2015. Available from: <http://podaac.jpl.nasa.gov/dataset/MUR-JPL-L4-GLOB-v4.1>. 22, 81, 97

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA).
Coastal Zone Color Scanner (CZCS). 2019. Available from:
<https://oceancolor.gsfc.nasa.gov/data/czcs/instrument/>. 12

NELSON, N. B.; PREZELIN, B. B.; BIDIGARE, R. R. Phytoplankton light
absorption and the package effect in California coastal waters. **Marine Ecology
Progress Series**, v. 94, n. 3, p. 217–227, 1993. ISSN 01718630. 60

OLSON, R. J.; SOSIK, H. M. A submersible imaging-in-flow instrument to analyze
nano-and microplankton: Imaging FlowCytobot. **Limnology and
Oceanography: Methods**, v. 5, n. 6, p. 195–203, 2007. ISSN 15415856. 5

PALACZ, A. P.; JOHN, M. A. S.; BREWIN, R. J.; HIRATA, T.; GREGG, W. W.
Distribution of phytoplankton functional types in high-nitrate, low-chlorophyll
waters in a new diagnostic ecological indicator model. **Biogeosciences**, v. 10,
n. 11, p. 7553–7574, 2013. ISSN 17264170. 11

PIOLA, A. R.; CAMPOS, E. J.; MÖLLER, O. O.; CHARO, M.; MARTINEZ, C.
Subtropical shelf front off eastern South America. **Journal of Geophysical
Research: Oceans**, v. 105, n. C3, p. 6565–6578, 2000. ISSN 21699291. 50, 54, 82

PIOLA, A. R.; MÖLLER, O. O.; GUERRERO, R. A.; CAMPOS, E. J. Variability
of the subtropical shelf front off eastern South America: winter 2003 and summer
2004. **Continental Shelf Research**, v. 28, n. 13, p. 1639–1648, 2008. ISSN
02784343. 50, 51, 110, 115

PRIEUR, L. Louis Prieur. **Limonology And Oceanography**, v. 26, n. 4, p.
671–689, 1981. ISSN 19395590. Available from:
<http://www.jstor.org/stable/10.2307/2836033>. 6

PRIMPAS, I.; KARYDIS, M. Scaling the trophic index (TRIX) in oligotrophic
marine environments. **Environmental Monitoring and Assessment**, v. 178,
n. 1-4, p. 257–269, 2011. ISSN 01676369. 57

RAITSOS, D. E.; LAVENDER, S. J.; MARAVELIAS, C. D.; HARALABOUS, J.;
RICHARDSON, A. J.; REID, P. C. Identifying four phytoplankton functional
types from space: an ecological approach. **Limnology and Oceanography**, 2008.
ISSN 00243590. 2, 11

REMBAUVILLE, M.; BRIGGS, N.; ARDYNA, M.; UITZ, J.; CATALA, P.;
PENKERC'H, C.; POTEAU, A.; CLAUSTRE, H.; BLAIN, S. Plankton
assemblage estimated with BGC-Argo floats in the Southern Ocean: implications

for seasonal successions and particle export. **Journal of Geophysical Research: Oceans**, v. 122, p. 8278–8292, 2017. 126

RIAL, P.; GARRIDO, J. L.; JAÉN, D.; RODRÍGUEZ, F. Pigment composition in three Dinophysis species (Dinophyceae) and the associated cultures of Mesodinium rubrum and Teleaulax amphioxeia. **Journal of Plankton Research**, v. 35, n. 2, p. 433–437, 2013. ISSN 01427873. 82

RIBEIRO, C. G.; SANTOS, A. L.; MARIE, D.; BRANDINI, F. P.; VAULOT, D. Small eukaryotic phytoplankton communities in tropical waters off Brazil are dominated by symbioses between Haptophyta and nitrogen-fixing cyanobacteria. **ISME Journal**, v. 12, n. 5, p. 1360–1374, 2018. ISSN 17517370. Available from: <http://dx.doi.org/10.1038/s41396-018-0050-z>. 5

RIBEIRO, C. G.; SANTOS, A. L. dos; MARIE, D.; PELLIZARI, V. H.; BRANDINI, F. P.; VAULOT, D. Pico and nanoplankton abundance and carbon stocks along the Brazilian Bight. **PeerJ**, v. 2016, n. 11, p. 1–20, 2016. ISSN 21678359. 50, 87

RIEBESELL, U.; ZONDERVAN, I.; ROST, B.; TORTELL, P. D.; ZEEBE, R. E.; MOREL, F. M. Reduced calcification of marine plankton in response to increased atmospheric CO2. **Nature**, v. 407, n. 6802, p. 364–366, 2000. ISSN 00280836. 6

ROY, S.; LIEWELLYN, C.; EGALAND, E. S.; JOHNSEN, G. **Phytoplankton pigments: chemotaxonomy and applications in oceanography**. [S.l.]: Cambridge University Press, 2011. ISSN 1098-6596. ISBN 9788578110796. 5

RUDORFF, N. d. M.; KAMPEL, M. Orbital remote sensing of phytoplankton functional types: A new review. **International Journal of Remote Sensing**, v. 33, n. 6, p. 1967–1990, 2012. ISSN 13665901. 2, 86

SABINE, C. L.; FEELY, R. A.; GRUBER, N.; KEY, R. M.; LEE, K.; BULLISTER, J. L.; WANNINKHOF, R.; WONG, C. S.; WALLACE, D. W.; TILBROOK, B.; MILLERO, F. J.; PENG, T. H.; KOZYR, A.; ONO, T.; RIOS, A. F. The oceanic sink for anthropogenic CO2. **Science**, v. 305, n. 5682, p. 367–371, 2004. ISSN 00368075. 49

SARKER, S.; YADAV, A. K.; HOSSAIN, M. S.; CHOWDHURY, S. R.; KABIR, M. A.; SHARIFUZZAMAN, S. M. The drivers of diatom in subtropical coastal waters: a Bayesian modelling approach. **Journal of Sea Research**, v. 163, p. 101915, 2020. ISSN 13851101. Available from: <https://doi.org/10.1016/j.seares.2020.101915>. 91

SATHYENDRANATH, S.; BREWIN, B.; MUELLER, D.; DOERFFER, R.; KRASEMANN, H.; MELIN, F.; BROCKMANN, C.; FOMFERRA, N.; PETERS, M.; GRANT, M.; STEINMETZ, F.; DESCHAMPS, P. Y.; SWINTON, J.; SMYTH, T.; WERDELL, J.; FRANZ, B.; MARITORENA, S.; DEVRED, E.; LEE, Z.; HU, C.; REGNER, P. Ocean colour climate change initiative - approach and initial results. **International Geoscience and Remote Sensing Symposium (IGARSS)**, p. 2024–2027, 2012. 13

SATHYENDRANATH, S.; BREWIN, R. J.; JACKSON, T.; MÉLIN, F.; PLATT, T. Ocean-colour products for climate-change studies: what are their ideal characteristics? **Remote Sensing of Environment**, v. 203, p. 125–138, 2017. ISSN 00344257. Available from: <https://doi.org/10.1016/j.rse.2017.04.017>. 13, 16, 45, 88, 125, 126

SATHYENDRANATH, S.; COTA, G.; STUART, V.; MAASS, H.; PLATT, T. Remote sensing of phytoplankton pigments: a comparison of empirical and theoretical approaches. **International Journal of Remote Sensing**, v. 22, n. 2-3, p. 249–273, 2001. ISSN 01431161. 15, 85, 86, 91

SATHYENDRANATH, S.; LAZZARA, L.; PRIEUR, L. Variations in the spectral values of specific absorption of phytoplankton. **Limnology and Oceanography**, v. 32, n. 2, p. 403–415, 1987. ISSN 19395590. 59

SATHYENDRANATH, S.; STUART, V.; NAIR, A.; OKA, K.; NAKANE, T.; BOUMAN, H.; FORGET, M. H.; MAASS, H.; PLATT, T. Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. **Marine Ecology Progress Series**, v. 383, p. 73–84, 2009. ISSN 01718630. 64

SCHLÜTER, L.; MØHLENBERG, F.; HAVSKUM, H.; LARSEN, S. The use of phytoplankton pigments for identifying and quantifying phytoplankton groups in coastal areas: testing the influence of light and nutrients on pigment/chlorophyll a ratios. **Marine Ecology Progress Series**, v. 192, p. 49–63, 2000. ISSN 01718630. 49

SHERRARD, N. J.; NIMMO, M.; LLEWELLYN, C. A. Combining HPLC pigment markers and ecological similarity indices to assess phytoplankton community structure : an environmental tool for eutrophication? **Science of the Total Environment**, v. 361, n. 1-3, p. 97–110, 2006. 55

SHOAF, W. T.; LIUM, B. W. Improved extraction of chlorophyll a and b from algae using dimethyl sulfoxide. **Limnology and Oceanography**, v. 21, n. 6, p.

926–928, 1976. Available from: <https://aslopubs.onlinelibrary.wiley.com/doi/abs/10.4319/lo.1976.21.6.0926>. 92, 93

SICKO-GOAD, L. M.; SCHELSKE, C. L.; STOERMER, E. F. Estimation of intracellular carbon and silica content of diatoms from natural assemblages using morphometric techniques. **Limnology and Oceanography**, v. 29, n. 6, p. 1170–1178, 1984. ISSN 19395590. 85

SIEBURTH, J.; SMETACEK, V.; LENZ, J. Pelagic ecosystem structure : heterotrophic compartments of the plankton and their relationship to plankton size fractions. **Limnology and Oceanography**, v. 23, n. 6, p. 1256–1263, 1978. 6

SIEGEL, D. A.; PETERSON, P.; MCGILLICUDDY, D. J.; MARITORENA, S.; NELSON, N. B. Bio-optical footprints created by mesoscale eddies in the Sargasso Sea. **Geophysical Research Letters**, v. 38, n. 13, p. 1–6, 2011. ISSN 00948276. 51

SILVEIRA, I. C. A. da; SCHMID, A. C. K.; CAMPOS, E. J. D.; GODOI, S. S. de; IKEDA, Y. A corrente do Brasil ao largo da costa leste brasileira. **Revista Brasileira de Oceanografia**, v. 48, n. 2, p. 171–183, 2000. Available from: <http://www.scielo.br/pdf/bjoce/v48n2/08.pdf>. 54

SOETAERT, K.; MEERSCHE, K. van den; OEVELEN, D. van. **limSolve: solving linear inverse models**. [S.l.: s.n.], 2009. 94

SOUTELINO, R. G.; GANGOPADHYAY, A.; SILVEIRA, I. C. da. The roles of vertical shear and topography on the eddy formation near the site of origin of the Brazil Current. **Continental Shelf Research**, v. 70, p. 46–60, 2013. ISSN 02784343. Available from: <http://dx.doi.org/10.1016/j.csr.2013.10.001>. 86

STOCK, A.; SUBRAMANIAM, A. Accuracy of empirical satellite algorithms for mapping phytoplankton diagnostic pigments in the open ocean: a supervised learning perspective. **Frontiers in Marine Science**, v. 7, 2020. ISSN 22967745. 17, 20, 22, 46, 47

STRAMSKI, D.; SCIANDRA, A.; CLAUSTRE, H. Effects of temperature, nitrogen, and light limitation on the optical properties of the marine diatom Thalassiosira pseudonana. **Limnology and Oceanography**, v. 47, n. 2, p. 392–403, 2002. ISSN 00243590. 22

STRICKLAND, J.; PARSONS, T. Determination of dissolved oxygen. In: STRICKLAND, J.; PARSONS, T. (Ed.). **A practical handbook of seawater analysis**. Ottawa: Fisheries Research Board of Canada, Bulletin 167, 1968. p. 71–75. 57

SUN, D.; HUAN, Y.; QIU, Z.; HU, C.; WANG, S.; HE, Y. Remote-sensing estimation of phytoplankton size classes from GOCI satellite measurements in Bohai Sea and Yellow Sea. **Journal of Geophysical Research: Oceans**, v. 122, n. 10, p. 8309–8325, 2017. ISSN 21699291. 15

SUN, D.; HUAN, Y.; WANG, S.; QIU, Z.; LING, Z.; MAO, Z.; HE, Y. Remote sensing of spatial and temporal patterns of phytoplankton assemblages in the Bohai Sea, Yellow Sea, and east China sea. **Water Research**, v. 157, p. 119–133, 2019. ISSN 18792448. Available from: <https://doi.org/10.1016/j.watres.2019.03.081>. 15, 91, 120

SUN, J.; LIU, D. Geometric models for calculating cell biovolume and surface area for phytoplankton. **Journal of Plankton Research**, v. 25, n. 11, p. 1331–1346, 2003. 55

SUSINI, S. M.; LOPES, R. M.; TEIXEIRA, C.; GAETA, S. A.; ABE, D. S. Size-fractionation of primary production and phytoplankton biomass on inshore waters of the Ubatuba region, Brazil. **Publicacao Especial do Instituto Oceanografico PP - São Paulo**, v. 11, p. 153–162, 1995. 50

TASSAN, S.; FERRARI, G. An alternative approach to absorption measurements of aquatic particles retained on filters. **Liminology and Oceanography**, v. 40, n. 8, p. 1358–1368, 1995. 56

TASSAN, S.; FERRARI, M. A sensitivity analysis of the 'Transmittance-Reflectance' method for measuring light absorption by aquatic particles. **Journal of Plankton Research**, v. 24, n. 8, p. 757–774, 2002. 56

TEIXEIRA, L. R. **Megaprojetos no litoral norte paulista: o papel dos grandes empreendimentos de infraestrutura na transformação regional**. PhD Thesis (PhD in Ambiente e Sociedade) — Universidade Estadual de Campinas, 2013. Available from: <http://repositorio.unicamp.br/handle/REPOSIP/279975>. 3

TURNER, K. J.; MOUW, C. B.; HYDE, K. J.; MORSE, R.; CIOCHETTO, A. B. Optimization and assessment of phytoplankton size class algorithms for ocean

color data on the Northeast U.S. continental shelf. **Remote Sensing of Environment**, v. 267, p. 112729, 2021. ISSN 00344257. Available from: <https://doi.org/10.1016/j.rse.2021.112729>. xv, 2, 15, 22, 33, 35, 36, 37, 39, 44, 46, 47, 91, 99, 100

UITZ, J.; CLAUSTRE, H.; MOREL, A.; HOOKER, S. B. Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll. **Journal of Geophysical Research: Oceans**, 2006. ISSN 21699291. xv, xvi, xvii, 2, 6, 7, 15, 23, 33, 34, 36, 45, 51, 58, 67, 68, 71, 72, 82, 83, 84, 86, 87, 91, 95, 101

UTERMÖHL, H. Zur vevookommung der quantitativen phytoplankton methodik. **Mitteilungen Internationale Vereiningung fuer Theoretische und Angewandte Limnologie**, n. 9, p. 1–38, 1958. 5, 55

VANDENMEERSCHE, K.; SOETAERT, K.; VANOEVELEN, D. xsample(): an R Function for sampling linear inverse problems. **Journal of Statistical Software, Code Snippets**, v. 30, n. 1, p. 1–15, 2009. ISSN 1548-7660. Available from: <http://www.jstatsoft.org/v30/c01>. 94

VANHEUKELEM, L.; HOOKER, S. B. The importance of a quality assurance plan for method validation and minimizing uncertainties in the HPLC analysis of phytoplankton pigments. In: ROY, S.; LLEWELLYN, C. A.; EGELAND, E. S. (Ed.). **Phytoplankton Ppgments: characterization. chemotaxonomy and applications in oceanography**. Cambridge: Cambridge University Press, 2011. p. 195–242. ISBN 9780511732263. 55

VANTREPOTTE, V.; LOISEL, H.; DESSAILLY, D.; MÉRIAUX, X. Optical classification of contrasted coastal waters. **Remote Sensing of Environment**, v. 123, p. 306–323, 2012. ISSN 00344257. 15

VASCONCELLOS, M.; GASALLA, M. A. Fisheries catches and the carrying capacity of marine ecosystems in southern Brazil. **Fisheries Research**, v. 50, n. 3, p. 279–295, 2001. ISSN 01657836. 50

VIDUSSI, F.; CLAUSTRE, H.; MANCA, B. B.; LUCHETTA, A.; JEAN-CLAUDE, M. Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea During winter. **Journal of Geophysical Research**, v. 106, p. 19939–19956, 2001. ISSN 1078-8956. 2, 7, 19

VILLAC, M. C.; CABRAL-NORONHA, V. D. P. A.; PINTO, T. d. O. The phytoplankton biodiversity of the coast of the state of São Paulo, Brazil. **Biota Neotropica**, v. 8, n. 3, p. 151–173, 2008. ISSN 16760603. 83

VOLLENWEIDER, R. A.; GIOVANARDI, F.; MONTANARI, G.; RINALDI, A. Characterization of the trophic conditions of marine coastal waters with special reference to the NW Adriatic Sea: proposal for a trophic scale, turbidity and generalized water quality index. **Environmetrics**, v. 9, n. 3, p. 329–357, 1998. ISSN 11804009. 57

WELSCHMEYER, N. A. Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. **Limnology and Oceanography**, v. 39, n. 8, p. 1985–1992, 1994. ISSN 00223840. 92, 93

WERDELL, P. J.; BAILEY, S.; FARGION, G.; PIETRAS, C.; KNOBELSPIESSE, K.; FELDMAN, G.; MCCLAIN, C. Unique data repository facilitates ocean color satellite validation. **EOS 84**, p. 377–387, 2003. Available from: <`https: //oceancolor.gsfc.nasa.gov/staff/jeremy/werdell_et_al_2003_eos.pdf`>. 17

WERDELL, P. J.; FRANZ, B. A.; BAILEY, S. W.; FELDMAN, G. C.; BOSS, E.; BRANDO, V. E.; DOWELL, M.; HIRATA, T.; LAVENDER, S. J.; LEE, Z.; LOISEL, H.; MARITORENA, S.; MÉLIN, F.; MOORE, T. S.; SMYTH, T. J.; ANTOINE, D.; DEVRED, E.; D'ANDON, O. H. F.; MANGIN, A. Generalized ocean color inversion model for retrieving marine inherent optical properties. **Applied Optics**, v. 52, n. 10, p. 2019, 2013. ISSN 0003-6935. Available from: <`https://www.osapublishing.org/abstract.cfm?URI=ao-52-10-2019`>. 96, 97

WERDELL, P. J.; ROESLER, C. S.; GOES, J. I. Discrimination of phytoplankton functional groups using an ocean reflectance inversion model. **Applied Optics**, v. 53, n. 22, p. 4833, 2014. ISSN 1559-128X. 11

WESTBERRY, T. K.; SIEGEL, D. A.; SUBRAMANIAM, A. An improved bio-optical model for the remote sensing of Trichodesmium spp. blooms. **Journal of Geophysical Research: Oceans**, v. 110, n. 6, p. 1–11, 2005. ISSN 21699291. 11

WILTSHIRE, K. H.; BOERSMA, M.; CARSTENS, K.; KRABERG, A. C.; PETERS, S.; SCHARFE, M. Control of phytoplankton in a shelf sea: determination of the main drivers based on the Helgoland Roads Time Series.

Journal of Sea Research, v. 105, p. 42–52, 2015. ISSN 13851101. Available from: <http://dx.doi.org/10.1016/j.seares.2015.06.022>. 91

ZAPATA, M.; JEFFREY, S. W.; WRIGHT, S. W.; RODRÍGUEZ, F.; GARRIDO, J. L.; CLEMENTSON, L. Photosynthetic pigments in 37 species (65 strains) of Haptophyta: Implications for oceanography and chemotaxonomy. **Marine Ecology Progress Series**, v. 270, p. 83–102, 2004. ISSN 01718630. 49

ZHANG, Y.; WANG, G.; SATHYENDRANATH, S.; XU, W.; XIAO, Y.; JIANG, L. Retrieval of phytoplankton pigment composition from their in vivo absorption spectra. **Remote Sensing**, v. 13, n. 24, 2021. ISSN 20724292. 121

APPENDIX A - REGRESSION OF CHLOROPHYLL-*a* BY FLUO-ROMETRY AND HPLC FOR CHAPTER 5

In order to obtain the chlorophyll-*a* estimated by a fluorometric method which would meet the concentrations estimated by HPLC, we did a regression in the $log_{10}space$ between the chlorophyll-*a* measured by fluorometry and the total chlorophyll-*a* measured by HPLC. And using the regression parameters, we corrected the chlorophyll-*a* by fluorometry. It is possible to observe in Figure A.1, that the fluorometric method is consistently larger than the HPLC values in Antares-Ubatuba station. When this regression was performed not all the analysis used in the study were finished, since it was in an intermediary moment. As these parameters were the ones used on the correction of the fluorometry Chl*a* they were reported as used. A regression including the remaining data presented similar results (n=41, R²=0.87, a=-0.207, b=0.886).

Figure A.1 - Regression analysis applied to correct the chlorophyll-*a* measured by fluorometry to meet the total chlorophyll-*a* measured by HPLC.
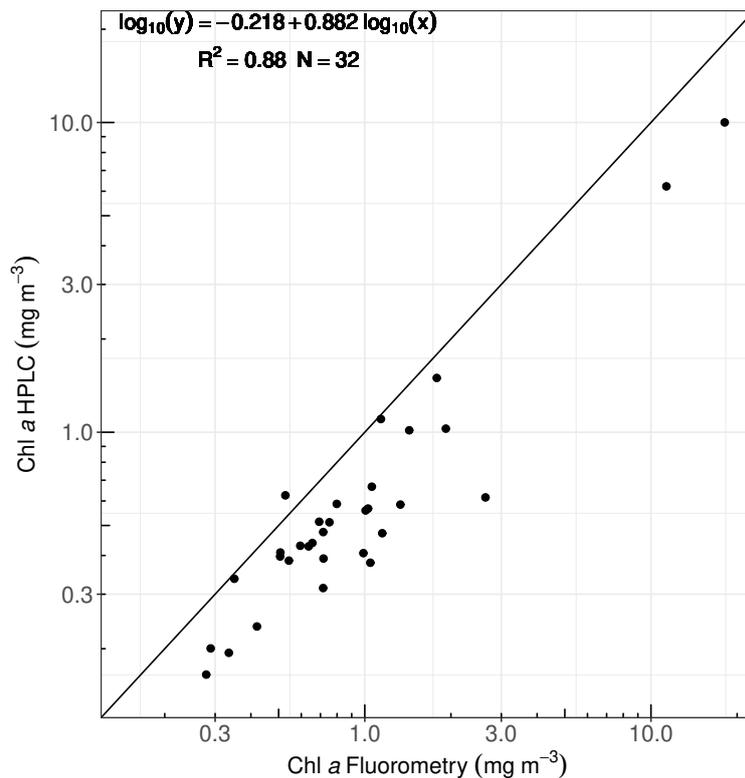
Solid black line is the 1:1.

SOURCE: Author's production.

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.