

FINAL REPORT

Assessing the inter-annual **Physiological** response of
phytoplankton to **Global** warming using long-term satellite
observations

PhysioGlob

Fellow: Marco Bellacicco

Supervisor: Dr. Salvatore Marullo

Host Institute: ENEA, Climate Modelling Laboratory, Centro Ricerche Frascati, Via Enrico
Fermi, 45, 00044 Frascati, Italy

Prepared by
Approved by
Reference
Date of Issue

[Marco Bellacicco]
[Dr. Salvatore Marullo]
FR_<PhysioGlob>
2021-01-29

23 **Table of Contents:**

24 **1 Executive summary**

25 **2 Objectives and workplan**

26 **3 Work performed**

27 **4 Conclusions, Recommendations and Challenges**

28 **5 Essential References**

29 **6 Publications**

30

31

32 1 EXECUTIVE SUMMARY

33 Oceans constitute nearly the 70% of Earth's surface and hold an important population of
34 aquatic microscopic plant-like organism, the so-called phytoplankton. These organisms are
35 considered to be responsible for approximately 50% of the planetary primary production
36 [Behrenfeld *et al.*, 2006] and are at the base of the trophic chain. Large scale factors such as
37 climate, ocean circulation, and mainly, anthropogenic activities, affect phytoplankton
38 biomass and its spatio-temporal variability. Therefore, due to their relevant role in the ocean,
39 it is critical to complement their knowledge. Phytoplankton are also defined as a sort of
40 sentinels of changes in the ocean, because they rapidly respond to environment perturbations.
41 Light, nutrients and temperature are the most important agents that influence phytoplankton
42 production [Behrenfeld *et al.*, 2008]. Phytoplankton cells respond to the change of these
43 variables with different physiological strategies which aims to increase the efficiency of light
44 capturing, photosynthetic capacity, growth and persistence. There are two main types of
45 phytoplankton response: photoadaptation and photoacclimation. The former describes
46 changes that might happen at genotype level, and are expected to occur at a long evolutionary
47 time-scale. The latter is a cellular process enables phytoplankton to change (*i.e.*, increase or
48 reduction) the intracellular chlorophyll-*a* concentration (Chl) which can be observed and
49 quantified also by using space-borne observations [Moore *et al.*, 2006; Dubinsky and
50 Stambler, 2009]. The photoacclimation is commonly described in terms of variation of the
51 ratio between Chl and phytoplankton carbon (C_{phyto}) the well-known Chl: C_{phyto} ratio. Even if
52 in the last twenty-years, several studies have shown that ocean warming causes an expansion
53 of the low-Chl and low-productivity areas (*e.g.* subtropical gyres) [Martinez *et al.*, 2009],
54 two main questions are not fully addressed: (i) which is the physiological response – in terms
55 of temporal oscillations (or temporal cycles) – of phytoplankton to global warming/climate
56 change on both global and regional scales? (ii) which are the main drivers of the
57 phytoplankton changing and of the physiological signal oscillations? Summarizing,
58 **PhysioGlob is a multi-disciplinary project that wants to study the inter-annual**
59 **physiological response of phytoplankton to global warming using long-term satellite**
60 **observations (*i.e.* entire ESA OC-CCI time-series) through the use of Chl: C_{phyto} ratio.**

61 The key parameter in PhysioGlob is the phytoplankton carbon biomass proxy, C_{phyto} , based
62 on the particulate backscattering coefficient (b_{bp} , λ). b_{bp} is related to the particles
63 concentration in seawater, on their size distribution, refractive index, shape and structure
64 [Twardowski *et al.*, 2001; Slade and Boss, 2015]. Former research suggested that b_{bp} is mostly
65 influenced by submicron non-algal particles [Stramski *et al.*, 2004]. However, it has been
66 recently shown that most of b_{bp} is due to particles with equivalent diameters between 1 and
67 10 μm [Organelli *et al.*, 2018], thus including the contribution of phytoplankton cells and
68 supporting thus the use of b_{bp} for C_{phyto} computation [Behrenfeld *et al.*, 2005; Martinez-
69 Vicente *et al.*, 2017]. One of the most used and applied algorithm for b_{bp} retrievals from space
70 is the Quasi-Analytical Algorithm (QAA), that is also the selected algorithm in the ESA
71 Ocean Color (OC) – Climate Change Initiative (CCI) framework. The first critical issue is
72 thus to evaluate the efficiency of the QAA-based b_{bp} retrievals over the global ocean by using
73 a large in-situ dataset. In the first year, **PhysioGlob demonstrated that high efficiency of**
74 **QAA in b_{bp} detection by using both in-situ and satellite observations. In addition,**

75 **another important result found is the necessity to include the Raman Scattering**
76 **correction on ESA OC-CCI R_{rs} prior the application of QAA [Pitarch et al., 2020]. C_{phyto}**
77 **is also strictly connected to a specific coefficient, the *background* contribution of non-algal**
78 **particles (NAP) to the total b_{bp} (hereafter b_{bp}^k), historically defined as a single constant value**
79 **[Behrenfeld et al., 2005]. **PhysioGlob demonstrated that b_{bp}^k varies both in space and****
80 **time on a global scale by using a large in-situ dataset [Bellacicco et al., 2019]. It derives**
81 **that the inclusion of its spatio-temporal variability is mandatory in order to refine C_{phyto} from**
82 **space. During the second year, a refined algorithm of C_{phyto} from space was developed.**
83 **The C_{phyto} algorithm, based on satellite Chl and b_{bp} , shows higher accuracy in respect**
84 **to the available in-situ C_{phyto} global database either if some caveats are found and have**
85 **to be taken into account for its correct use [Bellacicco et al., 2020]. After these steps, it**
86 **was, consequently, possible to go through the multi-annual physiological responses (i.e.**
87 **trends and other oscillatory modes) of phytoplankton to global warming using long-term**
88 **satellite data. The database of Chl, b_{bp} and C_{phyto} is surely suitable for mean and variability**
89 **studies, Earth System Model (ESM) assessment, as well as for detecting extreme climate**
90 **events. Several peer-reviewed papers have used Chl within the ESA CCI Ocean Colour**
91 **database in connection with ESM. However, to define linear and not-linear trends as well**
92 **as multi-year oscillatory components, the use of satellite-derived Chl probably suffers**
93 **from differences in sensor combination and related-processing along the entire time-**
94 **series. Indeed, it has already been highlighted how time series resulting from merged R_{rs}**
95 **measurements might be slightly impacted by changes in the combination of operating**
96 **satellites from which the Chl has been derived [Garnesson et al., 2019]. In addition, it has**
97 **been argued that changes in the optical properties of the sea due to climate change might**
98 **impact the performances of retrieval ocean colour algorithms along the time series**
99 **[Dutkiewicz et al., 2019]. All these critical features, on Chl together with the reported**
100 **caveats of satellite C_{phyto} , partially prevented the use of these satellite data to observe**
101 **and characterize the inter-annual oscillations of a derived parameter as the Chl: C_{phyto}**
102 **with sufficient accuracy and efficiency.**

103 2 OBJECTIVES AND WORKPLAN

104 Oceans constitute nearly the 70% of Earth's surface and hold an important population of
105 aquatic microscopic plant-like organism, the so-called phytoplankton. Phytoplankton
106 organisms are considered to be responsible for approximately 50% of the planetary primary
107 production and are at the base of the trophic chain [Behrenfeld et al., 2006, 2015]. They also
108 capture carbon dioxide which dissolve into the sea from the atmosphere through the well-
109 known photosynthesis process; these organisms are thus fundamental actress in the global
110 carbon cycle. In addition, phytoplankton are defined as sentinels of changes in the ocean
111 because they can rapidly respond to environmental perturbations. Several factors such as
112 climate, ocean circulation and anthropogenic activities, can affect the phytoplankton biomass
113 and its spatio-temporal distribution [Behrenfeld et al., 2015] and due to their relevant role in
114 the ocean is important to complement their knowledge (e.g., distribution, ecology,
115 physiology). With the ongoing climate change, phytoplankton dynamics are thus
116 consequently expected to change.

117 Light, nutrients and temperature are the most important variables that lead phytoplankton
118 production and define the so-called “Integrated Growth Environment” [Behrenfeld *et al.*,
119 2002, 2008]. Phytoplankton organisms respond to fluctuations in light, nutrients and
120 temperature with different physiological strategies which aims to increase the efficiency of
121 light capturing, photosynthetic capacity, growth and persistence [Geider *et al.*, 1997, 1998;
122 MacIntyre *et al.*, 2002]. In such a context, there are two different types of phytoplankton
123 responses to environmental variations: (i) photoadaptation and (ii) photoacclimation. The
124 former describes the changes that might happen at genotype level, and are expected to occur
125 at a long evolutionary time-scale [Moore *et al.*, 2006]. The latter identifies the “short-term”
126 phenotypic response to changes in irradiance conditions at cellular level [Moore *et al.*, 2006].
127 This process consists mainly in the regulation of intra-cellular pigment amounts (*e.g.*,
128 chlorophyll-*a*) and other components of the photosynthetic machinery (*e.g.*, electron
129 transport chain, photosystem I and II concentration, and their efficiency) in order to optimize
130 the photosynthesis and growth [Dubinsky and Stambler, 2009].

131 The most important and easily observable effect due to photoacclimation is the variation
132 (*e.g.*, increase or reduction) of the main intra-cellular photosynthetic pigment concentration,
133 the chlorophyll-*a* (hereafter Chl; mg m⁻³). The photoacclimation-induced cellular changes
134 can be observed on the time scale ranging from minutes to seasons [MacIntyre *et al.*, 2000;
135 Halsey and Jones, 2015], and can therefore be detected and quantified using space-borne
136 observations. This physiological adaptation can be accounted in terms of variation of the ratio
137 between Chl and phytoplankton carbon (C_{phyto}), the well-known Chl:C_{phyto} ratio [MacIntyre
138 *et al.*, 2002], a variation that is consistent with the physiological adjustment to environmental
139 factors. It is generally accepted that Chl:C_{phyto} is also an important property in evaluating
140 phytoplankton community dynamics and its seasonality [Laws and Bannister, 1980; Halsey
141 and Jones, 2015]. Despite the natural phytoplankton population variability (*e.g.*, health of the
142 cells, life stages and so on), here, all phytoplankton cells are assumed to behave in the same
143 way. In order to optimize photosynthesis, phytoplankton cells need to produce an amount of
144 Chl that strictly depends on nutrient concentration, light availability and temperature. Over
145 the annual cycle, changes on these parameters have an impact on the phytoplankton
146 abundance and on the cellular Chl concentration in seawater. However, Chl and C_{phyto} do not
147 covary always over the different temporal cycles (*e.g.*, seasonal, year) [Behrenfeld *et al.*,
148 2004, Bellacicco *et al.*, 2016]. Indeed, even if, historically, in satellite oceanography, Chl has
149 been used as the only available proxy of algal biomass concentration [Siegel *et al.*, 2013], its
150 retrieval does not take into account the physiological adjustments that phytoplankton
151 undergoes in response to changes in environmental conditions. The use of satellite Chl alone
152 does not allow to distinguish between an intracellular and community dynamics variations
153 (Figure 1) [Bellacicco *et al.*, 2016]. For this ambiguity, Chl cannot be considered *always* the
154 best indicator for phytoplankton biomass concentration [Behrenfeld *et al.*, 2008; Bellacicco
155 *et al.*, 2016].

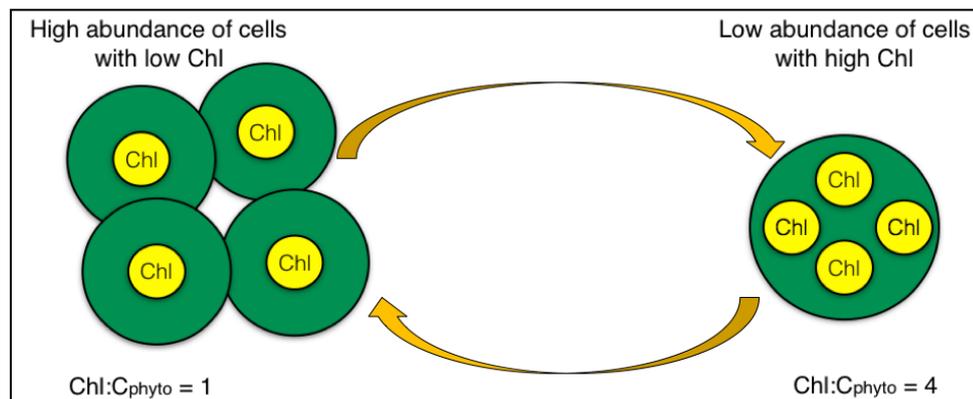
156 Behrenfeld *et al.*, [2005] developed an alternative algal biomass concentration proxy,
157 C_{phyto} (mg m⁻³);, trying to exclude therefore the physiological effects that can be find in the
158 use of Chl. The C_{phyto} is based on the particulate backscattering coefficient, *b*_{bp} (λ) which is
159 related to the particles concentration in seawater, on their size distribution, refractive index,

160 shape and structure [Twardowski et al., 2001; Slade and Boss, 2015; Organelli et al., 2018].
161 Former research suggested that b_{bp} is mostly influenced by submicron non-algal particles
162 [Morel and Ahn, 1991; Stramski and Kiefer, 1991; Stramski et al., 2004]. However, it has
163 been recently shown that most of b_{bp} is due to particles with equivalent diameters between 1
164 and 10 μm [Organelli et al., 2018], thus including the contribution of phytoplankton cells
165 and supporting the use of b_{bp} for the retrieval of: (i) particulate organic concentration (POC)
166 [Loisel et al., 2001; Thomalla et al., 2017], (ii) particle size distribution (PSD) [Kostadinov
167 et al., 2009, 2016] and (iii) C_{phyto} [Behrenfeld et al., 2005; Bellacicco et al., 2016; Martinez-
168 Vicente et al., 2017] that, as previously mentioned, is the key parameter for phytoplankton
169 physiology studies. It is thus clear that efficiency in the b_{bp} retrievals from space is a crucial
170 issue for ocean biology and global ocean carbon estimations.

171 Nowadays, in the European Space Agency (ESA) Ocean Colour (OC) - Climate Change
172 Initiative (CCI) framework, the Quasi-Analytical Algorithm (QAA) [Lee et al., 2002, 2013,
173 2014] is the selected algorithm to retrieve b_{bp} . Specifically, the ESA OC-CCI project aims at
174 creating a long-term, consistent, uncertainty-characterized time-series of ocean colour
175 products for use in climate-change studies [Brewin et al., 2015; Sathyendranath et al., 2019].

176 In such a context, while in the case of Chl the uncertainties are fully provided, the satellite
177 b_{bp} products lack such information (see also the Product User Guide on [http://www.esa-](http://www.esa-oceancolour-cci.org)
178 [oceancolour-cci.org](http://www.esa-oceancolour-cci.org)) that is also crucial for POC and C_{phyto} . This absence of statistical
179 assessment is influenced by the paucity of a sufficient number of in-situ observations for the
180 determination of uncertainties. It derives that a validation and, in case, an “update” of QAA
181 can be potentially impactful in order to enhance the performance of b_{bp} detection from space
182 and the resulting C_{phyto} computation.

183



184

185 **Figure 1:** Example of the ambiguity in the use of satellite Chl to detect phytoplankton abundance in seawater.
186 On the left, high abundance of cells with low Chl content; on the right, low abundance of cells but with the
187 same Chl concentration of the left conditions. It derives that by using satellite Chl, we cannot distinguish
188 between the left and right situations which have a different community dynamics and related description (4 cells
189 vs 1 single cell).

190 As mentioned above, the temperature is one of the main environmental forcing that affects
191 phytoplankton production. It has been well documented that ocean warming, mostly due to
192 anthropogenic activities, determines an expansion of the low-Chl and low-productivity areas,

193 such as the oligotrophic gyres [Martinez et al., 2009; Boyce et al., 2010]. Satellite
194 observations revealed an inverse relationship between inter-annual and inter-decadal changes
195 in Sea Surface Temperature (SST) and surface phytoplankton Chl concentration [Behrenfeld
196 et al., 2015]. It implies that Chl tends to decrease when temperature increases, and *viceversa*.
197 Bellacicco et al., [2016] showed the impact of SST onto phytoplankton seasonality in the
198 Mediterranean Sea. Since Chl is used for phytoplankton biomass, with the ongoing ocean
199 warming, the increase of SST can determine a continuously decrease of phytoplankton stocks
200 and productivity in different areas of the ocean. Due to the explained impact of physiological
201 processes on Chl space-borne estimations, a deeper study on the relationship between both
202 Chl and the C_{phyto} with SST is needed. An analysis of trend and oscillatory modes of
203 parameters along multi-years of satellite observations should be an important step for the
204 comprehension of the future ocean productivity.

205 Additionally, the estimation and evaluation of the weight and changes of phytoplankton
206 physiological signal, $\text{Chl}:C_{\text{phyto}}$, based on long-term satellite observations may be decisive in
207 respect to one of the future research ESA challenges which is to study the response of the
208 marine ecosystem and associated ecosystem services to natural and anthropogenic changes
209 (*i.e.* global warming). Even if we are able to distinguish changes in phytoplankton biomass
210 from changes in physiology, we currently do not know the inter-annual variability of
211 physiological process, and we are also unable to describe how and where those signals have
212 been increased or decreased in the different areas of the ocean.

213 **The main scientific goal of PhysioGlob is thus to define the inter-annual oscillations**
214 **of physiological response of phytoplankton to global warming using long-term satellite**
215 **observations.** To this aim, some others **sub-goals are defined:**

- 216
- 217 1. To infer satellite retrieval uncertainties, and possible source of errors, on b_{bp} of the
218 satellite ESA OC-CCI time-series and, if possible, define an updated version of the QAA
219 by using in-situ data on a global scale;
 - 220 2. To develop/refine a b_{bp} -based phytoplankton carbon (C_{phyto}) satellite product based on the
221 updated version of QAA;
 - 222 3. To extract and study trend and oscillations of the physiological signal ($\text{Chl}:C_{\text{phyto}}$) and its
223 relationship with different physical forcing (*e.g.*, SST, nutrients, light) on a global ocean
224 scale.

225

226 Originally, three tasks were planned to achieve the main goals. After the first year, the
227 tasks were expanded for a total of **four tasks**. Below some details for each task with the
228 background and the gaps to be filled. **Figure 2 overviews the work plan and schedule of**
229 **the planned reports, ESA-meeting and expected scientific publications.** Table 1 reports
230 the main symbols and acronyms in PhysioGlob.

231 **Task #1:** *How accurately can be retrieved the b_{bp} from space on a global scale with the*
232 *QAA? Which are the major sources of errors in this retrieval?*

233 The QAA is a multi-level algorithm that concatenates a sequence of empirical, analytical
 234 and semi-analytical steps to retrieve spectral total non-water light absorption and particulate
 235 backscattering (a_{nw} and b_{bp}) and to decompose a_{nw} into its CDOM, algal and non-algal
 236 contributions [Lee *et al.*, 2002, 2014]. Specifically, on b_{bp} , some studies suggested some
 237 degree of b_{bp} overestimation by the QAA [Huot *et al.*, 2008; Brewin *et al.*, 2012], but their
 238 reference b_{bp} data were sub-products of Chl measurements. QAA estimations from satellite
 239 R_{rs} showed a bias of +16.4% with respect to in-situ b_{bp} for the Adriatic Sea [Melin *et al.*,
 240 2011]. Using the in-situ NOMAD dataset [Werdell and Bailey, 2005], a b_{bp} overestimation
 241 of +38% by the QAA with respect to the observed value was reported [Werdell *et al.*, 2013].
 242 Other results, based on in-situ matchups, showed a bias of +2.5–8.8% for the QAA-derived
 243 b_{bp} in Arctic waters, and of +9.5% to +16.4% in low-latitude waters [Zheng *et al.*, 2014].
 244 Pitarch *et al.* [2016] reported a slight underestimation within 10% in the Mediterranean Sea.
 245 Most recently, QAA-derived b_{bp} from different satellite sensors (*i.e.* MODIS, VIIRS, OLCI)
 246 showed good performance with respect to a large in-situ b_{bp} dataset collected from
 247 Biogeochemical (BGC)-Argo floats [Bisson *et al.*, 2019].

248 Nowadays, the uncertainties associated to QAA-based b_{bp} retrievals globally are poorly
 249 known. In order to provide a best-effort b_{bp} uncertainty assessment, **this task aims at**
 250 **evaluating the efficiency of QAA for global b_{bp} retrievals by using a large database of**
 251 **correspondent in-situ R_{rs} and b_{bp} data on a global scale.**

252 **Task #2:** *Is the background particulate backscattering coefficient variable in space and*
 253 *time? Or is it a constant value? Which is the best method for its computation?*

254 Here, the b_{bp} -based phytoplankton biomass model is presented. As mentioned in Section
 255 1, Behrenfeld *et al.* [2005] developed an alternative algal biomass concentration index, in
 256 terms of phytoplankton carbon, C_{phyto} (mg m^{-3}). C_{phyto} is based on the b_{bp} as follows:
 257

$$258 \quad C_{phyto} = [b_{bp}(\lambda) - b_{bp}^k(\lambda)] \cdot SF \quad [\text{Eq. 1}]$$

259 where λ is the wavelength, C_{phyto} being the phytoplankton carbon biomass expressed in mg
 260 C m^{-3} ; $b_{bp}(\lambda)$ is the particulate backscattering coefficient; $b_{bp}^k(\lambda)$ is the *background*
 261 contribution of non-algal particles (NAP) to the total b_{bp} , while SF is the scaling factor
 262 ($13,000 \text{ mg C m}^{-2}$) chosen to give satellite Chl: C_{phyto} values (average value of 0.010)
 263 consistent with laboratory results, and an average contribution of phytoplankton to total
 264 particulate organic carbon (around 30%) to be consistent with field estimates [Behrenfeld *et*
 265 *al.* 2005]. **The main issue in Eq. 1 is the computation of b_{bp}^k .**

266 In the original work, b_{bp}^k was estimated as fraction of the b_{bp} that does not covary with Chl,
 267 and estimated it as the *offset* of a linear regression between satellite-derived b_{bp} and Chl (*i.e.*,
 268 5 years of ocean color data) when Chl concentrations were greater than 0.14 mg m^{-3} . This
 269 offset was defined as the *background* of the backscattering due to NAP and refers only to a
 270 fraction of the total b_{bp} signal caused by NAP that thus does not covary with Chl (*i.e.*
 271 phytoplankton).
 272

273 In *Behrenfeld et al.*, [2005] (hereafter Be05), b_{bp}^k is assumed to be a constant value both
274 on space and time scales (*i.e.* $3.5 \cdot 10^{-4} \text{ m}^{-1}$). Be05 attributed it to “*a stable heterotrophic and*
275 *detrital component of the surface particle population and therefore independent of the*
276 *phytoplankton dynamics*”. Recently, *Bellacicco et al.* [2016] (hereafter Blc16) applied
277 Be05’s approach for distinct bioregions and seasons in the Mediterranean Sea, and showed
278 that b_{bp}^k has instead a marked regional and seasonal variability. Such a result thus confirmed
279 that the heterotrophic and detrital components at the sea surface are neither negligible nor
280 stable, but highly variable in seawater [*Siokou-Frangou et al.*, 2010]. The variability of the
281 b_{bp}^k by Blc16 was also later confirmed by *Bellacicco et al.*, [2018] for the global ocean
282 (hereafter Blc18). Indeed, Blc18 highlighted two distinct oceanic areas: the productive sub-
283 polar North Atlantic Ocean, where b_{bp}^k and particle biomass (*i.e.* phytoplankton cells) are
284 anti-correlated; and the Southern Ocean, where b_{bp}^k signal is mainly driven by inorganic
285 particles, such as algal coccoliths [*Balch et al.*, 2016, 2018], bubbles or foam that occur in
286 the stormy seas [*Stramski et al.*, 2004].

287 **Task #2 aims at evaluating which is the best method for b_{bp}^k estimations and if b_{bp}^k has**
288 **a significant spatio-temporal variability that has to be accounted for in the C_{phyto}**
289 **computation from space. To this specific aim, a large in-situ database of correspondent**
290 **b_{bp} and Chl measurements is used.**

291 **Task #3: Definition of a refined b_{bp} -based phytoplankton carbon biomass and production**
292 **of daily satellite phytoplankton carbon biomass time-series (1997-2019).**

293 In this task, the aim is to refine daily C_{phyto} satellite products taking into account the results
294 found in Tasks #1 and #2: the possible correction on QAA for satellite b_{bp} retrievals together
295 with the spatial and temporal b_{bp}^k variability. A comparison with the previous *Behrenfeld et*
296 *al.*, [2005], *Bellacicco et al.*, [2018] approaches is planned. Another semi-empirical
297 algorithm for C_{phyto} can be evaluated. This started from the relationship between in-situ flow
298 cytometry-based carbon and b_{bp} at 470nm [*Martinez-Vicente et al.*, 2013, 2017].

299 **Task #3 has the goal to produce daily C_{phyto} satellite maps, from 1997 to 2019, based**
300 **on the “revised” QAA for b_{bp} retrievals coupled with the correct method for**
301 **b_{bp}^k estimations and, if exists, its spatio-temporal variations.**

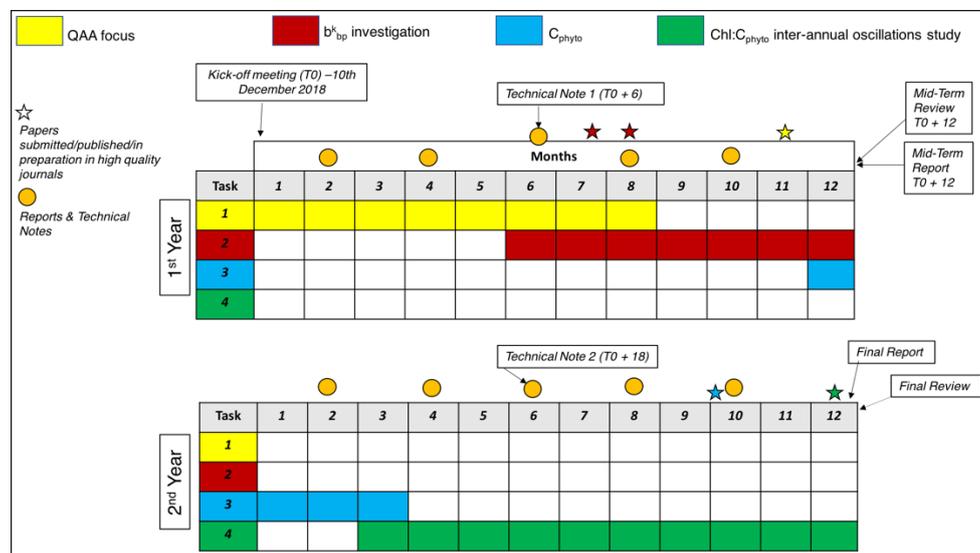
302 **Task #4: This task is referred to address the question of possible shared mechanisms of**
303 **variability between C_{phyto} , $\text{Chl}:C_{\text{phyto}}$, SST (and others physical forcings) as derived from**
304 **multi-years ESA OC-CCI data time-series ranging from seasonal to inter-annual scales.**

305 To identify variability in this dataset, we rely here on multi-channel (or multivariate)
306 singular spectrum analysis (M-SSA), an advanced method for the analysis of spatio-temporal
307 datasets which provides an efficient and robust tool to extract dynamics from relatively short,
308 noisy time-series [*Ghil et al.*, 2002]. **The goal is to classify components of the single (or**
309 **coupled) time-series into trends, oscillatory modes, and noise.** Having in mind that, by
310 using M-SSA, the trends do not need to be linear and that the oscillations can be amplitude-
311 and phase- modulated.

312 As a first analysis, a single time-series could be studied in separate M-SSA analyses to
 313 evaluate similarities among the inter-annual variabilities of the Chl:C_{phyto}, SST, and other
 314 parameters such as mixed layer depth (MLD), light availability. Then, we will proceed with
 315 a joint M-SSA analysis of the time-series to better understand the spatio-temporal structure
 316 associated with inter-annual variability in the Chl:C_{phyto}, or phytoplankton indices, and global
 317 ocean temperature fields. This coupled analysis will also help in addressing the question to
 318 which extent the inter-annual oscillatory modes found in the Chl:C_{phyto} can be attributed to
 319 its response to inter-annual variability in SST fields.

320 A similar approach has already been used by *Groth et al.* [2017] to study the inter-annual
 321 variability of the ocean temperature field in association with wind stress forcing. Here, the
 322 idea is essentially to follow a similar approach but with different variables: Chl:C_{phyto} against
 323 SST, instead of SST versus wind stress.

324



325

326 **Figure 2:** Overview of the work including ESA-meetings, published (stars with same color of the related Task)
 327 and expected scientific papers (triangles), bi-monthly reports and technical notes (orange circles).

328

C _{phyto}	Phytoplankton Carbon biomass (mg m ⁻³)
Chl	Chlorophyll- <i>a</i> (mg m ⁻³)
Chl:C _{phyto}	Chlorophyll- <i>a</i> to Phytoplankton Carbon ratio
<i>b</i> _{bp}	Particulate backscattering coefficient (m ⁻¹)
<i>b</i> ^k _{bp}	<i>Background</i> particulate backscattering coefficient by non- algal particles (m ⁻¹)
R _{rs}	Remote Sensing Reflectance (sr ⁻¹)
ESA	European Space Agency
OC-CCI	Ocean Color - Climate Change Initiative
QAA	Quasi-Analytical Algorithm
SSA	Singular Spectral Analysis
M-SSA	Multi-Channel Singular Spectral Analysis

329

Table 1: Symbols and acronyms definition in PhysioGlob.

330 3 WORK PERFORMED

331 *Task #1*

332 Here, the context, method and the performed worked about Task #1 are presented. The
333 main goal of this task is to evaluate the efficiency of QAA algorithm for global b_{bp} retrievals
334 by using a large database of correspondent in-situ R_{rs} and b_{bp} data. For more details see
335 *Pitarch et al. [2020] (see Section 6)*.

336 *Scientific context*

337 Retrieval of water inherent optical properties (IOPs) from both field and ocean colour
338 radiometry is at the base of several biogeochemical and physical oceanographic studies
339 [*Behrenfeld et al. 2005; Bellacicco et al., 2016*]. IOPs of algal and non-algal particles can be
340 derived from remote sensing reflectance spectra (R_{rs} ; units of sr^{-1}) by using appropriate
341 algorithms [*Lee et al., 2002; Brewin et al., 2015; Loisel et al., 2018*]. Among IOPs, the
342 particulate optical backscattering coefficient (b_{bp} ; in m^{-1}) is related to the particle
343 concentration in seawater, on their size distribution, refractive index, shape and structure
344 [*Twardowski et al., 2001; Slade and Boss, 2015; Organelli et al., 2018*]. Former research
345 suggested that b_{bp} is mostly influenced by submicron non-algal particles [*Morel et al., 1991;*
346 *Stramski et al., 2004*]. However, it has been recently shown that most of b_{bp} is due to particles
347 with equivalent diameters between 1 and 10 μm [*Organelli et al., 2018*], thus including the
348 contribution of phytoplankton cell and supporting the use of b_{bp} for the retrieval of: (i)
349 particulate organic concentration (POC) [*Loisel et al., 2001; Thomalla et al., 2017*]; (ii)
350 particle size distribution [*Kostadinov et al., 2009, 2016*]; and (iii) phytoplankton carbon
351 biomass concentration (C_{phyto} ; $mg\ m^{-3}$) [*Behrenfeld et al., 2015; Bellacicco et al., 2018*], a
352 key parameter also for phytoplankton physiology studies [*Behrenfeld et al., 2005; Siegel et*
353 *al., 2013; Halsey and Jones, 2015*]. Efficiency in the b_{bp} retrieval is crucial for ocean biology
354 and global ocean carbon estimations.

355 On one hand, radiative transfer theory provides the link between b_{bp} and optical
356 radiometry [*Mobley, 1994*]. Therefore, inversion algorithms for b_{bp} detection from optical
357 radiometry can be developed. In particular, the QAA [*Lee et al., 2002*] is a multi-level
358 algorithm that concatenates a sequence of empirical, analytical, and semi-analytical steps to
359 retrieve spectral total non-water light absorption and backscattering (a_{nw} and b_{bp}) first and to
360 decompose a_{nw} into its CDOM, algal and non-algal contributions. Specifically, about b_{bp} ,
361 some studies suggested some degree of b_{bp} overestimation by the QAA [*Huot et al., 2008;*
362 *Brewin et al., 2012*], but their reference b_{bp} data were sub-products of chlorophyll-a (Chl)
363 measurements. QAA estimations from satellite R_{rs} showed a bias of +16.4% with respect to
364 in situ b_{bp} for the Adriatic Sea [*Melin et al., 2011*]. Using the in-situ NOMAD dataset
365 [*Werdell et al., 2005*], a b_{bp} overestimation of +38% by the QAA with respect to the observed
366 value was reported [*Werdell et al., 2013*]. Other results, based on in situ matchups, showed
367 a bias of +2.5–8.8% for the QAA-derived b_{bp} in Arctic waters, and of +9.5% to +16.4% in
368 low-latitude waters [*Zheng et al., 2014*]. *Pitarch et al. [2016]* reported a slight
369 underestimation within 10% in the Mediterranean Sea. Most recently, QAA-derived b_{bp} from
370 different satellite sensors (i.e., MODIS, VIIRS, OLCI) showed good performance with

371 respect to a large in situ b_{bp} dataset collected from biogeochemical (BGC)-Argo floats
372 [Bisson *et al.*, 2019].

373 Currently, in the European Space Agency (ESA) Ocean Colour (OC) Climate Change
374 Initiative (CCI), QAA is the selected algorithm to retrieve b_{bp} . Specifically, the ESA OC-CCI
375 project aims at creating a long-term, consistent, uncertainty-characterized time series of
376 ocean colour products, for use in climate-change studies [Sathyendranath *et al.*, 2020]. In
377 such a context, while in the case of Chl the uncertainties are fully provided, the b_{bp} satellite
378 products lack such information that is also crucial for POC and C_{phyto} estimations [Bellacicco
379 *et al.*, 2019]. This absence of statistical assessment is influenced by the paucity of a sufficient
380 number of in situ observations for the determination of uncertainties.

381 Nowadays, the uncertainties associated to QAA-based b_{bp} retrievals globally are not
382 known. In order to provide a best-effort b_{bp} uncertainty assessment, this work aims at
383 evaluating the efficiency of QAA for global b_{bp} retrievals by using a large database of
384 corresponding in situ R_{rs} and b_{bp} data ($N = 2881$). In details, we use the updated version of
385 the recent in situ global bio-optical dataset [Valente *et al.*, 2019] together with field
386 measurements from the BOUSSOLE buoy [Antoine *et al.*, 2006] and two different
387 oceanographic cruises in the Tyrrhenian and Adriatic Seas [Pitarch *et al.*, 2016]. Unlike
388 previous studies, here, the QAA performance is considered at multiple bands that further
389 allow the evaluation of the b_{bp} spectral slope retrievals against in situ measurements. The
390 goals of this paper are thus: (i) to define the accuracy of the QAA for b_{bp} retrievals using only
391 in situ R_{rs} data; (ii) to validate the CCI R_{rs} with in situ corresponding data; and (iii) to evaluate
392 the performance of the QAA using satellite CCI R_{rs} as input data.

393 Methods

394 *Assessment of the Quasi-Analytical Algorithm (QAA)*

395 The original algorithm [Lee *et al.*, 2002] has undergone many updates and developments
396 by several researchers. The QAA version here used is based on the algorithm for the CCI
397 bands which is currently integrated in the SeaWiFS data analysis system (SeaDAS) [Lee *et al.*,
398 2014].

399 The sub-surface R_{rs} (named r_{rs}) is calculated as $r_{rs} = R_{rs}/(0.52 + 1.7R_{rs})$ and modeled
400 as a function of the IOPs as: $r_{rs} = g_0u + g_1u^2$, with $u = b_b/(a + b_b)$, $g_0 = 0.089$ and $g_1 =$
401 0.1245 . This approach provided good results in the Mediterranean Sea in case of oligotrophic
402 and coastal waters [Pitarch *et al.*, 2016; Bracaglia *et al.*, 2019].

403 The QAA uses an empirical inversion of R_{rs} to retrieve absorption and then it solves total
404 backscattering (b_b) analytically. b_{bp} is calculated by subtraction of pure seawater
405 backscattering (b_{bw}) for an average temperature of 14 °C and an average salinity of 38 PSU
406 [Zhang *et al.*, 2009]. b_{bp} is first estimated at a reference wavelength of $\lambda = 555$ nm and then
407 the calculation is extended to other wavelengths by assuming a power law $b_{bp} =$
408 $b_{bp}(\lambda_0)(\lambda/\lambda_0)^{-\eta}$ for the b_{bp} with a spectral slope:

409

$$\eta = p_1 \left[1 - p_2 \exp \left(-p_3 \frac{r_{rs}(443)}{r_{rs}(555)} \right) \right]$$

410 Equation above is widely used for QAA retrievals of b_{bp} at multiple wavelengths.
 411 Nevertheless, we use the in situ dataset presented here to evaluate the accuracy of analytical
 412 η . The functional form of Equation above is used and the default numerical coefficients $p1 =$
 413 2.0 , $p2 = 1.2$ and $p3 = 0.9$ [Lee *et al.*, 2014] are replaced by unknown variables established
 414 by non-linear regression. To this aim, we used the iterative bi-square method, which
 415 minimizes a weighted sum of squared errors, where the weight given to each data point
 416 decreases with the distance from the fitted curve [Huber and Ronchetti, 1981]. This
 417 procedure makes the curve sensitive to the bulk of the data and the effect of outliers is
 418 reduced. The error function is minimized through the trust region algorithm [Moré *et al.*,
 419 1983]. In addition, the 95% confidence prediction bounds are also computed.

420 It is known that for oligotrophic waters the Raman scattering plays a significant role that
 421 is not accounted for in the semi-analytical R_{rs} modelling. Therefore, a pertinent question is
 422 how much this phenomenon affects the semi-analytical b_{bp} retrievals. With this aim, Lee *et*
 423 *al.* [2013] developed empirical correction formulas to compensate the Raman scattering on
 424 the R_{rs} . Here, we assess the effects of this compensation on the difference between the in situ
 425 b_{bp} and R_{rs} -derived b_{bp} . The statistical assessment was also replicated in other different cases:
 426 (i) validation of satellite CCI R_{rs} against in situ R_{rs} ; and (ii) b_{bp} retrievals after the application
 427 of QAA to satellite R_{rs} (Raman corrected), were compared to in situ measurements at the
 428 different available wavelengths.

429 Estimated data y_i are compared to reference data x_i by using the following statistical
 430 indexes: relative bias (units of %), relative root-mean square error (RMS, units of %) and
 431 determination coefficient (r^2)

$$bias = 100 \frac{1}{N} \sum_{i=1}^N \frac{y_i - x_i}{x_i}$$

$$RMS = 100 \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{y_i - x_i}{x_i} \right)^2}$$

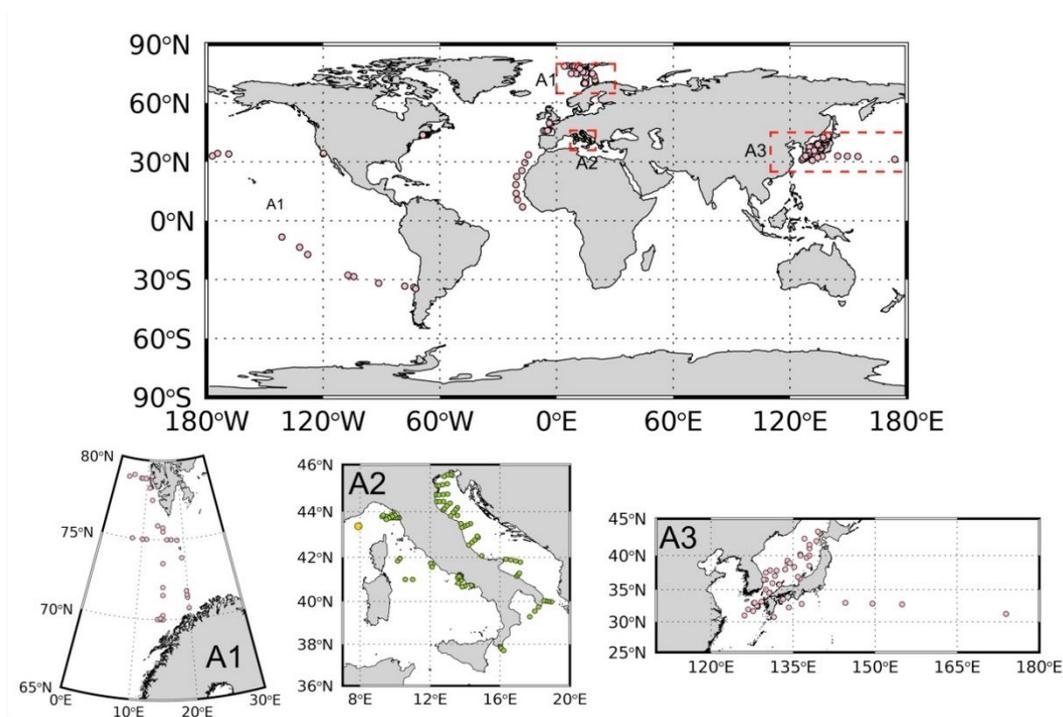
$$r^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$$

432 Data

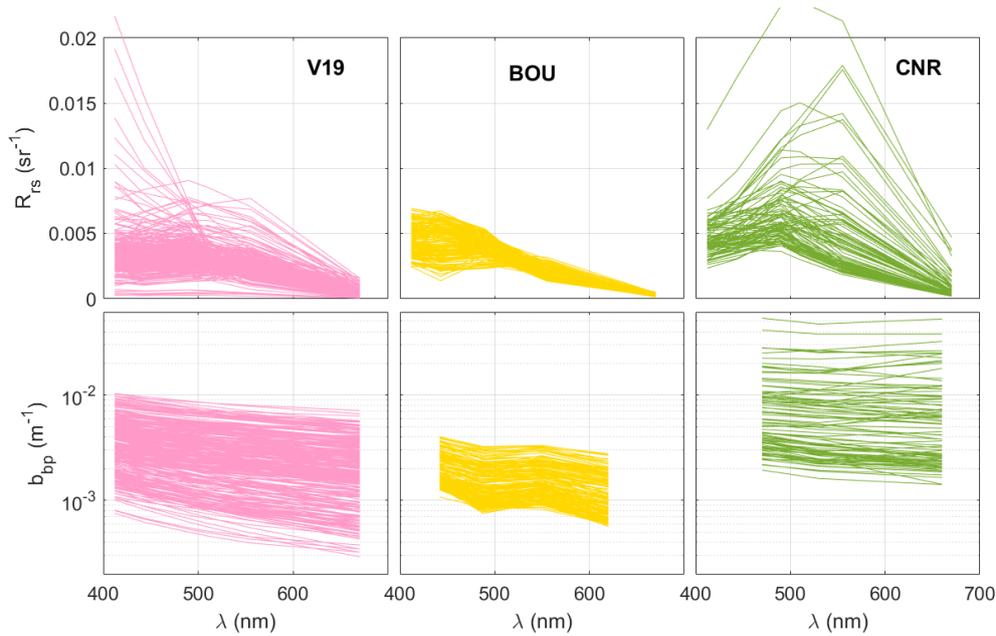
433 *In Situ Data*

434 The in situ database is composed of three distinct datasets containing multi-spectral R_{rs}
 435 and b_{bp} : the recently updated global in situ database [Valente *et al.*, 2019] (hereafter V19
 436 dataset), an in situ dataset collected by the National Research Council (CNR) during two field
 437 campaigns in the Mediterranean Italian Sea (hereafter CNR dataset) [Pitarch *et al.*, 2016]
 438 and the time-series of data acquired by the BOUSSOLE buoy in the northwestern
 439 Mediterranean Sea (hereafter BOU dataset) [Antoine *et al.*, 2006, Antoine *et al.*, 2008;
 440 hereafter. The three in situ databases were quality-checked as described below. All the R_{rs}

441 data were band-shifted to the CCI bands (those of the NASA SeaWiFS instrument, namely
 442 412, 443, 490, 510, 555, and 670 nm). The band-shifting procedure [Melin *et al.*, 2015] is a
 443 technique to compensate small band differences in multispectral R_{rs} data. It takes into account
 444 the spectral shape of the absorption and scattering that contribute to R_{rs} and constitutes a
 445 more accurate approach than a simple linear interpolation. Considering every wavelength an
 446 independent measurement, the final dataset accounts for a total of $N = 2881$ R_{rs} and b_{bp} co-
 447 located measurements around the global ocean (Figure 3). As shown in Figure 4, the total R_{rs}
 448 and b_{bp} spectra cover from oligotrophic open-ocean to more eutrophic coastal waters as the
 449 range of R_{rs} and b_{bp} values vary between $0-0.02 \text{ sr}^{-1}$ and $10^{-4}-10^{-1} \text{ m}^{-1}$ respectively.
 450



451
 452 **Figure 3:** Geographical distribution of the in situ R_{rs} vs. b_{bp} matchups. Some areas (A1, A2, and A3) concentrate
 453 a high point density and are highlighted in zoomed maps. Pink, yellow, and green dots refer to V19, BOU, and
 454 CNR data, respectively.



455

456 **Figure 4:** R_{rs} and b_{bp} spectra for the three-different datasets: V19, BOU, and CNR. Pink, yellow, and green
 457 lines refer to V19, BOU, and CNR data, respectively.

458 *V19 Dataset*

459 R_{rs} and IOPs, aggregated within ± 6 nm, were downloaded. V19 is a global compilation of
 460 in situ data that was acquired from many sources (e.g., MOBY, AERONET-OC, SeaBASS,
 461 NOMAD, MERMAID, AMT, and many others), motivated by the validation of the ocean-
 462 colour products from the ESA OC-CCI products. Methodologies were implemented for
 463 homogenization, quality control and merging of all data. No changes were made to the
 464 original data, other than averaging of observations that were close in time and space,
 465 elimination of some points after quality control and conversion to a standard format [*Valente*
 466 *et al.*, 2019].

467 In this study, data were selected only if valid and corresponding R_{rs} and b_{bp} measurements
 468 at all CCI bands were available. Such condition determines a total of $N = 319$ matchups.
 469 Remaining minor b_{bp} wavelength mismatches were removed by linear interpolation to the
 470 CCI bands. Although V19 is a merged dataset from multiple datasets, the condition we set
 471 for the matchup left data that were originally from the NOMAD dataset only.

472 *BOU Dataset*

473 The BOUSSOLE (BOUee pour l'acquiSition d'une Série Optique a Long termE) project
 474 started in 1999, and its activities are developed on a site located in the northwestern
 475 Mediterranean Sea ($7^{\circ}54'E$, $43^{\circ}22'N$, Figure 3, panel A2). Essential information about the
 476 site characteristics, the measurement platform, and the instrumentation was previously
 477 documented [*Antoine et al.*, 2006; *Antoine et al.*, 2008 a, b]. The b_{bp} data were collected at 9
 478 m nominal depth with a Hobilabs Hydrosat-4 (442, 488, 550, and 620 nm) and processed as
 479 in *Antoine et al.* [2011]. In addition, a quality control on b_{bp} was applied that required a
 480 spectral b_{bp} slope, calculated from every pair of two consecutive bands, within given bounds
 481 (more than -1 and less than 6). R_{rs} data were derived with a set of Satlantic 200-series

482 multispectral radiometers [Organelli *et al.*, 2016 and references therein]. The R_{rs} is available
483 at a varying number of the following bands, depending on the time period: 412.5, 442.5, 490,
484 510, 555, 560, 665, 670, and 681.25 nm. Since the application of the QAA requires R_{rs} at
485 443, 490, 555, and 670 nm, only R_{rs} records whose native bands matched those needed by
486 the QAA algorithms were selected (within a ± 6 nm range). Data at 412.5 nm and 442.5 nm
487 were band-shifted to 412 nm and 443 nm [Melin *et al.*, 2015], respectively. In the green
488 region, if the R_{rs} at 555 nm was available, it was directly sampled and the R_{rs} at 560 nm was
489 ignored. If the R_{rs} at 560 nm was available when the R_{rs} at 555 nm was missing, the R_{rs} at
490 560 nm was band-shifted to 555 nm. Similarly, in the red region, between the R_{rs} at 665 nm
491 and 670 nm, preference to R_{rs} at 670 nm was given. R_{rs} data at 681.25 nm was not considered
492 for the analysis. Data was generally available within two hours from the local noon. The time
493 series at sub-daily resolution were reduced by calculating the daily medians.

494 *CNR Dataset*

495 Data belong to two field campaigns conducted in 2013 and 2015 in Italian seas,
496 encompassing a high optical range between open and coastal waters. Measurements were
497 performed between 8:30 h and 16:00 h UTC. IOPs and R_{rs} were collected sequentially at each
498 station, with a maximum delay of ~ 1 h and ship drift of maximum few hundreds of meters.

499 Backscattering (b_b) was measured with an ECO-VSF3, manufactured by WET Labs, Inc.,
500 at the wavelengths 470, 530, and 660 nm. This instrument measures the volume scattering
501 function at three backward angles and calculates b_b by integration of a polynomial fit. Final
502 data are the result of a binning across the first optical depth.

503 Radiometry was performed with OCR-507 radiometers, manufactured by Satlantic, Inc.,
504 measuring at the center bands 412, 443, 490, 510, 556, 665, and 865 nm. In-water upwelling
505 radiance at nadir (L_u) sensor was mounted onto a free-falling T-shaped structure, with the
506 multicast technique. Above-water downwelling irradiance (E_s) data were collected by a
507 reference sensor, mounted at the top of the ship's deck. R_{rs} was computed using the SERDA
508 software developed at CNR. All the R_{rs} data were band-shifted to the CCI bands for
509 consistency with the satellite R_{rs} . Further details about this dataset are provided in Pitarch *et al.*
510 [2016].

511 *Satellite ESA OC-CCI R_{rs} Data*

512 The ESA OC-CCI version 4.0 global daily R_{rs} data at 4 km resolution for the period 1997–
513 2017 were downloaded [Jackson *et al.*, 2019]. CCI products are the result of the merging of
514 SeaWiFS, MERIS, MODIS, and VIIRS data in which the inter-sensor biases are removed
515 [Melin *et al.*, 2016]. This version 4.0 includes the latest NASA reprocessing R2018.0 that
516 mostly accounts for the aging of the MODIS sensor. ESA OC-CCI provides the daily R_{rs} data
517 and associated uncertainty maps in terms of bias and RMS, which were generated with a
518 procedure that included comparison to in situ data and optical water type analysis [Jackson
519 *et al.*, 2019].

520 In this work, a conservative extraction procedure was followed, in which the center R_{rs}
521 data within a 3×3 pixel box was extracted only if all the 9 pixels were not flagged, therefore
522 minimizing possible land border, cloud or other environmental contaminations, and obtaining

523 the highest quality of matchups. Finally, for each single R_{rs} , the bias was also extracted and
 524 then compensated pixel-by-pixel.

525 Results

526 *QAA Performance for b_{bp} Retrievals from In Situ Data*

527 QAA-retrieved b_{bp} from in situ R_{rs} is here compared to in situ measured b_{bp} . Comparisons
 528 are made at the native bands of each in situ b_{bp} instrument for the cases of BOU and CNR
 529 and at the CCI bands for V19. Statistics are also presented for each band and dataset. A first
 530 assessment consists of applying the QAA without performing the compensation for Raman
 531 scattering. Here, results show a general overestimation of around +43.4% for V19 (Table 2)
 532 that is not significant given the overall noise expressed by the RMS (125%). This high RMS
 533 is the likely consequence of the different protocols, instrumental and geophysical noises
 534 affecting all single contributors to the V19 dataset (Table 2). In the case of the BOU dataset,
 535 an overall overestimation of +49.2% is found for all the bands which is statistically
 536 significant given the related RMS (58.7%). On the other hand, the QAA applied to the CNR
 537 dataset showed the highest performances, with a bias of +3.3% and a RMS below 23%.

538 **Table 2:** Statistical descriptors of the difference between the QAA-derived b_{bp} and in situ b_{bp} for each dataset,
 539 without Raman scattering compensation.

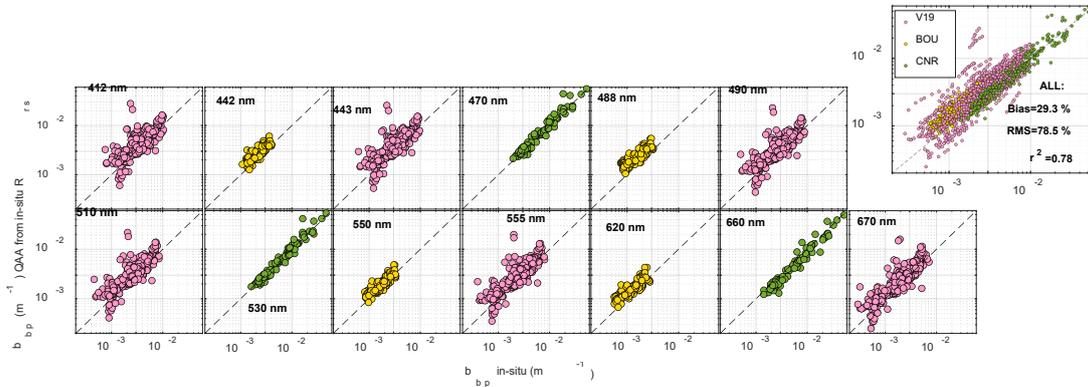
	Band (nm)	Bias (%)	RMS (%)	r^2	N
V19	412	40.3	128.4	0.35	319
	443	42.7	129.4	0.37	319
	490	44.5	127.8	0.41	319
	510	45.0	127.1	0.42	319
	555	45.2	124.2	0.44	319
	670	43.1	114.2	0.47	319
	All	43.4	125.3	0.43	1914
BOU	442	44.5	50.7	0.73	172
	488	71.3	79.2	0.73	172
	550	29.0	36.5	0.78	172
	620	52.0	60.2	0.73	172
	All	49.2	58.7	0.75	688
CNR	470	11.8	25.1	0.88	93
	530	7.7	22.8	0.89	93
	660	-9.6	20.7	0.93	93
	All	3.3	22.9	0.88	279

540 To understand the importance of the Raman scattering correction in semi-analytical b_{bp}
 541 retrievals, the analysis is repeated with corrected R_{rs} [Lee et al., 2013]. The application of the
 542 Raman scattering correction reduces both bias and RMS nearly for all the b_{bp} at all bands
 543 (Table 3 and Figure 5). Indeed, for the V19 dataset, the bias decreases to 12% with respect
 544 to the retrievals obtained without correction of the Raman scattering (Table 2). The RMS
 545 reduction is around 34%. For the BOU data, the RMS and bias improve of about 11% and
 546 12%, respectively. In the case of the CNR dataset, statistics show a modest increase in
 547 accuracy except for $\lambda = 660$ nm, which is likely influenced by chlorophyll-a fluorescence.
 548 Although fluorescence peaks at around $\lambda = 660$ nm, the ECO-VSF 3 sensor, used to collect
 549 the CNR dataset, has a full width at half maximum (FWHM) of about 20 to 30 nm, so a
 550 fluorescence interference may not be excluded.

551 Overall, these results are somewhat expected as the Raman scattering correction produces
 552 a smaller effect in coastal waters [Lee et al., 2014], which represent a significant part of the
 553 CNR dataset with respect to the two other datasets (Figure 5). The overall statistics are in
 554 agreement with previous comparisons that showed negligible biases over noise at global scale
 555 [Brewin et al., 2015] and at regional level [Pitarch et al., 2016]. Results in this section
 556 highlight the importance of applying the Raman scattering correction to the source R_{rs} prior
 557 to semi-analytical b_{bp} retrieval in order to increase the accuracy.
 558

559 **Table 3:** Statistical descriptors of the difference between the b_{bp} -QAA derived and in situ b_{bp} for each dataset
 560 with Raman scattering compensation.

	Band (nm)	Bias (%)	RMS (%)	r^2	N
V19	412	28.5	94.6	0.45	319
	443	30.7	95.0	0.47	319
	490	32.2	93.4	0.50	319
	510	32.6	92.8	0.51	319
	555	32.7	90.4	0.52	319
	670	30.7	83.1	0.54	319
	All	31.2	91.6	0.52	1914
BOU	442	33.0	40.1	0.73	172
	488	57.2	64.8	0.73	172
	550	18.2	27.1	0.78	172
	620	39.0	47.8	0.73	172
	All	37.0	47.0	0.75	688
CNR	470	6.5	22.6	0.88	93
	530	2.5	21.3	0.89	93
	660	-14.2	23.0	0.93	93
	All	-1.73	22.3	0.89	279

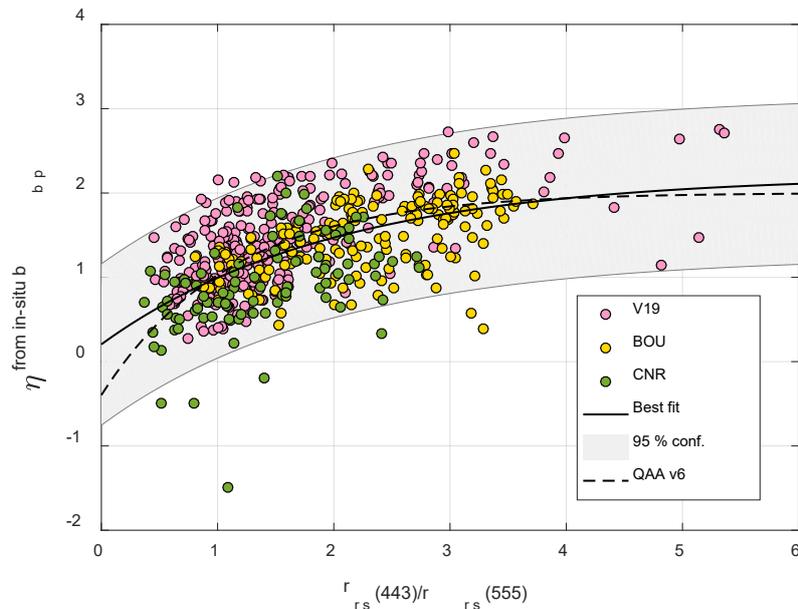


561 **Figure 5:** Scatter plots of QAA-derived b_{bp} vs. in situ b_{bp} data for each wavelength and dataset considered and
 562 for the merged dataset. Raman correction is applied to R_{rs} . The dashed line represents the 1:1 ratio. Pink, yellow
 563 and green dots refer to V19, BOU and CNR data, respectively.
 564

565 Estimation of the b_{bp} Spectral Slope from R_{rs} Data

566 The in-situ dataset described above is used to assess the proposed relationship in the QAA
 567 and perform a model to data fit that is compared to the common QAA v6 equation [Lee et
 568 al., 2014]. Figure 6 shows a comparison of the independent variable (i.e., the blue-to-green
 569 band ratio $r_{rs}(443)/r_{rs}(555)$) with respect to η derived from the in situ b_{bp} . Fitting a functional
 570 form of mother equation returns a curve ($p_1 = 2.2$, $p_2 = 0.9$ and $p_3 = 0.5$) and a 95% prediction
 571 interval, which is around ± 1 wide, caused by the high scatter of the data cloud. The difference

572 between η computed here and the one derived via QAA is much smaller than the width of
 573 the prediction interval, thus making them equivalent for prediction purposes. Therefore, by
 574 the principle of parsimony, the operational η functional form (dashed line in Figure 6)
 575 remains valid. However, one must keep in mind that the low predictive value of this
 576 relationship may result in b_{bp} extrapolations to bands outside the reference one (usually 555
 577 nm) that accumulate significant errors. In particular, within a worst-case scenario, an error in
 578 η estimation, $\Delta\eta = 1$, will lead to a $\sim 26\%$ error when extrapolating b_{bp} from 555 nm to 412
 579 nm.



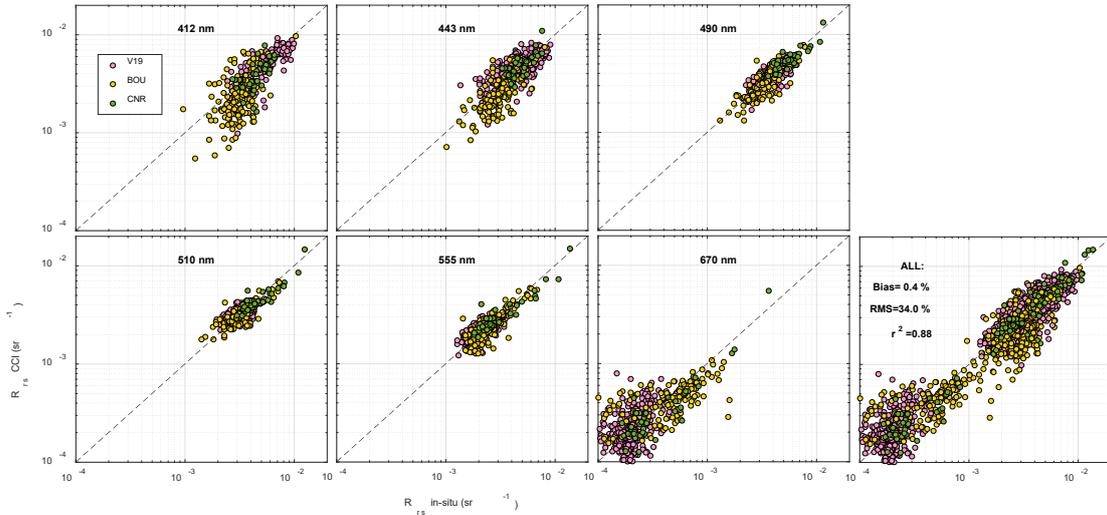
580
 581 **Figure 6:** η calculation considering all the in situ data available: V19 (pink dots), BOU (yellow dots), and CNR
 582 (green dots). The solid curve is the best fit of original equation to all the data ($p_1 = 2.2$, $p_2 = 0.9$ and $p_3 = 0.5$).
 583 The 95% confidence prediction bounds are represented by the grey shaded area. The dashed curve is the η
 584 estimation from R_{rs} as defined in original Equation. Pink, yellow, and green dots refer to V19, BOU, and CNR
 585 data, respectively.

586 *Validation of CCI R_{rs}*

587 Prior to applying to satellite data an algorithm that has been developed with in situ data,
 588 assessing the quality of the satellite R_{rs} with respect to in situ measured R_{rs} is desirable in
 589 order to identify possible biases. Therefore, this section uses the in situ R_{rs} contained in the
 590 three datasets to evaluate the CCI R_{rs} . There is a total of 882 matchups for V19, 581 for BOU,
 591 and 252 for CNR. Good agreement between in situ values and the CCI R_{rs} products is found
 592 (Figure 7, Table 4) at all wavelengths, rather consistently with other previous results [*Brewin*
 593 *et al.*, 2015]. Overall, all datasets display similar performance, with negligible biases with
 594 respect to the overall noise expressed by the RMS. In the case of $\lambda = 670$, increased RMS is
 595 mostly due to the low values R_{rs} , except for CNR, that contains a higher data range. It is
 596 concluded that the CCI R_{rs} do not require adjustments at the studied wavelengths.

597 The magnitude of this RMS expresses a high bound for the overall uncertainty of the R_{rs}
 598 product as it is a measure of the errors in the comparison experiment, including those within
 599 the in situ data. The fraction of this error which is attributable to the satellite data only is
 600 likely to be much lower. To have a measure of this fraction, a comparison to global in situ

601 dataset with a traceable uncertainty budget would be desirable, though such option is
 602 presently not available.



603
 604 **Figure 7:** Scatter plots of CCI R_{rs} versus in situ R_{rs} for the six different wavelengths. The dashed line represents
 605 the 1:1 ratio. Pink, yellow, and green dots refer to V19, BOU, and CNR data, respectively.

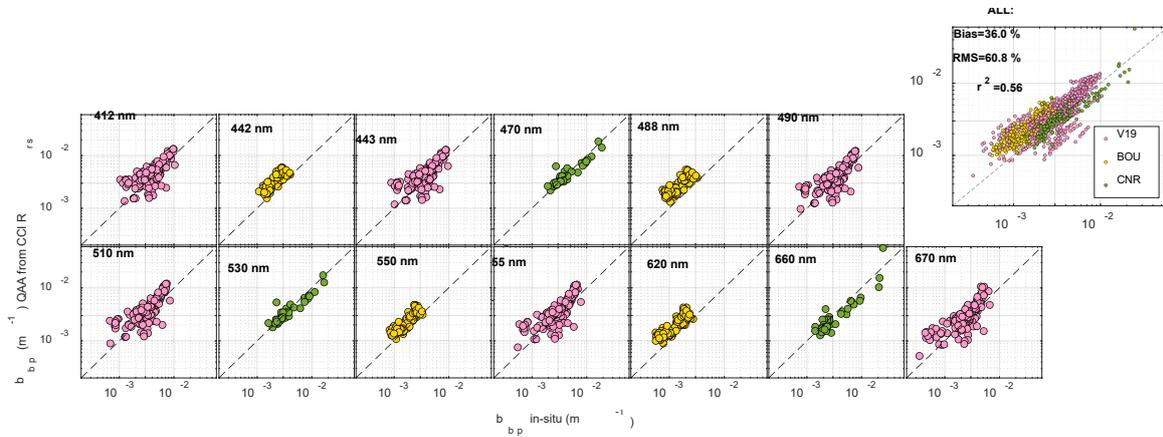
606 **Table 4:** Statistical descriptors of the difference between satellite CCI R_{rs} and in situ R_{rs} for each dataset.

	Band (nm)	Bias (%)	RMS (%)	r^2	N
V19	412	-19.6	42.7	0.37	147
	443	-16.9	30.6	0.53	147
	490	-5.0	19.3	0.66	147
	510	-0.4	15.3	0.73	147
	555	-4.6	18.7	0.78	147
	670	28.4	117.9	0.47	147
	All	-3.0	54.2	0.73	882
BOU	412	-4.0	22.5	0.50	96
	443	-3.7	23.9	0.63	97
	490	-1.9	11.1	0.66	97
	510	-6.4	11.9	0.47	97
	555	9.5	16.0	0.64	97
	670	24.2	49.5	0.31	97
	All	3.0	26.0	0.89	581
CNR	412	-10.7	24.8	0.42	42
	443	2.6	18.2	0.53	42
	490	-0.4	13.2	0.75	42
	510	-3.7	14.9	0.81	42
	555	0.9	19.9	0.88	42
	670	-4.9	83.1	0.90	42
	All	-2.7	21.9	0.87	252

607 *QAA Performance for b_{bp} Retrievals from CCI Data*

608 After assessing the quality of the QAA retrievals with in situ b_{bp} and the quality of the
 609 CCI R_{rs} respect to in situ R_{rs} , the QAA is applied to the CCI R_{rs} to retrieve the b_{bp} that are
 610 then compared to the in situ data. In agreement with our findings, CCI R_{rs} are corrected for
 611 Raman scattering. Results of this comparison are shown in Figure 8 and Table 5. For V19,
 612 biases are not significant (less than 30%) in comparison of RMS values (less than 60%). On
 613 the other hand, similarly to the statistics derived from section before, QAA-derived b_{bp} , as

614 compared to the BOU data displays significant positive biases. Comparison with CNR data
 615 shows the highest performances, with bias of +2.7% and RMS of 48%. The conclusions from
 616 our analysis are consistent with previous comparisons to QAA, reporting negligible biases
 617 above noise level both at global and regional scales [*Melin et al., 2007, 2011; Bisson et al.,*
 618 2019;].



619
 620 **Figure 8:** Scatter plots of QAA derived b_{bp} from CCI Rrs versus in situ b_{bp} data for each wavelength and dataset
 621 considered and for the merged data set. The dashed line represents the 1:1 ratio. Pink, yellow, and green dots
 622 refer to V19, BOU, and CNR data, respectively.

623 **Table 5:** Statistical descriptors of the difference between the QAA-derived b_{bp} from satellite CCI Rrs and in situ
 624 b_{bp} for each dataset with Raman scattering compensation. Figure A4 provides a graphical representation of this
 625 table.

	Band (nm)	Bias (%)	RMS (%)	r^2	N
V19	412	24.2	51.8	0.66	147
	443	26.8	53.9	0.67	147
	490	29.1	56.0	0.68	147
	510	29.9	56.8	0.67	147
	555	31.0	58.1	0.67	147
	670	31.6	60.7	0.62	147
	All	28.8	56.3	0.68	882
BOU	442	56.6	62.7	0.67	97
	488	86.9	96.2	0.64	97
	550	41.9	50.2	0.70	97
	620	66.8	75.3	0.69	97
	All	63.1	73.1	0.69	388
CNR	470	10.1	52.9	0.48	42
	530	7.8	54.9	0.46	42
	660	-9.6	33.5	0.63	42
	All	2.7	48.1	0.50	126

626 **Task #2**

627 Here, the context, method and the performed worked about Task #2 are presented. The
 628 main goal of this task is to evaluate which is the best method for b_{bp}^k estimations and if b_{bp}^k
 629 coefficient has a significant spatio-temporal variability that has to be accounted for in the
 630 C_{phyto} computation from space-borne observations. For more details see *Bellacicco et al.,*
 631 [2019] (*see Section 6*).

632 **Scientific context**

633 In the ocean, the pool of non-algal particles (NAP) includes: (i) heterotrophic organisms
634 such as bacteria, micro-grazers and viruses, (ii) organic particles of detrital origin such as
635 faecal pellets and cell debris, (iii) mineral particles of both biogenic (e.g. calcite liths and
636 shells) and terrestrial origin (e.g. clays and sand), (iv) bubbles [Sosik *et al.*, 2008] and (v)
637 plastics. Understanding of the spatial and temporal dynamics of NAP in the open ocean can
638 improve estimations of carbon export and sequestration [Azam *et al.*, 1983; Bishop and Wood,
639 2009]. NAP can covary with phytoplankton abundance or accumulate regardless of algal
640 dynamics. In such a context, a possible way to monitor these particles and distinguish
641 between these two fractions is via their optical backscattering properties and relationship with
642 chlorophyll-a. Unfortunately, only a few studies have concerned the backscattering
643 properties of NAP up to date ($b_b\text{NAP}$; units of m^{-1}) [Cho and Azam, 1990; Morel and Ahn,
644 1990, 1991; Stramski and Kiefer, 1991], as a consequence of the difficulties in directly
645 measuring this optical coefficient. Indeed, optical backscattering sensors measure
646 backscattering of all particles suspended in seawater (b_{bp} ; units of m^{-1}) [Dall’Olmo *et al.*,
647 2009, 2012, Westberry *et al.*, 2010], which includes algal particles among the others. The
648 NAP signal cannot be separated from that of phytoplankton. However, total b_{bp} offer the great
649 advantage to be measured by satellite and in situ from Biogeochemical-Argo (aka BGC-
650 Argo) floats. Using b_{bp} we can thus observe the global ocean with high spatial and temporal
651 resolutions.

652 The first attempt to derive $b_b\text{NAP}$ in the open waters was by Behrenfeld *et al.* [2005]
653 (hereafter Be05) using five-years of ocean colour remote sensing data. They computed the
654 fraction of the b_{bp} that does not covary with phytoplankton chlorophyll-a concentration (Chl;
655 units of mg m^{-3}), and estimated it as the offset of a linear regression between satellite-derived
656 b_{bp} and Chl when Chl concentrations were $> 0.14 \text{ mg m}^{-3}$. This offset was defined as the
657 background of the $b_b\text{NAP}$ (hereafter b_{bp}^k ; units of m^{-1}) and refers only to a fraction of the
658 total b_{bp} signal caused by NAP that thus does not covary with Chl (i.e. phytoplankton).

659 In Be05, b_{bp}^k is assumed to be a constant value both in space and time (i.e. $3.5 \cdot 10^{-4} \text{ m}^{-1}$).
660 Be05 attributed it to “a stable heterotrophic and detrital component of the surface particle
661 population and therefore independent of the phytoplankton dynamics”. Recently, Bellacicco
662 *et al.* [2016] (hereafter Blc16) applied Be05’s approach for distinct bioregions and seasons
663 in the Mediterranean Sea, and showed that b_{bp}^k has instead a marked regional and seasonal
664 variability. Such a result thus confirmed that the heterotrophic and detrital components at the
665 sea surface are neither negligible nor stable, but highly variable in seawater [Siokou-Frangou
666 *et al.*, 2010]. These observations were consistent with field observations of Chl and b_{bp} from
667 the BOUSSOLE buoy in which the Chl- b_{bp} relationship was highly dependent on the season
668 of the area [Antoine *et al.*, 2011]. The variability of the b_{bp}^k by Blc16 was also later confirmed
669 by Bellacicco *et al.*, [2018] for the global ocean (hereafter Blc18). Indeed, Blc18 highlighted
670 two distinct oceanic areas: the productive sub-polar North Atlantic Ocean, where b_{bp}^k and
671 particle biomass (i.e. phytoplankton cells) are anti-correlated; and the Southern Ocean, where
672 b_{bp}^k signal is mainly driven by inorganic particles, such as algal coccoliths [Balch *et al.*, 2016,
673 2018], bubbles or foam that occur in the stormy seas [Stramski *et al.*, 2004]. However, ocean-
674 colour data used in these works are only sensitive to the surface layer. The increasing number

675 of BGC-Argo floats, equipped with b_{bp} sensors, can therefore expand the analysis to
676 underlying layers.

677 The relationship between b_{bp} and Chl is also influenced by phytoplankton specific
678 composition and diversity (e.g. size, shape, internal structure), physiology (e.g.,
679 photoacclimation) and the nature of NAP itself [Stramski *et al.*, 2004; Dall'Olmo *et al.*, 2009,
680 2012]. Therefore, an analytical fit between b_{bp} and Chl that includes these factors may
681 improve b_{bp}^k estimations. In such a context, Brewin *et al.*, [2012] (hereafter Br12) presented
682 a relationship between b_{bp} and Chl that accounted for modifications in phytoplankton size.
683 The model, based on surface in situ observations, included separated b_{bp} terms for small and
684 large cells that dominated the overall fit at different Chl ranges. This model also estimated
685 b_{bp}^k , as the offset of the fit between b_{bp} and Chl in clear waters where this relationship
686 converged to a flat value for low Chl values. The b_{bp}^k parameter was interpreted as a constant
687 background of NAP (e.g. heterotrophic bacteria, detritus, viruses, minerogenic particles),
688 possibly partly influenced by very small phytoplankton (e.g. prochlorophytes).

689 In this study, the Br12 model is applied to an extensive global dataset of Chlorophyll-a
690 fluorescence, here converted in Chl, and b_{bp} (700) measurements acquired from BGC-Argo
691 profiling floats. In detail, we estimate b_{bp}^k across different oceanic areas (i.e. from productive
692 to ultra-oligotrophic zones), months, and in two distinct layers of the water column: at the
693 surface and within the euphotic layer. To interpret our estimations of b_{bp}^k , we use as a
694 reference of the b_{bp}^k value in each region the median b_{bp} at 950 – 1000 meters also derived
695 from BGC-Argo observations. At these depths b_{bp} is entirely due to the fraction of NAP that
696 does not covary with Chl [Poteau *et al.*, 2017].

697 Method

698

699 b_{bp}^k estimation: the model

700

701 In this study, the model developed by Brewin *et al.*, [2012] is used to compute b_{bp}^k . The
702 b_{bp} is modeled as a function of Chl and takes into account the fractional contributions of small
703 and large phytoplankton, as follows:

$$704 \quad b_{bp} = C_1^m \cdot [b_{bp,1}^* - b_{bp,2}^*][1 - e^{-S_1 \cdot Chl}] + b_{bp,2}^* \cdot Chl + b_{bp}^k$$

705

706 where the subscript 1 and 2 refer to two populations of phytoplankton cells partitioned
707 according to size: 1 is for cells $< 20\mu\text{m}$ while 2 is for cells $> 20\mu\text{m}$; $b_{bp,1}^*$ and $b_{bp,2}^*$ refer to
708 the Chl-specific b_{bp} coefficients associated with environments dominated by the two
709 populations of phytoplankton; C_1^m and S_1 refer to the maximum Chl concentration population
710 1 can reach and the initial slope relating the Chl concentration of population 1 to total Chl,
711 respectively.

712 The term b_{bp}^k refers to the background b_{bp} coefficient. The general equation of the model
713 can be simplified as:

714

$$715 \quad b_{bp} = c \cdot [1 - e^{(-S_1 Chl)}] + b_{bp,2}^* \cdot Chl + b_{bp}^k,$$

716

717 in which $b_{bp,2}^*$ is the slope, b_{bp}^k is the intercept of the fit, while $c = C_1^m [b_{bp,1}^* - b_{bp,2}^*]$ and
 718 S_1 terms are the coefficients of the non-linear part of the model. The $b_{bp,2}^*$, b_{bp}^k , c and
 719 S_1 coefficients are found from fitting Equation above to b_{bp} and Chl data by using the iterative
 720 bi-square method. The initial guess for the four parameters are reported in Table S1 (see
 721 published paper). These values are in the range and order of magnitude of the values reported
 722 in *Brewin et al.*, [2012]. This model reduces to the Be05, Blc16 and Blc18 linear models if
 723 the non-linear term is discarded out, which would be the case where $b_{bp,1}^*$ and $b_{bp,2}^*$ tend to
 724 the same value. This model represents an evolution of the previous published model (i.e.
 725 Be05, Blc16 and Blc18) because of it takes into account the phytoplankton populations
 726 variability in the Chl- b_{bp} relationship and thus for b_{bp}^k estimations. In addition, the inclusion
 727 of the non-linear term introduces more flexibility reducing the fit errors for the areas here
 728 analyzed (see Figures S1 and S2 on the published paper).

729 The equation above is applied to each area (spatially-resolved with the temporal
 730 aggregation approach), and for every month (spatially- and temporal-resolved approach) for
 731 the two layers. The ratio between the b_{bp}^k value found in the surface and in the bottom layers,
 732 and analogously for the euphotic layer, enables understanding the difference between upper
 733 and deeper layers for each area of interest. It is computed as:

734

$$735 \widehat{b_{bp}^k} = \frac{b_{bp,surface}^k}{b_{bp,bottom}^k}$$

736

$$737 \widehat{b_{bp}^k} = \frac{b_{bp,euphotic}^k}{b_{bp,bottom}^k}$$

738

739 In addition to this ratio, $\overline{b_{bp}^k}$ is here defined as the fraction of the b_{bp}^k with respect to the
 740 median b_{bp} (in %) giving an understanding on the relationship between NAP and particle
 741 biomass in the different areas, and the layers, of the ocean:

$$742 \overline{b_{bp}^k} = \frac{b_{bp}^k}{b_{bp}}$$

743 *Model fit and statistics*

744 For all the computations, Chl measurements below the value of 0.01 mg m^{-3} are considered
 745 too noisy for a proper estimation of b_{bp}^k and are filtered out from the dataset. The model in
 746 Equation above is fitted to the data using the iterative bi-square method which minimizes a
 747 weighted sum of squared errors, where the weight given to each data point decreases with
 748 the distance from the fitted curve [*Huber*, 1981]. Therefore, the error function is sensitive to
 749 the bulk of the data and the effect of outliers is thus reduced. This error function is minimized
 750 through the Trust-Region algorithm [*Moré and Sorensen*, 1983] and the final fit estimate is
 751 found after a maximum of 400 iterations. For each b_{bp}^k the 95% confidence intervals and two-
 752 standard deviation as confidence limit ($\pm 2\sigma$) are computed. In order to assess the model

753 performance for the b_{bp}^k calculation, the root mean square (RMS; in m^{-1}) error between the
754 modeled- b_{bp} and measured- b_{bp} are computed. The RMS is calculated according to:
755

$$756 \quad \text{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N (b_{bp,modeled,i} - b_{bp,measured,i})^2}$$

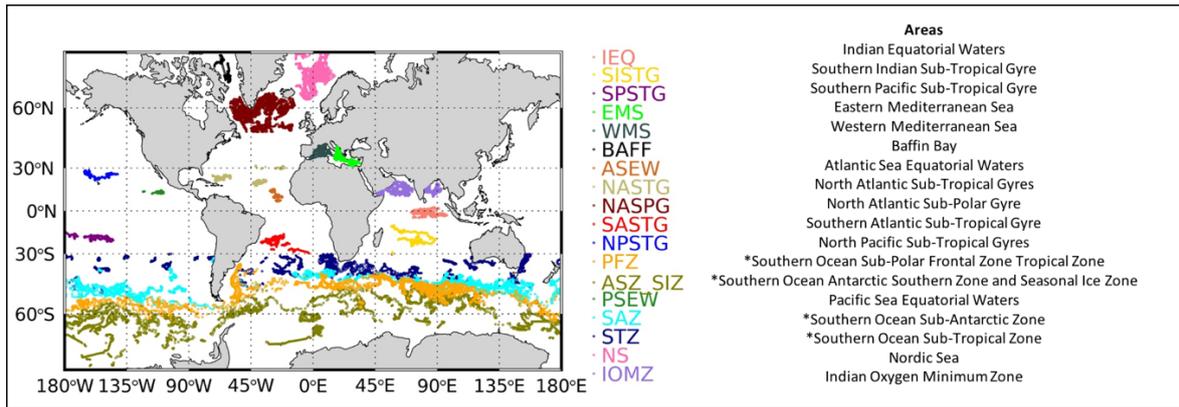
757

758 Data

759

760 An array of 425 BGC-Argo profiling floats was deployed around the World's oceans as
761 part of several national and international programs (<http://biogeochemical-argo.org>), and
762 collected data from 30/05/10 to 31/12/18 every one up to ten days. These floats acquired 0-
763 1000 m vertical profiles of pressure, temperature and salinity by a Seabird Scientific SBE 41
764 Conductivity-Temperature-Depth (CTD) sensor, Chlorophyll-a fluorescence (FChla;
765 excitation at 470 nm, emission at 695 nm) and the angular scattering function at 700 nm by
766 Seabird-WetLABS combo sensors (mostly FLBB, ECOTRIPLET, or MCOMS).
767 Chlorophyll-a fluorescence is then converted to Chl concentration (units of $mg\ m^{-3}$) and the
768 angular scattering to particulate optical backscattering coefficient b_{bp} (units of m^{-1}) (see
769 supplementary materials in the published paper). All the data were downloaded from the
770 Coriolis database (<ftp://ftp.ifremer.fr/ifremer/argo/dac/coriolis>) and quality controlled (see
771 supplementary material). The BGC-Argo floats (more than 35000 correspondent Chl and b_{bp}
772 data) over global ocean used in the present study are partitioned into 18 areas (Figure 9). The
773 dataset of Chl and b_{bp} here used, represents the update version of the databases BOPAD-prof
774 and BOPAD-surf by *Organelli et al.* [2017]. The depth of euphotic zone, Z_{eu} (units of m),
775 which is the depth where PAR reaches 1% of its surface value, was estimated from the Chl
776 profile through the iterative process described in *Morel and Maritorena* [2001].
777 Subsequently, the first optical depth, Z_{pd} (units of m), was calculated as $Z_{eu}/4.6$ [*Morel,*
778 1988]. Finally, for each profile, the mean and standard deviation of Chl and b_{bp} were
779 calculated within: i) the surface layer: the layer between sea surface and the first optical
780 depth; ii) the euphotic layer: the layer between sea surface and euphotic zone; and iii) the
781 bottom layer: the layer between 950 and 1000 m.

782



783

784

785

786

787

788

789

Figure 9: Geographical distribution of the BGC-Argo dataset on a global ocean scale. Each colour represents sampling areas and abbreviations. * indicates data acquired in four regions below 30°S which have been delineated by using temperature profiles [Gray *et al.*, 2018]: Sub-Tropical Zone (STZ) with a temperature at 100 m above 11°C; the Sub-Antarctic Zone (SAZ) with a temperature at 400 m below 5°C; the Polar Frontal Zone (PFZ) with the minimum temperature between 0 and 200 m above 2°C; the Antarctic Southern Zone and Seasonal Ice Zone (ASZ_SIZ) minimum temperature between 0 and 200 m below 2°C.

790

791 Results

792

793 *Global overview of b_{bp}^k*

794

795

796

797

798

799

800

801

802

803

804

805

806

Aggregated quality-controlled data within the surface layer for all areas and months (N=36067) are shown in Figure 10a. The b_{bp} coefficients increase with Chl but with relatively constant b_{bp} for low Chl values (Figure 10a). This behavior is consistent with previous observations by Behrenfeld *et al.* [2005] and Brewin *et al.* [2012], and is considered to be the consequence of two distinct oceanic conditions: “photoacclimation-dominance” and “biomass-dominance” of Chl signal. The former is typical of oligotrophic areas (e.g. subtropical gyres) where variability of Chl is uncoupled with biomass and the process of acclimation to light and nutrients drives Chl variations [Siegel *et al.*, 2013; Halsey and Jones, 2015; Barbieux *et al.*, 2018]. On the reverse, the latter case is typical of most productive areas where Chl and b_{bp} strongly covary [Dall’Olmo *et al.*, 2009, 2012; Westberry *et al.*, 2010]. The high Chl- b_{bp} co-variability is a clear indication that particles (and biomass) covary with phytoplankton abundance, while the physiological photoacclimation process playing a secondary role in determining the Chl variations.

807

808

809

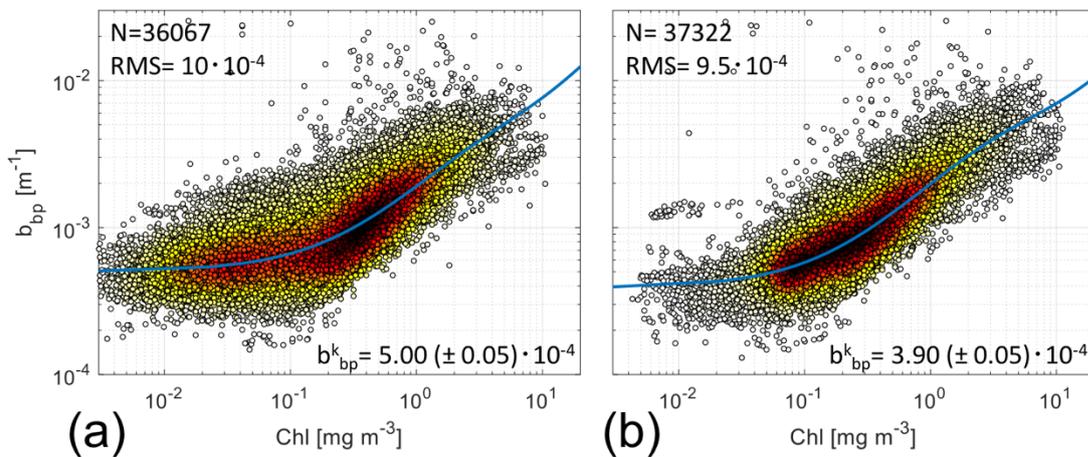
810

811

812

Here, the application of the Br12 model to these BGC-Argo data leads to a b_{bp}^k equal to $5.0 \cdot 10^{-4} \text{ m}^{-1}$ at the surface, a value higher than that found by Be05 ($3.5 \cdot 10^{-4} \text{ m}^{-1}$ at 443 nm). On the other hand, Br12 reported $7.0 \cdot 10^{-4} \text{ m}^{-1}$ for 470 nm and $5.6 \cdot 10^{-4} \text{ m}^{-1}$ at 526 nm. Blc18 found a median b_{bp}^k value equal to $9.5 \cdot 10^{-4} \text{ m}^{-1}$ based on 19-years of ocean colour data. These values are comparable as the spectral variability is limited in case of b_{bp} ($\pm 30\%$ between 443nm and 700nm when assuming b_{bp} decreasing as a power law with slope equal to 0.7). In

813 relative terms, our study shows that b_{bp}^k dominate within the surface layer as it accounts for
 814 $\pm 57\%$ of the total b_{bp} measured by all BGC-Argo floats, a remarkably high percentage. An
 815 increased Chl- b_{bp} co-variability is observed within the euphotic layer (Figure 10b; N=37322).
 816 The derived b_{bp}^k is not comparable to our estimates from the surface layer or from previous
 817 satellite observations because it includes deeper layers where there is high particle
 818 concentration, as for example oligotrophic areas such as the subtropical gyres and the eastern
 819 Mediterranean Sea [Volpe *et al.*, 2007; Barbieux *et al.*, 2018]. The first estimation of b_{bp}^k for
 820 this layer is a value of $3.9 \cdot 10^{-4} \text{ m}^{-1}$, and accounts for $\pm 45\%$ of the total b_{bp} , suggesting that in
 821 the euphotic layer NAP are more correlated to Chl than at the surface.



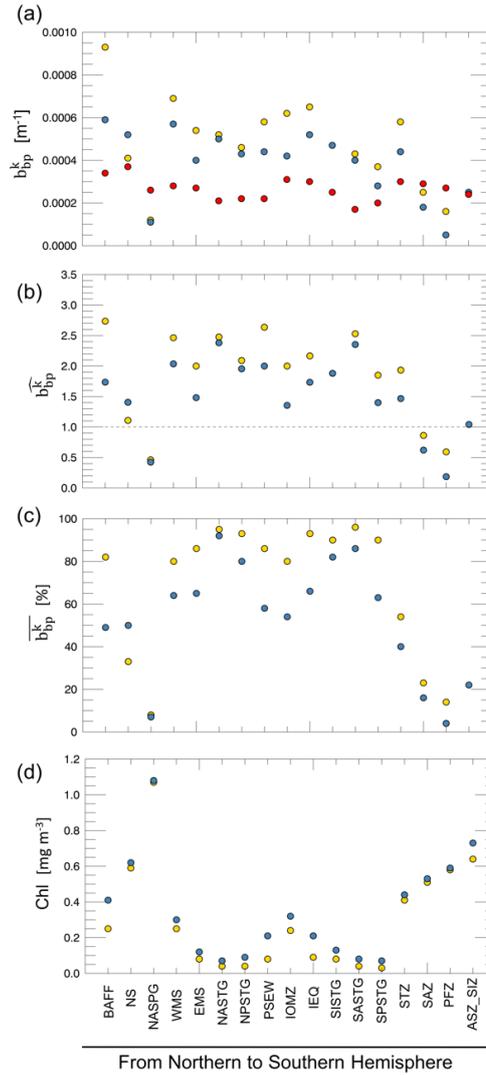
822
 823 **Figure 10:** Plot density between Chl and b_{bp} (700) within the surface layer (panel a) and the euphotic layer
 824 (panel b). Both panels include the number of observations (N) and the RMS (in m^{-1}). The b_{bp}^k estimation (in m^{-1})
 825 with two standard deviation as confidence limit ($\pm 2\sigma$) is also reported. Chl values $< 0.01 \text{ mg m}^{-3}$ are not
 826 included in the fit computations. The plots are presented in logarithmic scale in both axes though the fit has
 827 been calculated in linear scale. Dot density is indicated as colour from white (low density) to black (high
 828 density).

829 *Geographical distribution of b_{bp}^k*

830 Figure 11a shows b_{bp}^k estimations for the surface, euphotic and bottom layers within each
 831 geographical area sampled by BGC-Argo floats. In surface layer, the range of variability
 832 spans between 10^{-4} m^{-1} and 10^{-3} m^{-1} , consistent with global ocean-colour estimations
 833 [Bellacicco *et al.*, 2018]. Lower variability characterizes the euphotic layer (of a factor of
 834 ~ 6), from $\sim 1.0 \cdot 10^{-4} \text{ m}^{-1}$ to $6.0 \cdot 10^{-4} \text{ m}^{-1}$. For the bottom layer, variability is the lowest, between
 835 $2.0 \cdot 10^{-4} \text{ m}^{-1}$ and $4.0 \cdot 10^{-4} \text{ m}^{-1}$. The two upper layers display a latitudinal gradient, with a
 836 general b_{bp}^k decrease from northern to southern oceans. b_{bp}^k in the bottom layer does not show
 837 a clear geographical pattern and remains relatively constant across all sampled oceanic areas.

838 Figure 11b shows the $\widehat{b_{bp}^k}$ for each area, the ratios between the spatially-resolved b_{bp}^k found
 839 at the surface and euphotic layers with the estimation for the bottom layer. Globally, $\widehat{b_{bp}^k}$ is

840 higher in the upper layer than the at the bottom from mid- to low-latitudes, while b_{bp}^k at the
841 bottom is higher than at the surface in most productive seas such as the NASPG, SAZ, PFZ
842 and ASZ_SIZ areas [Uitz *et al.*, 2009; Alkire *et al.*, 2014; Artega *et al.*, 2018]. In these areas,
843 $\overline{b_{bp}^k}$ is only a small fraction of the total b_{bp} in surface waters (< 20%; Figure 11c) as a
844 consequence of the higher relative variability in the b_{bp} and phytoplankton abundance [Alkire
845 *et al.*, 2014]. In the NASPG, characterized by high phytoplankton biomass, $\overline{b_{bp}^k}$ is lower than
846 10%. It means that b_{bp} is more dominated by particles that covary with phytoplankton cells,
847 thus being more influenced by phytoplankton dynamics. In the Southern Ocean (i.e. STZ,
848 SAZ, PFZ and ASZ_SIZ areas), $\overline{b_{bp}^k}$ ranges from 15% (i.e. PFZ) to 60% (i.e. STZ) for surface
849 waters suggesting inorganic particles (e.g. coccoliths) can also drive the b_{bp}^k signal (Figure
850 11c). Indeed, coccoliths concentrations covary with b_{bp} because they scatter light with high
851 efficiency [Balch *et al.*, 2016; 2018]. The b_{bp}^k values, and their order of magnitude, are
852 consistent with measurements of b_{bp} from $CaCO_3$ reported in Balch *et al.* [2016] along the
853 Great Calcite Belt (GCB) (their Figure 2c). Thus, in these areas of the Southern Ocean, the
854 b_{bp}^k may be related to the coccolithophorids seasonality (i.e. skeleton compounds of no longer
855 living cells; b_{bp}^k is the b_{bp} when Chl is zero) [Balch *et al.*, 2016; 2018; Bellacicco *et al.*, 2018].
856 In less productive areas (e.g. EMS, IEQ, NASTG, SISTG, SASTG, SPSTG; Figure 11d), $\overline{b_{bp}^k}$
857 is greater than 80% at the surface layer, consistent with previous findings [Brewin *et al.*,
858 2012; Bellacicco *et al.*, 2018]. These areas are characterized by limited nutrients availability
859 determining low phytoplankton biomass, especially pico- and nano-phytoplankton
860 dominated communities [Bricaud *et al.*, 2004; Mignot *et al.*, 2014], which are rapidly
861 recycled in the surface layer thus supporting relatively high bacterial and detrital biomass.
862 For the euphotic layer, much of the b_{bp} can be related to phytoplankton biomass as highlighted
863 by a lower $\overline{b_{bp}^k}$ value of around 60%. This is the consequence of the subsurface chlorophyll
864 maximum (SCM) which is deeper in the subtropical gyres and oligotrophic seas as found by
865 Mignot *et al.*, [2014] and Barbioux *et al.*, [2019]. It determines that, at depth, there is an
866 increase of phytoplankton biomass and of NAP covarying with phytoplankton: the b_{bp}^k
867 coefficient indeed decreases from the surface to the euphotic layers (Figure 11a).



868

869 **Figure 11:** Geographical distribution of b_{bp}^k (in m^{-1}) in the three layers: surface (gold), euphotic (blue) and
870 bottom (red) (a). The $\widehat{b_{bp}^k}$ for the surface (gold) and euphotic (blue) layers for each area (b). The dashed line
871 indicates the case where $\widehat{b_{bp}^k}$ estimates between surface or euphotic layer with bottom layer are close to the same
872 value. Panel c shows the $\widehat{b_{bp}^k}$ (in %) for each area and layer (gold for surface layer; blue for euphotic layer). The
873 model performance, in terms of RMS (m^{-1}) and interval of confidence at 95% for each b_{bp}^k estimation is reported
874 in the supplementary information (see Figures S3, S4; Table S2; supplementary in the published paper). ASEA
875 area is not included in this analysis due to the low performance of the model and highest uncertainties in b_{bp}^k
876 assessment in both layer (for details see the supplementary materials). Note that the areas have been sorted from
877 the northern to the southern hemisphere. Panel d shows the mean Chl values for each region and layers (gold
878 for surface layer and blue for euphotic layer). See Figure 9 for locations and abbreviations.

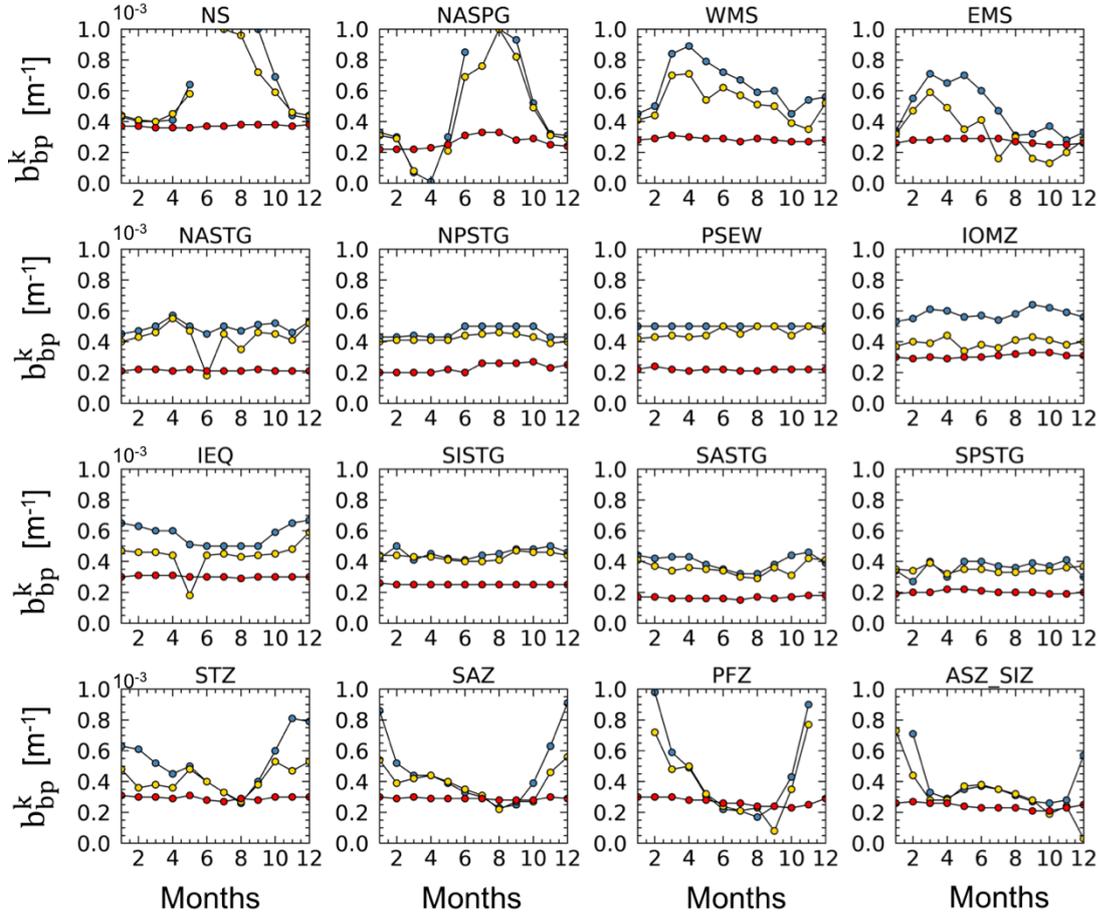
879 *Seasonal variability of b_{bp}^k*

880 The b_{bp}^k values within surface and euphotic layers show a clear seasonal cycle with
881 maxima during the productive periods ($b_{bp}^k > 5.0 \cdot 10^{-4}$) and minima during the low productive
882 periods ($b_{bp}^k < 4.0 \cdot 10^{-4}$) in all the areas outside the oligotrophic seas (e.g. NS, NASPG, WMS,
883 EMS, STZ, SAZ, PFZ, ASZ_SIZ) (Figure 12). In the NASPG, b_{bp}^k shows high values during

884 the well-known spring bloom and low values from December to April [*Briggs et al.*, 2011;
885 *Alkire et al.*, 2014; *Mignot et al.*, 2018]. In the Southern Ocean, and especially SAZ, PFZ and
886 ASZ_SIZ areas, b_{bp}^k shows the maxima values from December to April (i.e. period of bloom)
887 while the minima are detected in the period May-September. In the Mediterranean Sea (i.e.
888 WMS and EMS), the seasonal cycle varies within the sub-basins showing different amplitude
889 and shape, clearly linked to the regional trophic regimes. WMS shows b_{bp}^k values higher than
890 the eastern ones confirming the presence of a general decreasing eastward gradient for this
891 coefficient. In the western basin of Mediterranean Sea, deep-water formation dynamics
892 and/or the generally shallow nutricline results in a maximum value in April. On the contrary,
893 maxima generally occur earlier between February and March in the eastern Mediterranean
894 basin. These results confirm *Bellacicco et al.*, [2016] findings for this semi-enclosed basin.
895 In their work, b_{bp}^k was demonstrated to be variable both in space and time with a marked
896 seasonality in the different bio-regions of both the sub-basins. As shown by *Bellacicco et al.*,
897 [2016], periods characterized by lower b_{bp}^k (e.g. summer) are also associated with higher
898 variability and uncertainties in the estimations. This is valid for the b_{bp}^k both in the surface
899 and euphotic layers, and has to be taken into account in the interpretation of these results (see
900 Tables S3 and S4 in the supplementary materials of published paper).

901 The b_{bp}^k at the bottom layer shows a smoother seasonal cycle in respect to what occur in
902 the upper layers. As found by *Poteau et al.*, [2017], an annual cycle is only observed at the
903 Southern Ocean and sub-polar North Atlantic area, regions with the largest amplitude in the
904 seasonal cycles at the surface and euphotic layer (Figure 12) due to blooms of large
905 phytoplankton [*Alkire et al.*, 2014; *Barbieux et al.*, 2018]. *Poteau et al.*, [2017], indeed,
906 suggested that the b_{bp}^k at the depth can be mostly related to disaggregation of these large
907 settling particles.

908 The seasonal cycle of b_{bp}^k in the less productive seas for all the layers is low, suggesting
909 low NAP seasonal variations (e.g. detrital matter, heterotrophic bacteria, virus). The b_{bp}^k
910 estimation for each month appears to be nearly constant throughout the year (Figure 12) and
911 thus b_{bp} may be controlled mostly by b_{bp}^k , as highlighted also in Figure 11c.



912

913 **Figure 12:** Temporal variability of b_{bp}^k (in m^{-1}) for each area and all the three layers: surface (gold), euphotic
 914 (blue) and bottom (red). The model performance, in terms of RMS (m^{-1}) and interval of confidence at 95% for
 915 each monthly b_{bp}^k estimation, are reported in the supplementary materials (see Tables S3, S4 and S5). ASEW
 916 and BAFF areas are not included in the analysis due to the absence/limited number of observations that prevents
 917 the description of the annual cycle. See Figure 9 for locations and abbreviations.

918 **Task #3**

919 Here, the context, method and the performed worked about Task #3 are presented. The
 920 main goal of this task is to refine and developed the algorithm for the C_{phyto} detection from
 921 space-borne observations. For more details see *Bellaciccio et al.*, [2020].

922 **Scientific Context**

923 Phytoplankton is responsible for approximately half of the global primary production and
 924 is at the base of the marine food web [*Siegel et al.* 2014]. Phytoplankton is consequently a
 925 fundamental actor in the global carbon cycle [*Boyce et al.*, 2010]. Moreover, these organisms
 926 are regarded as sentinels of changes in the ocean because of their capacity to respond rapidly
 927 to environmental perturbations. Several factors, such as ocean circulation, anthropogenic
 928 activities, and climate affect phytoplankton abundance and distribution. In particular,
 929 phytoplankton spatio-temporal patterns are expected to vary with climate change [*Dutkiewicz*

930 *et al.*, 2019]. Monitoring the distribution of phytoplankton becomes crucial for a full
931 understanding of the oceanic biogeochemical cycles.

932 Ocean colour observations have significantly improved our capability of mapping
933 phytoplankton global distribution since the 1970s. Chlorophyll-a concentration (Chl; in mg
934 m^{-3}) is a consolidated proxy of algal biomass [O'Really *et al.*, 2019]. However, Chl variability
935 may not always correspond to actual changes in algal biomass but rather to cellular
936 physiological adjustments in response to environmental stressors such as light and nutrient
937 limitation [Laws and Bannister, 1980; Behrenfeld *et al.*, 2014]. Therefore, there has been an
938 increasing focus towards estimating phytoplankton biomass in units of Carbon (C_{phyto} ; mg C
939 m^{-3}). C_{phyto} has found application in studies of: (i) primary production [Behrenfeld *et al.*,
940 2005; Westberry *et al.*, 2008]; (ii) phytoplankton physiology [Behrenfeld *et al.*, 2008]; (iii)
941 phytoplankton growth/loss rates [Zhai *et al.*, 2008, 2010]; (iv) comparison with marine
942 ecosystem models [Dutkiewicz *et al.*, 2015]; (v) pools of carbon in the ocean (i.e. their stocks
943 and turnover rates); and (vi) ocean-atmosphere interactions on marine phytoplankton
944 feedbacks on atmospheric aerosols [Fossum *et al.*, 2018].

945 *Martinez-Vicente et al.* [2017] recently reviewed algorithms for deriving C_{phyto} from
946 satellite ocean colour observations. Three groups of algorithms are currently available: (i)
947 backscatter-based models developed using satellite [Behrenfeld *et al.*, 2005] or in-situ data
948 [Martinez-Vicente *et al.*, 2013; Graff *et al.*, 2015]; (ii) Chl-based models [Sathyendranath *et al.*
949 *et al.*, 2009]; and (iii) allometric conversion-based models [Kostadinov *et al.*, 2009, 2016; Roy
950 *et al.*, 2017]. Among these algorithms, those relying on the particulate optical backscattering
951 coefficient, b_{bp} (in m^{-1}), raised high interest because b_{bp} is related to the concentration and
952 composition of suspended particles in seawater as well as to their size and shape [Stramski *et al.*
953 *et al.*, 2004]. Though former research suggested that b_{bp} was mainly influenced by submicron
954 detrital particles, it has recently been demonstrated that most of b_{bp} signal in the surface
955 oligotrophic ocean is due to particles with equivalent spherical diameters between 1-10 μm
956 [Organelli *et al.*, 2018], thus reinforcing the role of phytoplankton as a source of the open-
957 ocean b_{bp} [Zhang *et al.*, 2020], and thus the usefulness of b_{bp} to retrieve C_{phyto} . However,
958 satellite b_{bp} -based algorithms show large errors in predicting C_{phyto} mainly because the
959 current formulation does not take into account the spatio-temporal changes of non-algal
960 particles (NAP) [Bellacicco *et al.*, 2019].

961 Here we refine the first and widely-applied satellite b_{bp} -based C_{phyto} model [Behrenfeld *et al.*
962 *et al.*, 2005] (hereafter Beh05) and make crucial arrangements to improve its performance.
963 Beh05 uses the linear relationship between Chl and b_{bp} to estimate the background
964 contribution of NAP to $b_{\text{bp}}(443)$, b_{bp}^{k} , which corresponds to b_{bp} when Chl is zero (i.e., the
965 intercept of the linear fit). Beh05 is based on the following equation:

966
967

$$C_{\text{phyto}} = [b_{\text{bp}}(443) - b_{\text{bp}}^{\text{k}}] \cdot \text{SF}$$

968

969 In Beh05, SF is the scaling factor used to convert b_{bp} into C_{phyto} and set to $13000 \text{ mg C m}^{-2}$.
970 b_{bp}^k is assumed to represent a stable surface population of heterotrophic and detrital particles
971 that is not strictly dependent on phytoplankton dynamics and physiology. Thus, in Beh05,
972 b_{bp}^k was assumed as a constant value ($3.5 \cdot 10^{-4} \text{ m}^{-1}$). However, NAP can include various types
973 of particles ranging from heterotrophic organisms such as bacteria, micrograzers, viruses, to
974 faecal pellets and cell debris, to mineral particles and bubbles [Stramski *et al.*, 2004] that
975 varies both in space and time. The nature, definition and evaluation of b_{bp}^k is still under debate
976 and has been investigated using in-situ [Brewin *et al.*, 2012; Bellacicco *et al.*, 2019; Zhang
977 *et al.*, 2020] and satellite data [Bellacicco *et al.*, 2016, 2018]. Understanding the spatial and
978 temporal dynamics of b_{bp}^k in open ocean will improve estimations of C_{phyto} , and more
979 generally our knowledge on the marine carbon fate.

980 Beh05 estimated b_{bp}^k as the offset (intercept) of the least-square regression between
981 monthly satellite-derived b_{bp} and Chl computed via GSM algorithm [Garver and Siegel,
982 1997]. The regression analysis used only Chl values higher than 0.14 mg m^{-3} to separate the
983 changes in Chl due to physiology (i.e., photoacclimation) from actual changes in biomass,
984 and thus in C_{phyto} . Bellacicco *et al.* [2018] used satellite data from the Mediterranean Sea to
985 estimate b_{bp}^k for various bioregions and seasons, and showed that b_{bp}^k varied regionally and
986 seasonally, thus rejecting the hypothesis of invariance of the heterotrophic and detrital
987 components at the sea surface [Bellacicco *et al.*, 2016]. Worldwide variability in b_{bp}^k was
988 observed [Bellacicco *et al.*, 2018], reporting a median global value of $9.5 \cdot 10^{-4} \text{ m}^{-1}$, nearly
989 threefold the one found by Beh05. The newly evaluated b_{bp}^k resulted in C_{phyto} that were half
990 those estimated by Beh05.

991 The Chl- b_{bp} relationship is influenced by phytoplankton specific composition and
992 diversity (e.g. size, shape, internal structure), physiology (e.g., photoacclimation) and the
993 nature of NAP itself [Dall'Olmo *et al.*, 2009, 2012]. For these reasons, Brewin *et al.* [2012]
994 (hereafter Bre12) developed an analytical non-linear fit between b_{bp} and Chl that accounted
995 for changes in phytoplankton cell size. In Bre12, the offset of the non-linear fit (i.e. b_{bp}^k)
996 between b_{bp} and Chl in clear waters was interpreted as a constant background of NAP or also
997 partly influenced by very small phytoplankton (e.g. prochlorophytes) in addition to NAP. In
998 Bre12, the reported b_{bp}^k values were of $7.0 \cdot 10^{-4} \text{ m}^{-1}$ at the wavelength of 470 nm and $5.6 \cdot 10^{-4}$
999 m^{-1} at 526 nm. More recently, Bre12 model was used with a large database of 0-1000 m
1000 depth profiles of Chl and b_{bp} acquired by the global BGC-Argo float array. b_{bp}^k was
1001 demonstrated to vary also over depth and to account only for a small fraction of total b_{bp} in
1002 productive areas, while being the main source of b_{bp} in oligotrophic waters [Bellacicco *et al.*,
1003 2019].

1004 The former evidence suggests that, if a realistic spatio-temporal variation of b_{bp}^k is
1005 introduced in the C_{phyto} algorithm, its precision and accuracy can be largely improved.

1006 Specifically, the hypothesis that b_{bp}^k varies in space and time is here applied monthly at
 1007 the scale of satellite pixel rather than using a unique b_{bp}^k value or pre-defined bioregions.
 1008 Algorithm outputs are validated against the only available, to the best of our knowledge, in-
 1009 situ dataset and compared to the performances of Beh05, Bel18 and Bre12 with single b_{bp}^k
 1010 values. The performance of the new algorithm is also compared to empirical approaches
 1011 [*Martinez-Vicente et al.*, 2013; *Graff et al.*, 2015]. The analysis was developed globally by
 1012 partitioning the data among Optical Water Class (OWC) [*Jackson et al.*, 2019] in order to
 1013 explore the efficiency and applicability of the different tested approaches.

1014 **Method**

1015 *Computation of b_{bp}^k*

1016
 1017 b_{bp}^k is estimated for every pixel and month by using the following basic relationship:

1018

1019

$$b_{bp}^k = b_{bp} - k \cdot Chl$$

1020

1021 with k being the slope of the least-square regression fit between daily Chl and b_{bp} , while
 1022 b_{bp}^k is the intercept of the fit [*Behrenfeld et al.*, 2005]. The term $k \cdot Chl$ thus identifies the
 1023 fraction of b_{bp} that covaries with Chl [*Behrenfeld et al.*, 2005, *Bellacicco et al.*, 2016, 2018,
 1024 2019]. A scheme of the algorithm implementation is shown in Figure S1 (see published
 1025 paper).

1026 We computed monthly climatological b_{bp}^k maps at 25 km by applying equation to daily
 1027 Chl and b_{bp} data within each month of the year from 1998 to 2019. The estimation of b_{bp}^k ,
 1028 and consequently of C_{phyto} , relies on good relationships between Chl and b_{bp} which then
 1029 constitutes the fundamental condition to exploit this method. To this aim, for each b_{bp}^k map,
 1030 the significance S (through the t-Student Test), the Pearson correlation coefficient (r ; not
 1031 shown) and the 1σ uncertainty maps (due to the linear fit between Chl and b_{bp}) were
 1032 computed to give an estimate of the robustness of the fit.

1033 We then applied for each pixel of every resulting monthly b_{bp}^k map an additional moving
 1034 average of 1000 km in order to remove any source of noise and smaller scale variability,
 1035 while conserving the large scale oceanic patterns [*Resplandy et al.*, 2019].

1036 Finally, the monthly b_{bp}^k maps computed with the new approach were interpolated to
 1037 obtain daily climatological b_{bp}^k maps at 25 km resolution. Daily pixel-based estimates are
 1038 then used to assess C_{phyto} by applying a revised Equation [*Behrenfeld et al.*, 2005; *Bellacicco*
 1039 *et al.*, 2016, 2018] such as:

1040

1041

$$C_{phyto} = [b_{bp} (443) - b_{bp}^k (443)] \cdot SF$$

1042

1043 In some pixels, C_{phyto} may show values less of, or close to 0. This occurred for pixels
 1044 where the Chl- b_{bp} relationship had a significance $S < 0.95$ and $r \leq 0$. In those pixels, b_{bp}^k may
 1045 be higher than the b_{bp} thus giving non-reliable, negatives, estimations of C_{phyto} (less than 0;

1046 e.g. in the subtropical gyres). For those areas, we applied a threshold of C_{phyto} equal to 0.13
 1047 mg C m^{-3} in order to highlight that C_{phyto} is expected to be low [Martinez-Vicente et al., 2017].
 1048 All match-up points falling in the areas where the Chl- b_{bp} relationship showed a significance
 1049 $S < 0.95$ and $r \leq 0$ were not included in the validation and algorithm performance analysis
 1050 (suppl. materials in the published paper). Such areas thus need to be interpreted and managed
 1051 with caution. For comparative matters, C_{phyto} was also computed following approaches of
 1052 Beh05, Bel18 and Bre12 with single b_{bp}^k values of $3.5 \cdot 10^{-4} \text{ m}^{-1}$, $9.5 \cdot 10^{-4} \text{ m}^{-1}$ and $7.0 \cdot 10^{-4} \text{ m}^{-1}$,
 1053 respectively. In addition, the performance of the new algorithm was compared to the
 1054 empirical algorithm of Martinez-Vicente et al. [2017] (hereafter MV17) and Graff et al.
 1055 [2015] (hereafter Gra15).

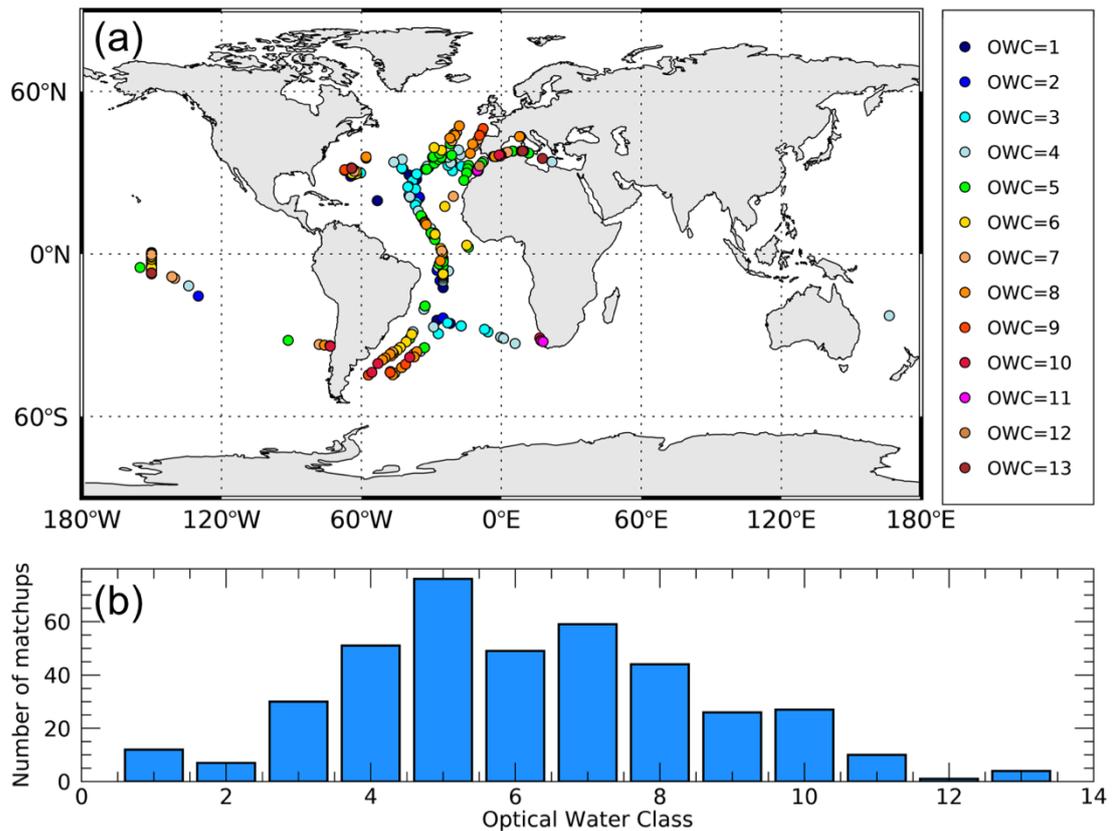
1056 *Statistical assessment*

1057 Estimated satellite C_{phyto} (y_i) computed with revised equation were compared to reference
 1058 in-situ data (x_i) by computing the bias (δ ; mg m^{-3}), the relative bias in percentage (∇ ; in %)
 1059 and the standard deviation of the difference (σ_{Δ} ; mg m^{-3}). The assessment was made for each
 1060 OWC individually as well as for all matchups at once:

$$1061 \quad \delta = \frac{1}{N} \sum_{i=1}^N (y_i - x_i)$$

$$1062 \quad \nabla = 100 * \frac{1}{N} \sum_{i=1}^N \frac{(y_i - x_i)}{x_i}$$

$$1063 \quad \sigma_{\Delta} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N [(y_i - x_i) - \overline{(y_i - x_i)}]^2}$$



1064

1065 **Figure 13:** Geographical distribution of the matchup dataset (N=396) used for the analysis. Each data is also
 1066 associated with the specific OWC (a). Number of matchups per OWC (b).

1067 **Data**

1068

1069 *Ocean color data*

1070

1071 The full European Space Agency (ESA) Ocean Colour – Climate Change Initiative (OC-
 1072 CCI) time series (1997-2019) of global daily R_{rs} and Chl version 4.2 data at 4 km resolution
 1073 was downloaded from the ESA OC-CCI FTP server (<https://esa-oceancolour-cci.org>;
 1074 <ftp://oceancolour.org/occci-v4.2/geographic/netcdf/daily/rrs/>). ESA OC-CCI data products
 1075 are the result of the merging of SeaWiFS, MERIS, MODIS and VIIRS observations in which
 1076 the inter-sensor biases are removed. The version 4.2 includes the latest NASA reprocessing
 1077 R2018.0 that mostly accounts for the ageing of the MODIS sensor. The ESA OC-CCI
 1078 provides the daily R_{rs} data and associated uncertainty maps in terms of bias and root mean
 1079 square error. In this work, for each daily R_{rs} , the bias was also extracted and then corrected
 1080 pixel-by-pixel, as recommended in the Product User Guide on the ESA CCI website.

1081 Chl was estimated with a blending of the OCI (as implemented by NASA, itself a
 1082 combination of CI and OC4), the OC5 and the OC3 algorithms. For each daily Chl value, the
 1083 bias was also extracted and compensated at the pixel scale. Daily b_{bp} maps at 443 nm were
 1084 produced from daily R_{rs} for the same period (1997-2019) by applying the Quasi-Analytical
 1085 Algorithm (QAA) [Lee *et al.*, 2002] with prior correction for Raman scattering [Pitarch *et*
 1086 *al.*, 2020]. The accuracy of the QAA in retrieving b_{bp} was fully demonstrated. In this work,

1087 the QAA algorithm is thus selected for its high efficiency in the b_{bp} retrievals as well as to
1088 show consistency and coherency with the OC-CCI programme in which the QAA is the
1089 designated algorithm [Brewin *et al.*, 2015, Sathyendranath *et al.*, 2020].

1090 The daily datasets were then under-sampled to 25 km resolution to resolve only the
1091 broader oceanographic scales of variability.

1092

1093 *In-situ reference data*

1094

1095 The in-situ C_{phyto} database [Martinez-Vicente *et al.*, 2017] is a compilation of data acquired
1096 from many sources (e.g., MAREDAT, AMT, etc.) for a total of $N=557$ data points and consists
1097 of carbon biomass of picophytoplankton organisms (*i.e.*, cell size $< 2\mu m$). It was downloaded
1098 from <http://www.zenodo.org> (*doi: 10.5281/zenodo.1067229*). For complete details about the
1099 in-situ dataset see [Martinez-Vicente *et al.*, 2017]. Only pixels with good ($S \geq 0.95$ and $r > 0$)
1100 satellite relationship between Chl and b_{bp} were retained so that the originally 557 data points
1101 decreased to a total of 396 matchups (Figure 13). The final matchup database encompassed
1102 from oligotrophic to mesotrophic waters (Chl from 0.035 to 3.13 $mg\ m^{-3}$ and C_{phyto} from 1.80
1103 to 60.25 $mg\ C\ m^{-3}$) and OWCs from 1 to 13. OWC from 1 to 6, corresponding to less productive
1104 waters, represented the 56% of the in-situ data.

1105

1106 **Results**

1107 *Algorithm Performance for C_{phyto} Retrievals*

1108 Results and statistics of the comparison between satellite C_{phyto} estimates obtained with
1109 our algorithm and the previously published ones, and in-situ C_{phyto} are reported in Table 6
1110 and Figure 14. Overall, algorithms perform with ∇ values spanning from 23.7% to 203%.
1111 The algorithm here developed shows the smallest bias (∇ of 23.7%). On the other hand, the
1112 algorithms based on a constant b_{bp}^k value have lower performance, while Gra15, Beh05 and
1113 MV17 showed a systematic overestimation for low C_{phyto} values (Table 6 and Figure 14).

1114 The b_{bp} -based algorithms that use a single b_{bp}^k constant value (Beh05, Bre12 and Bel18)
1115 show similar σ_{Δ} . Specifically, Bre12 has a b_{bp}^k twice as high as the value used in Beh05,
1116 while Bel18 has a value nearly three times larger than the Beh05 value. ∇ ranges from 23.7%
1117 (Bel18), 80.1% (Bre12), up to 130.3% (Beh05) giving different performance in relation to
1118 the selection of the single value.

1119 In the case of Beh05, a source of discrepancy leading to such a bias may be linked to the
1120 b_{bp} input from the algorithm as also highlighted in Bellacicco *et al.* [2016]. Indeed, the
1121 original equation was derived using the GSM algorithm [Garver and Siegel, 2002] for
1122 obtaining a relationship between Chl and b_{bp} and thus for the b_{bp}^k estimation, instead of the
1123 QAA, here applied. In addition, it has recently been found that Raman scattering (here
1124 corrected prior the QAA application) plays a fundamental role for the retrievals of b_{bp} in very
1125 clear waters because, if not corrected, it generates overestimation of b_{bp} [Pitarch *et al.*, 2016;
1126 Westberry *et al.*, 2013].

1127 However, these results highlight how the use of a single b_{bp}^k determine a strong
1128 overestimation of satellite C_{phyto} , in line with the results of MV17 (Table 6 and Figure 14).
1129 The comparison between our approach and previously published models underlines the

1130 necessity to account for the spatio-temporal variability of the b_{bp}^k for satellite-based C_{phyto}
1131 calculation. However, because of Bre12 and Bel18 show relatively low bias, such approaches
1132 might be used alternatively to our new approach here proposed as it may occur in the most
1133 oligotrophic waters (e.g. subtropical gyres areas). Note that in Bre12, the b_{bp}^k was at 470 nm
1134 so that the overestimation on C_{phyto} can decrease if the b_{bp}^k is reported to the band of 443 nm.

1135 The empirical algorithms of Gra15 and MV17 give the worst results in terms of ∇ . In case
1136 of Gra15, this is the consequence of the few in-situ measurements used for the algorithm
1137 definition, located in the Equatorial Pacific Ocean and in the Atlantic Ocean. Extrapolation
1138 of their algorithm, developed at 470 nm, to 443 nm, might have had some effect as well.
1139 Regarding MV17, the results found are consistent with what reported in their Tables 3 and 5.

1140 MV17 reported that the geographical distribution of the in-situ data, the median C_{phyto}
1141 concentrations and also the carbon-to-chlorophyll ratios are in agreement with previous
1142 observations in oligotrophic oceanic conditions [*Graff et al., 2015; Maranon et al., 2014*].
1143 The fact that the in-situ C_{phyto} dataset used here contains only data from picophytoplankton
1144 population must be taken into account for the interpretation of the algorithm evaluation. In
1145 most oligotrophic waters (OWCs from 1 to 6), δ shows the lowest values for all the
1146 algorithms, with a maximum value of 14.4 mg m^{-3} (Table 6). However, when analyzing the
1147 ∇ index, the C_{phyto} computed by considering the spatio-temporal variability of b_{bp}^k is the most
1148 efficient (Table 6) except for OWCs 5 and 6 where Bel18 shows the best performance.
1149 Indeed, within OWCs where picophytoplankton is expected to be dominating (OWCs from
1150 1 to 6), the performance of majority of the tested algorithms generally improves (Table 6)
1151 but for Gra15 and MV17. As expected, there is a decrease in the accuracy for all the
1152 algorithms in the OWCs from 7 to 13. This could be influenced by the dominance of nano-
1153 and micro-phytoplankton (cell size $> 2\mu\text{m}$) in these waters, while the picophytoplankton
1154 contribute in low relative proportions [*Martinez-Vicente et al., 2017*]. Since most of the b_{bp}
1155 signal is due to particles with equivalent diameters larger than that those associated with
1156 picophytoplankton [*Organelli et al., 2018*], a decreasing performance of the b_{bp} -based
1157 algorithms in water dominated by larger cells (e.g. nano- and micro-phytoplankton) was
1158 expected (Table 6).

1159
1160
1161
1162

Table 6: Basic statistics for each single approach. For each single approach are reported the bias (δ ; in mg m^{-3}), the standard deviation (σ_{Δ} ; in mg m^{-3}) and the relative percentage bias (∇ ; in %). Note that the OWCs 1 and 2 are grouped in one single class to increase number of observations available for the statistics. This is also replicated for the OWCs 11-12 and 13.

OWCs	Number of obs.	This Study			Bel18			Gra15			Beh05			Bre12			MV17		
		δ	σ_{Δ}	∇	δ	σ_{Δ}	∇	δ	σ_{Δ}	∇	δ	σ_{Δ}	∇	δ	σ_{Δ}	∇	δ	σ_{Δ}	∇
1:2	19	-1.9	5.0	5.0	-2.6	6.2	8.8	9.2	6.0	259.4	5.2	6.2	174.9	0.7	6.2	78.0	8.5	7.2	250.0
3	30	-1.8	3.4	-13.9	-2.8	3.4	-27.3	9.1	3.3	211.8	5.0	3.4	129.8	0.5	3.4	38.2	7.9	3.9	189.2
4	51	-1.7	7.6	18.2	-2.1	8.5	31.3	9.4	8.6	193.1	5.7	8.5	141.1	1.2	8.5	77.0	10.7	10.6	213.6
5	76	-1.6	6.6	8.1	-3.2	7.9	5.9	8.2	7.7	122.8	4.6	7.9	85.8	0.03	7.9	39.2	9.4	9.2	142.6
6	49	1.5	7.9	38.7	-0.2	9.3	26.9	11.0	9.0	131.4	7.6	9.3	99.7	3.1	9.3	57.2	14.4	11.4	163.0
7	59	-1.9	8.4	-0.4	0.2	9.5	13.7	11.3	9.0	102.9	8.0	9.5	76.0	3.5	9.5	39.7	15.2	12.4	132.3
8	44	-0.5	10.1	38.8	5.2	10.3	73.6	15.8	9.9	164.99	13.0	10.3	139.9	8.5	10.3	101.3	23.3	13.1	220.0
9	26	1.9	11.8	39.2	13.8	12.3	113.4	23.6	12.2	181.5	21.6	12.3	166.5	17.1	12.3	135.5	36.4	14.6	259.1
10	27	4.6	16.5	103.4	17.2	17.6	206.2	26.7	17.0	278.2	25.0	17.6	266.2	20.4	17.6	231.2	41.1	22.6	396.0
11:13	15	3.2	14.0	51.7	12.2	15.2	148.8	22.6	15.2	275.9	20.0	15.2	241.4	15.5	15.2	187.3	31.5	30.4	355.6
1:6	225	-1.0	6.8	14.0	-2.2	7.9	12.0	9.3	7.7	164.0	5.6	7.9	114.7	1.1	7.9	54.8	10.7	9.6	178.4
7:13	171	0.5	11.8	36.6	7.3	13.7	86.5	17.8	13.0	173.8	15.1	13.7	150.7	10.6	13.7	113.3	26.0	18.4	235.4
All	396	-0.4	9.2	23.7	1.9	11.8	44.2	13.0	11.1	168.2	9.7	11.8	130.3	5.2	11.8	80.1	17.3	16.0	203.0

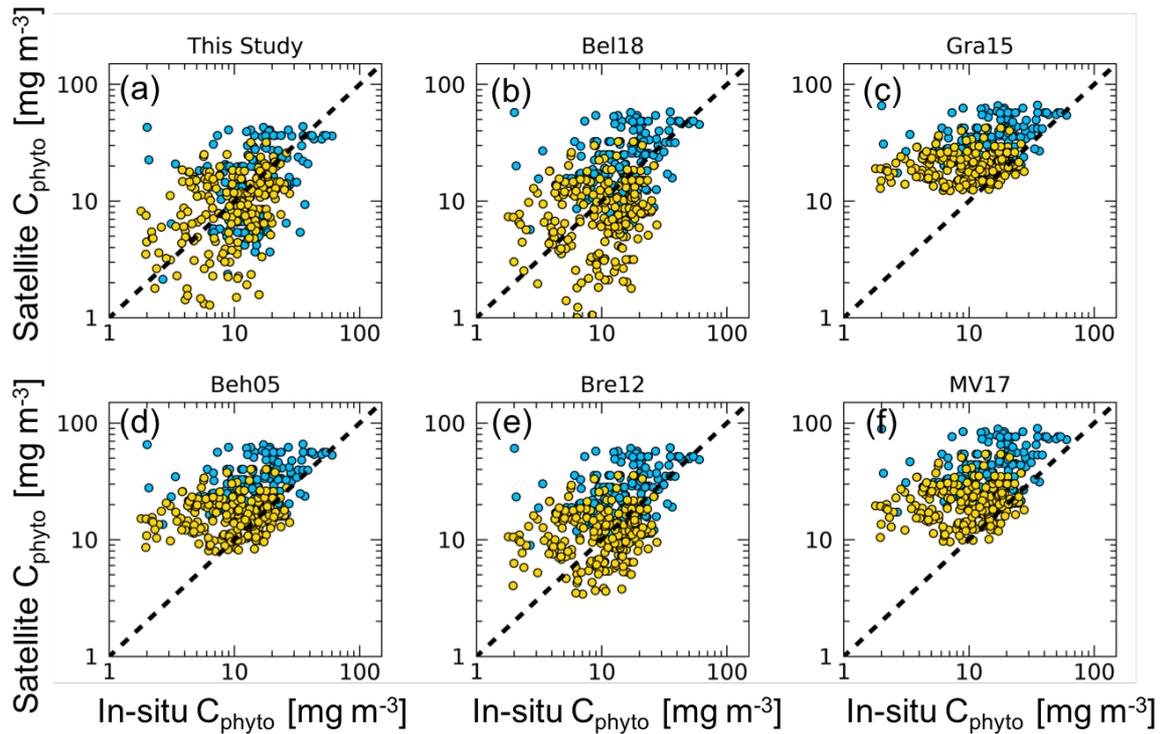


Figure 14: Match-ups between satellite C_{phyto} versus in-situ C_{phyto} data for each algorithm ($N=396$). The dashed line represents the 1:1 ratio. Gold dots represent in-situ C_{phyto} data in correspondence of OWCs 1-6; blue dots represent in-situ C_{phyto} data for OWCs 7-13. Statistics are reported in Table 6.

Spatial and Temporal Distribution of C_{phyto}

The global mean climatology of C_{phyto} (Figure 15a) agrees with the expected geographical distribution [Martinez-Vicente *et al.*, 2017]. C_{phyto} highs are found in the most productive regions such as the high latitudes regions and coastal areas as well as in upwelling systems such as off Mauritania and Perù, the Benguela and Kuroshio currents, and along the equatorial belt of the Pacific Ocean (Figure 15a). C_{phyto} lows are found in mid-latitude areas such as the subtropical gyres or the Eastern Mediterranean Sea and Black Sea. C_{phyto} varies within the same order of magnitude of recent estimates done by satellite data [Martinez-Vicente *et al.*, 2017] and biogeochemical models [Arteaga *et al.*, 2016]. The two main oceanic regimes as defined by Siegel *et al.* [2005] are also highlighted from our analysis: “photoacclimation-dominance” and “biomass-dominance”. The former is typical of oligotrophic areas (e.g. subtropical gyres) where the variability of Chl is uncoupled with biomass and the process of acclimation to light and nutrients drives Chl variations [Siegel *et al.*, 2005]. In these areas, C_{phyto} shows the lowest mean values ($<10 \text{ mg C m}^{-3}$). On the other hand, the latter is typical of most productive areas where Chl and b_{bp} strongly covary [Westberry *et al.*, 2010; Dall’Olmo *et al.*, 2012]. The high Chl- b_{bp} co-variability indicates that particles (and biomass) covary with phytoplankton abundance, while the physiological photoacclimation process plays a secondary role in determining the Chl variations. E.g. in the North Atlantic and Southern Oceans, C_{phyto} shows the highest mean values ($>30 \text{ mg C m}^{-3}$).

Living Planet Fellowship

The amplitude of seasonal variations (i.e., interannual variability) of C_{phyto} is shown in Figure 15b. Most of the variability is observed at high latitudes (e.g. Baltic Sea, Irish Sea, Norwegian Sea, North Sea), in the North Pacific and North Atlantic Oceans, as well as in the Southern Ocean, as σ_{std} values are higher than 12 mg C m^{-3} . The North Atlantic and Southern Oceans are characterized by high seasonal variations due to intense productive months followed by unproductive periods [Westberry *et al.*, 2016; Balch *et al.*, 2018] (supplementary materials on published paper). In the North Atlantic Ocean, C_{phyto} starts to increase in May reaching the maximum values in June and July ($> 50 \text{ mg C m}^{-3}$) following the spring bloom. Then, C_{phyto} decreases to the lowest values for the remaining of the year. In the Southern Ocean, C_{phyto} follows the typical seasonal cycle with high values during the most productive months which occur from November to March ($> 50 \text{ mg C m}^{-3}$). Highs are mostly observed in the Antarctic Circumpolar Current (ACC) and the Patagonia Shelf. Later on, there is a decrease in C_{phyto} from May to September (supplementary materials on published paper). Additionally, the northwestern Indian Ocean, known to be an Oxygen Minimum Zone, as well as the Benguela upwelling system also show significant variability in C_{phyto} , with σ_{std} greater than 12 mg C m^{-3} . Conversely, mid-latitude areas (e.g. in the subtropical gyres), display much lower variability, with σ_{std} less than 2 mg C m^{-3} .

Figure S14 (supplementary materials on published paper) clearly shows how the single b_{bp}^k of Beh05 is located at the lowest limit of possible values, confirming what reported in Bellacicco *et al.* [2018], while the Bel18 and Bre12 single values are representative of a wider area. In the subtropical gyres, b_{bp}^k does not significantly deviate from the coefficient found by Beh05, but in the most productive regions (e.g., North Atlantic and Southern Oceans, high latitudes seas and coastal upwelling regions), they largely differ with great impact on C_{phyto} estimations. This discrepancy in C_{phyto} is still valid in case of the use of Bel18 and Bre12 single values, but for other areas (e.g. in the subtropical gyres or in the Mediterranean Sea), the use of the single value of Bel18 or Bre12 yields an underestimation of C_{phyto} , while in the productive Arctic and Southern Oceans C_{phyto} is generally overestimated. This is also indicated in the seasonal b_{bp}^k variation as reported in the Figures S2-S13 (see for details the published paper). In the North Atlantic Ocean, b_{bp}^k shows high values during the spring bloom and low values from December to April. However, in this area, the b_{bp}^k is only a small fraction of the entire b_{bp} signal as the consequence of the high variability in the b_{bp} and phytoplankton abundance thus following the phytoplankton dynamics. In the Southern Ocean, and especially the region influenced by the ACC and Patagonia Shelf, b_{bp}^k shows highs from December to April (i.e., due to coccoliths) and lows in the period May-September in accordance also with the C_{phyto} seasonal cycle [Bellacicco *et al.*, 2019]. In less productive seas, the seasonal cycle of b_{bp}^k is smooth, suggesting low NAP variations. In these areas, the seasonal b_{bp}^k change is weak throughout the year but also dominates the b_{bp} signal [Bellacicco *et al.*, 2019]. These areas are characterized by limited nutrients availability determining low C_{phyto} , mostly of picophytoplankton class, high bacterial and detrital concentrations (Heywood *et al.*, 2006; Terran *et al.*, 2006; Grob *et al.*, 2007; Organelli *et al.*, 2020). It follows that the use of a varying

Living Planet Fellowship

b_{bp}^k enables accounting for its seasonal variations in the C_{phyto} computation, giving the model more flexibility in relation to the change of the biogeochemical and trophic conditions.

Caveats of the C_{phyto} algorithm

Three main caveats of the algorithm have to be pointed out. The first one is that C_{phyto} assumes a constant scaling factor, SF, equal to 13,000 mg C m⁻² following *Behrenfeld et al.* [2005]. This can be an additional possible source of error in the satellite retrievals of C_{phyto} . Therefore, one important future effort should be to investigate a refined scaling factor relating b_{bp} to C_{phyto} coupled with the b_{bp}^k space-time variability as presented in this work. With this perspective, additional laboratory work should be done to evaluate if change in SF values can affect the C_{phyto} estimations.

The second caveat of the algorithm is that it relies on a tight relationship between b_{bp} and Chl that is influenced by the algorithms (including the atmospheric correction) used for Chl and b_{bp} retrievals as well as environmental conditions. In some areas of the subtropical gyres, the Chl- b_{bp} relationship showed $S < 0.95$ (Figure S2-S13 on published paper) and r smaller than or close to 0 as also reported by *Bellacchio et al.* [2018]. The main reason of such relationship may be the photoacclimation process, known to be dominant in such areas over the year that introduces variability in Chl that is unrelated to biomass. In these areas, our refined method must be applied with caution. With this perspective, one future challenge is to solve this limitation in b_{bp}^k computation, and consequently, in the satellite C_{phyto} . Indeed, as the subtropical gyres cover about 60% of the global ocean, computation of C_{phyto} and Chl: C_{phyto} can be of great importance to understand the oceanic biogeochemical cycles. Conversely, the remaining areas of global ocean (e.g. North Atlantic Ocean) present significant positive r and highly significant S (Figure S2-S13; see published paper) throughout the year indicating that either b_{bp} or Chl can be used for determining the phytoplankton dynamics and distribution, as expected in open-ocean waters. In these areas, the algorithm performance is robust.

The third caveat is that the algorithm validation is restricted only to an in-situ C_{phyto} data associated with picophytoplankton carbon. It determines that the algorithm performance has to be interpreted with cautions in those areas where other phytoplankton size classes dominate. This means that, currently, it is not possible to define the accuracy for OWCs from 7 to 13. Indeed, Table 6 shows that there is a decrease in the accuracy for all the algorithms in the OWCs where nano- and micro-phytoplankton dominates. With this perspective, one future need is to improve the in-situ C_{phyto} dataset with new measurements representative of all the phytoplankton size classes.

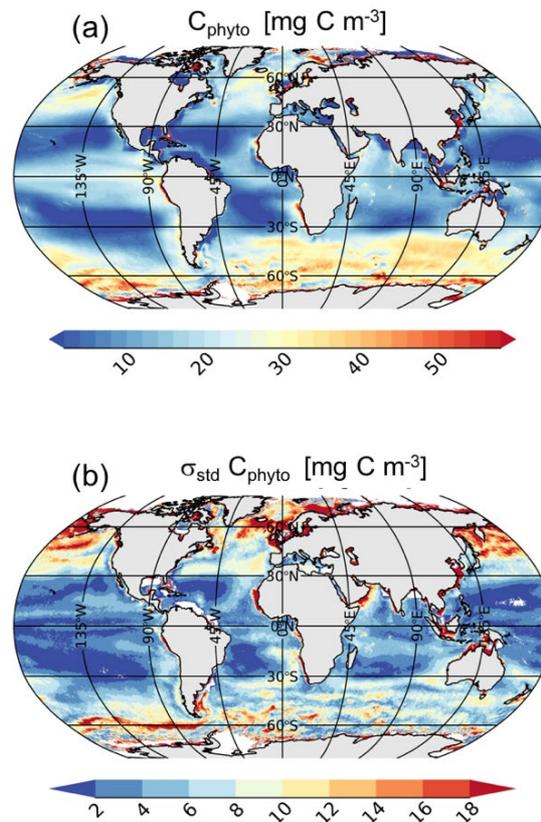


Figure 15: Annual global mean (a) and standard deviation (b) of C_{phyto} . Note that at high latitudes, the number of satellite observations used for the mean and standard deviation computations are lower than those used at low or mid-latitudes due to the winter darkness (i.e. high sun zenith angle).

Task #4

Here, the context, method and the performed worked about Task #4 are presented. The main goal of this task is to classify components of the single (or coupled) time-series into trends, oscillatory modes, and noise.

Scientific context

As discussed above, the temperature is one of the main environmental forcing that affects phytoplankton production. It has been well documented that ocean warming, mostly due to anthropogenic activities, determines an expansion of the low-Chl and low-productivity areas, such as the oligotrophic gyres [Martinez *et al.*, 2009; Boyce *et al.*, 2010]. Satellite observations revealed an inverse relationship between inter-annual and inter-decadal changes in Sea Surface Temperature (SST) and surface phytoplankton Chl concentration [Behrenfeld *et al.*, 2015]. It implies that Chl tends to decrease when temperature increases, and *viceversa*. Bellacicco *et al.*, [2016] showed the impact of SST onto phytoplankton seasonality in the Mediterranean Sea. Since Chl is used for phytoplankton biomass, with the ongoing ocean warming, the increase of

Living Planet Fellowship

SST can determine a continuously decrease of phytoplankton stocks and productivity in different areas of the ocean. Due to the explained impact of physiological processes on Chl space-borne estimations [Siegel et al., 2014], a deeper study on the relationship between both Chl and the C_{phyto} with SST is needed. An analysis of trend and oscillatory modes of parameters along multi-years of satellite observations should be an important step for the comprehension of the future ocean productivity.

Additionally, the estimation and evaluation of the weight and changes of phytoplankton physiological signal, $\text{Chl}:C_{\text{phyto}}$, based on long-term satellite observations may be decisive in respect to one of the future research ESA challenges which is to study the response of the marine ecosystem and associated ecosystem services to natural and anthropogenic changes (*i.e.* global warming). Even if we are able to distinguish changes in phytoplankton biomass from changes in physiology [Behrenfeld et al., 2015], we currently do not know the inter-annual variability of physiological process and its temporal cycles, and we are also unable to describe how and where those signals have been increased or decreased in the different areas of the ocean.

Data

The full European Space Agency (ESA) Ocean Colour – Climate Change Initiative (OC-CCI) time series (1997-2019) of global daily R_{rs} and Chl version 4.2 data at 4 km resolution was downloaded from the ESA OC-CCI FTP server (<https://esa-oceancolour-cci.org>; <ftp://oceancolour.org/occci-v4.2/geographic/netcdf/daily/rrs/>). ESA OC-CCI data products are the result of the merging of SeaWiFS, MERIS, MODIS and VIIRS observations in which the inter-sensor biases are removed. The version 4.2 includes the latest NASA reprocessing R2018.0 that mostly accounts for the ageing of the MODIS sensor. The ESA OC-CCI provides the daily R_{rs} data and associated uncertainty maps in terms of bias and root mean square error. In this work, for each daily R_{rs} , the bias was also extracted and then corrected pixel-by-pixel, as recommended in the Product User Guide on the ESA CCI website. Chl was estimated with a blending of the OCI (as implemented by NASA, itself a combination of CI and OC4), the OC5 and the OC3 algorithms. For each daily Chl value, the bias was also extracted and compensated at the pixel scale.

Daily b_{bp} maps at 443 nm were computed from daily R_{rs} for the same period (1997-2019) by applying the Quasi-Analytical Algorithm (QAA) [Lee et al., 2002] with prior correction for Raman scattering [Pitarch et al., 2020; Bellacicco et al., 2020]. The accuracy of the QAA in retrieving b_{bp} was fully demonstrated. In this work, the QAA algorithm is thus selected for its high efficiency in the b_{bp} retrievals as well as to show consistency and coherency with the OC-CCI programme in which the QAA is the designated algorithm [Brewin et al., 2015, Sathyendranath et al., 2020].

Then, daily C_{phyto} maps were produced from daily b_{bp} for the same period (1997-2019) following the method developed by Bellacicco et al. [2020]. The datasets were then under-sampled to 25 km and 100km resolution to resolve only the broader oceanographic scales of variability.

Living Planet Fellowship

Lastly, monthly data of Secchi Disk depth (z_{SD}) at same resolution were used to complement the information from the ocean colour data for the specific period of 1998-2018. For more details about the dataset see *Pitarch et al.* [2021].

The full ESA CCI SST time series (1997-2019) at 4 km resolution was downloaded from the ESA CCI website (<https://climate.esa.int/en/projects/sea-surface-temperature/>), and then remapped at same resolution of ocean colour data (25km and 100km).

In this task, the period of analysis is restricted to 1998-2018 that is the temporal window common to all the satellite products here used.

Method

Gaps Filling

The ocean colour data contained gaps. Gaps were filled using a Multi-Channel Singular Spectral Analysis (M-SSA) technique which is a non-parametric spectral estimation method relying on data only [*Kondrashov and Ghil*, 2006; 2010]. This technique is not based on a priori parametrized family of probability distribution. The method uses temporal correlation to fill in the missing data and represents a generalization of the *Beckers and Rixen* [2003] spatial empirical orthogonal functions-(EOFs) based reconstruction. *Kondrashov and Ghil* [2006] have demonstrated that an increased number of gaps yields the same effect as an increase of the noise in the measurements. Two different inputs are required to apply M-SSA for satellite field reconstruction: window-length (W) and components (M). Both depend on the characteristics of the time-series, and need to be accurately defined to avoid any bias in the reconstructed fields. The W represents the length of the sliding window (expressed in number of observations) used in the M-SSA in order to identify the leading components of the time-series [*Ghil et al.*, 2002; *Kondrashov and Ghil*, 2006; 2010]. Diversely, M is the number of eigen-functions used for signal reconstruction. Here, we applied the M-SSA to monthly ocean colour data (Chl, b_{bp} , C_{phyto} and z_{SD}) at 100km resolution. SST dataset is excluded due to it does not contain gaps. The M-SSA was applied on using $W=20$ and $M=5$ following the recommendations listed in *Ghil et al.* [2002] and *Kondrashov et al.* [2005, 2010] (Figure 16). These settings are compatible with the properties of the time series hereby analyzed, taking into account long, mid- and short-term variations. For more details about the mathematical equations and theoretical principles at the base of the M-SSA method, see *Ghil et al.* [2002] and *Kondrashov et al.* [2006, 2010].

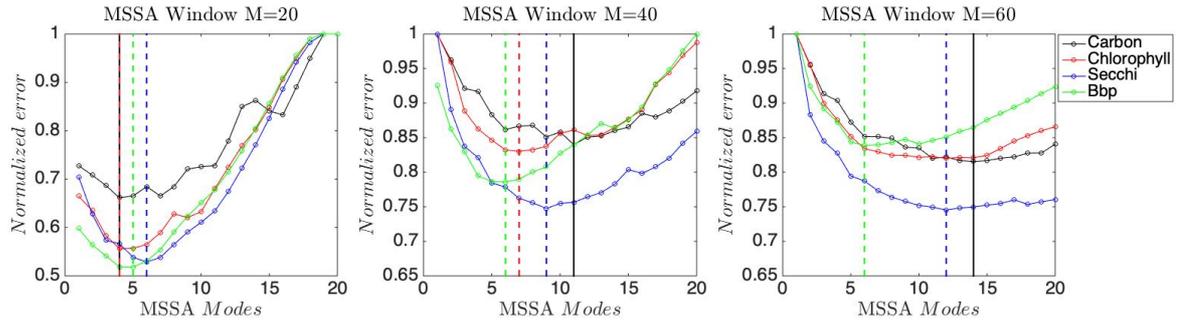


Figure 16: MSSA normalized error in relation to the selection of models (M) and windows (W). The $W=20$ and $M=5$ gives the best results in terms of *minimum* error in the reconstruction.

Trends and oscillatory modes computations

As already argued, the primary goal of the ESA OC CCI project is the construction of a time series for the calculation of oceanic trends and oscillations. Here, we used the lower spatial resolution product at 100 km resolution over the global ocean. A trend analysis on a global scale is performed. Seasonal decomposition was made pixel-wise with the publicly available BEAST code [Zhao *et al.*, 2019]. The Sen's slope [Sen, 1968] of the de-seasonalized signal was calculated and the result was tested for significance at 95 % [Mann, 1945; Kendall, 1975]. In order to extract information about the different oscillatory modes (i.e., annual, seasonal, noise), M-SSA technique was applied following Marullo *et al.* [2011] and Groth *et al.* [2017].

Results

Overview of ocean colour and SST datasets (1998-2018)

The first step about is to perform a qualitative description of the datasets. Figure 17 shows that the use of satellite-derived Chl may suffers from differences in sensor combination and related-processing along the entire time-series. Indeed, it has already been highlighted how time series resulting from merged R_{rs} measurements might be slightly impacted by changes in the combination of operating satellites from which the Chl has been derived as argued in Garnesson *et al.* [2019] (Table 7).

Period	Mission
January 1998 - February 2002	SeaWiFS
February/June 2002 - December 2010	SeaWiFS, MERIS, MODIS-Aqua
January 2011 - December 2011	MERIS, MODIS-Aqua
January 2012 - April 2012	MERIS, MODIS-Aqua, VIIRS
April 2012 - December 2019	MODIS-Aqua, VIIRS

Living Planet Fellowship

Table 7: Time periods of the satellite missions used to generate the ESA CCI Ocean Colour database version 4.2 at 0.042° horizontal resolution. Please note that the SeaWiFS mission started in September 1997 but for evaluation purposes we considered here only complete years (from 1998 to 2019).

This rapid Chl decrease from 2010 could be also explained by: impact of coastal pixels, different flags used in processing chains and different atmospheric corrections. However, when coastal pixels (here pictured with white colour; Figure 18) are removed, the rapid Chl increases from 2002-2004 does not occur, while the rapid Chl decreases from 2010 still persists either if with a lower drop (Figure 18b). These peculiar temporal dynamics do not occur in case of C_{phyto} or Z_{SD} . This critical issue has to be investigated in the next future in order to explain which are the main reasons at the base of such increases/decreases (e.g. processing chain, atmospheric correction, sensors merging, algorithms, climate change). On the reverse, SST time series confirms the well-known temporal dynamics [Pisano *et al.* 2020]. This first analysis determined that there are some doubts in the use of Chl data to study inter-annual oscillations. However, next step has been a spectral analysis to detect and characterize the main recurrent cycles in the time series.

Figure 19 shows one example of the general pattern observed for all ocean colour products (i.e. Chl, C_{phyto} or Z_{SD}). C_{phyto} dataset shows no significant cycles except annual, seasonal and coupled modes (e.g. 12, 6, 3 months) (Figure 19). All the other lower-term cycles can be defined as noise (peaks are under the red dots in Figure 19). It follows that it is crucial to understand which is the impact of the annual cycle for all of these parameters. Figure 20 shows the percent of variance explained by annual cycle (i.e. 12 months/1 year). There is a general consistence between ocean colour products. Z_{SD} shows values of variance greater than 70% at low and mid latitudes with respect to b_{bp} or C_{phyto} across many areas. Chl shows high seasonality in case of Mediterranean Sea, North Atlantic Ocean and low latitudes. In such areas, the annual cycle explains over the 70% of variance. In the subtropical gyres, the annual cycle has a reduce impact, but it is well known that in those areas there is not a marked and recurrent seasonal cycle. The subtropical gyres are sometimes defined as ocean deserts [Irwin *et al.*, 2009]. In addition to all these features, one has to take in mind that C_{phyto} product has caveats that prevent its use in case of oligotrophic areas; e.g. subtropical gyres (see Task #3). **The combination of both Chl (i.e. unexplained rapid increase/decreases) and C_{phyto} (i.e. three caveats; Task 3) limits plus the absence of recurrent significant cycles prevent an accurate and useful analysis on multi-annual oscillations of the Chl: C_{phyto} ratio against SST.** However, other important new insights can be derived from such time-series: the trends specifically in the period 1998-2018.

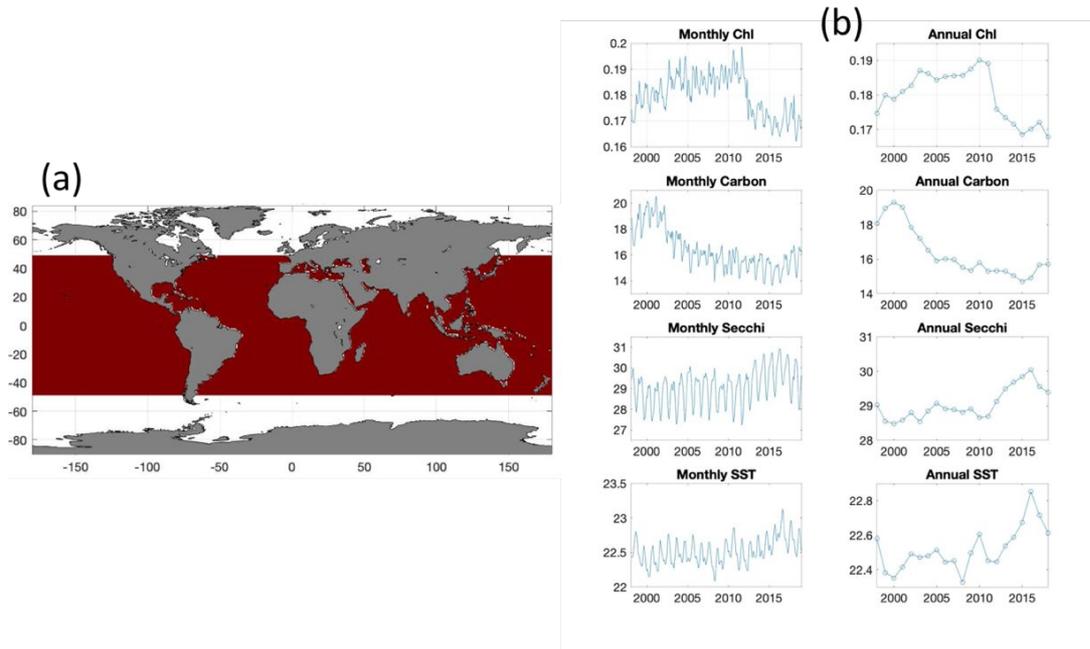


Figure 17: Area of study (a); monthly and annual satellite time-series (b). High and low latitudes pixels were removed from the analysis due to the winter darkness (i.e. high sun zenith angle).

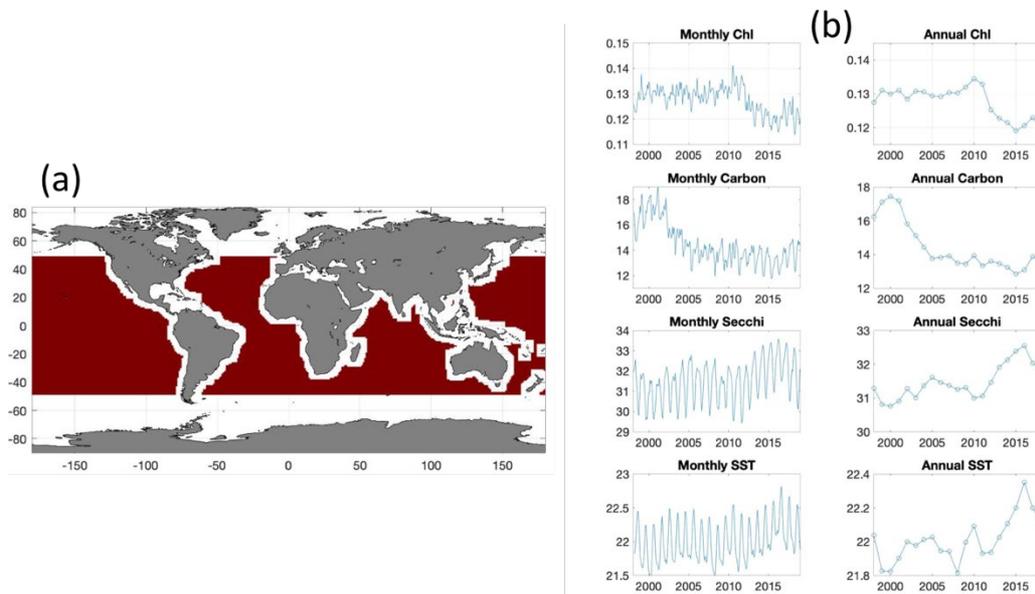


Figure 18: Area of study without coastal pixels, depicted in white (a); monthly and annual satellite time-series without coastal pixels (b). High and low latitudes pixels were removed from the analysis due to the winter darkness (i.e. high sun zenith angle).

Living Planet Fellowship

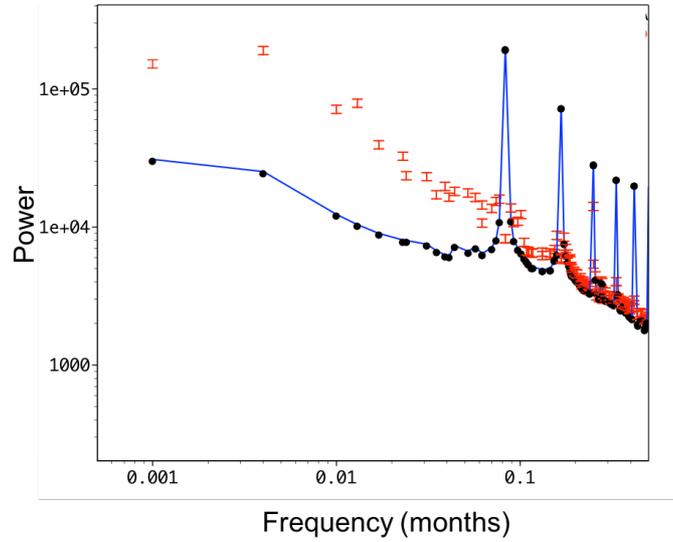


Figure 19: M-SSA eigenvalues spectra of C_{phyto} . Red dots indicate the red noise signal.

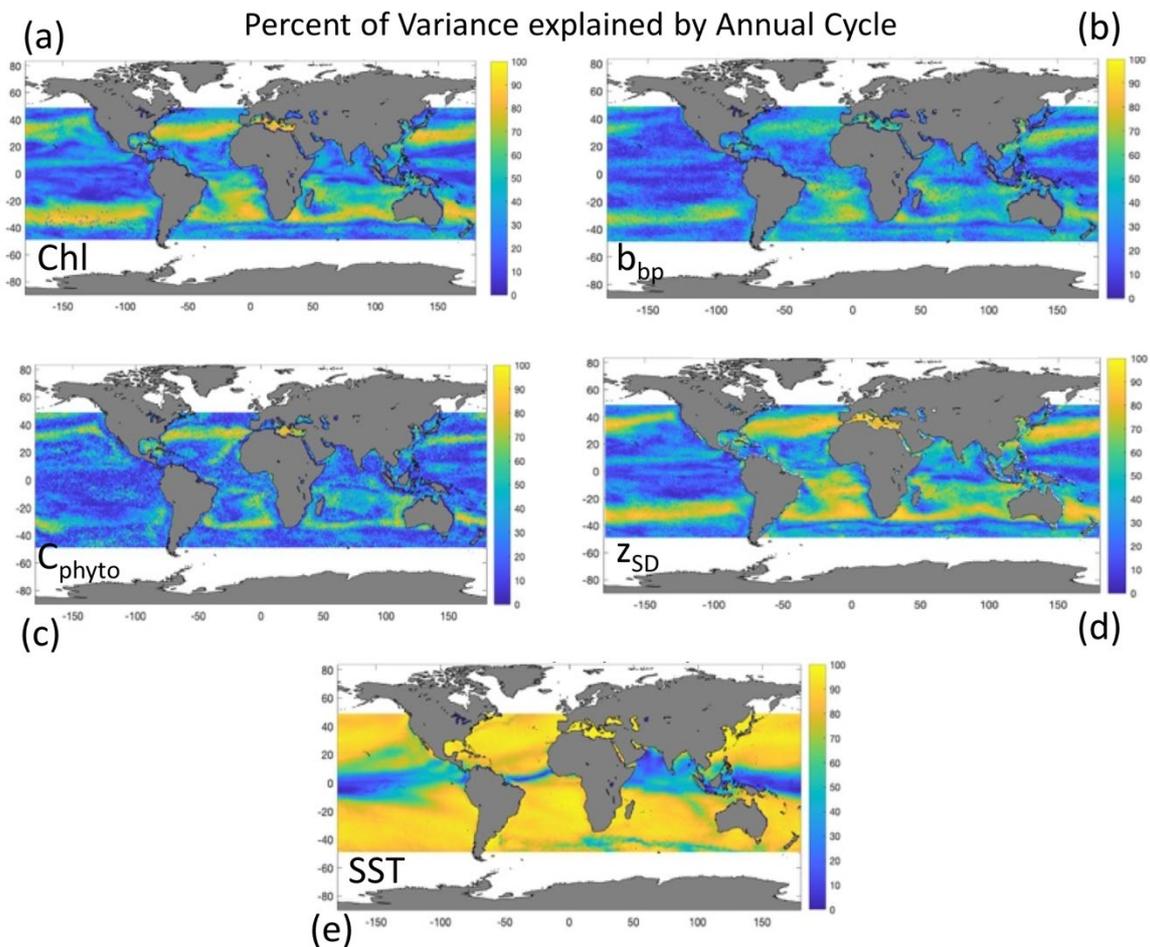


Figure 20: Percent of variance explained by the annual cycles: a) Chl; b) b_{bp} ; c) C_{phyto} ; d) Z_{SD} ; SST. High and low latitudes pixels were removed from the analysis due to the winter darkness (i.e. high sun zenith angle).

Trend of ocean colour and SST data (1998-2018)

Trends of Chl are shown in Figures 21. Pixels where significance is lower than 95% were masked out (in white). Local Chl trends vary in the range $\pm 0.01 \text{ mg m}^{-3} \text{ yr}^{-1}$. Note that a negative trend means a reduction of Chl each year, while positive trend means an increase. Negative trends are mainly observed in Tropical areas ($0\text{-}20^\circ \text{ N/S}$) and at the equator in the Pacific Ocean. Significant positive trends are found at high latitudes, and especially in the Northern Hemisphere as well as in the Southern Ocean (e.g. Patagonia Shelf). No significant trends characterize subtropical areas. This evaluation shows similar features to those found by *Gregg and Rousseaux* [2014] and references therein based on SeaWiFS and MODIS observations between 1998 and 2012, though with some differences across the northern Hemisphere, especially in the North Atlantic Ocean, and on the magnitude of the trends. *Gregg and Rousseaux* [2014] did not include Arctic Sea in their analysis, and a longer time series as evaluated here can significantly have an impact on the estimates of the trends at the global scale.

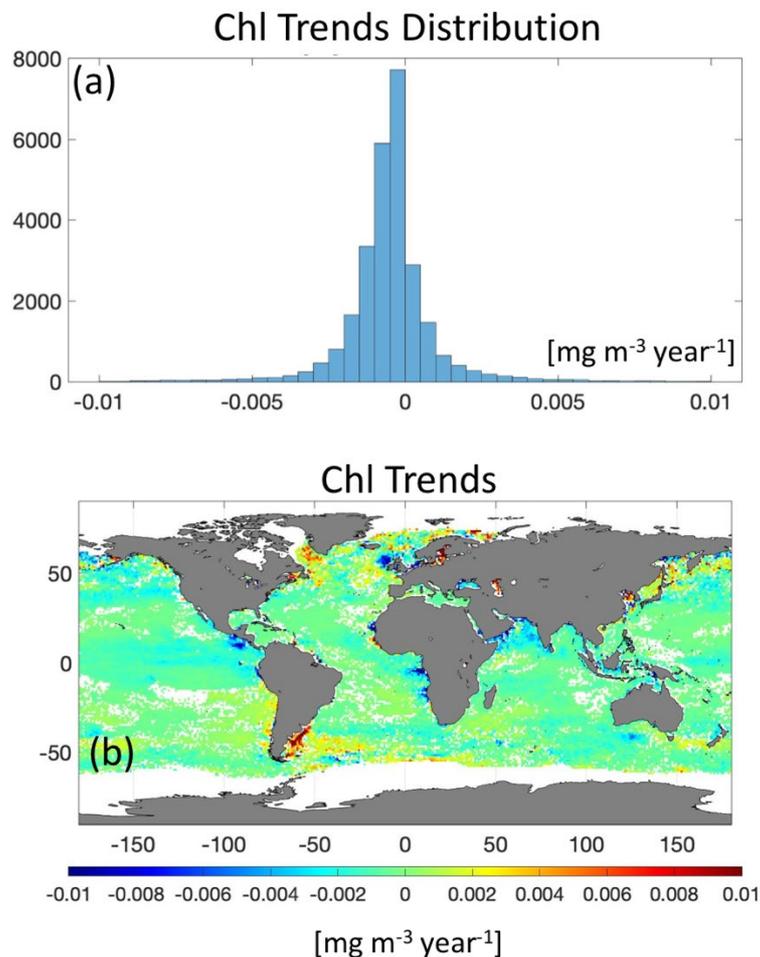


Figure 21: Statistical distribution of Chl trends (a) and trend map of Chl for the ESA CCI Ocean Colour database version 4.2, using monthly values from 1998 to 2018 at 0.1° horizontal resolution (b). Trend values are significant with $p\text{-value}\% \geq 95\%$ (i.e., $p\text{-value} < 0.05$). Pixels with $p\text{-value}\% < 95\%$ were masked out (white pixels).

Living Planet Fellowship

The variable Chl included in the ESA CCI database is less responsive in capturing ESM's trends than other Ocean Colour products [Dutkiewicz *et al.*, 2019]. It has been argued that current Chl retrieval algorithms might not adequately take into account possible modifications in water optical properties due to climate change [Dutkiewicz *et al.*, 2019]. Trends might also be affected by changes in satellite combinations from which the Chl product has been derived [Hammond *et al.*, 2018; Garnesson *et al.*, 2019]. However, the ESA CCI Ocean Colour Chl dataset has been recommended for detecting trends in phytoplankton phenology [Henson *et al.*, 2017]. These issues have to be taken into account also for the other satellite-derived ocean colour products.

Figure 22 shows trends in C_{phyto} . Pixels where statistical significance is lower than 95% were masked out as in case of Chl and successive satellite products. Local trends vary in the range $\pm 1.5 \text{ mg m}^{-3} \text{ yr}^{-1}$. Note that a negative trend means a reduction of C_{phyto} each year, while positive trend means an increase. Significant negative trends are mainly observed in mid (i.e. North Atlantic and Pacific Oceans) and high latitudes (i.e. Arctic and Southern Oceans). The Indian sector of the Southern Ocean also shows significant negative trends of C_{phyto} . No trends characterize tropical and subtropical areas. This evaluation shows different features with respect to Chl. Indeed, Chl distribution show a general absence of trends, while the C_{phyto} distribution shows a general negative trend as confirmed by the Figure 22a. Because satellite C_{phyto} is a recent satellite products, **there is actually a necessity to increase the in-situ data to improve its robustness and resolve the actual caveats.**

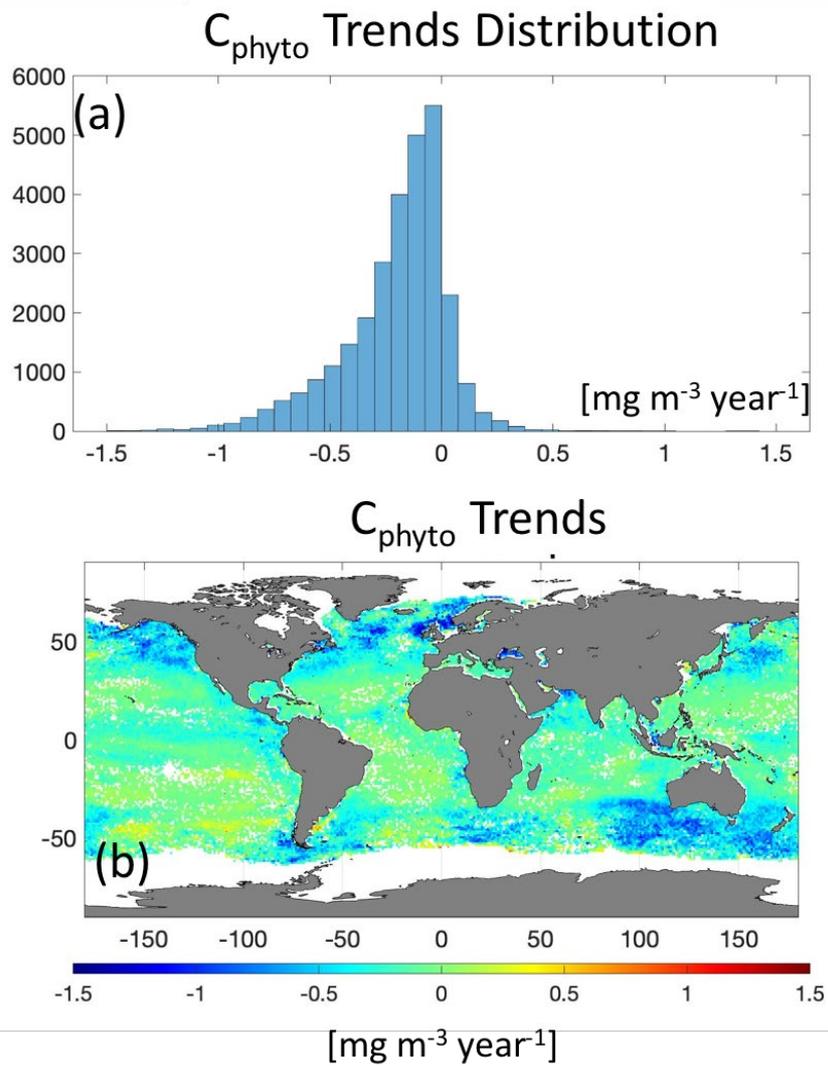


Figure 22: Statistical distribution of C_{phyto} trends (a) and map (b) using monthly values from 1998 to 2018 at 0.1° horizontal resolution (b). Trend values are significant with $p\text{-value} \geq 95\%$ (i.e., $p\text{-value} < 0.05$). Pixels with $p\text{-value} < 95\%$ were masked out (white pixels).

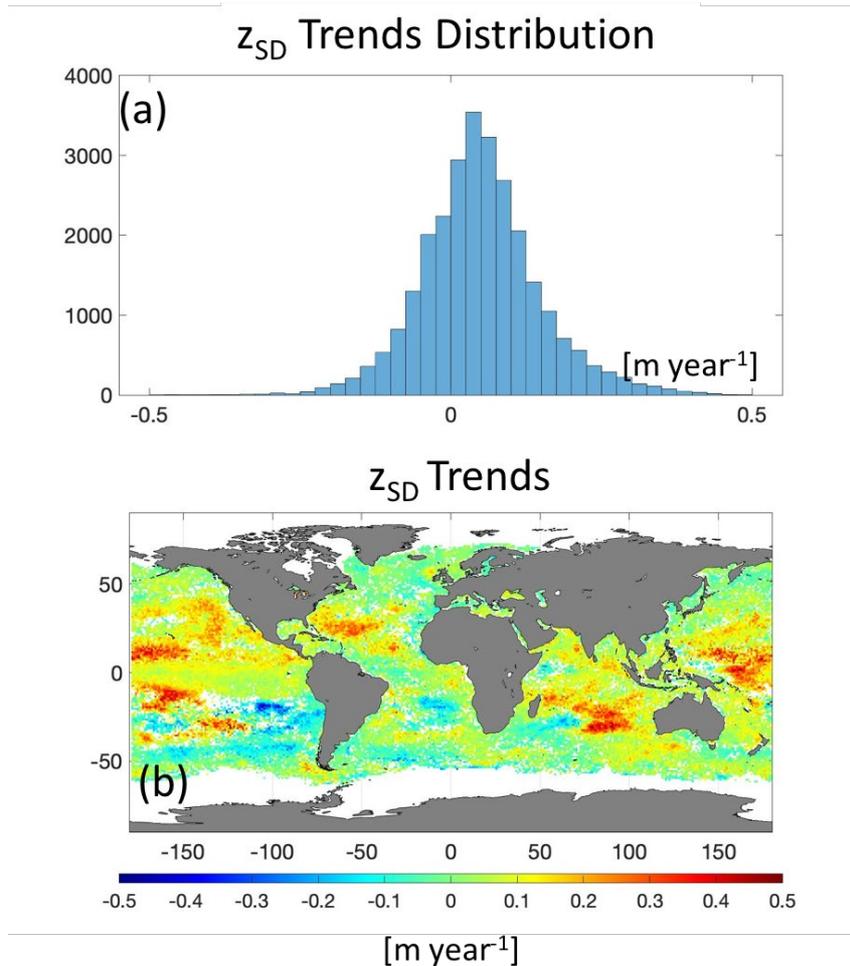


Figure 23: Statistical distribution of z_{SD} trend (a) and map of z_{SD} using monthly values from 1998 to 2018 at 0.1° horizontal resolution (b). Trend values are significant with $p\text{-value}\% \geq 95\%$ (i.e., $p\text{-value} < 0.05$). Pixels with $p\text{-value}\% < 95\%$ were masked out (white pixels). For more details about z_{SD} see Pitarch et al. [2021].

Figure 23 shows trends in z_{SD} . Local trends vary in the range $\pm 0.5 \text{ m yr}^{-1}$. Note that a negative trend means a reduction of z_{SD} (less transparency) each year, while positive trend means an increase (more transparency). Significant positive trends are mainly observed in mid (i.e. North Atlantic and Pacific Oceans) and low latitudes (i.e. subtropical gyres). The Indian Ocean shows significant positive trends of z_{SD} . No trends characterize many areas. This evaluation shows different features with respect to Chl and C_{phyto} . There is a general positive trend as also highlighted by the upper panel. It means that there is a general increment of transparency over the global ocean, and especially in the subtropical gyres as in the North Atlantic Ocean. Future challenge aims to understand such increase.

Figure 24 shows SST trends over the global ocean. As in case of z_{SD} , a general increase of SST is detected, especially in case of Arctic Ocean and Pacific Ocean. Future challenges need to connect the differences trends of ocean colour products (i.e. z_{SD}) and SST in order to understand and explain the observed patterns.

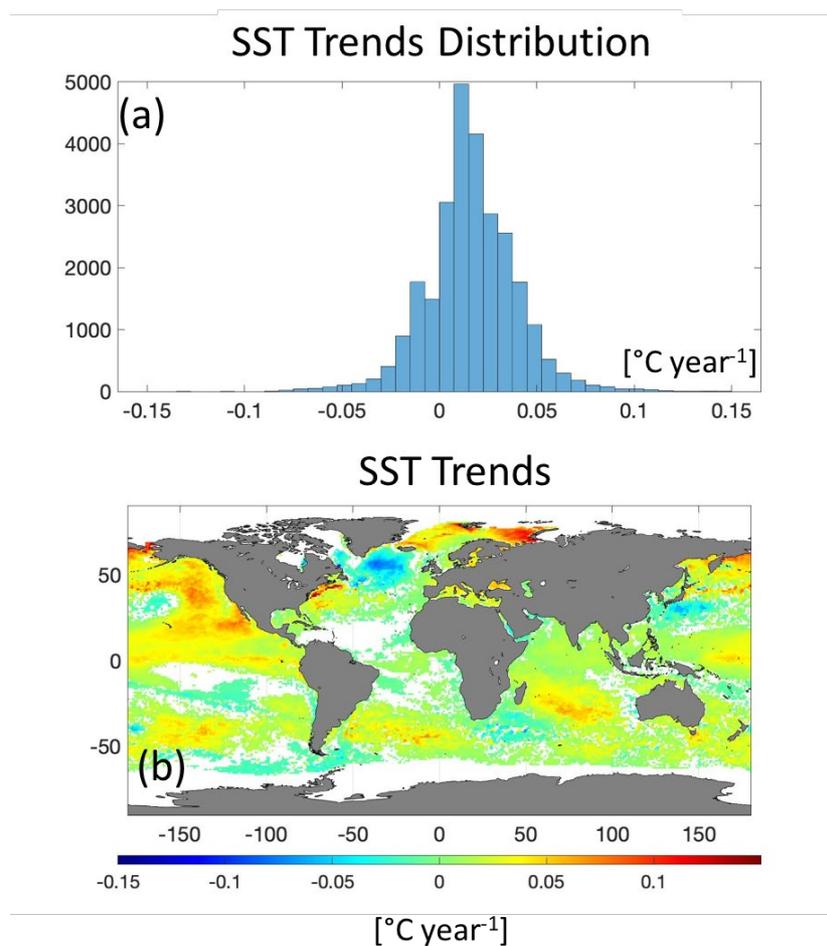


Figure 24: Statistical distribution of SST (a); SST trend for the ESA CCI dataset, using monthly values from 1998 to 2018 at 0.1° horizontal resolution (b). Trend values are significant with $p\text{-value} \geq 95\%$ (i.e., $p\text{-value} < 0.05$). Pixels with $p\text{-value} < 95\%$ were masked out (white pixels).

4 CONCLUSIONS, RECOMMENDATIONS AND CHALLENGES

[*Note that in bold are highlighted the future challenges and urgency for ocean colour products and studies].

Task #1

The main findings of the task and their relevance for ocean color studies are summarized here:

1. Raman scattering compensation of R_{rs} prior to the application of the QAA significantly reduces errors in the retrieval of b_{bp} with respect to in situ b_{bp} . Inclusion of this processing step in operational schemes is recommended.
2. The QAA-derived b_{bp} from in situ radiometry has negligible biases with respect to in situ b_{bp} .
3. $CCIR_{rs}$ shows low biases but higher RMS differences with respect to in situ data, that could be excessive for the monitoring of natural change over short periods. Here, the standardization of in situ radiometry protocols is highly encouraged [IOCCG, 2019], in

order to reduce the errors when in situ datasets formed by multiple contributors are merged and used for R_{rs} matchup analysis.

4. In part as a consequence of the findings above, QAA-derived b_{bp} from CCI R_{rs} displays negligible biases respect to in situ b_{bp} , with moderately low RMS errors.
5. In situ radiometry-derived spectral backscattering slope (η) has low predictive value as compared to η derived from b_{bp} matchups. In this context, the impact of using the best fitted curve instead of the widely-used expression is negligible, thus validating the application of the latter without its retuning.

Notwithstanding these results, **one future challenge should be to evaluate the impact of two other sources of inelastic scattering before the application of QAA on R_{rs}** : (i) red fluorescence, caused by chlorophyll, that usually plays an important role around the peak close to 685 nm; and (ii) the blue fluorescence, caused by CDOM, that can be relevant close to the peak at 425 nm [Haltrin *et al.*, 1997].

In addition, **there is the need of increasing the amount of spatial and spectral coverage of high-quality in situ b_{bp} observations**. As of today, available multispectral b_{bp} is limited to a small number of ship-borne data, or longer datasets but in fixed points (i.e., buoy). On the other hand, Biogeochemical-Argo floats cover large areas but their data are mainly given at a single band. Therefore, there is need to significantly increase the amount of b_{bp} data at multiple bands, seasons and geographical regions. **New technological developments on autonomous platforms will aid to enhance data density across many water types, to extend the CCI uncertainty derivation approach to b_{bp} as well, thus allowing the mapping of uncertainties for every b_{bp} product**.

Lastly, in situ b_{bp} measurements lag behind the standards on protocols and uncertainty characterization with respect to other quantities such as the radiometry [IOCCG, 2019]. **Only when in situ uncertainty-characterized datasets, from instrument characterization to deployment [Dall'Olmo *et al.*, 2012], become available, more detailed algorithm validation could be performed and this will help to better evaluate the influence of optically active constituents (e.g., CDOM, chlorophyll).**

Task #2

The main findings of the task and their relevance for ocean color studies are summarized here.

In this work, an extensive global dataset of Chl and b_{bp} (700) measurements acquired from Biogeochemical-Argo (BGC-Argo) profiling floats was analyzed. Specifically, we investigated and describe the spatial, vertical and temporal variability of b_{bp}^k at global scale. The main results are:

1. b_{bp}^k shows a similar order of magnitude in both surface and euphotic layers, as previously published works based on ocean-colour data: ranging between 10^{-4} and 10^{-3} m^{-1} .

Living Planet Fellowship

2. In the surface layer, the b_{bp}^k increase from southern to the northern hemisphere, confirming what was found by *Bellacicco et al.* [2018] using ocean colour data.
3. In the surface layer of most productive areas (*e.g.* NASPG), the b_{bp}^k is only a small fraction of the total b_{bp} (< 20%), while in the oligotrophic waters, b_{bp}^k is the main contributor to the total b_{bp} (> 80%). In the euphotic layer of the oligotrophic areas, the b_{bp}^k has a lower contribution to the total b_{bp} (average value of 65%).
4. In the surface and euphotic layers, the b_{bp}^k shows strong seasonal variability in the main productive areas of the global ocean, such as NASPG and the Southern Ocean areas. b_{bp}^k has instead a weak temporal variability in the low productivity areas, such as the subtropical gyres. This is valid also for the b_{bp}^k estimations at the bottom layer.

The b_{bp}^k is a key parameter for satellite estimations of phytoplankton biomass in terms of carbon [*Behrenfeld et al.*, 2005, 2016; *Bellacicco et al.*, 2016, 2018, 2019; *Martinez-Vicente et al.*, 2017; *Westberry et al.*, 2008, 2016]. Recently, *Bellacicco et al.*, [2018] highlighted the difference (of around a factor of 2) in the phytoplankton carbon biomass estimation from space by using a b_{bp}^k variable in space, rather than a single value. Consequently, inclusion of this reported spatial-temporal and depth variations of b_{bp}^k into phytoplankton carbon models may help to improve their predictions from remote sensing data [*Martinez-Vicente et al.*, 2017] but also from BGC-Argo floats [*Mignot et al.*, 2014, 2018].

Remote optical-based predictions and interpretation of phytoplankton carbon models would also benefit from a better understanding of NAP composition and which particles generate the b_{bp} signal across the world's oceans. Indeed, submicron detrital particles have long been considered as the main source of b_{bp} [*Stramski et al.*, 2004]. However, *Organelli et al.* [2018] has highlighted that b_{bp} is mainly due to particles with diameters between 1-10 μm which may also include NAP and aggregates. This latter study thus opens the way to new questions on the sources of the open-ocean b_{bp} signal that are critical to improving our interpretation of open-ocean b_{bp} .

Future research challenges should therefore be directed to: **(i)** understand the drivers of the observed spatio-temporal variability and explore the composition of NAP across the world's oceans and how it influences the b_{bp} and b_{bp}^k signal; **(ii)** study the impact on biogeochemistry of b_{bp}^k , *e.g.* on the particles assemblage in different ocean trophic regimes (*i.e.* subpolar, subtropical); **(iii)** include b_{bp}^k spatial and temporal variability into phytoplankton carbon estimations from space and its connections with phytoplankton physiology; and most importantly **(iv)**: advance technology for (autonomous) optical measurements of NAP directly, for example by exploiting the birefringence properties of mineral particles such as calcite compounds [*Guay and Bishop*, 2002; *Bishop and Wood*, 2009], and acquire spectral angular scattering to better understand the influence of bubbles and plastics [*Zhang et al.*, 1998; *Twardowski et al.*, 2012].

Living Planet Fellowship

Task #3

The main findings of the task and their relevance for ocean color studies are summarized here.

A revisited version of the original b_{bp} -based algorithm for C_{phyto} retrievals from space [Behrenfeld *et al.*, 2005] has been proposed. The spatio-temporal variations of the background backscattering coefficient of non-algal particles (b_{bp}^k) is the crucial part of this refinement which builds upon a series of recent literature [Bellacicco *et al.*, 2016, 2018, 2019]. The main findings are:

1. The new C_{phyto} algorithm proposed here performs better than any previously published model, with a relative error of 24% with respect to a reference in-situ dataset.
2. The new algorithm shows the lowest error in C_{phyto} (14.0%) across most of the OWCs where the picophytoplankton population dominates. On the contrary, the highest errors (36.6%) occur in the OWCs 7-13 where larger phytoplankton cells are supposed to dominate.

However, we acknowledge that the refined algorithm, here presented, could be improved by addressing the caveats mentioned above. **Enlarge the databases of quality controlled and freely accessible in-situ C_{phyto} at the different size-classes (pico-, nano-, micro-), measured simultaneously with an array of optical variables (especially b_{bp}) could help to reduce uncertainties in the C_{phyto} retrievals from space. Improving the accuracy of satellite C_{phyto} could be of great importance to adding new knowledge on the contribution of the phytoplankton to particulate organic carbon as well as for the validation of ocean primary productivity and biogeochemical models at a wider scale [Westberry *et al.*, 2008; Dutkiewicz *et al.*, 2015].**

Task #4

The main findings of the task and their relevance for ocean color studies are summarized here.

The database of Chl, b_{bp} , C_{phyto} and z_{SD} is surely suitable for mean and variability studies, Earth System Model (ESM) assessment, as well as for detecting extreme climate events (e.g. ENSO). However, **to define linear and not-linear trends as well as multi-year oscillatory components, different critical issues are extrapolated. The use of satellite-derived Chl suffers from differences in sensor combination and related-processing along the entire time-series. Indeed, it has already been highlighted how time series resulting from merged R_{rs} measurements might be slightly impacted by changes in the combination of operating satellites from which the Chl has been derived [Garneison *et al.*, 2019].** In addition, it has been argued that **changes in the optical properties of the sea due to climate change might impact the performances of retrieval algorithms along the time series [Dutkiewicz *et al.*, 2019].** Moreover, **satellite C_{phyto} has critical caveats (e.g., absence of a large in-situ dataset for its validation) that prevent its use for multi-annual variability studies.** Such issues together with a reduced time series length (only 22 years) prevented their use to observe and characterize the multi-annual oscillations of derived parameter as the Chl: C_{phyto} with the correct

and sufficient accuracy and efficiency. Lastly, ocean colour products show the annual cycle as the main temporal oscillations. Future challenge has to be take into account these main findings in order to improve satellite products and derived $\text{Chl:C}_{\text{phyto}}$, and then going thought its relationship with SST data. This could help to shed light on the ongoing global warming and its impact on marine ecosystems.

5 ESSENTIAL REFERENCES

1. Antoine, D.; Chami, M.; Claustre, H.; d'Ortenzio, F.; Morel, A.; Bécu, G.; Gentili, B.; Louis, F.; Ras, J.; Roussier, E. BOUSSOLE: a joint CNRS-INSU, ESA, CNES, and NASA ocean color calibration and validation activity. 2006.
2. Antoine, D.; Guevel, P.; Desté, J.-F.; Bécu, G.; Louis, F.; Scott, A.J.; Bardey, P. (2008). The “BOUSSOLE” Buoy—A New Transparent-to-Swell Taut Mooring Dedicated to Marine Optics: Design, Tests, and Performance at Sea. *Journal of Atmospheric and Oceanic Technology*, 25, 968-989.
3. Antoine, D.; d'Ortenzio, F.; Hooker, S.B.; Bécu, G.; Gentili, B.; Tailliez, D.; Scott, A.J. (2008). Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean Sea (BOUSSOLE project). *Journal of Geophysical Research: Oceans*, 113.
4. Balch, W. M., Bates, N. R., Lam, P. J., Twining, B. S., Rosengard, S. Z., Bowler, B. C., ... & Rauschenberg, S. (2016). Factors regulating the Great Calcite Belt in the Southern Ocean and its biogeochemical significance. *Global Biogeochemical Cycles*, 30(8), 1124-1144.
5. Balch, W. M. (2018). The ecology, biogeochemistry, and optical properties of coccolithophores. *Annual review of marine science*, 10, 71-98.
6. Behrenfeld, M. J., Marañón, E., Siegel, D. A., & Hooker, S. B. (2002). Photoacclimation and nutrient-based model of light-saturated photosynthesis for quantifying oceanic primary production. *Marine Ecology Progress Series*, 228, 103-117.
7. Behrenfeld, M. J., and Boss, E. (2003). The beam attenuation to chlorophyll ratio: an optical index of phytoplankton physiology in the surface ocean? *Deep Sea Research Part I: Oceanographic Research Papers*, 50(12), 1537-1549.
8. Behrenfeld, M. J., Prasil, O., Babin, M., Bruyant, F. (2004). In search of a physiological basis for covariations in light-limited and light-saturated photosynthesis. *J. Phycol.* 40 (1), 4–25.
9. Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M. (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Global Biogeochemical Cycles*, 19(1).
10. Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., ... & Boss, E. S. (2006). Climate-driven trends in contemporary ocean productivity. *Nature*, 444(7120), 752.
11. Behrenfeld, M. J., Halsey, K. H., and Milligan, A. J. (2008). Evolved physiological responses of phytoplankton to their integrated growth environment. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1504), 2687–703.
12. Behrenfeld, M. J., O'Malley, R. T., Boss, E. S., Westberry, T. K., Graff, J. R., Halsey, K. H., Milligan, A. J., Siegel, D. and Brown, M. B. (2015). Revaluating ocean warming impacts on global phytoplankton. *Nature Climate Change*.

Living Planet Fellowship

13. **Bellacicco, M., Volpe, G., Colella, S., Pitarch, J., & Santoleri, R. (2016).** Influence of photoacclimation on the phytoplankton seasonal cycle in the Mediterranean Sea as seen by satellite. *Remote Sensing of Environment*, 184, 595-604.
14. **Bellacicco, M., Volpe, G., Briggs, N., Brando, V., Pitarch, J., Landolfi, A., ... & Santoleri, R. (2018).** Global Distribution of Non-algal Particles From Ocean Color Data and Implications for Phytoplankton Biomass Detection. *Geophysical Research Letters*, 45(15), 7672-7682.
15. **Bellacicco, M., Vellucci, V., Scardi, M., Barbieux, M., Marullo, S., & D'Ortenzio, F. (2019).** Quantifying the Impact of Linear Regression Model in Deriving Bio-Optical Relationships: The Implications on Ocean Carbon Estimations. *Sensors*, 19(13), 3032.
16. **Bellacicco, M., Cornec, M., Organelli, E., Brewin, R. J. W., Neukermans, G., Volpe, G., ... & Marullo, S. (2019).** Global Variability of Optical Backscattering by Non-algal particles From a Biogeochemical-Argo Data Set. *Geophysical Research Letters*, 46(16), 9767-9776.
17. **Bishop, J. K., & Wood, T. J. (2009).** Year-round observations of carbon biomass and flux variability in the Southern Ocean. *Global Biogeochemical Cycles*, 23(2).
18. **Bisson, K.M.; Boss, E.; Westberry, T.K.; Behrenfeld, M.J. (2019).** Evaluating satellite estimates of particulate backscatter in the global open ocean using autonomous profiling floats. *Opt. Express*, 27, 30191-30203.
19. **Boyce, D. G., Lewis, M. R., & Worm, B. (2010).** Global phytoplankton decline over the past century. *Nature*, 466(7306), 591.
20. **Bracaglia, M.; Volpe, G.; Colella, S.; Santoleri, R.; Braga, F.; Brando, V.E. (2019).** Using overlapping VIIRS scenes to observe short term variations in particulate matter in the coastal environment. *Remote Sensing of Environment*, 233, 111367
21. **Brewin, R.J.W.; Dall'Olmo, G.; Sathyendranath, S.; Hardman-Mountford, N.J. (2012).** Particle backscattering as a function of chlorophyll and phytoplankton size structure in the open-ocean. *Opt. Express*, 20, 17632-17652.
22. **Brewin, R.J.W.; Sathyendranath, S.; Müller, D.; Brockmann, C.; Deschamps, P.-Y.; Devred, E.; Doerffer, R.; Fomferra, N.; Franz, B.; Grant, M., et al. (2015).** The Ocean Colour Climate Change Initiative: III. A round-robin comparison on in-water bio-optical algorithms. *Remote Sensing of Environment*, 162, 271-294.
23. **Dubinsky, Z., Stambler, N. (2009).** Photoacclimation processes in phytoplankton: mechanisms, consequences, and applications. *Aquat. Microb. Ecol.* 56, 163–176.
24. **Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1997).** 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*, 187-200.
25. **Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1998).** A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. *Limnology and oceanography*, 43(4), 679-694
26. **Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., ... & Yiou, P. (2002).** Advanced spectral methods for climatic time series. *Reviews of geophysics*, 40(1), 3-1.
27. **Groth, A., Feliks, Y., Kondrashov, D., & Ghil, M. (2017).** Interannual variability in the North Atlantic ocean's temperature field and its association with the wind stress forcing. *Journal of Climate*, 30(7), 2655-2678.
28. **Guay, C. K., & Bishop, J. K. (2002).** A rapid birefringence method for measuring suspended CaCO₃ concentrations in seawater. *Deep Sea Research Part I: Oceanographic Research Papers*, 49(1), 197-210.

Living Planet Fellowship

29. Halsey, K. H., & Jones, B. M. (2015). Phytoplankton strategies for photosynthetic energy allocation. *Annual review of marine science*, 7, 265-297.
30. Huber, P. J. (1981). *Robust Statistics*. Hoboken, NJ: John Wiley & Sons, Inc.
31. Huot, Y.; Morel, A.; Twardowski, M.S.; Stramski, D.; Reynolds, R.A. (2008). Particle optical backscattering along a chlorophyll gradient in the upper layer of the eastern South Pacific Ocean. *Biogeosciences*, 5, 495-507.
32. Jackson, T.; Chuprin, A.; Sathyendranath, S.; Grant, M.; Zühlke, M.; Dingle, J.; Storm, T.; Boettcher, M.; Fomferra, N. Ocean Colour Climate Change Initiative (OC_CCI) –Interim Phase. Product User Guide. D3.4 PUG; 2019.
33. Kostadinov, T.S.; Siegel, D.A.; Maritorea, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *Journal of Geophysical Research: Oceans*, 114.
34. Kostadinov, T.S.; Milutinović, S.; Marinov, I.; Cabré, A. (2016). Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. *Ocean Sci.*, 12, 561-575, doi:10.5194/os-12-561-2016.
35. Laws, E.A., Bannister, T.T. (1980). Nutrient-and light-limited growth of *Thalassiosira fluviatilis* in continuous culture, with implications for phytoplankton growth in the ocean. *Limnol. Oceanogr.* 25 (3), 457–473.
36. Lee, Z., Carder, K. L., & Arnone, R. A. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Applied optics*, 41(27), 5755-5772.
37. Lee, Z.; Hu, C.; Shang, S.; Du, K.; Lewis, M.; Arnone, R.; Brewin, R. (2013). Penetration of UV-visible solar radiation in the global oceans: Insights from ocean color remote sensing. *Journal of Geophysical Research: Oceans*, 118, 4241-4255.
38. Lee, Z., & Huot, Y. (2014). On the non-closure of particle backscattering coefficient in oligotrophic oceans. *Optics Express*, 22(23), 29223-29233.
39. Lee, Z. Update of the Quasi-Analytical Algorithm (QAA_v6). Available online: http://www.ioccg.org/groups/Software_OCA/QAA_v6_2014209.pdf.
40. Loisel, H.; Bosc, E.; Stramski, D.; Oubelkheir, K.; Deschamps, P.-Y. (2001). Seasonal variability of the backscattering coefficient in the Mediterranean Sea based on satellite SeaWiFS imagery. *Geophysical Research Letters*, 28, 4203-4206.
41. MacIntyre, H.L., Kana, T.M., Geider, R.J., (2000). The effect of water motion on short-term rates of photosynthesis by marine phytoplankton. *Trends Plant Sci.* 5 (1), 12–17.
42. MacIntyre, H., Kana, T. (2002). Photoacclimation of photosynthesis irradiance response curve and photosynthetic pigments in microalgae and cyanobacteria. *Journal of Phycology*, 38(1), 17-38.
43. Martinez, E., Antoine, D., D’Ortenzio, F., & Gentili, B. (2009). Climate-driven basin-scale decadal oscillations of oceanic phytoplankton. *Science*, 326(5957), 1253-1256.
44. Martinez-Vicente, V., Dall’Olmo, G., Tarran, G., Boss, E., & Sathyendranath, S. (2013). Optical backscattering is correlated with phytoplankton carbon across the Atlantic Ocean. *Geophysical Research Letters*, 40(6), 1154-1158.
45. Martínez-Vicente, V., Evers-King, H., Roy, S., Kostadinov, T. S., Tarran, G. A., Graff, J. R., ... & Röttgers, R. (2017). Intercomparison of ocean color algorithms for picophytoplankton carbon in the ocean. *Frontiers in Marine Science*, 4, 378.
46. Mélin, F. (2011). Comparison of SeaWiFS and MODIS time series of inherent optical properties for the Adriatic Sea. *Ocean Sci.*, 7, 351-361.
47. Mélin, F.; Chuprin, A.; Grant, M.; Jackson, T.; Sathyendranath, S. Ocean Colour Climate Change Initiative (OC_CCI) - Phase Two. Ocean Colour Data Bias Correction and Merging, D2.6. 2016, 35.

Living Planet Fellowship

48. Mélin, F.; Sclép, G. (2015). Band shifting for ocean color multi-spectral reflectance data. *Opt. Express*, 23, 2262-2279.
49. Mignot, A., Ferrari, R., & Claustre, H. (2018). Floats with bio-optical sensors reveal what processes trigger the North Atlantic bloom. *Nature communications*, 9(1), 190.
50. Moré, J. J., & Sorensen, D. C. (1983). Computing a trust region step. *SIAM Journal on Scientific and Statistical Computing*, 4(3), 553-572.
51. Moore, C.M., Suggett, D.J., Hickman, A.E., Kim, Y.-N., Tweddle, J.F., Sharples, J., Geider, R.J., Holligan, P.M., (2006). Phytoplankton photoacclimation and photoadaptation in response to environmental gradients in a shelf sea. *Limnol. Oceanogr.* 51 (2), 936–949.
52. Morel, A. (1988). Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). *Journal of Geophysical Research: Oceans*, 93(C9), 10749-10768.
53. Morel, A., & Ahn, Y. H. (1991). Optics of heterotrophic nanoflagellates and ciliates: A tentative assessment of their scattering role in oceanic waters compared to those of bacterial and algal cells. *Journal of Marine Research*, 49(1), 177-202.
54. Morel, A., & Maritorena, S. (2001). Bio-optical properties of oceanic waters: A reappraisal. *Journal of Geophysical Research: Oceans*, 106(C4), 7163-7180.
55. Organelli, E., Barbieux, M., Claustre, H., Schmechtig, C., Poteau, A., Bricaud, A., ... & Leymarie, E. (2017). Two databases derived from BGC-Argo float measurements for marine biogeochemical and bio-optical applications. *Earth System Science Data*, 9, 861-880.
56. Organelli, E.; Dall’Olmo, G.; Brewin, R.J.W.; Tarran, G.A.; Boss, E.; Bricaud, A. (2018). The open-ocean missing backscattering is in the structural complexity of particles. *Nature Communications*, 9, 5439, doi:10.1038/s41467-018-07814-6.
57. Pitarch, J., **Bellacicco, M.**, Volpe, G., Colella, S., & Santoleri, R. (2016). Use of the quasi-analytical algorithm to retrieve backscattering from in-situ data in the Mediterranean Sea. *Remote Sensing Letters*, 7(6), 591-600.
58. Poteau, A., Boss, E., & Claustre, H. (2017). Particulate concentration and seasonal dynamics in the mesopelagic ocean based on the backscattering coefficient measured with Biogeochemical-Argo floats. *Geophysical Research Letters*, 44(13), 6933-6939.
59. Sathyendranath, S.; Brewin, R.J.W.; Brockmann, C.; Brotas, V.; Calton, B.; Chuprin, A.; Cipollini, P.; Couto, A.B.; Dingle, J.; Doerffer, R., et al. (2019). An Ocean-Colour Time Series for Use in Climate Studies: The Experience of the Ocean-Colour Climate Change Initiative (OC-CCI). *Sensors*, 19, 4285.
60. Siegel, D. A., Behrenfeld, M. J., Maritorena, S., McClain, C. R., Antoine, D., Bailey, S. W., ... & Eplee Jr, R. E. (2013). Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission. *Remote Sensing of Environment*, 135, 77-91.
61. Siokou-Frangou, I., Christaki, U., Mazzocchi, M. G., Montresor, M., Ribera d’Alcalà, M., Vaqué, D., & Zingone, A. (2010). Plankton in the open Mediterranean Sea: a review.
62. Slade, W.H.; Boss, E. (2015). Spectral attenuation and backscattering as indicators of average particle size. *Appl. Opt.*, 54, 7264-7277.
63. Stramski, D.; Kiefer, D.A. (1991). Light scattering by microorganisms in the open ocean. *Progress in Oceanography*, 28, 343-383,
64. Stramski, D.; Boss, E.; Bogucki, D.; Voss, K.J. (2004). The role of seawater constituents in light backscattering in the ocean. *Progress in Oceanography*, 61, 27-56
65. Thomalla, S.J.; Ogunkoya, A.G.; Vichi, M.; Swart, S. (2017). Using Optical Sensors on Gliders to Estimate Phytoplankton Carbon Concentrations and Chlorophyll-to-Carbon Ratios in the Southern Ocean. *Frontiers in Marine Science*, 4, doi:10.3389/fmars.2017.00034.

66. Twardowski, M.S.; Boss, E.; Macdonald, J.B.; Pegau, W.S.; Barnard, A.H.; Zaneveld, J.R.V. (2001). A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in case I and case II waters. *Journal of Geophysical Research: Oceans*, 106, 14129-14142.
67. Twardowski, M., Zhang, X., Vagle, S., Sullivan, J., Freeman, S., Czerski, H., ... & Kattawar, G. (2012). The optical volume scattering function in a surf zone inverted to derive sediment and bubble particle subpopulations. *Journal of Geophysical Research: Oceans*, 117(C7).
68. Werdell, P.J.; Bailey, S.W. (2005). An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, 98, 122-140.
69. Werdell, P.J.; Franz, B.A.; Lefler, J.T.; Robinson, W.D.; Boss, E. (2013). Retrieving marine inherent optical properties from satellites using temperature and salinity-dependent backscattering by seawater. *Opt. Express*, 21, 32611-32622, doi:10.1364/OE.21.032611.
70. Valente, A.; Sathyendranath, S.; Brotas, V.; Groom, S.; Grant, M.; Taberner, M.; Antoine, D.; Arnone, R.; Balch, W.M.; Barker, K., et al. (2019). A compilation of global bio-optical in situ data for ocean-colour satellite applications – version two. *Earth Syst. Sci. Data*, 11, 1037-1068.
71. Westberry, T., Behrenfeld, M. J., Siegel, D. A., & Boss, E. (2008). Carbon-based primary productivity modeling with vertically resolved photoacclimation. *Global Biogeochemical Cycles*, 22(2).
72. Westberry, T.K.; Boss, E.; Lee, Z. (2013). Influence of Raman scattering on ocean color inversion models. *Appl. Opt.*, 52, 5552-5561.
73. Westberry, T. K., Schultz, P., Behrenfeld, M. J., Dunne, J. P., Hiscock, M. R., Maritorena, S., ... & Siegel, D. A. (2016). Annual cycles of phytoplankton biomass in the subarctic Atlantic and Pacific Ocean. *Global Biogeochemical Cycles*, 30(2), 175-190.
74. Zhang, X., Lewis, M., & Johnson, B. (1998). Influence of bubbles on scattering of light in the ocean. *Applied Optics*, 37(27), 6525-6536.
75. Zhang, X.; Hu, L.; He, M.-X. Scattering by pure seawater. (2009). Effect of salinity. *Opt. Express*, 17, 5698-5710, doi:10.1364/OE.17.005698.
76. Zheng, G.; Stramski, D.; Reynolds, R.A. (2014). Evaluation of the Quasi-Analytical Algorithm for estimating the inherent optical properties of seawater from ocean color: Comparison of Arctic and lower-latitude waters. *Remote Sensing of Environment*, 155, 194-209.

6 PUBLICATIONS

Papers published within PhysioGlob:

1. Leonelli, F. E., **Bellacicco, M.**, Pitarch, J., Marullo, S. The North Atlantic Ocean is getting bluer and bluer (*in preparation*).
2. J. Pitarch, **M. Bellacicco**, S. Marullo, H.J. van der Woerd (2021). Global maps of Forel-Ule index, hue angle and Secchi disk depth derived from twenty-one years of monthly ESA-OC-CCI data, *Earth System Science Data*, 1-17 (*accepted*).
3. **Bellacicco, M.***, Pitarch, J., Organelli, E., Martinez-Vicente, V., Volpe, G., & Marullo, S. (2020). Improving the Retrieval of Carbon-Based Phytoplankton

Living Planet Fellowship

Biomass from Satellite Ocean Colour Observations. *Remote Sensing*, 12(21), 3640. (*corresponding author).

4. Pitarch, J., **Bellacicco, M.***, Organelli, E., Volpe, G., Colella, S., Vellucci, V and Marullo, S. (2020). Retrieval of particulate backscattering using field and satellite radiometry: assessment of the QAA algorithm. *Remote Sensing*, 12 (1), 77. (*corresponding author).
5. **Bellacicco, M.***, Cornec, M., Organelli, E., Brewin, R., Neukermans, G., Volpe, G., Barbieux, M., Poteau, A., Schmechtig, C., D’Ortenzio, F., Marullo, S., Claustre, H. and Pitarch, J. (2019). Global variability of optical backscattering by non-algal particles from a Biogeochemical-Argo dataset. *Geophysical Research Letters*, 46 (16), 9767-9776. (*corresponding author).
6. **Bellacicco, M.***, Vellucci, V., Scardi, M., Barbieux, M., Marullo, S., & D’Ortenzio, F. (2019). Quantifying the Impact of Linear Regression Model in Deriving Bio-Optical Relationships: The Implications on Ocean Carbon Estimations. *Sensors*, 19(13), 3032. (*corresponding author).

Other papers published:

1. Karam Mansour, Stefano Decesari, **Marco Bellacicco**, Salvatore Marullo, Rosalia Santoleri, Paolo Bonasoni, Maria Cristina Facchini, Jurgita Ovadnevaite, Darius Ceburnis, Colin O’Dowd, Matteo Rinaldi (2020). Linking Oceanic Biological Activity to Aerosol Chemical Composition and Cloud-Relevant Properties over the North Atlantic. *Journal of Geophysical Research: Atmospheres*. doi: 10.1029/2019JD032246.
2. Mansour, K., Decesari, S., **Bellacicco, M.**, Marullo, S., Santoleri, R., Bonasoni, P., Facchini, M. C., Ovadnevaite, J., Ceburnis, D., O’Dowd, C. and Rinaldi, M. (2019). Particulate methanesulfonic acid over the central Mediterranean Sea: Source region identification and relationship with phytoplankton activity. *Atmospheric Research*, 104837.
3. **Bellacicco, M.***, Vellucci, V., D’Ortenzio, F. and Antoine, D. (2019). Discerning dominant temporal patterns of bio-optical properties in the northwestern Mediterranean Sea (BOUSSOLE site). *Deep-Sea Research: Part I*, 148, 12-24. (*corresponding author).

Conferences:

1. 17 – 21 May 2021: 52nd Liege colloquium at Liege (Belgium). “Trends of color and transparency of the North Sea’s water, from in-situ and satellite data”. Pitarch J., Thewes D., Wollschlaeger J., Zielinski O., Stanev E., Pisano A., **Bellacicco M.**, Marullo S., Falcini F., Colella S., Sammartino M. (poster presentation).
2. 16 – 21 February 2020: Ocean Sciences Meeting 2020 at San Diego (USA). “Sea Surface Temperature and Air-Sea Interaction in the Mediterranean Region: Direct measurements, satellite estimates and model based assessments”. S. Marullo, G. di

Living Planet Fellowship

Sarra, V. Artale, **M. Bellacicco**, J. Pitarch, D. Sferlazzo, C. Yang, G. Pace, R. Santoleri (poster presentation).

3. 16 – 21 February 2020: Ocean Sciences Meeting 2020 at San Diego (USA). “*Global trends of transparency and color based on merged multi-sensor satellite data*”. J. Pitarch, **M. Bellacicco**, V. de Toma and S. Marullo (poster presentation).
4. 16 – 21 February 2020: Ocean Sciences Meeting 2020 at San Diego (USA). “*Towards the comprehension of the inter-annual physiological response of phytoplankton to global warming as seen from space: a refined phytoplankton biomass proxy*”. **M. Bellacicco**, J. Pitarch and S. Marullo (poster presentation).