# Phytoplankton composition from sPACE: requirements, opportunities, and challenges

Ivona Cetinić [a,b,1], Cecile S. Rousseaux [b,*,1], Ian T. Carroll [c,b], Alison P. Chase [d], Sasha J. Kramer [e], P. Jeremy Werdell [b], David. A. Siegel [f], Heidi M. Dierssen [g], Dylan Catlett [h], Aimee Neeley [b,i], Inia M. Soto Ramos [a,b], Jennifer L. Wolny [j], Natasha Sadoff [b,i], Erin Urquhart [b,i], Toby K. Westberry [k], Dariusz Stramski [l], Nima Pahlevan [m,i], Bridget N. Seegers [a,b], Emerson Sirk [b,i], Priscila Kienteca Lange [n,o], Ryan A. Vandermeulen [b,l,#], Jason R. Graff [k], James G. Allen [p], Peter Gaube [d], Lachlan I. W. McKinna [q], S.Morgaine McKibben [r], Caren E. Binding [s], Violeta Sanjuan Calzado [c,b], Michael Sayers [t]

[a] GESTAR II, Morgan State University, Baltimore, Maryland, USA

[b] Ocean Ecology Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

[c] GESTAR II - University of Maryland Baltimore County, Baltimore, Maryland, USA

[d] Applied Physics Laboratory, University of Washington, Seattle, Washington, USA

[e] Monterey Bay Aquarium Research Institute, Moss Landing, California, USA

[f] Earth Research Institute, University of California, Santa Barbara, Santa Barbara, California, USA

[g] Department of Marine Sciences, University of Connecticut, Groton, Connecticut, USA

[h] Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

[i] Science Systems and Applications, Inc., Greenbelt, Maryland, USA

[j] Office of Regulatory Science, Center for Food Safety and Applied Nutrition, US Food and Drug Administration, College Park, Maryland, USA

[k] Department of Botany and Plant Pathology, 2071 SW Campus Way, Oregon State University, Corvallis, OR, USA

[l] Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093-0238, USA

24   [m] Terrestrial Information Systems Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland,

25   USA

26   [n] Departamento de Meteorologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

27   [o] Blue Marble Space Institute of Science, Seattle, USA

28   [p] Department of Oceanography, School of Ocean and Earth Science and Technology, University of Hawai'i

29   at Mānoa, Honolulu, Hawai'i, USA

30   [q] 3Go2Q Pty Ltd., Sunshine Coast, Queensland, Australia

31   [r] NASA Postdoctoral Program/NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

32   [s] Water Science and Technology Directorate, Environment and Climate Change Canada, Burlington,

33   Ontario, Canada

34   [t] Michigan Tech Research Institute, Michigan Technological University. Ann Arbor, Michigan, USA

35   [#]Present address: National Marine Fisheries Service, National Oceanic and Atmospheric Administration,

36   Silver Spring, MD, USA

37   * Corresponding author

38   [1] Ivona Cetinić and Cecile S. Rousseaux contributed equally

39

# Abstract

41   Ocean color satellites have provided a synoptic view of global phytoplankton for over 25 years

42   through surface measurements of the concentration of chlorophyll a. While remote sensing of

43   ocean color has revolutionized our understanding of phytoplankton and their role in the oceanic

44   and freshwater ecosystems, it is important to consider both total phytoplankton biomass and

45   changes in phytoplankton community composition in order to fully understand the dynamics of

46   the aquatic ecosystems. With the upcoming launch of NASA's Plankton, Aerosol, Clouds, ocean

47    Ecosystem (PACE) mission, we will be entering into a new era of global hyperspectral data, and

48    with it, increased capabilities to monitor phytoplankton diversity from space. In this paper, we

49    analyze the needs of the user community, review existing approaches for detecting

50    phytoplankton community composition in situ and from space, and highlight the benefits that

51    the PACE mission will bring. Using this three-pronged approach, we highlight the challenges and

52    gaps to be addressed by the community going forward, while offering a vision of what global

53    phytoplankton community composition will look like through the "eyes" of PACE.

54    ## 1. Background

55    Phytoplankton are microscopic, photosynthetic organisms that inhabit all sunlit waters and

56    represent the first level of the aquatic food web. They play a vital role in the global ecosystem

57    where phytoplankton convert carbon dioxide into carbohydrates and oxygen through the process

58    of photosynthesis. The primary photosynthetic pigment in most phytoplankton species is

59    Chlorophyll a (*Chl a*), which has been commonly used as a proxy for phytoplankton biomass in

60    oceanographic research and monitoring programs for almost 100 years (Harvey 1934). Global

61    ocean color satellites (SeaWiFS, MODIS, VIIRS) have provided continuous global datasets of

62    aquatic *Chl a* since 1997 (O'Reilly et al. 1998), generating insight into spatial and temporal changes

63    in phytoplankton biomass (Behrenfeld et al. 2006; Gregg and Rousseaux 2014; Gregg et al. 2017;

64    McClain et al. 2022; Siegel et al. 2013). Despite the ubiquity of *Chl a*, taxonomic diversity in

65    phytoplankton is wide-ranging, spanning across eukaryotic and eubacterial domains. This

66    taxonomic diversity underpins vast morphological diversity among phytoplankton (e.g., Beardall

67    et al. 2009), including variation in pigment composition (e.g., Jeffrey et al. 2011), environmental

68    requirements (e.g., for light, nutrients, and temperature), trophic strategies (e.g., Mitra et al.

69 2016), and in the roles phytoplankton play in aquatic biogeochemistry (e.g., Dutkiewicz et al.

70 2020).

71 Knowledge of phytoplankton community composition (PCC) and its spatial and temporal

72 variability is vital for our understanding of the many aspects of aquatic ecosystems, as well as the

73 services those ecosystems provide. It is critical to assess water quality (including harmful algal

74 blooms - HABs) and functioning of the higher trophic levels including fisheries worldwide.

75 Knowledge about PCC can be used to manage and support the aquaculture industry, including to

76 infer optimal site selection and sustainable aquaculture development[1] (Anderson et al. 2009;

77 2019; Johnson et al. 2019; Snyder et al. 2017), to prevent economic losses due to the impact of

78 HABs (Smith and Bernard 2020; Wolny et al. 2020), or to support monitoring of potential impacts

79 that aquaculture might have on aquatic ecosystems (e.g., eutrophication; Gowen 1994) or on

80 natural phytoplankton population estimates (Schaeffer et al. 2015; Frieder et al. 2022).

81 Additionally, PCC can enhance the parameterization of end-to-end ecosystem models designed

82 to simulate biogeochemical, ecological, fishery, management, and socio-economic processes

83 within marine ecosystems (Turner et al. 2021; Caracappa et al. 2022).

84 Furthermore, to understand the effects of climate variability on aquatic biogeochemistry, the role

85 of the oceans as a sink of anthropogenic carbon, and the various effects of human pressure on

86 water resources, it is imperative that we improve our characterization of the food web's

87 foundation—the phytoplankton community—across the continuum of aquatic ecosystems.

88 Understanding the role that different phytoplankton groups play in the export of carbon,

---

[1] ShellGIS, an application developed to predict growth of numerous aquaculture species using remote sensing inputs, http://www.shellgis.com/examples/TFWMidMaine.html. ShellGIS team is part of the PACE early adopter program.

89    specifically in regard to ballasting, export depths and scales of remineralization (Cram et al. 2018;

90    Guidi et al. 2015), is crucial to our understanding of contemporary and future oceans. Currently,

91    there is also much debate as to the future of the ocean's role in absorbing anthropogenic carbon,

92    as well as the role that biology, specifically different phytoplankton types, will play in different

93    climate scenarios. Earth System Models (ESMs) are used to understand and forecast the role of

94    oceans in the global carbon budget, including the role of oceans in future climate scenarios (IPCC[2],

95    CMIP[3], etc). For most ESMs, understanding the role that biology will have in future carbon

96    sequestration relies on our understanding of oceanic PCC, in part due to the relationship between

97    phytoplankton size and sinking rate (i.e., Guidi et al. 2015; Henson et al. 2022).

98    Whether categorized by taxa, function, or size, the particular mix of phytoplankton detectable by

99    in situ sampling is discernably variable across regions, seasons, and conditions; underscoring a

100   strong need for studying what we collectively call PCC through ocean remote sensing. During the

101   past 25 years, numerous studies have shown that space-based instruments can reach beyond

102   estimates of total phytoplankton biomass and quantify certain aspects of PCC (e.g., Bracher et al.

103   2017; IOCCG 2014). Such insights on PCC have mostly relied on *Chl a* or multi-band phytoplankton

104   absorption spectra (see Mouw et al. 2017 and references within), information that inherently

105   creates large uncertainties from the inverse problem standpoint, i.e., solving for various

106   components of PCC from a limited number of observations (Defoin-Platel and Chami 2007; Sydor

107   et al. 2004).

---

[2] IPCC stands for Intergovernmental Panel on Climate Change, https://www.ipcc.ch/
[3] CMIP stands for World Climate Research Programme Coupled Model Intercomparison Project, https://www.wcrp-climate.org/wgcm-cmip

108    As the needs of aquatic research surpassed the capabilities of current ocean color satellites, the

109    community responded with various instrument concepts, materialized in specific calls by the

110    National Research Council Earth Sciences Decadal Survey (National Research Council 2007) and

111    National Aeronautics and Space Administration (NASA, NASA 2010) among others (McClain et al.

112    2022; Muller-Karger et al. 2018). The need to resolve taxonomic components of PCC is one of the

113    main underlying motivations for NASA's Plankton, Aerosol, Cloud, ocean Ecology (PACE) mission

114    (Werdell et al. 2019). The PACE mission, to be launched no earlier than January 2024, will respond

115    to community needs for a highly calibrated, ocean-focused hyperspectral radiometer combined

116    with multi-angle polarimeters that satisfy multiple scientific needs, including better resolution of

117    phytoplankton diversity (McClain et al. 2022; Werdell et al. 2019).

118    PACE is the first global ocean and atmosphere hyperspectral mission (Figure 1). PACE will collect

119    ocean color imagery over a period of one-to-two days from a sun-synchronous polar orbit at 676.5

120    km (inclination of 98°), with an ascending equatorial crossing time at 1 pm. The spacecraft will

121    host three instruments: a hyperspectral imaging radiometer named OCI (for Ocean Color

122    Instrument, developed by NASA Goddard Space Flight Center (GSFC)) and two polarimeters,

123    named SPEXone (for Spectro-polarimeter for Planetary Exploration, developed by collaborators

124    at Space Research Organization of the Netherlands, Hasekamp et al. 2019) and HARP2 (for Hyper

125    Angular Research Polarimeter, developed by the Earth and Space Institute of the University of

126    Maryland Baltimore County, Martins et al. 2002).

127    The OCI has hyperspectral capabilities that continuously span from the ultraviolet to near-

128    infrared, with nominal spectral steps of 2.5 nm and average bandwidths of ~5 nm across a

129    spectral range of 340-890 nm. The spectral steps decrease to 1.25 nm (with the same ~5 nm

130    bandwidths) in the spectral regions of chlorophyll a fluorescence and oxygen A and B band

131    absorption, resulting in roughly 230 total wavelengths of information (Table 1). Additional bands

132    in the near- and shortwave-infrared support heritage atmospheric products and will also bring

133    improvements to atmospheric correction in optically complex waters over large freshwater and

134    coastal estuaries (Frouin et al. 2019; Ibrahim et al. 2018). The spatial resolution of OCI imagery

135    will be ~1 km$^2$ at nadir with a swath width of 2663 km (supporting global coverage once every

136    two days, given geometry constraints applied to atmospheric correction). In practice, pixels sizes

137    at the center of the swath will be ~1.1 km$^2$ given OCI's 20-deg tilt. Due to the specific optical

138    design, time-delayed integration, lunar, spectral calibrations, and tilt mechanism OCI will produce

139    high-quality data over a very high dynamic range of radiances from highly absorbing waters to

140    ecosystems rich in inorganics.

141    The role of the two PACE polarimeters, SPEXone and HARP2 (Table 1), is primarily to improve our

142    understanding of cloud and atmospheric aerosol properties (Chowdhary et al. 2019; Remer et al.

143    2019; Sayer et al. 2022). Improved characterization of the atmospheric properties will enhance

144    the retrieval of ocean surface reflectances, and there is hope that information about the

145    microphysical properties of particles from polarimeters may also prove useful for phytoplankton

146    studies (Jamet et al. 2019). SPEXOne is a hyperspectral polarimeter (385-770 nm, 5 nm steps),

147    measuring light at 5 viewing angles, collecting information in narrow 100 km swath at 2.5 km$^2$.

148    HARP2 is a multispectral (nominal 441, 549, 665, 866 nm), hyper-angular instrument, with a wide

149    swath that matches OCI and a ground sampling distance of 3 km$^2$.

150    By identifying the needs, challenges, and future of estimating PCC from space using hyperspectral

151    instruments, this paper contributes to a decades-long community effort to improve remotely-

152  sensed phytoplankton datasets (e.g., Bracher et al. 2017; Bracher et al. 2022; IOCCG 2014; Mouw

153  et al. 2017). After evaluating user community needs (Section 2) and identifying several

154  applications that will benefit from improved PCC estimates (Section 3), we outline current

155  approaches to identifying PCC using in situ data, a critical step in developing PCC retrieval

156  algorithms (Section 4). We then describe the existing approaches that PACE will build on to derive

157  PCC from space (Section 5). Finally, we conclude the paper by identifying some of the existing

158  challenges and gaps, address what PCC will mean in the era of PACE, and discuss how PACE will

159  allow the community to advance understanding of the role that phytoplankton diversity plays on

160  earth and how it is likely to change in the future.

161  ## 2. Community Needs

162  There is a consensus across the ocean color user community that better resolution of

163  phytoplankton diversity from space will improve our knowledge and understanding of marine

164  ecosystems (e.g., Kavanaugh et al. 2021), ocean health (e.g., Smith and Bernard 2020), shifts in

165  communities across freshwaters (Rasconi et al. 2015; Verbeek et al. 2018), and the ocean's effects

166  on the global climate (IPCC 2019). However, the needed resolution (taxonomic or otherwise) of

167  phytoplankton diversity varies widely within the user community.

168  A request for input was directed to the PACE Community of Practice[4] (CoP) to identify PCC

169  taxonomic resolution needs. The PACE CoP is a diverse group of users composed of modelers and

170  empirical researchers, academic and government scientists, local to international decision-

---

[4] https://pace.oceansciences.org/app_community.htm

171  makers, and industry professionals within various sectors, including air quality and atmospheric,

172  terrestrial, and marine/ocean applications (Figure 2).

173  Responses indicate that data consumers most need resolution down to a few traditionally

174  recognized "taxonomic groups" (e.g., diatoms, dinoflagellates), followed by plankton pigments,

175  and then groups defined by functionality (Figure 3). A large portion of questionnaire respondents

176  would be satisfied with "whatever taxonomic grouping is available" beyond *Chl a* alone.

177  Approximately 25% of the users prioritized species-level resolution, with 10% of users relying on

178  PACE to resolve specific species (such as the HAB forming dinoflagellate *Alexandrium*). While this

179  short questionnaire provides preliminary insights into the needs of the PACE user community, it

180  does not provide information on how these needs relate to each subcommunity.

181  The above-mentioned needs are likely driven by the spatiotemporal constraints of each

182  respondent's focus, ranging from local to global spatial scales, and from daily, monthly,

183  interannual, and decadal scales. In a similar way, user-specific capabilities to access and

184  manipulate the data may also drive the differences in needs across the PACE CoP.

185  ## 3. Applications

186  A deeper look into the type of applications that use data on phytoplankton diversity (and would

187  benefit from remote sensing data on PCC) highlights the complexities that come with defining

188  PCC and the associated temporal, spatial, and diversity resolution. Hereafter, we present some

189  examples of applications that range from regional to global and the level of diversity resolution

190  associated with them.

191  Monitoring of local and regional water quality, including forecasting HABs, is a critical societal

192  application that requires "local" knowledge and data (e.g., Lekki et al. 2019; Smith and Bernard

193    2020). Algal blooms occur in coastal and freshwater systems worldwide and include a variety of

194    harmful species that can impact wildlife, pet, livestock, and human health (Lundholm et al. 2009;

195    Michalak 2016). To address the need for products capable of detecting, quantifying,

196    characterizing, classifying, and being used in forecasting applications will require finely tuned

197    algorithms for specific HAB (or any other) algal type (Schaeffer et al. 2015, Bernard et al. 2021),

198    at an appropriate temporal and spatial resolution to capture the transient nature of these events.

199    PACE, similarly to the majority of heritage ocean color missions, will have adequate temporal

200    resolution, and some limitations with regard to the spatial resolution to detect developing HABs

201    in small water bodies. Therefore, synergies and cross-product generation with other missions

202    with smaller spatial resolution but less optimal temporal resolution (e.g., NASA's upcoming

203    Surface Biology and Geology (SBG[5]), the European Space Agency's Copernicus Hyperspectral

204    Imaging Mission for the Environment (CHIME[6]), could be combined to alleviate this weakness.

205    Lastly, low latency (i.e., the time between data collection and availability) is also critical for timely

206    public-safety decisions in response to HABs, which requires near-real time processing of the PCC

207    data product.

208    The fisheries industry has also increasingly integrated phytoplankton data into management

209    activities for the purpose of: (1) fish stock assessment, (2) harvesting by identifying suitable fishing

210    zones, and (3) fisheries management (Forget et al. 2009). The appropriate PCC definition for

211    fisheries applications could range from single species (HABs) to more functional or taxonomic

---

[5] SBG stands for Surface Biology and Geology - https://sbg.jpl.nasa.gov/

[6] CHIME stands for the Copernicus Hyperspectral Imaging Mission for the Environment - https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Going_hyperspectral_for_CHIME

212    information to support fish stock assessment and management to promote best fishing

213    practices[7].

214    At the global scale, satellite-derived PCC products can also be used to improve climate

215    applications and research (Figure 4). Using existing satellite observations of total phytoplankton

216    biomass, natural climate variability, such as El Niño, has been shown to directly affect

217    phytoplankton composition, with changes trickling up the food web to the higher tropic levels

218    (e.g., due to the change in nutrient dynamics, Fisher et al. 2015; Franz et al. 2021; McCabe et al.

219    2016; McKibben et al. 2017; Racault et al. 2017; Rousseaux and Gregg 2012). In a similar way, PCC

220    products can help indicate effects of the anthropogenic driven changes in aquatic ecosystems.

221    This has been demonstrated using long term in situ datasets (e.g., Rivero-Calle et al. 2015) and

222    models (Anderson et al. 2021; Cael et al. 2022a; Dutkiewicz et al. 2013). While prior analysis

223    suggested natural variability compounded climate trends in current chlorophyll ocean color

224    records (Henson et al. 2010), recent research is suggesting that ocean color itself (i.e., remote

225    sensing reflectances) is demonstrating significant, global climate trends in 20 yearlong records

226    (Cael et al. 2022c). PCC records from remote sensing may assist with understanding observed

227    trends and interpreting those trends in reference to the whole ecosystem structure and role that

228    oceanic ecosystem plays in global carbon cycle. On more local scales, in polar regions, a rapidly

229    shifting landscape (due to the unprecedented ice loss) is modifying the local ecosystems, and

230    composition and seasonality of phytoplankton (Eayrs et al. 2021; Meier et al. 2021: Nardell et al.

231    2023). Understanding of such close coupling of PCC, sea ice, polar food webs, and carbon flux

---

[7] pezCA, an application developed by the Federación Costarricense de Pesca (FECOP) in Costa Rica, that combines policy and real time satellite based products to identify potentially favorable fishing areas, https://pezca.org/.
pezCA team is part of the PACE early adopter program.

232    (Flexas et al. 2022; Neeley et al. 2018; Schofield et al. 2018) will benefit greatly from continuous

233    monitoring of PCCs from space.

234    Numerical models, including simple inverse and food web models as well as more complex ESMs,

235    provide an opportunity to integrate satellite PCC data as well as in situ data in a global context

236    (Dinauer et al. 2022; IOCCG 2020; Siegel et al. 2023). However, data assimilation is often limited

237    by the availability of data with appropriate units and levels of detail, such as PCC categories, and

238    quantified uncertainties. As detailed by Le Quéré et al. (2010), not only is phytoplankton biomass

239    (usually resolved into functional, taxonomical, and/or size groups) important for modeling

240    applications, but so are parameters such as growth rates and export rates. Moreover, sufficient

241    continuity of measurement (including the various seasons, see also Dutkiewicz et al. (2020)) at

242    the global scale is required to study the feedbacks between climate and ocean biogeochemistry.

243    Assimilation of data on PCC to improve those models will decrease uncertainty in the role of the

244    ocean in the global carbon budget (and decrease the societal and monetary impact of such

245    uncertainties e.g., Bontempi et al. submitted). Importantly, reported uncertainties are crucial for

246    improving models through parametrization or data assimilation, or informing ocean color science.

247    These uncertainties will also enable better understanding of regional, temporal bias, and

248    instrument-based bias, especially in merged products originating from multiple instruments

249    (Dutkiewicz et al. 2020; Gregg et al. 2017).

250    Lastly, society will have to continue to respond to the growing challenges associated with climate

251    change. It is becoming increasingly evident that controlling future Earth warming to well below

252    +2.0°C (preferably within +1.5°C), as is the aim of the 2016 Paris Agreement, will require

253    deployment of large-scale technologies to reduce carbon dioxide emissions (Fuss et al. 2014; IPCC

254    2021). Several ocean-based Carbon Dioxide Removal (CDR) methods have been proposed to curb

255    on-going global warming (GESAMP 2019; National Academies of Sciences 2022). Some of the CDR

256    approaches that rely on the ocean, such as ocean iron fertilization or seaweed cultivation, can

257    dramatically alter upper ocean ecosystems and phytoplankton communities (National Academies

258    of Sciences 2022). An important part of the monitoring, reporting, and verification process of CDR

259    will be understanding the impact of these technologies on the ocean, including the assessment

260    of effects on PCC, net primary production, and carbon export.  Remote sensing observations, such

261    as those from PACE, will also be key to quantifying the success of ocean CDR strategies,

262    monitoring their effects on ocean ecosystems (e.g., potential taxonomic shifts in PCC leading to

263    harmful algal blooms), and providing data needed for ESMs to improve our understanding of the

264    role these perturbations have on ocean ecosystems and the global carbon cycle.

265    ## 4. In situ methods

266    Many methods exist to quantify phytoplankton composition for mixed natural communities.

267    While no individual method captures the vast diversity of marine phytoplankton across multiple

268    dimensions (taxonomy, size, morphology, genetics, metabolism, etc.), many of these methods

269    provide specific information that is useful to describe PCC across observations and support future

270    PCC models. Some methods (HPLC derived pigments) have been used extensively to develop and

271    validate satellite remote sensing approaches for PCC. Other methods (imaging-in-flow cytometry,

272    DNA metabarcoding) offer globally distributed datasets with high taxonomic resolution that have

273    great potential for future PACE PCC applications. Here, we review the general approaches for

274    many common methods of phytoplankton observation and summarize their notable strengths

275    and weaknesses, particularly with respect to remote sensing approaches (Table 2).

## 4.1. Light Microscopy

Microscopy phytoplankton enumeration and biovolume calculation are a common approach for development and validation of ocean color PCC algorithms, especially for HAB detection (Pan et al. 2011; Soto et al. 2015; Wolny et al. 2020). In combination with cell volume to carbon models (e.g., Menden-Deuer and Lessard 2000; Worden et al. 2004), this method offers a simple route to phytoplankton carbon biomass. Standard compound and inverted microscopes are commonly used for the visualization of phytoplankton communities from water samples, offering taxonomic classification for nano- to micro size ranges in live or preserved samples (Karlson et al. 2010). The light microscopy methods used vary with size and concentration of phytoplankton within samples and on the accuracy and taxonomic resolution needed when determining PCC. A small volume of sample (1 ml or less) is suitable for counting nano- and microphytoplankton (>2 μm) samples with a high concentration of cells (Godhe et al. 2007; LeGresley and McDermott 2010; McAlice 1971). However, when cells are in low concentrations or samples represent a more diverse community structure, a greater sample volume (<25 mL) is needed to adequately capture both small-sized and numerous cells, as well as large-sized and rare cells (Edler and Elbrächter 2010; Haas and Marshall 1989).

There are notable limitations to microscopic approaches. Due to the small volume of material examined for some samples, certain taxa may be over- or under-represented. Settling chamber approaches that allow for larger sampling volume and better optical resolution, resulting in a larger phytoplankton size range, are time-demanding (hours/days) in comparison to the small volume approach (minutes, Elder and Elbrächter 2010). Uncertainties in these methods can be highly variable, depending on the optical material (LeGresley and McDermott 2010), fixatives,

298    enumeration method, sample volume, and assumptions used to determine biovolume and

299    resulting carbon content (Vuorio et al. 2007; Willén 1976). However, regardless of the approach

300    used, the greatest limiting factors (and additional sources of uncertainty) are analyst skill level,

301    their experience, and time required for sample analysis (see Clayton et al. 2022 and references

302    within).

303    ## 4.2. HPLC phytoplankton pigments

304    High performance liquid chromatography (HPLC) allows for the direct quantification of a number

305    of different phytoplankton pigments. HPLC pigments are currently one of the most widespread

306    approaches for characterizing PCC in ocean color studies: samples have been collected

307    throughout the global ocean, at time series observatories and along transect cruises, across

308    depths and over seasons (e.g., Kramer and Siegel 2019; Uitz et al. 2006). HPLC methods for

309    quantification of pigment concentrations have also been highly standardized, with successive

310    quality control efforts to ensure consistency between measurements (Hooker et al. 2012; van

311    Heukelem and Hooker 2011).

312    Pigment-based taxonomy is fairly low resolution and depends heavily on the HPLC pigment library

313    used to determine PCC (e.g., Catlett and Siegel 2018). Existing methods to characterize PCC from

314    HPLC pigments typically rely on biomarker pigments to separate phytoplankton groups based on

315    broad taxonomic association (i.e., diatoms with fucoxanthin from dinoflagellates with peridinin)

316    or size classes. Any pigment-based PCC method has to make necessary assumptions, as most

317    pigments are shared between phytoplankton groups and are not unambiguous biomarkers

318    (Jeffrey et al. 2011 and references therein). PCC methods using HPLC pigments are further

319    complicated by environmental and physiological variations that impact pigment production and

320     expression (e.g., Catlett et al. 2022; Dierssen et al. 2015; Henriksen et al. 2002; Neeley et al. 2022;

321     Zapata et al. 2004). Given these variations in pigment concentration and composition, statistical

322     methods that make assumptions about constant relationships between taxa abundances and

323     pigment ratios as well as the neglect of co-linearity among pigments must be used cautiously for

324     evaluating PCC (e.g., CHEMTAX; Mackey et al. 1996).

325     Phytoplankton pigment absorption directly impacts the shape and magnitude of remote sensing

326     reflectance spectra; thus, HPLC pigments are an ideal measurement for development and

327     validation of ocean color algorithms (e.g., Chase et al. 2017; Kramer et al. 2022; Torrecilla et al.

328     2011; Uitz et al. 2015). Studies that compare HPLC pigments with other, higher-resolution

329     taxonomic approaches have variable success, depending on the ecosystem and the taxonomic

330     resolution of the comparison (Chase et al. 2022; Havskum et al. 2004; Lin et al. 2019; Pan et al.

331     2011). More comparisons between HPLC pigments and other methods using larger datasets will

332     allow for better constraint of pigment-based PCC, and thus more accurate ocean color algorithms

333     using PACE data.

334     4.3. Flow cytometry

335     Flow cytometry (FCM) can enumerate particles and plankton from the size of viruses to cells

336     greater than 50 μm, depending on the instrument used and its configuration, and the whole size

337     spectrum can be captured when combined with complementary methods (microscopy, imaging

338     in-flow cytometry, e.g., Chase et al. 2020; Haëntjens et al. 2022). Recent evolution in FCM

339     technology can also provide continuous or near-continuous monitoring of phytoplankton

340     community dynamics and structure, and other ecological studies (e.g., Hunter-Cevera et al. 2021;

341     Swalwell et al. 2011). FCM delivers cells single-file in sheath fluid past an excitation laser(s) and a

342    set of optical detectors, allowing investigators to identify specific groups of plankton based on

343    their fluorescence and scattering properties. Additionally, some flow cytometry instruments have

344    capability of sorting different cell types (based on their fluorescence and scattering properties) in

345    separate streams, allowing for more detailed analysis (e.g., phytoplankton group specific carbon

346    content, Graff et al 2012; Casey et al. 2013). Samples can be run "live" or preserved (Lepesteur et

347    al. 1993). Published relationships can then be used to convert cell size (in μm) to cell biovolume

348    (e.g., Calvo-Díaz and Morán 2006) and carbon (e.g., Menden-Deuer and Lessard 2000; Worden et

349    al. 2004). Fluorescence emission from the excitation of pigments can be used to distinguish

350    specific phytoplankton populations for some groups, such as the pico-sized cyanobacteria

351    *Synechococcus*, which produces an orange fluorescence signal when the phycoerythrin in the cell

352    is excited by a blue laser, distinguishing them from other small cells (Olson et al. 1988).

353    Fluorescent dyes may also be used to enumerate different groups of cells without specific marker

354    pigments (Rose et al. 2004).

355    Uncertainty in PCC data from FCM can come from both the sample collection and data analysis.

356    Studies have shown that fixatives can impact the autofluorescence signal or cause cell loss and

357    cell shrinkage by phytoplankton cells if samples are stored for long periods of time, e.g., >1 month

358    (Marie et al. 2014; Sato et al. 2006; Vaulot et al. 1989). The instrument calibration method

359    determines the model used to derive cell size from fluorescence or scattering. Uncertainties also

360    arise during data analysis when discrete populations of cells are separated (sometimes

361    subjectively by operator) based on fluorescence and scattering parameters. Nevertheless, FCM is

362    an effective approach that can provide enumeration of pico and nano-sized cells, which are

363   important, novel datasets that support development of PCC algorithms (e.g., Lange et al. 2020;

364   Kramer et al. 2020; Thyssen et al. 2015).

## 4.4.Imaging-in-flow cytometry

366   Imaging-in-flow instruments have revolutionized collection and analysis of aquatic samples for

367   the qualification and quantification of PCC. These approaches combine flow cytometric

368   enumeration with the digital photography of particles that can be used for taxonomic

369   identification. A number of different imaging-in-flow instruments, such as the Imaging Flow

370   Cytobot (Olson and Sosik 2007), Underwater Vision Profiler (UVP, Picheral et al. 2010), and

371   FlowCAM (Sieracki et al. 1998), have been developed over the last two decades and capture

372   different regions of the size spectrum (Lombard et al. 2019). Unlike standard microscopy, these

373   instruments allow for automated, rapid collection of plankton and particle images over high

374   spatial and temporal resolution that can be classified and enumerated (sometimes iteratively)

375   after collection using image recognition by machine learning.

376   An imaging in-flow system typically contains an objective, an excitation laser, and fluorescence

377   detectors to detect *Chl* a fluorescence and light scatter. A water sample is drawn through a flow

378   cell and past the optical package, where particles that are excited by the laser get magnified by

379   the objective (like a microscope) and imaged by a color or monochrome camera. The images are

380   interpreted based on morphological characteristics and/or size dimensions that may be used by

381   a machine learning classifier or manually attributed to specific types of particles, plankton, or

382   phytoplankton. As with every method, there are challenges that must be considered. Imaging-in-

383   flow systems typically measure the larger range of phytoplankton (e.g., >~6μm for the IFCB). The

384   large number of images that are collected creates challenges regarding the time needed to

385   complete automated and manual classification, and in the interpretation and sharing of results

386   (Durden et al. 2017; Kerr et al. 2020). Advances in machine learning techniques and other

387   automated approaches (Orenstein et al. 2022), in particular deep learning such as convolutional

388   neural networks, will increase the accuracy of automated classifications and reduce the analysis

389   time, thereby increasing throughput of PCC data. The scientific and environmental monitoring

390   communities using these methods are coming together to overcome these challenges not only

391   for image post-processing, but also dataset creation and taxonomic training augmentation

392   (Clayton et al. 2022, Neeley et al. 2021). Overall, imaging-in-flow methods present great potential

393   for describing PCC in situ, particularly for PACE applications, given the relative ease of sampling

394   and high taxonomic resolution.

395   ## 4.5. DNA metabarcoding

396   Numerous molecular approaches are available to quantify phytoplankton diversity, population

397   dynamics, and PCC (Johnson and Martiny 2015). DNA meta-barcoding is now one of the most

398   widely used methods to provide a holistic view of PCC including diverse members of the

399   phytoplankton community. DNA meta-barcoding refers to targeted amplicon sequencing of

400   highly conserved, hypervariable "barcode" genes. These data are usually compositional (i.e.,

401   estimate proportions rather than concentrations or counts), which complicates analysis and

402   interpretation of PCC (Aitchison 1982; Gloor et al. 2017), but some exciting approaches have been

403   developed recently to remove this constraint (Lin et al. 2019; Satinsky et al. 2013).

404   The taxonomic resolution of DNA meta-barcoding varies from division- to species-level depending

405   on a number of factors including the analysis workflow or the specific Amplicon Sequence

406   Variants (ASV). Uncertainty in relative sequence abundances is introduced by both wet lab and

407 bioinformatic procedures and is difficult to quantify (Catlett et al. 2020; Yeh et al. 2021). In

408 general, nucleic acid sequencing data are generated through a series of complex biochemical

409 reactions that make it difficult to evaluate and constrain analytical uncertainty. Consistent use of

410 positive and negative controls to ensure reproducibility in sequence analysis has only recently

411 become common practice (Bradley et al. 2016; Parada et al. 2016), but is still omitted by many

412 investigators. Interpretation of relative (or absolute) sequence abundances is not straightforward

413 in practice, as the number of barcode gene copies per cell or unit biomass can vary across taxa

414 depending on the chosen barcode gene (Gong and Marchetti 2019; Zhu et al. 2005). Despite these

415 caveats, several recent studies have demonstrated that DNA meta-barcoding workflows can

416 provide accurate and precise estimates of the relative sequence abundances of most

417 phytoplankton (Catlett et al. 2020; Yeh et al. 2021). Some barcode genes also provide relative

418 sequence abundances that scale roughly with cell size, biovolume or biomass proportions (de

419 Vargas et al. 2015; Godhe et al. 2008; Zhu et al. 2019).

420 DNA meta-barcoding data have not been employed in direct validation of satellite algorithms to

421 the best of our knowledge, but these studies suggest high potential for their use in efforts to

422 validate the next generation of PCC algorithms developed for PACE (Catlett et al. 2022). Despite

423 some methodological challenges, the ability to sample a nearly comprehensive range of size

424 classes, the detailed resolution of taxonomic and functional diversity, and the growing

425 appreciation for the quantitative potential of well-validated nucleic acid sequencing workflows

426 make these methods primed for in situ PCC quantification.

## 4.6. Phytoplankton optics

Phytoplankton, either due to their size, or morphological (external or internal) characteristics have a taxa-specific contribution to the scattering of the light in the water column (Organelli et al. 2018; Poulin et al. 2018; Stramski et al. 2001; Stramski and Kiefer 1991; Whitmire et al. 2010), and that signal is present in the $R_{rs}(\lambda)$ as well. Ocean color measurements can provide rudimentary estimates of the slope of the particulate backscattering or particulate attenuation spectra through the inversion of $R_{rs}(\lambda)$ (e.g., Roesler and Boss 2003; Loisel et al. 2018). Particulate backscattering itself, which is readily retrievable from in situ or satellite derived $R_{rs}(\lambda)$ (e.g., Werdell et al. 2013) can be used to infer phytoplankton size composition (Kostadinov et al. 2022). Backscattering can be affected by internal and external morphological characteristics that impact the shape of the slope (see Organelli et al. 2018), whereas particulate attenuation is mostly dependent on particle size. Increased values of attenuation spectral slope are associated with particle populations that have higher proportions of smaller size particles, whereas the inverse is true for backscattering slopes (e.g., Boss et al. 2001). This approach is highly applicable to the open ocean where changes in particle population are driven by PCC change and could offer an additional dimension in information when distinguishing the phytoplankton with different mineral components (such as silica and calcium carbonate).

Absorption of the phytoplankton is to the first extent driven by the pigments present in the cell. Chlorophylls, and other (sometimes taxa-specific) cellular pigments (see 4.2 and references therein) determine the shape of the absorption spectra (Mobley 2022 and references within). Additionally, pigment packaging and the structure of protein-pigment complexes will have a significant effect on the phytoplankton absorption spectra. Shape, peak, and width of the specific

449 absorption features can be used as a tool to detect different taxa from in situ measurements of

450 absorption (e.g., Chase et al. 2013). While some of these features (e.g., Chlorophyll peaks) are

451 visible in $R_{rs}(\lambda)$ and therefore easy to relate to pigment concentrations, ocean reflectance

452 inversion models (e.g., GIOP – Werdell et al. 2013) are often used to derive phytoplankton

453 absorption from the $R_{rs}(\lambda)$, that is ultimately used to infer PCC distribution.

## 5. Phytoplankton composition from space

### 5.1. Heritage multispectral PCC remote sensing algorithms

456 Heritage approaches for deriving phytoplankton composition from ocean color typically exploit

457 relationships generated from pairing in situ phytoplankton data with either in situ $R_{rs}(\lambda)$ or top-

458 of-atmosphere (TOA) satellite radiometry. Until recently, satellite and in situ radiometry

459 measurements were mostly collected at multispectral resolution. Therefore, the methods

460 developed to derive PCC relied on the multispectral information ($R_{rs}(\lambda)$) and could be applied to

461 SeaWiFS, MODIS, and multispectral ocean color missions operated by other space agencies

462 (Alvain et al. 2005; Alvain et al. 2008; Ben Mustapha et al. 2014; Sathyendranath et al. 2004;

463 Werdell et al. 2014; Westberry et al. 2005). In general, their applicability and success (and

464 associated uncertainties) depend heavily on the in situ training datasets and the formulated

465 methodology. Recent reviews have demonstrated that existing ocean color PCC approaches agree

466 on a global scale, but disagree considerably on regional scales (e.g., Bracher et al. 2017; IOCCG

467 2014; Kostadinov et al. 2017; Mouw et al. 2017). These reviews also identified numerous gaps in

468 ocean color-based PCC (including mismatch between in situ, satellite and model data, the lack of

469 uncertainty estimates for the satellite data, and the spectral limitation of existing sensors) that

470 advanced ocean color missions such as PACE will be able to address. Looking towards PACE ocean

471    color and polarimetry data, we can draw from the knowledge gained from these previously

472    developed methods, which address global-scale data products as well as more regional taxon or

473    phytoplankton group specific products.

474    The breadth of methods for detecting PCC from space includes approaches that rely on total

475    biomass (e.g., Hirata et al. 2011), or other proxies for abundance, rather than spectral shape or

476    features, as well as methods that rely on single or multiple ocean color products; $R_{rs}(\lambda)$, carrying

477    information from both scattering and absorption components (e.g., Morel and Prieur 1977; Xi et

478    al. 2021); phytoplankton absorption (Brewin et al. 2011b; Bricaud et al. 2007 and similar, derived

479    from $R_{rs}(\lambda)$ via semi analytical models such as Werdell et al. 2013); or backscattering (Kostadinov

480    et al. 2009). Different mathematical and statistical approaches (e.g., Gaussian decomposition,

481    principal component analysis, derivatives, machine learning etc.) have been used to extract the

482    information on PCCs from the abovementioned datasets. A more comprehensive list of

483    algorithms is presented in Table 3. The types of PCC retrievals that result from previously

484    developed methods are highly variable. Certain algorithms focus on the retrieval of single taxa,

485    while others retrieve multiple taxa or fractional contributions of multiple taxa. These retrievals

486    are generally reported as probability of detection, as the dominating portion of biomass (usually

487    *Chl a*), as the proportion of the total *Chl a* concentration, or in concentration units such as biomass

488    (e.g., cells) per volume. In contrast, phytoplankton size classes (PSCs) are used to define the

489    heterogeneity in size distribution of phytoplankton communities; mainly reporting three size

490    classes: the pico-, nano-, and microphytoplankton (<2 μm, 2-20, and >20 μm, respectively). Some

491    approaches rely on the backscattering to retrieve PSCs, and in combination with modeling,

492    retrieve the particulate and phytoplankton size distribution, expressed as % of total biomass or

493 total phytoplankton carbon (Kostadinov 2016; Kostadinov et al. 2022). Other methods estimate

494 the contributions of each of these three classes to total *Chl a*, through assigning HPLC-determined

495 accessory pigments to taxonomic groups, and the groups to size classes (as per Vidussi et al. 2001,

496 Uitz et al. 2006). Numerous revisions and evaluations of the "Diagnostic Pigment Analysis"

497 approach to estimating PSCs from accessory pigment concentrations have been published and

498 provide a body of literature for reference regarding the development and validation of PSC

499 algorithms from space. Ultimately, these accessory pigments themselves can be used to estimate

500 PCC (Table 3).

501 Identification of phytoplankton composition down to the species level is possible in some cases

502 for certain taxa with unique optical properties. One example is the cosmopolitan coccolithophore

503 species *Emiliania huxleyi*, which when present at high concentrations (blooms) in the surface

504 layer, can profoundly impact the optical properties of the upper ocean (Balch et al., 1991; Balch,

505 2018). This detection is not based on the absorption properties of the phytoplankton, but rather

506 on its intense scattering properties (Neukermans and Fournier, 2018), which lead to "milky white"

507 seas observable even in broadband satellite sensors like AVHRR (Loveday and Smyth, 2018).

508 Similarly, algorithms have been developed to detect the blooms of dinoflagellate *Karenia brevis*

509 due to its consistently lower backscattering properties compared to other types of blooms found

510 in the same region (Craig et al., 2006; Cannizzaro et al., 2008; Soto et al., 2015). The addition of

511 vacuoles or intracellular spaces that create high scattering also has a substantial effect on $R_{rs}(\lambda)$

512 even at relatively low biomass and has led to approaches for discerning cyanobacteria (Matthews

513 et al., 2012; Matthews and Bernard, 2013, Schaeffer et al. 2015). Combination of increased

514 scattering due to the intercellular gas vesicles and pigment specific signal in $R_{rs}(\lambda)$ was the base

515   of early algorithms used to detect cyanobacteria *Trichodesmium* from space (Subramaniam et al.

516   2001; Subramaniam and Carpenter 1994). More recent approaches rely on the dense surface

517   accumulations (sea slicks, Hu et al. 2010; McKinna et al. 2011), or rely on *Trichodesmium* specific

518   bio-optical properties (Dupouy et al. 2011; Westberry et al. 2005; see review in McKinna 2015).

## 5.2. Anticipated benefits of the PACE instruments

### 5.2.1.   Hyperspectral ocean color

521   Hyperspectral $R_{rs}(\lambda)$ from PACE will provide, by definition, additional information content over

522   existing multispectral data and their associated algorithms. Several studies have addressed

523   optimal spectral resolutions and/or band placement for radiometric measurements, starting with

524   work from Lee et al. in 2007. Wolanin et al. (2016) demonstrated that band placement

525   requirements depend on both the method as well as the target phytoplankton groups to be

526   retrieved, and they suggest that hyperspectral (vs. any lower spectral resolution) observations

527   are ideal. Torrecilla et al. (2011) used field data of concurrent phytoplankton absorption and

528   remote-sensing reflectance measurements and demonstrated the advantage of high-spectral

529   resolution data over multi-spectra data for discriminating phytoplankton pigment assemblages in

530   open ocean, including the benefits of spectral derivative analysis. Vandermeulen et al. (2017),

531   used a database of in situ hyperspectral reflectance measurements from a wide range of water

532   types to demonstrate that a spectral interval of 5 nm is optimal to separate differently absorbing

533   phytoplankton groups while also accounting for measurement uncertainties. Further, Kramer et

534   al. (2022) found a similar result where the performance of statistical models for predicting

535   phytoplankton pigment concentrations was greatly reduced above a spectral interval of 5 nm.

536   Hence, PACE radiometry can be used to separate the contribution or fine scale spectral features

537 of phytoplankton and large-scale spectral features of the other oceanic constituents (>100 nm,

538 e.g., CDOM, backscattering) as recently demonstrated by Kramer et al. (2022).

539 Hyperspectral phytoplankton absorption and $R_{rs}(\lambda)$ spectra are more effective than multispectral

540 data during derivative, Gaussian, and clustering analyses for pigment assemblage discrimination

541 and size-based, phytoplankton community composition assessment (Chase et al. 2017; Kramer et

542 al. 2022; Lange et al. 2020; Roelke et al. 1999; Torrecilla et al. 2011; Uitz et al. 2015). This finding

543 reflects the relative similarity of spectral absorption of different phytoplankton pigments and

544 groups (e.g, Garver et al. 1994; Mao et al. 2010), which necessitates the use of optical information

545 at a high spectral resolution to discern subtle differences in spectral absorption and reflectance

546 that are the result of differently absorbing phytoplankton pigments. By focusing on changes on

547 small spectral scales some of these approaches minimize the source of noise in absolute signals

548 (usually associated with atmospheric correction approaches – see Ibrahim et al. 2018).

549 Existing algorithms that take advantage of hyperspectral $R_{rs}(\lambda)$, have demonstrated their

550 capability on data collected by previously flown hyperspectral instruments; either by quantifying

551 *Chl a* or phytoplankton absorption to specifying taxonomic groups. Studies based on the

552 Hyperspectral Imager for the Coastal Ocean (HICO) (Lucke et al. 2011) have successfully

553 demonstrated the use of hyperspectral $R_{rs}(\lambda)$ to observe a phytoplankton bloom in Monterey Bay,

554 CA (Ryan et al. 2014), to differentiate a red tide ciliate bloom (*Mesodinium rubrum*) in Long Island

555 Sound, NY at high spectral and spatial resolution using unique yellow fluorescence features

556 (Dierssen et al. 2015), and to map PCC in coastal China using machine/transfer learning approach

557 (Zhu et al. 2019). Recent publications have also shown high-quality retrievals of *Chl a*,

558 phycocyanin, and phytoplankton absorption spectra from HICO imarrgery over freshwater and

559 coastal ecosystems (Gitelson et al. 2011; O'Shea et al. 2021; Pahlevan et al. 2021), and

560 discrimination of cyanobacteria monospecific blooms in large (Wynne et al. 2008) and small lakes

561 (Kudela et al. 2015). In the open ocean, the PhytoDOAS method, which makes use of

562 hyperspectral $R_{rs}(\lambda)$, has been applied to SCIAMACHY (SCanning Imaging Absorption

563 spectroMeter for Atmospheric ChartographY) to discern blooms of cyanobacteria,

564 coccolithophores, and diatoms (Bracher et al. 2009; Sadeghi et al. 2012).

565 The OCI on PACE will be able to measure ultraviolet (UV) wavelengths at the same high spectral

566 resolution as the visible portion of the spectrum. Information from the UV part of the ocean

567 color signal will bring additional means to constrain the phytoplankton community structure.

568 Numerous phytoplankton species produce specific MAAs of individual spectral characteristics

569 (Llewellyn and Airs 2010), with additional UV absorbing compounds (Jeffrey et al. 1999) that

570 could impact the remote sensing reflectance and be traced back to either species composition

571 or community-specific response to environmental conditions (e.g., polar regions, Ha et al. 2018).

572 Using in situ measured $R_{rs}(\lambda)$ Kahru and Mitchell (1998) demonstrated that the presence of a

573 HAB species, *Lingulodinium* polyedra could be distinguished from diatom-dominated

574 populations by the absorption of Mycosporine-like amino acids (MAAs), compounds produced

575 by dinoflagellates to protect organelles from harmful UV rays. This approach was recently

576 validated using a UV band from JAXA's GCOM-C satellite to track the spatial and temporal

577 distribution of this specific HAB (Kahru et al. 2021).

578     5.2.2. Polarimetry

579 PACE will carry two polarimeters, HARP2 and SPEXone, targeting atmospheric properties.

580 Measurements of atmospheric properties by these instruments will indirectly advance the

581 retrieval of PCC from hyperspectral OCI information by improving atmospheric correction through

582 better characterization of aerosols. Regarding the measurement of oceanic properties, these

583 instruments are limited by their mode of operation; collecting information on coarser spatial

584 resolution (3-4 km vs 1 km for OCI) and measuring approaches that maximize the number of

585 scattering angles. However, a similar instrument, ESA's POLDER (Polarization and Directionality

586 of the Earth's Reflectances), has been previously used in combination with modeling to offer

587 information about characteristics of oceanic bulk particle composition (Loisel et al. 2008). Similar

588 approaches could be used to derive additional information from PACE's polarimetric data,

589 separating organic from inorganic particles, inferring particle size distribution and other particle

590 properties that can be used in better retrievals of PCC and constrain uncertainties associated with

591 phytoplankton products (see Jamet et al. (2019) and references within). However, it is important

592 to note the basic need to first better understand the Muller matrix (i.e., inherent polarization

593 properties) of different marine particles/phytoplankton as a prerequisite to potential meaningful

594 interpretation and use of polarization properties of natural light fields, such as polarized water-

595 leaving light derived from PACE's polarimeters.

596 5.3. Other approaches to deriving PCC from space

597 In addition to the development of PCC algorithms that solely rely on optical properties of

598 phytoplankton, the scientific community has developed various approaches that use non-optical

599 proxies including model outputs or results from analysis of lower-level data (e.g., variables

600 derived from multiple measurements[8]) to derive PCC. These data are referred to as Level-4 data

601 products and include satellite-based models, machine learning approaches that could incorporate

---

[8] earthdata.nasa.gov

602 ancillary data, and Earth System Models. Such approaches have the advantage of being able to

603 integrate in situ, model, and/or satellite data to gain insights that would not be achievable with

604 the use of phytoplankton optical properties alone and can provide information on variables that

605 might not currently have designated algorithms (e.g., Anderson et al. 2009; Hill et al. 2020).

606 Several satellite-based models have emerged over the years that result in phytoplankton

607 taxonomic composition as well as phytoplankton size information (e.g., Xi et al. 2021). Hirata et

608 al. (2011), for example, used satellite *Chl a* to derive the fraction of both phytoplankton

609 composition and size fraction. California Harmful Algae Risk Mapping (C-HARM) is routinely (since

610 2014) using a combination of multispectral $R_{rs}(\lambda)$, remote sensed *Chl a*, sea surface temperature

611 and ROMS modeling to forecast distribution of *Pseudo-nitzchia* spp. and associated toxins

612 (Anderson et al. 2019). By combining hyperspectral in situ $R_{rs}(\lambda)$ and sea surface temperature in

613 a principal component analysis framework, Lange et al. (2020) derived cell counts for several

614 smaller phytoplankton groups across the Atlantic, demonstrating improved performance for such

615 a combined approach over a multispectral one (Figure 5).

616 Machine learning approaches to estimate phytoplankton composition have also been developed

617 in the last decades. Raitsos et al. (2008) developed an artificial neural network that incorporates

618 ecological and geographical information (e.g., longitude, latitude, season) with ocean color

619 products (e.g., *Chl a*, normalized water-leaving radiance, PAR), bio-optical characteristics, and

620 remotely sensed physical parameters (e.g., SST, wind stress). Using this approach, they were able

621 to discriminate four major phytoplankton functional types based on probability of occurrence

622 (diatoms, dinoflagellates, coccolithophores, and silicoflagellates) with an accuracy of more than

623 70%. Palacz et al. (2013) used another approach that relies on an artificial neural network to

624    simulate the global distribution of PCC with a focus on diatoms and coccolithophores in the high

625    nutrient low chlorophyll (HNLC) regions. A recent approach by Chase et al. (2022) utilizes *Chl a* in

626    combination with environmental variables (sea surface temperature and salinity) to retrieve

627    diatom carbon biomass in the North Atlantic. This approach is a clear demonstration of

628    improvement in PCC retrieval due to the inclusion of environmental variables as well as imaging

629    in-flow cytometry data to define PCC during algorithm development (Figure 6). These approaches

630    become even more relevant in coastal and freshwater systems where phytoplankton absorption

631    features in $R_{rs}(\lambda)$ are masked by strong absorption and backscattering by other optically relevant

632    constituents, hence, auxiliary environmental and physical information could further constrain the

633    solution space enabling high-quality PCC retrievals.

634    Finally, ESMs integrate in situ and satellite data through parametrization, forcing, and/or

635    assimilation. ESMs can provide global coverage (no gaps due to high solar zenith angle, clouds,

636    polar night, etc.) and can provide information on components of the ocean biogeochemical cycle

637    that cannot currently be derived from satellite data alone (IOCCG 2020). These models do

638    encompass the diversity of phytoplankton as either functionality (e.g., diatoms, cyanobacteria,

639    diazotrophs, etc; Bopp et al. 2005; Gregg and Casey 2007) or size classes (e.g., Ward et al. 2012;

640    Ward and Follows 2016). ESMs make assumptions that are aligned with our current

641    understanding of phytoplankton dynamics (i.e., nutrient uptake, growth rate, sinking rate etc.).

642    Some models also assimilate satellite ocean color products (e.g., Gregg and Casey 2007; Jones et

643    al. 2016; Shulman et al. 2013) in models including PCC as functional groups (e.g., Ciavatta et al.

644    2018; Skákala et al. 2018) or size classes (e.g., Xiao and Friedrichs 2014).

645     5.4. Challenges of deriving PCC from space

646     In the previous sections, we highlighted approaches that are currently used to detect the

647     dominant phytoplankton groups in the world aquatic ecosystem from ocean color imagery. In

648     vast areas of open ocean, non-phytoplankton particles (i.e., detritus, heterotrophic bacteria) are

649     more abundant than phytoplankton. While all these particulate (and dissolved) components of

650     the open oceanic systems contribute to the ocean color, the strong absorption by phytoplankton

651     pigments dominates the contribution to the ocean color signature, which has been identified as

652     a promising signal in terms of identifying groups of phytoplankton (Alvain et al. 2008; Devred et

653     al. 2006). As algal particle concentrations increase, however, scattering by phytoplankton can

654     dominate the phytoplankton absorption properties and therefore lead to inaccuracies in

655     measured optical properties (Brewin et al., 2017). As demonstrated by Garver at al. (1994) both

656     pigment assemblages and pigment packaging effects (self-shading effect in large cells) contribute

657     to the observed similarities in absorption spectra among phytoplankton, making it challenging to

658     differentiate between different species solely based on their spectral characteristics. They found

659     that more than 99% of the variance in the particulate absorption spectra was related to the

660     biomass, and less than 0.5% was related to the presence of auxiliary pigments (Garver et al.,

661     1994). The combined effects of assemblage effective cell diameter and phytoplankton biomass,

662     together with non-algal optical contributors, are not easily interpreted from $R_{rs}(\lambda)$ as these

663     quantities have ambiguous effects on the bulk optics (Evers-King et al., 2014). Recent information

664     content studies have shown that the spectral signatures of absorption by different phytoplankton

665     groups are similar within the uncertainty of the measurement mainly because of the considerable

666     overlap in pigment composition, and thus spectral absorption, between different groups spatially

667 and temporally across aquatic ecosystems (Cael et al. 2020). Also, correlation among reflectances

668 can decrease information content in hyperspectral data, limiting the capability of derivation of

669 independent parameters (Cael et al. 2022b).

670 Using the hyperspectral reflectance spectra from UV to NIR (320 to 715 nm) incorporates

671 significantly more information (than traditional multispectral) including backscattering of the

672 phytoplankton and associated particles, as well as fluorescence information, and may yield

673 further taxonomic resolution. New methods to differentiate fractional phytoplankton

674 composition from reflectance detected six different groups globally including diatoms,

675 dinoflagellates, haptophytes, green algae, prokaryotes, and *Prochlorococcus* (Xi et al. 2020).

676 However, such approaches may not apply to all coastal and inland waters. In the North Sea,

677 Castagna et al. (2021) found that blooms of *Phaeocystis globosa* are synchronous with those of

678 the diatom *Pseudo-nitzschia delicatissima*, both harmful bloom-forming species with similar

679 pigmentation and optical properties.

680 Remote sensing algorithms and biogeochemical models can be derived and tuned for the regional

681 or local phytoplankton groups down to specific taxa, if they are known to occur in an area and do

682 not have similar optical properties compared to other local taxa (examples listed in chapter 5.1

683 and Table 3). For example, in South African waters, five different probabilistic indicators of

684 harmful algal blooms were retrieved relevant to the aquaculture industry including waters with

685 high and low density of dinoflagellate, *Pseudo-nitzschia* dominated waters, as well as waters with

686 mixed assemblages of high and moderate concentrations (Smith and Bernard 2020). We have

687 only begun to assess the hyperspectral scattering and fluorescence properties that may also aid

688 in differentiating different types and stages of blooms, as well as relationships to seasonal trends

689 and other remotely sensed quantities including polarization parameters, temperature,

690 photosynthetically available radiation, and salinity (see Section 4.2.3).

691 Some studies have developed hyperspectral techniques to differentiate over 20 different species

692 of phytoplankton (e.g., Zhu et al., 2019). While the statistical approaches used in such studies are

693 robust, the lack of validation data across the breadth of space and time hinders the widespread

694 application of such tuned approaches. Our approach to treatment of error and uncertainties in

695 these applications determines the difference between retrieving four to five phytoplankton

696 groups rather than >60 (Cael et al. 2020). A recent study using a large HPLC pigment dataset

697 confirmed that only a limited number of phytoplankton groups (~4) may be differentiated globally

698 including cyanobacteria, diatoms/dinoflagellates, haptophytes, and green algae (Kramer and

699 Siegel 2019). However, they also demonstrated that, on regional scales, pigment association

700 varies, suggesting that regional algorithms could resolve up to 6 taxonomic groups. Data from the

701 Santa Barbara Channel, California revealed that around five phytoplankton pigment communities,

702 which are covarying assemblages of phytoplankton groups, could be differentiated based on their

703 spectral properties (Catlett and Siegel, 2018; Catlett et al., 2021).

704 Another challenge of deriving PCC from space is that the vertical structure of phytoplankton in

705 the ocean is not always homogenous as many of the models are assuming; phytoplankton layers

706 are often found across the oceanic ecosystems. Depth on which specific phytoplankton is, as well

707 as a thickness and number of layers with different PCCs will heavily influence observed $R_{rs}(\lambda)$ at

708 the surface, as that signal represent optically weighted contribution of all components. These

709 vertical distributions are not simple to resolve either from perspective of total biomass (Gordon

710 and McCluney 1975; Morel and Berthon 1989; Stramski and Stramska 2005), fluorescence signal

711    (Erickson et al. 2019) size structure (Uitz et al. 2006), or PCC (Lange et al. 2018; Werdell et al.

712    2014).

713    Another challenge requiring attention is the mismatch in the products that some of the remote

714    sensing PCC approaches produce and in situ measurements. Above mentioned approaches that

715    output % of biomass (chlorophyll or carbon) to specific size class or group, are hard to validate,

716    as in situ datasets for those parameters are rare, or modeled (carrying lot of assumptions,

717    therefore errors), or unavailable. In those situations, validation is carried through comparison

718    with in situ measurements that explain PCC in different units (see section 4), leading to additional

719    uncertainties that are often unaccounted for. Furthermore there is a spatial and temporal

720    mismatch between in situ and remote sensing data. The current approach, where a 5×5 pixel

721    (nominally) box is centered on a pixel closest to an in situ data point, has shown success at

722    validating indices of total phytoplankton biomass on open ocean scales. A key reason for this is

723    the high dynamic range of *Chl a* concentration (and by proxy, total phytoplankton biomass) in the

724    ocean, which may not be the case for individual phytoplankton taxonomic groups. Issues of sub-

725    pixel variability, however, are not trivial in heterogeneous and dynamic ocean regions, including

726    coastal areas that are likely to have strong tidal influences, river runoff, stratification, sediment

727    resuspension, and other biophysical interactions that create patchy water mass distributions

728    (e.g., Aurin et al. 2013; Ryan et al. 2005). In addition, diel variability in phytoplankton processes

729    causes changes in cell size consistent with patterns of cell growth during the light period and cell

730    division late in the day (Sosik et al. 2003) that is visible in optical properties, including $R_{rs}(\lambda)$ (e.g.,

731    Briggs et al. 2018; Claustre et al. 2002; Concha et al. 2019; Gernez et al. 2011; Henderikx Freitas

732    et al. 2020; Poulin et al. 2018; Stramski et al. 1995; Stramski and Reynolds 1993). In regions with

733 considerable spatiotemporal variability, for example, matching the seasonal climatology can be a

734 useful tool to evaluate and validate different algorithms (Henderikx Freitas and Dierssen 2019).

735 Going forward, thoughtful approaches that take in consideration the spatiotemporal variability

736 of PCC, phytoplankton growth stages, and associated bio-optical retrievals should be applied to

737 the validation of these remote sensing products.

## 6. Phytoplankton from PACE

739 In this review we presented a three-pronged approach to define PCC from PACE; as information

740 that is needed by the users, definable by in-situ methodology, and detectable from satellites

741 (Figure 7). In the preceding sections we defined current capabilities to resolve phytoplankton

742 taxonomy from in situ and space-based observations. We described the strengths and

743 weaknesses of each approach. We also described the availability of various techniques for

744 phytoplankton enumeration and identification, from pigments to DNA to microscopy to imaging

745 in flow approaches. The current suite of satellite PCC algorithms is largely only capable of deriving

746 size classes, with a few that can discriminate unique taxonomic classes. Now the question is: what

747 is possible with PACE? Although we would like to satisfy every user's needs, we must be realistic

748 in our assumptions or expectations, and limitations that are outside of the sphere of technological

749 capabilities of PACE instruments and our methods in-situ.

### 6.1. In situ data requirements

751 Even with PACE's technological advancements, its potential to characterize the phytoplankton

752 community will still depend on the availability of pigment and taxonomic data—that is, how well

753 PCC is assessed in situ and subsequently used in algorithm development and validation.

754 Previously, most PCC algorithms developed for application to satellite data have relied on

755 phytoplankton accessory pigment concentrations to define the phytoplankton groups of interest.

756 A challenge arises in that the biomass of a given phytoplankton group cannot be directly defined

757 using accessory pigment concentrations, as there are assumptions and approximations made

758 when defining groups via pigment proxy (Jeffrey et al. 2011). Diversifying the types of in situ data

759 used to define PCC (see Section 4) as well as use of multiple types of in situ taxonomic approaches

760 will not only will improve the development of robust PACE PCC algorithms, but it will be critical

761 as we go forward, as each data type has its advantages and limitations (see Catlett et al. 2022).

762 Additionally, the limited number of validation exercises have been ad-hoc, and to ensure the

763 quality and fidelity of the PCCs produced from PACE, validation must be done in a continuous

764 standardized fashion (PACE mission 2020).

765 The first steps for validation of PACE algorithms will be using solid foundation built upon heritage

766 ocean color missions such as SeaWiFS, MODIS, MERIS, and VIIRS. From these missions, we have

767 learned that the development, usability, and reliability of ocean color data depends on the

768 availability of quality field data from (and prior to) launch date through the final days of the

769 mission. Here we define "quality field data" as that which have been collected and processed

770 following vetted community protocols, and deposited in an open access long-term data

771 repository, especially those following up-to-date international data standards. The need for

772 calibration and validation of optical oceanographic data during the SeaWiFS era set the stage for

773 development of the SeaWiFS Bio-optical Archive and Storage System (SeaBASS), which today is

774 one of the largest data repositories for optical oceanographic data. SeaBASS is part of NASA's

775 Ocean Biology Distributed Active Archive Center (OB.DAAC), operating under NASA's Earth

776 Observing System Data and Information System (EOSDIS). In addition to serving the ocean color

community as a long-term data repository and data distribution (via OB.DAAC), SeaBASS is

responsible for compiling and curating data used for calibration and validation activities of ocean

color satellite missions, including PACE. Most of the radiometric and bio-optical data used for

validation is shared via NASA NOMAD (NASA bio-Optical Marine Algorithm Dataset), which is a

publicly available, global, high-quality in situ bio-optical data set for use in ocean color algorithm

development and satellite data product validation activities.

When it comes to measurements of phytoplankton and community composition, SeaBASS has

been primarily limited to phytoplankton pigments and relatively few flow cytometry and imaging

datasets. As HPLC based *Chl a* has historically been the preferred validation data type for ocean

color *Chl a,* community has developed standard protocols, and has participated in numerous

round robin comparisons, that allowed standardization of not only *Chl a*, but the whole HPLC

pigment suite across international community (e.g., SeaHARRE round robin series, Hooker et al.

2012). In preparation for PACE and to expand phytoplankton data availability beyond pigments,

SeaBASS implemented community protocols outlined in Neeley et al. (2021) for standardizing

image data collected using imaging-in-flow instruments, such as the IFCB, UVP and FlowCAM, and

associated metadata and other documentation. Standards and best practices for other taxonomic

datasets, such as traditional microscopy and flow cytometry are either published or underway

(Neeley et al. 2023). Incorporating molecular taxonomy datasets would be the next step, as

currently, these data are not easily accessible for validation or dataset-building purposes. In

parallel, community accepted protocols are being developed for a whole suite of other PACE

products, including hyperspectral remote sensing reflectance (Zibordi et al. 2019). As these new

798    datatypes are added to the validation pipeline, they will be included into future versions of

799    NOMAD datasets.

800    Currently, in situ data on phytoplankton concentration and composition are still heavily biased

801    towards some regions and/or are not directly available (Thompson and Carstensen 2022). To

802    allow for the above-mentioned applications, it is critical that we have data that allow us to

803    calibrate/validate satellite products, as well as to validate ESM outputs (Dierssen et al. 2020). As

804    stated by Thompson and Carstensen (2022), this network needs to be carefully planned in order

805    to provide the information that we will need to understand and manage climate variability and

806    change. Several recent grassroots initiatives address the geographic biases in SeaBASS (or

807    NOMAD) towards ocean data by sharing a large database of hyperspectral $R_{rs}(\lambda)$ (> 7500+) along

808    with their co-located *Chl a*, CDOM, and SPM across global freshwater and coastal (Lehman et al.

809    2023), and coastal and open ocean ecosystems (Casey et al. 2020) but with no accompanying PCC

810    datasets.

811    Finally, error propagation and attribution of uncertainties to PCC products are paramount. This

812    information will determine whether an algorithm cab be applied effectively and/or broadly. While

813    uncertainty has long been a topic of discussion in the ocean color community, only recently have

814    significant strides been made to improve uncertainty estimation and error propagation to the

815    measured or modeled products (IOCCG 2019; Kostakis et al. 2021; McKinna et al. 2019). With

816    more complex algorithms, such as the ones used for estimation of PCC, uncertainty associated

817    with underlying ocean color observations (Ibrahim et al. 2022; Zhang et al. 2022) accumulates

818    through each step in the algorithm hierarchy (Siegel et al. 2023). Ancillary data present sources

819    of uncertainty, as well as the theoretical assumptions or empirical data used to develop the PCC

820 algorithm. As PACE has a requirement to estimate uncertainty at the level of $R_{rs}(\lambda)$, such

821 cumulative uncertainty should be calculated and reported with the PCC products and used in

822 future validation exercises in combination with contemporary validation metrics (McKinna et al.

823 2021). Such uncertainty propagation exercises can be helpful when developing future PCC

824 algorithms, to elucidate the interpretation of (e.g., climate driven) trends in PCC observations

825 from space. In addition, having uncertainty as part of the harmful algal blooms (HABs) detection

826 report is crucial for water quality managers, especially when it comes to determining the potential

827 public health risks and implementing appropriate responses.

828 ## 6.2. Data accessibility and open science

829 Chl *a*, as a heritage ocean color product, sets a high bar when it comes to data objectives and

830 potential recognition as an Essential Ocean Variable (EOV, see Muller-Karger et al. 2018).

831 Wilkinson et al. (2016) describe widely-accepted "FAIR" data objectives, listed below with a

832 description of the relevant OB.DAAC practice:

833 ● Findable: total *Chl a* estimates from multiple satellites and sensors are indexed on the

834 Earthdata Search website, including spatial and temporal filters.

835 ● Accessible: single or multiple total *Chl a* datasets may be downloaded on-demand by users

836 or applications that authenticate with a free Earthdata account.

837 ● Interoperable: total *Chl a* datasets are delivered as NetCDF files, which embed metadata

838 in a vocabulary controlled by the Climate and Forecast metadata conventions.

839 ● Reusable: in addition to being in the public domain, total *Chl a* datasets include metadata

840 relevant to the ocean color research (e.g., variable units, sensor calibration parameters,

841 and references to protocols and algorithms) that cover all aspects of the data provenance.

842    PACE ocean color products, including PCC, will be distributed via OB.DAAC, as soon as the data

843    are available. However, it is important to understand that data will be considered "provisional",

844    meaning that they will not be validated (or validation is in progress) and quality may not be

845    optimal. Assessment of the PACE data quality, i.e., validation, will be done in several steps,

846    starting with remote sensing reflectance, followed by *Chl a* (and other products in ocean color

847    standard suite), and ultimately PCC (when data is available).

848    PACE data products for PCC are on the path to achieve each of the FAIR principles; despite being

849    a more complicated variable than total *Chl a* (e.g., PCC is multivariate), there is nothing

850    structurally novel about these products. New components of the vocabulary needed for metadata

851    about phytoplankton composition (Neeley et al., 2021) demonstrate that the ocean color

852    research community is actively engaged in establishing standards for interoperability. The units

853    of PCC variables are well defined, ranging from *Chl a* or carbon density to total or relative

854    abundance. Standard methods will have to be defined so that taxa can be pooled into genera or

855    class (or higher taxonomic levels) depending on the users and goal outcomes. One of the biggest

856    challenges, owing to the types of models that produce L4 data products containing PCC variables,

857    will be the complete description of data source and methodology. Variables estimated through

858    semi-analytical algorithms are relatively easy to document, using equations or open-source

859    software and a short list of parameters. It is more difficult to document the provenance of variable

860    estimates that get introduced after a model-fitting procedure, such as in supervised machine

861    learning. In this case, both the fitted, or trained, model itself and any in situ data used while

862    training must both be FAIR (e.g., Schoening et al. 2022). OB.DAAC and SeaBASS are well suited

863    repositories to support the  data-intensive model-fitting procedures reliance on in situ data.

864  The NASA Science Mission Directorate (SMD) objectives for conducting supported missions,

865  including PACE, through open science practices extends beyond open data. The topic of

866  inclusivity, creating new pathways for end users to become involved in science, is also relevant to

867  open science[9]. While the majority of academia and government entities have computational

868  capacity to download and process large datasets, many of those in the broader PACE CoP may

869  have challenges working with a large number of files in scientific data formats containing multiple

870  parameters, like $R_{rs}(\lambda)$, pigments, and light attenuation, and metadata on calibration, validation,

871  and other provenance. For this reason, the PACE Mission will develop and offer trainings,

872  tutorials, data "recipes", and other forms of community support to ensure that users know how

873  to appropriately access and utilize PACE data products. Further, the PACE team will continue its

874  work with the PACE CoP to identify further training needs and gaps, and gather feedback post-

875  launch on continued challenges and barriers and the development of new products. The PACE

876  team, for example, is working towards making a new type of merged "Water Quality Product"

877  that will highlight specific products for a wide range of users. Such a merged and simplified data

878  product, drawing from different parameters, will be adaptable and can grow as new algorithms

879  come online. Finally, a PACE Community of Potential will also be convened; versus (and in addition

880  to) the CoP, which is made up of more technical users. The Community of Potential will target

881  individuals or groups (within or outside of PACE CoP) who are unfamiliar with satellite data

882  products and PACE capabilities but may be able to leverage and benefit from PACE data products.

883  Members of the Community of Potential may benefit from additional support on remote sensing

884  at large, and the benefits and applications that PACE specifically could offer. For example, white

---

[9] https://science.nasa.gov/open-science-overview

885 papers with case studies may be developed so that new users can understand possible

886 applications and uses of PACE data. In addition, certain new conferences or venues could be

887 identified and targeted where users are less familiar with NASA remote sensing data. Other

888 efforts to engage with new and existing users – whether for awareness raising, training, or some

889 combination – will be considered as they are identified.

890 Steps to reduce financial barriers associated with research on PCC include provisioning of cloud-

891 computing services with timely access to PCC data products, that should not only rely on NASA

892 based providers. The Giovanni (10.1029/2007EO020003) and Google Earth Engine (Ugur et al

893 2021) systems are in-browser applications coupled to cloud computing facilities that provide low-

894 barrier entries to analysis of earth observation variables. The inclusion of new variables on

895 phytoplankton communities in these or other comparable systems is included in the design

896 objectives for PCC data products. Training on software and online data tools for new users will be

897 integral to the mission.

## 898 7. Conclusion

899 In this paper we have identified the strengths and weaknesses of existing approaches and

900 presented how PACE will address some of the remaining gaps and challenges for quantifying

901 PCC. The knowledge gained from this mission will rely on the existence of appropriate validation

902 products (e.g., quality, coverage, diversity of products available, error and uncertainties

903 requirements), the use of the various approaches currently available (in situ, satellite, and

904 model products) as well as the distribution of data products following FAIR principles including

905 training of the community and developing user friendly data products. Improved

906 characterization of PCC is key to understanding aquatic ecosystems, our ocean's health, and the

907 ocean's effects on global climate. From local applications, such as water quality assessments

908 that are focused on specific phytoplankton taxa, to global applications such as the role that

909 ocean plays in carbon sequestration, the needs for information on phytoplankton composition

910 requires the community to work together to develop the next generation of data products. This

911 is especially critical when it comes to improving our understanding of the impact that climate

912 change has on the ocean and developing effective management strategies for the ocean and

913 other aquatic resources, both today and in the future.

## Acknowledgments

921 Credit author statement

922 **Ivona Cetinić**: Conceptualization, Writing - Original Draft, Writing - Review & Editing, Visualization

923 **Cecile S. Rousseaux:** Conceptualization, Writing - Original Draft, Writing - Review & Editing,

924 Visualization **Ian T. Carroll** Conceptualization, Writing - Original Draft, Writing - Review & Editing,

925 Visualization **Alison P. Chase**: Conceptualization, Writing - Original Draft, Writing - Review &

926 Editing, Visualization **Sasha J. Kramer**: Conceptualization, Writing - Original Draft, Writing -

927 Review & Editing, Visualization **P. Jeremy Werdell** Conceptualization, Writing - Original Draft,

928 Writing - Review & Editing **David. A. Siegel**: Conceptualization, Writing - Original Draft, **Heidi M.**

929 **Dierssen** Conceptualization, Writing - Original Draft, **Dylan Catlett:** Conceptualization, Writing -

930 Original Draft , **Aimee Neeley**: Conceptualization, Writing - Original Draft , **Inia M. Soto Ramos**:

931 Conceptualization, Writing - Original Draft, **Jennifer L. Wolny**: Conceptualization, Writing -

932 Original Draft **Natasha Sadoff:** Writing - Original Draft  **Erin Urquhart:** Writing - Original Draft,

933 **Toby K. Westberry**: Writing - Review & Editing, **Dariusz Stramski:** Writing - Review & Editing,

934 **Nima Pahlevan**: Writing - Review & Editing, **Bridget N. Seegers**: Writing - Review & Editing

935 **Emerson Sirk:** Writing - Review & Editing, **Priscila Kienteca Lange**: Writing - Review & Editing ,

936 **Ryan A. Vandermeulen**: Writing - Review & Editing, **Jason R. Graff**: Writing - Review & Editing,

937 **James G. Allen**: Writing - Review & Editing, **Peter Gaube**: Writing - Review & Editing, **Lachlan I.**

938 **W. McKinna:** Writing - Review & Editing, **S.Morgaine McKibben**: Writing - Review & Editing, **Caren**

939 **E. Binding**: Writing - Review & Editing **Violeta Sanjuan Calzado**: Writing - Review & Editing,

940 **Michael Sayers:** Writing - Review & Editing

941

942 Funding

961

## References

963 Aguirre-Gómez, R., Boxall, S., & Weeks, A. (2001). Detecting photosynthetic algal pigments in natural
964 populations using a high-spectral-resolution spectroradiometer. *International Journal of Remote Sensing,*
965 *22*, 2867-2884
966 Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society:*
967 *Series B (Methodological), 44*, 139-160
968 Alvain, S., Moulin, C., Dandonneau, Y., & Bréon, F.-M. (2005). Remote sensing of phytoplankton groups in
969 case 1 waters from global SeaWiFS imagery. *Deep Sea Research Part I: Oceanographic Research Papers,*
970 *52*, 1989-2004
971 Alvain, S., Moulin, C., Dandonneau, Y., & Loisel, H. (2008). Seasonal distribution and succession of
972 dominant phytoplankton groups in the global ocean: A satellite view. *Global Biogeochemical Cycles, 22*
973 Anderson, C.R., Siegel, D.A., Kudela, R.M., & Brzezinski, M.A. (2009). Empirical models of toxigenic
974 Pseudo-nitzschia blooms: potential use as a remote detection tool in the Santa Barbara Channel. *Harmful*
975 *Algae*, 8, 478-492
976 Anderson, C.R., Berdalet, E., Kudela, R.M., Cusack, C.K., Silke, J., O'Rourke, E., Dugan, D., McCammon, M.,
977 Newton, J.A., Moore, S.K. & Paige, K., (2019). Scaling up from regional case studies to a global harmful
978 algal bloom observing system. *Frontiers in Marine Science*, 9
979 Anderson, S.I., Barton, A.D., Clayton, S., Dutkiewicz, S., & Rynearson, T.A. (2021). Marine phytoplankton
980 functional types exhibit diverse responses to thermal change. *Nature Communications, 12*, 6413
981 Aurin, D., Mannino, A., & Franz, B. (2013). Spatially resolving ocean color and sediment dispersion in
982 river plumes, coastal systems, and continental shelf waters. *Remote Sensing of Environment, 137*, 212-
983 225

Behrenfeld, M.J., O'Malley, R.T., Siegel, D.A., McClain, C.R., Sarmiento, J.L., G.C., F., Milligan, A.J., Falkowski, P.G., Letelier, R.M., & Boss, E.S. (2006). Climate-driven trends in contemporary ocean productivity. Nature, 444, 752-755

Beardall, J., Allen, D., Bragg, J., Finkel, Z.V., Flynn, K.J., Quigg, A., Rees, T.A.V., Richardson, A., & Raven, J.A. (2009). Allometry and stoichiometry of unicellular, colonial and multicellular phytoplankton. *New Phytologist, 181*, 295-309

Ben Mustapha, Z., Alvain, S., Jamet, C., Loisel, H., & Dessailly, D. (2014). Automatic classification of water-leaving radiance anomalies from global SeaWiFS imagery: Application to the detection of phytoplankton groups in open ocean waters. *Remote Sensing of Environment, 146*, 97-112

Bernard, S., Kudela, R. M., Robertson Lain, L., & Pitcher, G. (2021). Observation of harmful algal blooms with ocean colour radiometry. International Ocean-Colour Coordinating Group (IOCCG) Report Series, No. 20, Dartmouth, Canada. http://dx.doi.org/10.25607/OBP-1042Bopp, L., Aumont, O., Cadule, P., Alvain, S., & Gehlen, M. (2005). Response of diatoms distribution to global warming and potential implications: A global model study. *Geophysical Research Letters, 32*, n/a-n/a

Boss, E., M.S. Twardowski, & Herring, S. (2001). Shape of particulate beam attenuation spectrum and its inversion to obtain the shape of the particle size distribution. *Applied Optics, 40*, 4885-4893

Bracher, A., Bouman, H.A., Brewin, R.J.W., Bricaud, A., Brotas, V., Ciotti, A.M., Clementson, L., Devred, E., Di Cicco, A., Dutkiewicz, S., Hardman-Mountford, N.J., Hickman, A.E., Hieronymi, M., Hirata, T., Losa, S.N., Mouw, C.B., Organelli, E., Raitsos, D.E., Uitz, J., Vogt, M., & Wolanin, A. (2017). Obtaining Phytoplankton Diversity from Ocean Color: A Scientific Roadmap for Future Development. *Frontiers in Marine Science, 4*

Bracher, A., Brewin, R.J.W., Ciotti, A.M., Clementson, L.A., Hirata, T., Kostadinov, T.S., Mouw, C.B., & Organelli, E. (2022). Chapter 7 - Applications of satellite remote sensing technology to the analysis of phytoplankton community structure on large scales. In L.A. Clementson, R.S. Eriksen, & A. Willis (Eds.), *Advances in Phytoplankton Ecology* (pp. 217-244): Elsevier

Bracher, A., Taylor, M.H., Taylor, B., Dinter, T., Röttgers, R., & Steinmetz, F. (2015). Using empirical orthogonal functions derived from remote-sensing reflectance for the prediction of phytoplankton pigment concentrations. *Ocean Science, 11*, 139-158

Bracher, A., Vountas, M., Dinter, T., Burrows, J., Röttgers, R., & Peeken, I. (2009). Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences, 6*, 751-764

Bradley, I.M., Pinto, A.J., Guest, J.S., & Voordouw, G. (2016). Design and Evaluation of Illumina MiSeq-Compatible, 18S rRNA Gene-Specific Primers for Improved Characterization of Mixed Phototrophic Communities. *Applied and Environmental Microbiology, 82*, 5878-5891

Brewin, R.J.W., Ciavatta, S., Sathyendranath, S., Jackson, T., Tilstone, G., Curran, K., Airs, R.L., Cummings, D., Brotas, V., Organelli, E., Dall'Olmo, G., & Raitsos, D.E. (2017). Uncertainty in Ocean-Color Estimates of Chlorophyll for Phytoplankton Groups. *Frontiers in Marine Science, 4*

Brewin, R.J., Hardman-Mountford, N.J., Lavender, S.J., Raitsos, D.E., Hirata, T., Uitz, J., Devred, E., Bricaud, A., Ciotti, A., & Gentili, B. (2011a). An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sensing of Environment, 115*, 325-339

Brewin, R.J., Sathyendranath, S., Hirata, T., Lavender, S.J., Barciela, R.M., & Hardman-Mountford, N.J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecological Modelling, 221*, 1472-1483

Brewin, R.J., Sathyendranath, S., Jackson, T., Barlow, R., Brotas, V., Airs, R., & Lamont, T. (2015). Influence of light in the mixed-layer on the parameters of a three-component model of phytoplankton size class. *Remote Sensing of Environment, 168*, 437-450

Brewin, R.J.W., Devred, E., Sathyendranath, S., Lavender, S.J., & Hardman-Mountford, N.J. (2011b). Model of phytoplankton absorption based on three size classes. *Applied Optics, 50*, 4535-4549

Bricaud, A., Ciotti, A.M., & Gentili, B. (2012). Spatial-temporal variations in phytoplankton size and colored detrital matter absorption at global and regional scales, as derived from twelve years of SeaWiFS data (1998–2009). *Global Biogeochemical Cycles, 26*, GB1010

Bricaud, A., Mejia, C., Blondeau-Patissier, D., Claustre, H., Crepon, M., & Thiria, S. (2007). Retrieval of pigment concentrations and size structure of algal populations from their absorption spectra using multilayered perceptrons. *Applied Optics, 46*, 1251-1260

Briggs, N., Guðmundsson, K., Cetinić, I., D'Asaro, E., Rehm, E., Lee, C., & Perry, M.J. (2018). A multi-method autonomous assessment of primary productivity and export efficiency in the springtime North Atlantic. *Biogeosciences, 15*, 4515-4532

Brown, C.W., & Yoder, J.A. (1994). Coccolithophorid blooms in the global ocean. *Journal of Geophysical Research: Oceans, 99*, 7467-7482

Cael, B., Chase, A., & Boss, E. (2020). Information content of absorption spectra and implications for ocean color inversion. *Applied Optics, 59*, 3971-3984

Cael, B.B., Begouen Demeaux, C., Henson, S., Stock, C.A., Taboada, F.G., John, J.G., & Barton, A.D. (2022a). Marine Ecosystem Changepoints Spread Under Ocean Warming in an Earth System Model. *Journal of Geophysical Research: Biogeosciences, 127*, e2021JG006571

Cael, B.B., Bisson, K., Boss, E., Dutkiewicz, S., & Henson, S. (2022c). Global climate change trends detected in indicators of ocean ecology. *EarthArXiv*

Cael, B. B., Bisson, K.M, Boss, E. S. & Erickson, Z. (2022b). Extracting information from ocean color. *Authorea*.

Calvo-Díaz, A., & Morán, X.A.G. (2006). Seasonal dynamics of picoplankton in shelf waters of the southern Bay of Biscay. *Aquatic Microbial Ecology, 42*, 159-174

Caracappa, J. C., Beet, A., Gaichas, S., Gamble, R. J., Hyde, K. J., Large, S. I., ... & Saba, V. S. (2022). A northeast United States Atlantis marine ecosystem model with ocean reanalysis and ocean color forcing. *Ecological Modelling*, 471, 110038.

Castagna, A., Dierssen, H., Organelli, E., Bogorad, M., Mortelmans, J., Vyverman, W., & Sabbe, K. (2021). Optical Detection of Harmful Algal Blooms in the Belgian Coastal Zone: A Cautionary Tale of Chlorophyll c3. *Frontiers in Marine Science, 8*

Casey, J. R., Aucan, J. P., Goldberg, S. R., & Lomas, M. W. (2013). Changes in partitioning of carbon amongst photosynthetic pico-and nano-plankton groups in the Sargasso Sea in response to changes in the North Atlantic Oscillation. *Deep Sea Research Part II: Topical Studies in Oceanography*, *93*, 58-70.

Casey, K. A., Rousseaux, C. S., Gregg, W. W., Boss, E., Chase, A. P., Craig, S. E., ... & Maritorena, S. (2020). A global compilation of in situ aquatic high spectral resolution inherent and apparent optical property data for remote sensing applications. *Earth system science data*, *12*, 1123-1139

Catlett, D., Matson, P.G., Carlson, C.A., Wilbanks, E.G., Siegel, D.A., & Iglesias-Rodriguez, M.D. (2020). Evaluation of accuracy and precision in an amplicon sequencing workflow for marine protist communities. *Limnology and Oceanography: Methods, 18*, 20-40

Catlett, D., Siegel, D.A., Matson, P.G., Wear, E.K., Carlson, C.A., Lankiewicz, T.S., & Iglesias-Rodriguez, M.D. (2022). Integrating phytoplankton pigment and DNA meta-barcoding observations to determine phytoplankton composition in the coastal ocean. *In revision at Limnology and Oceanography*

Catlett, D., Siegel, D.A., Simons, R.D., Guillocheau, N., Henderikx-Freitas, F., & Thomas, C.S. (2021). Diagnosing seasonal to multi-decadal phytoplankton group dynamics in a highly productive coastal ecosystem. *Progress in Oceanography, 197*, 102637

Catlett, D.S., & Siegel, D.A. (2018). Phytoplankton Pigment Communities Can be Modeled Using Unique Relationships With Spectral Absorption Signatures in a Dynamic Coastal Environment. *Journal of Geophysical Research: Oceans, 123*, 246-264

Chase, A., Boss, E., Zaneveld, R., Bricaud, A., Claustre, H., Ras, J., Dall'Olmo, G., & Westberry, T.K. (2013). Decomposition of in situ particulate absorption spectra. *Methods in Oceanography, 7*, 110-124

Chase, A.P., Boss, E., Cetinić, I., & Slade, W. (2017). Estimation of Phytoplankton Accessory Pigments From Hyperspectral Reflectance Spectra: Toward a Global Algorithm. *Journal of Geophysical Research: Oceans, 122*, 9725-9743

Chase, A.P., Boss, E.S., Haëntjens, N., Culhane, E., Roesler, C., & Karp-Boss, L. (2022). Plankton Imagery Data Inform Satellite-Based Estimates of Diatom Carbon. *Geophysical Research Letters, 49*, e2022GL098076

Chase, A.P., Kramer, S.J., Haëntjens, N., Boss, E.S., Karp-Boss, L., Edmondson, M., & Graff, J.R. (2020). Evaluation of diagnostic pigments to estimate phytoplankton size classes. *Limnology and Oceanography: Methods, 18*, 570-584

Chowdhary, J., Zhai, P.-W., Boss, E., Dierssen, H., Frouin, R., Ibrahim, A., Lee, Z., Remer, L.A., Twardowski, M., Xu, F., Zhang, X., Ottaviani, M., Espinosa, W.R., & Ramon, D. (2019). Modeling Atmosphere-Ocean Radiative Transfer: A PACE Mission Perspective. *Frontiers in Earth Science*, 7

Ciavatta, S., Brewin, R., Skakala, J., Polimene, L., de Mora, L., Artioli, Y., & Allen, J.I. (2018). Assimilation of ocean-color plankton functional types to improve marine ecosystem simulations. *Journal of Geophysical Research: Oceans, 123*, 834-854

Ciotti, A.M., & Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. *Limnology and Oceanography: Methods, 4*, 237-253

Claustre, H., Bricaud, A., Babin, M., Bruyant, F., Guillou, L., Le Gall, F., Marie, D., & Partensky, F. (2002). Diel variations in Prochlorococcus optical properties. *Limnology and Oceanography, 47*, 1637-1647

Clayton, S., Gibala-Smith, L., Mogatas, K., Flores-Vargas, C., Marciniak, K., Wigginton, M., & Mulholland, M.R. (2022). Imaging Technologies Build Capacity and Accessibility in Phytoplankton Species Identification Expertise for Research and Monitoring: Lessons Learned During the COVID-19 Pandemic. *Frontiers in Microbiology, 13*, 823109

Concha, J., Mannino, A., Franz, B., & Kim, W. (2019). Uncertainties in the Geostationary Ocean Color Imager (GOCI) Remote Sensing Reflectance for Assessing Diurnal Variability of Biogeochemical Processes. *Remote Sensing, 11*, 295

Craig, S.E., Lohrenz, S.E., Lee, Z.P., Mahoney, K.L., Kirkpatrick, G.J., Schofield, O.M., & Steward, R.G. (2006). Use of hyperspectral remote sensing reflectance for detection and assessment of the harmful alga, Karenia brevis. *Applied Optics, 45*, 5414-5425

Cram, J. A., Weber, T., Leung, S. W., McDonnell, A. M., Liang, J. H., & Deutsch, C. (2018). The role of particle size, ballast, temperature, and oxygen in the sinking flux to the deep sea. *Global Biogeochemical Cycles*, 32, 858-876

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., …& Karsenti, E. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science, 348*, 1261605

Defoin-Platel, M., & Chami, M. (2007). How ambiguous is the inverse problem of ocean color in coastal waters? *Journal of Geophysical Research: Oceans, 112*

Devred, E., Sathyendranath, S., Stuart, V., Maass, H., Ulloa, O., & Platt, T. (2006). A two-component model of phytoplankton absorption in the open ocean: Theory and applications. *Journal of Geophysical Research, 111*

Devred, E., Sathyendranath, S., Stuart, V., & Platt, T. (2011). A three component classification of phytoplankton absorption spectra: Application to ocean-color data. *Remote Sensing of Environment, 115*, 2255-2266

Di Cicco, A., Sammartino, M., Marullo, S., & Santoleri, R. (2017). Regional Empirical Algorithms for an Improved Identification of Phytoplankton Functional Types and Size Classes in the Mediterranean Sea Using Satellite Data. *Frontiers in Marine Science, 4*

Dierssen, H., Bracher, A., Brando, V., Loisel, H., & Ruddick, K. (2020). Data needs for hyperspectral detection of algal diversity across the globe. *Oceanography, 33*, 74-79

Dierssen, H., McManus, G.B., Chlus, A., Qiu, D., Gao, B.-C., & Lin, S. (2015). Space station image captures a red tide ciliate bloom at high spectral and spatial resolution. *Proceedings of the National Academy of Sciences, 112*, 14783-14787

Dinauer, A., Laufkötter, C., Doney, S.C., & Joos, F. (2022). What Controls the Large-Scale Efficiency of Carbon Transfer Through the Ocean's Mesopelagic Zone? Insights From a New, Mechanistic Model (MSPACMAM). *Global Biogeochemical Cycles, 36*, e2021GB007131

Dupouy, C., Benielli-Gary, D., Neveux, J., Dandonneau, Y., & Westberry, T.K. (2011). An algorithm for detecting Trichodesmium surface blooms in the South Western Tropical Pacific. Biogeosciences, 8, 3631-3647

Durden, J.M., Luo, J.Y., Alexander, H., Flanagan, A.M., & Grossmann, L. (2017). Integrating "Big Data" into Aquatic Ecology: Challenges and Opportunities. *Limnology and Oceanography Bulletin, 26*, 101-108

Dutkiewicz, S., Cermeno, P., Jahn, O., Follows, M.J., Hickman, A.E., Taniguchi, D.A.A., & Ward, B.A. (2020). Dimensions of marine phytoplankton diversity. *Biogeosciences, 17*, 609-634

Dutkiewicz, S., Scott, J.R., & Follows, M.J. (2013). Winners and losers: Ecological and biogeochemical changes in a warming ocean. *Global Biogeochemical Cycles, 27*, 463-477

Eayrs, C., Li, X., Raphael, M.N., & Holland, D.M. (2021). Rapid decline in Antarctic sea ice in recent years hints at future change. *Nature Geoscience, 14*, 460-464

Elder, L., & Elbrächter, M. (2010). The Utermöhl method for quantitative phytoplankton analysis. In B. Karlson, C. Cusack, & E. Bresnan (Eds.), *Microscopic and molecular methods for quantitative phytoplankton analysis*. PARIS: UNESCO

Erickson, Z. K., Frankenberg, C., Thompson, D. R., Thompson, A. F., & Gierach, M. (2019). Remote sensing of chlorophyll fluorescence in the ocean using imaging spectrometry: toward a vertical profile of fluorescence. *Geophysical Research Letters*, 46, 1571-1579.

Fisher, J.L., Peterson, W.T., & Rykaczewski, R.R. (2015). The impact of El Niño events on the pelagic food chain in the northern California Current. *Global Change Biology, 21*, 4401-4414

Flexas, M.M., Thompson, A.F., Schodlok, M.P., Zhang, H., & Speer, K. (2022). Antarctic Peninsula warming triggers enhanced basal melt rates throughout West Antarctica. *Science Advances, 8*, eabj9134

Forget, M.-H.l.n., Carignan, R., & Hudon, C. (2009). Influence of diel cycles of respiration, chlorophyll, and photosynthetic parameters on the summer metabolic balance of temperate lakes and rivers. *Canadian Journal of Fisheries & Aquatic Sciences, 66*, 1048-1058

Franz, B.A., Cetinic, I., Scott, J.P., Siegel, A., & Westberry, T.K. (2021). Global ocean phytoplankton [in "State of the Climate in 2020"]. *Bulletin of the American Meteorological Society, 102*, S179 - S193

Frieder, C.A., Yan, C., Chamecki, M., Dauhajre, D., McWilliams, J.C., Infante, J., McPherson, M.L., Kudela, R.M., Kessouri, F., Sutula, M., Arzeno-Soltero, I.B., & Davis, K.A. (2022). A Macroalgal Cultivation Modeling System (MACMODS): Evaluating the Role of Physical-Biological Coupling on Nutrients and Farm Yield. *Frontiers in Marine Science, 9*

Frouin, R.J., Franz, B.A., Ibrahim, A., Knobelspiesse, K., Ahmad, Z., Cairns, B., Chowdhary, J., Dierssen, H.M., Tan, J., Dubovik, O., Huang, X., Davis, A.B., Kalashnikova, O., Thompson, D.R., Remer, L.A., Boss, E., Coddington, O., Deschamps, P.-Y., Gao, B.-C., Gross, L., Hasekamp, O., Omar, A., Pelletier, B., Ramon, D., Steinmetz, F., & Zhai, P.-W. (2019). Atmospheric Correction of Satellite Ocean-Color Imagery During the PACE Era. *Frontiers in Earth Science, 7*

Fujiwara, A., Hirawake, T., Suzuki, K., & Saitoh, S.-I. (2011). Remote sensing of size structure of phytoplankton communities using optical properties of the Chukchi and Bering Sea shelf region. *Biogeosciences, 8*, 3567

Fuss, S., Canadell, J.G., Peters, G.P., Tavoni, M., Andrew, R.M., Ciais, P., Jackson, R.B., Jones, C.D., Kraxner, F., Nakicenovic, N., Le Quéré, C., Raupach, M.R., Sharifi, A., Smith, P., & Yamagata, Y. (2014). Betting on negative emissions. *Nature Climate Change, 4*, 850-853

Garver, S.A., Siegel, D.A., & Mitchell, B.G. (1994). Variability in near-surface particulate absorption spectra: What can a satellite ocean color imager see? *Limnology and Oceanography, 39*, 1349-1367

GESAMP (2019). Guidelines for the monitoring and assessment of plastic litter in the ocean-GESAMP reports and studies no. 99. In P. Kershaw, A. Turra, & F. Galgani (Eds.), *GESAMP Reports and Studies* (p. 130): IMO/FAO/UNESCO-IOC/UNIDO/WMO/IAEA/UN/UNEP/UNDP/ISA Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection

Gernez,P., Reynolds, R. A., & Stramski, D. (2014). Within-day variability of particulate organic carbon and remote-sensing reflectance during a bloom of Phaeocystis antarctica in the Ross Sea, Antarctica. *International Journal of Remote Sensing*, 35, 454–477.

Gitelson, A.A., Gao, B.-C., Li, R.-R., Berdnikov, S., & Saprygin, V. (2011). Estimation of chlorophyll-a concentration in productive turbid waters using a Hyperspectral Imager for the Coastal Ocean—the Azov Sea case study. *Environmental Research Letters, 6*, 024023

Gittings, J.A., Brewin, R.J., Raitsos, D.E., Kheireddine, M., Ouhssain, M., Jones, B.H., & Hoteit, I. (2019). Remotely sensing phytoplankton size structure in the Red Sea. *Remote Sensing of Environment, 234*, 111387

Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., & Egozcue, J.J. (2017). Microbiome Datasets Are Compositional: And This Is Not Optional. *Frontiers in Microbiology, 8*

Godhe, A., Asplund, M.E., Härnström, K., Saravanan, V., Tyagi, A., & Karunasagar, I. (2008). Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. *Applied and Environmental Microbiology, 74*, 7174-7182

Godhe, A., Cusack, C., Pedersen, J., Andersen, P., Anderson, D.M., Bresnan, E., Cembella, A., Dahl, E., Diercks, S., Elbrächter, M., Edler, L., Galluzzi, L., Gescher, C., Gladstone, M., Karlson, B., Kulis, D., LeGresley, M., Lindahl, O., Marin, R., McDermott, G., Medlin, L.K., Naustvoll, L.-J., Penna, A., & Töbe, K. (2007). Intercalibration of classical and molecular techniques for identification of Alexandrium fundyense (Dinophyceae) and estimation of cell densities. *Harmful Algae, 6*, 56-72

Gong, W., & Marchetti, A. (2019). Estimation of 18S Gene Copy Number in Marine Eukaryotic Plankton Using a Next-Generation Sequencing Approach. *Frontiers in Marine Science, 6*

Gordon, H. R., & McCluney, W. R. (1975). Estimation of the depth of sunlight penetration in the sea for remote sensing. *Applied optics, 14,* 413-416

Gowen, R.J. (1994). Managing eutrophication associated with aquaculture development. *Journal of Applied Ichthyology, 10*, 242-257

Graff, J. R., Milligan, A. J., & Behrenfeld, M. J., (2012). The measurement of phytoplankton biomass using flow-cytometric sorting and elemental analysis of carbon. *Limnology and Oceanography Methods*, *10*

Gregg, W.W., & Casey, N.W. (2007). Modeling coccolithophores in the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography, 54*, 447-477

Gregg, W.W., & Rousseaux, C.S. (2014). Decadal trends in global pelagic ocean chlorophyll: A new assessment integrating multiple satellites, in situ data, and models. *Journal of Geophysical Research: Oceans, 119*, 5921-5933

Gregg, W.W., Rousseaux, C.S., & Franz, B.A. (2017). Global trends in ocean phytoplankton: a new assessment using revised ocean colour data. *Remote Sensing Letters, 8*, 1102-1111

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

1219 Ha, S.-Y., Min, J.-O., Joo, H., Kim, M.-S., Kang, S.-H., & Shin, K.-H. (2018). Synthesis of mycosporine-like
1220 amino acids by a size-fractionated marine phytoplankton community of the arctic beaufort sea. *Journal*
1221 *of Photochemistry and Photobiology B: Biology, 188*, 87-94
1222 Haas, L.W., & Marshall, H.G. (1989). A comparison of the utermohl and epifluorescent microscopic
1223 techniques for quantifying natural picophytoplankton
1224 Haëntjens, N., Boss, E.S., Graff, J.R., Chase, A.P., & Karp-Boss, L. (2022). Phytoplankton size distributions
1225 in the western North Atlantic and their seasonal variability. *Limnology and Oceanography, 67*, 1865-1878
1226 Harvey, H.W. (1934). Measurement of Phytoplankton Population. *Journal of the Marine Biological*
1227 *Association of the United Kingdom, 19*, 761-773
1228 Hasekamp, O.P., Fu, G., Rusli, S.P., Wu, L., Di Noia, A., Brugh, J.a.d., Landgraf, J., Martijn Smit, J., Rietjens,
1229 J., & van Amerongen, A. (2019). Aerosol measurements by SPEXone on the NASA PACE mission: expected
1230 retrieval capabilities. *Journal of Quantitative Spectroscopy and Radiative Transfer, 227*, 170-184
1231 Havskum, H., Schlüter, L., Scharek, R., Berdalet, E., & Jacquet, S. (2004). Routine quantification of
1232 phytoplankton groups—microscopy or pigment analyses? *Marine Ecology Progress Series, 273*, 31-42
1233 Henderikx Freitas, F., & Dierssen, H.M. (2019). Evaluating the seasonal and decadal performance of red
1234 band difference algorithms for chlorophyll in an optically complex estuary with winter and summer
1235 blooms. *Remote Sensing of Environment, 231*, 111228
1236 Henderikx Freitas, F., Dugenne, M., Ribalet, F., Hynes, A., Barone, B., Karl, D.M., & White, A.E. (2020).
1237 Diel variability of bulk optical properties associated with the growth and division of small phytoplankton
1238 in the North Pacific Subtropical Gyre. *Applied Optics, 59*, 6702-6716
1239 Henriksen, P., Riemann, B., Kaas, H., Sørenson, H.M., & Sørenson, H.L. (2002). Effects of nutrient-
1240 limitation and irradiance on marine phytoplankton pigments. *Journal of Plankton Research, 24*, 835-858
1241 Henson, S.A., Laufkötter, C., Leung, S., Giering, S.L.C., Palevsky, H.I., & Cavan, E.L. (2022). Uncertain
1242 response of ocean biological carbon export in a changing world. *Nature Geoscience, 15*, 248-254
1243 Henson, S.A., Sarmiento, J.L., Dunne, J.P., Bopp, L., Lima, I., Doney, S.C., John, J., & Beaulieu, C. (2010).
1244 Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity.
1245 *Biogeosciences, 7*, 621-640
1246 Hirata, T., Aiken, J., Hardman-Mountford, N., Smyth, T., & Barlow, R. (2008). An absorption model to
1247 determine phytoplankton size classes from satellite ocean colour. *Remote Sensing of Environment, 112*,
1248 3153-3159
1249 Hirata, T., Hardman-Mountford, N., Brewin, R., Aiken, J., Barlow, R., Suzuki, K., Isada, T., Howell, E.,
1250 Hashioka, T., & Noguchi-Aita, M. (2011). Synoptic relationships between surface Chlorophyll-a and
1251 diagnostic pigments specific to phytoplankton functional types. *Biogeosciences, 8*, 311
1252 Hoepffner, N., & Sathyendranath, S. (1991). Effect of pigment composition on absorption properties of
1253 phytoplankton. *Mar. Ecol. Prog. Ser, 73*, l-23
1254 Hoepffner, N., & Sathyendranath, S. (1993). Determination of the Major Groups of Phytoplankton
1255 Pigments from the Absorption-Spectra of Total Particulate Matter. *Journal of Geophysical Research-*
1256 *Oceans, 98*, 22789-22803
1257 Hooker, S.B., Clementson, L., Thomas, C.S., Schlüter, L., Allerup, M., Ras, J., Claustre, H., Normandeau, C.,
1258 Cullen, J.J., Kienast, M., Kozlowski, W., Vernet, M., Chakraborty, S., Lohrenz, S.E., Tuel, M., Redalje, D.,
1259 Cartaxana, P., Mendes, C.R., Brotas, V., Prabhu Matondkar, S.G., Parab, S.G., Neeley, A., & Egeland, E.S.
1260 (2012). The Fifth SeaWiFS HPLC Analysis Round-Robin Experiment (SeaHARRE-5). In, *NASA Technical*
1261 *Reports* (pp. 1-108). Greenbelt, Maryland: NASA Goddard Space Flight Center
1262 Hu, C., Cannizzaro, J., Carder, K.L., Muller-Karger, F.E., & Hardy, R. (2010). Remote detection of
1263 Trichodesmium blooms in optically complex coastal waters: Examples with MODIS full-spectral data.
1264 Remote Sensing of Environment, 114, 2048-2058

1265 Hunter-Cevera, K.R., Hamilton, B.R., Neubert, M.G., & Sosik, H.M. (2021). Seasonal environmental
1266 variability drives microdiversity within a coastal Synechococcus population. *Environmental Microbiology,*
1267 *23*, 4689-4705

1268 Ibrahim, A., Franz, B., Ahmad, Z., Healy, R., Knobelspiesse, K., Gao, B.-C., Proctor, C., & Zhai, P.-W. (2018).
1269 Atmospheric correction for hyperspectral ocean color retrieval with application to the Hyperspectral
1270 Imager for the Coastal Ocean (HICO). *Remote Sensing of Environment, 204*, 60-75

1271 Ibrahim, A., Franz, B.A., Sayer, A.M., Knobelspiesse, K., Zhang, M., Bailey, S.W., McKinna, L.I.W., Gao, M.,
1272 & Werdell, P.J. (2022). Optimal estimation framework for ocean color atmospheric correction and pixel-
1273 level uncertainty quantification. *Applied Optics, 61*, 6453-6475

1274 IOCCG (2014). *Phytoplankton Functional Types from Space*. Dartmouth, Canada: IOCCG

1275 IOCCG (2019). *Uncertainties in Ocean Colour Remote Sensing*. Dartmouth, Canada: IOCCG

1276 IOCCG (2020). *Synergy between Ocean Colour and Biogeochemical/Ecosystem Models*. Dartmouth,
1277 Canada: IOCCG

1278 IPCC (2019). Special Report on the Ocean and Cryosphere in a Changing Climate. In H.-O. Pörtner, D.C.
1279 Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A.
1280 Okem, J. Petzold, B. Rama, & N.M. Weyer (Eds.). Cambridge, United Kingdom and New York, NY, USA

1281 IPCC (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the
1282 Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In V. Masson-Delmotte, P.
1283 Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K.
1284 Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.).
1285 Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press

1286 Isada, T., Hirawake, T., Kobayashi, T., Nosaka, Y., Natsuike, M., Imai, I., Suzuki, K., & Saitoh, S.-I. (2015).
1287 Hyperspectral optical discrimination of phytoplankton community structure in Funka Bay and its
1288 implications for ocean color remote sensing of diatoms. *Remote Sensing of Environment, 159*, 134-151

1289 Jamet, C., Ibrahim, A., Ahmad, Z., Angelini, F., Babin, M., Behrenfeld, M.J., Boss, E., Cairns, B., Churnside,
1290 J., Chowdhary, J., Davis, A.B., Dionisi, D., Duforêt-Gaurier, L., Franz, B., Frouin, R., Gao, M., Gray, D.,
1291 Hasekamp, O., He, X., Hostetler, C., Kalashnikova, O.V., Knobelspiesse, K., Lacour, L., Loisel, H., Martins,
1292 V., Rehm, E., Remer, L., Sanhaj, I., Stamnes, K., Stamnes, S., Victori, S., Werdell, J., & Zhai, P.-W. (2019).
1293 Going Beyond Standard Ocean Color Observations: Lidar and Polarimetry. *Frontiers in Marine Science, 6*

1294 Jeffrey, S., MacTavish, H., Dunlap, W., Vesk, M., & Groenewoud, K. (1999). Occurrence of UVA-and UVB-
1295 absorbing compounds in 152 species (206 strains) of marine microalgae. *Marine Ecology Progress Series,*
1296 *189*, 35-51

1297 Jeffrey, S.W., Wright, S.W., & Zapata, M. (2011). Microalgal classes and their signature pigments. In S.
1298 Roy, C.A. Llewellyn, E.S. Egeland, & G. Johnsen (Eds.), *Phytoplankton Pigments: Characterization,*
1299 *Chemotaxonomy, and Application in Oceanography* (pp. 3-77). Cambridge, United Kingdom: Cambridge
1300 University Press

1301 Johnson, T.R., Beard, K., Brady, D.C., Byron, C.J., Cleaver, C., Duffy, K., Keeney, N., Kimble, M., Miller, M.,
1302 Moeykens, S., Teisl, M., van Walsum, G.P., & Yuan, J. (2019). A Social-Ecological System Framework for
1303 Marine Aquaculture Research. *Sustainability, 11*, 2522

1304 Johnson, Z.I., & Martiny, A.C. (2015). Techniques for Quantifying Phytoplankton Biodiversity. *Annual*
1305 *Review of Marine Science, 7*, 299-324

1306 Jones, E.M., Baird, M.E., Mongin, M., Parslow, J., Skerratt, J., Lovell, J., Margvelashvili, N., Matear, R.J.,
1307 Wild-Allen, K., & Robson, B. (2016). Use of remote-sensing reflectance to constrain a data assimilating
1308 marine biogeochemical model of the Great Barrier Reef. *Biogeosciences, 13*, 6441-6469

1309 Kahru, M., Anderson, C., Barton, A.D., Carter, M.L., Catlett, D., Send, U., Sosik, H.M., Weiss, E.L., &
1310 Mitchell, B.G. (2021). Satellite detection of dinoflagellate blooms off California by UV reflectance ratios.
1311 *Elem Sci Anth, 9*, 00157

Kahru, M., & Mitchell, B. (1998). Spectral reflectance and absorption of a massive red tide off southern California. *Journal of Geophysical Research, 103*, 21601-21609

Karlson, B., Godhe, A., Cusack, C., & Bresnan, E. (2010). Introduction to methods for quantitative phytoplankton analysis. *Microscopic and molecular methods for quantitative phytoplankton analysis*, IOC Manual and Guides No. 55, UNESCO, Paris. 110p.

Kavanaugh, M. T., Bell, T., Catlett, D., Cimino, M. A., Doney, S. C., Klajbor, W., Messie, M., Montes, E., Muller Karger F.E., Otis, D., Santora, J.A., Schroeder, I.D., Trinanes, J., & Siegel, D. A. (2021). Satellite remote sensing and the marine biodiversity observation network. *Oceanography*, 34(2), 62-79.

Kerr, T., Clark, J.R., Fileman, E.S., Widdicombe, C.E., & Pugeault, N. (2020). Collaborative Deep Learning Models to Handle Class Imbalance in FlowCam Plankton Imagery. *IEEE Access, 8*, 170013-170032

Kirkpatrick, G.J., Millie, D.F., Moline, M.A., & Schofield, O. (2000). Optical discrimination of a phytoplankton species in natural mixed populations. *Limnology and Oceanography, 45*, 467-471

Kostadinov, T., Siegel, D., & Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *Journal of Geophysical Research: Oceans, 114*

Kostadinov, T., Siegel, D., & Maritorena, S. (2010). Global variability of phytoplankton functional types from space: assessment via the particle size distribution. *Biogeosciences, 7*, 3239-3257

Kostadinov, T.S. (2016). Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. *Ocean Science, 12*, 561

Kostadinov, T.S., Cabré, A., Vedantham, H., Marinov, I., Bracher, A., Brewin, R.J., Bricaud, A., Hirata, T., Hirawake, T., & Hardman-Mountford, N.J. (2017). Inter-comparison of phytoplankton functional type phenology metrics derived from ocean color algorithms and earth system models. *Remote Sensing of Environment, 190*, 162-177

Kostadinov, T.S., Robertson Lain, L., Kong, C.E., Zhang, X., Maritorena, S., Bernard, S., Loisel, H., Jorge, D.S.F., Kochetkova, E., Roy, S., Jonsson, B., Martinez-Vicente, V., & Sathyendranath, S. (2022). Ocean Color Algorithm for the Retrieval of the Particle Size Distribution and Carbon-Based Phytoplankton Size Classes Using a Two-Component Coated-Spheres Backscattering Model. *EGUsphere, 2022*, 1-38

Kostakis, I., Twardowski, M., Roesler, C., Röttgers, R., Stramski, D., McKee, D., Tonizzo, A., & Drapeau, S. (2021). Hyperspectral optical absorption closure experiment in complex coastal waters. *Limnology and Oceanography: Methods, 19*, 589-625

Kramer, S.J., Roesler, C.S., & Sosik, H.M. (2018). Bio-optical discrimination of diatoms from other phytoplankton in the surface ocean: Evaluation and refinement of a model for the Northwest Atlantic. *Remote Sensing of Environment, 217*, 126-143

Kramer, S.J., & Siegel, D.A. (2019). How Can Phytoplankton Pigments Be Best Used to Characterize Surface Ocean Phytoplankton Groups for Ocean Color Remote Sensing Algorithms? *Journal of Geophysical Research: Oceans, 124*, 7557-7574

Kramer, S.J., Siegel, D.A., & Graff J.R. (2020). Phytoplankton Community Composition Determined From Co-variability Among Phytoplankton Pigments From the NAAMES Field Campaign. *Frontiers in Marine Science, 7*

Kramer, S.J., Siegel, D.A., Maritorena, S., & Catlett, D. (2022). Modeling surface ocean phytoplankton pigments from hyperspectral remote sensing reflectance on global scales. *Remote Sensing of Environment, 270*, 112879

Kudela, R.M., Palacios, S.L., Austerberry, D.C., Accorsi, E.K., Guild, L.S., & Torres-Perez, J. (2015). Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. *Remote Sensing of Environment, 167*, 196-205

Lange, P. K., Brewin, R. J., Dall'Olmo, G., Tarran, G. A., Sathyendranath, S., Zubkov, M., & Bouman, H. A. (2018). Scratching beneath the surface: A model to predict the vertical distribution of Prochlorococcus using remote sensing. *Remote Sensing*, 10, 847

Lange, P.K., Jeremy Werdell, P., Erickson, Z.K., Dall'Olmo, G., Brewin, R.J.W., Zubkov, M.V., Tarran, G.A., Bouman, H.A., Slade, W.H., Craig, S.E., Poulton, N.J., Bracher, A., Lomas, M.W., & Cetinić, I. (2020). Radiometric approach for the detection of picophytoplankton assemblages across oceanic fronts. *Optics Express, 28*, 25682-25705

Lee, Z., Carder, K., Arnone, R. and He, M. (2007). Determination of primary spectral bands for remote sensing of aquatic environments. *Sensors*, 7,3428-3441

Le Quéré, C., Takahashi, T., Buitenhuis, E.T., Rödenbeck, C., & Sutherland, S.C. (2010). Impact of climate change and variability on the global oceanic sink of CO2. *Global Biogeochemical Cycles, 24*, n/a-n/a

LeGresley, M., & McDermott, G. (2010). Counting chamber methods for quantitative phytoplankton analysis-haemocytometer, Palmer-Maloney cell and Sedgewick-Rafter cell. *UNESCO (IOC manuals and guides)*, 25-30

Lehmann, M.K., Gurlin, D., Pahlevan, N., Alikas, K., … & Yue, L. (2023). GLORIA - A globally representative hyperspectral in situ dataset for optical sensing of water quality. *Scientific Data, 10.*

Lepesteur, M., Martin, J. M., & Fleury, A. (1993). A comparative study of different preservation methods for phytoplankton cell analysis by flow cytometry. *Marine Ecology Progress Series, 93,*55–63.

Lekki, J., Ruberg, S., Binding, C., Anderson, R., & Vander Woude, A. (2019). Airborne hyperspectral and satellite imaging of harmful algal blooms in the Great Lakes Region: Successes in sensing algal blooms. *Journal of Great Lakes Research, 45*, 405-412

Li, Z., Li, L., Song, K., & Cassar, N. (2013). Estimation of Phytoplankton Size Fractions Based on Spectral Features of Remote Sensing Ocean Color Data. *Journal of Geophysical Research: Oceans, 118*, n/a-n/a

Lin, Y., Gifford, S., Ducklow, H., Schofield, O., & Cassar, N. (2019). Towards quantitative microbiome community profiling using internal standards. *Applied and Environmental Microbiology, 85*, 1-14

Liu, Y., Boss, E., Chase, A., Xi, H., Zhang, X., Röttgers, R., Pan, Y., & Bracher, A. (2019). Retrieval of phytoplankton pigments from underway spectrophotometry in the Fram Strait. *Remote Sensing, 11*, 318

Llewellyn, C.A., & Airs, R.L. (2010). Distribution and abundance of MAAs in 33 species of microalgae across 13 classes. *Marine Drugs, 8*, 1273-1291

Lohrenz, S.E., Weidemann, A.D., & Tuel, M. (2003). Phytoplankton spectral absorption as influenced by community size structure and pigment composition. *Journal of Plankton Research, 25*, 35-61

Loisel, H., Duforet, L., Dessailly, D., Chami, M., & Dubuisson, P. (2008). Investigation of the variations in the water leaving polarized reflectance from the POLDER satellite data over two biogeochemical contrasted oceanic areas. *Optics Express, 16*, 12905-12918

Loisel, H., Stramski, D., Dessailly, D., Jamet, C., Li, L., & Reynolds, R.A. (2018). An inverse model for estimating the optical absorption and backscattering coefficients of seawater from remote-sensing reflectance over a broad range of oceanic and coastal marine environments. *Journal of Geophysical Research: Oceans, 123*, 2141-2171

Lombard, F., Boss, E., Waite, A.M., Vogt, M., Uitz, J., Stemmann, L., Sosik, H.M., Schulz, J., Romagnan, J.-B., Picheral, M., Pearlman, J., Ohman, M.D., Niehoff, B., Möller, K.O., Miloslavich, P., Lara-Lpez, A., Kudela, R., Lopes, R.M., Kiko, R., Karp-Boss, L., Jaffe, J.S., Iversen, M.H., Irisson, J.-O., Fennel, K., Hauss, H., Guidi, L., Gorsky, G., Giering, S.L.C., Gaube, P., Gallager, S., Dubelaar, G., Cowen, R.K., Carlotti, F., Briseño-Avena, C., Berline, L., Benoit-Bird, K., Bax, N., Batten, S., Ayata, S.D., Artigas, L.F., & Appeltans, W. (2019). Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Frontiers in Marine Science, 6*

Losa, S.N., Soppa, M.A., Dinter, T., Wolanin, A., Brewin, R.J.W., Bricaud, A., Oelker, J., Peeken, I., Gentili, B., Rozanov, V., & Bracher, A. (2017). Synergistic Exploitation of Hyper- and Multi-Spectral Precursor Sentinel Measurements to Determine Phytoplankton Functional Types (SynSenPFT). *Frontiers in Marine Science, 4*

Lubac, B., Loisel, H., Guiselin, N., Astoreca, R., Artigas, L.F., & Meriaux, X. (2008). Hyperspectral and multispectral ocean color inversions to detect Phaeocystis globosa blooms in coastal waters. *Journal of Geophysical Research-Oceans, 113*

Lucke, R.L., Corson, M., McGlothlin, N.R., Butcher, S.D., Wood, D.L., Korwan, D.R., Li, R.R., Snyder, W.A., Davis, C.O., & Chen, D.T. (2011). Hyperspectral Imager for the Coastal Ocean: instrument description and first images. *Applied Optics, 50*, 1501-1516

Lundholm, N., Churro, C., Fraga, S., Hoppenrath, M., Iwataki, M., Larsen, J., Mertens, K., Moestrup, Ø., & Zingone, A. (2009). IOC-UNESCO Taxonomic Reference List of Harmful Micro Algae. Accessed at https://www.marinespecies.org/hab on 2023-03-14.

Mackey, M.D., Mackey, D.J., Higgins, H.W., & Wright, S.W. (1996). CHEMTAX - A program for estimating class abundances from chemical markers: Application to HPLC measurements of phytoplankton. *Marine Ecology-Progress Series, 144*, 265-283

Mao, Z., Stuart, V., Pan, D., Chen, J., Gong, F., Huang, H., & Zhu, Q. (2010). Effects of phytoplankton species composition on absorption spectra and modeled hyperspectral reflectance. *Ecological Informatics, 5*, 359-366

Marie, D., Rigaut-Jalabert, F., & Vaulot, D. (2014). An improved protocol for flow cytometry analysis of phytoplankton cultures and natural samples. *Cytometry Part A, 85*, 962-968

Martins, J.V., Tanré, D., Remer, L., Kaufman, Y., Mattoo, S., & Levy, R. (2002). MODIS cloud screening for remote sensing of aerosols over oceans using spatial variability. *Geophysical Research Letters, 29*, MOD4-1-MOD4-4

McAlice, B.J. (1971). Phytoplankton Sampling with the Sedgwick-Rafter Cell. *Limnology and Oceanography, 16*, 19-28

McCabe, R.M., Hickey, B.M., Kudela, R.M., Lefebvre, K.A., Adams, N.G., Bill, B.D., Gulland, F.M.D., Thomson, R.E., Cochlan, W.P., & Trainer, V.L. (2016). An unprecedented coastwide toxic algal bloom linked to anomalous ocean conditions. *Geophysical Research Letters, 43*, 10,366-310,376

McClain, C.R., Franz, B.A., & Werdell, P.J. (2022). Genesis and Evolution of NASA's Satellite Ocean Color Program. *Frontiers in Remote Sensing, 3*

McKibben, S. M., Peterson, W., Wood, A. M., Trainer, V. L., Hunter, M., & White, A. E. (2017). Climatic regulation of the neurotoxin domoic acid. *Proceedings of the National Academy of Sciences*, 114, 239-244

McKinna, L.I. (2015). Three decades of ocean-color remote-sensing Trichodesmium spp. in the World's oceans: A review. Progress in Oceanograpy, 131, 177-199

McKinna, L.I.W., Cetinić, I., Chase, A.P., & Werdell, P.J. (2019). Approach for Propagating Radiometric Data Uncertainties Through NASA Ocean Color Algorithms. *Frontiers in Earth Science, 7*

McKinna, L.I.W., Cetinić, I., & Werdell, P.J. (2021). Development and Validation of an Empirical Ocean Color Algorithm with Uncertainties: A Case Study with the Particulate Backscattering Coefficient. *Journal of Geophysical Research: Oceans, 126*, e2021JC017231

McKinna, L.I.W., Furnas, M.J., & Ridd, P.V. (2011). A simple, binary classification algorithm for the detection of Trichodesmium spp. within the Great Barrier Reef using MODIS imagery. Limnology and Oceanography: Methods, 9, 50-66

Meier, W.N., Perovich, D., Farrell, S., Haas, C., Hendricks, S., Petty, A.A., Webster, M., Divine, D., Gerland, S., Kaleschke, L., Ricker, R., Steer, A., Tian-Kunze, X., Tschudi, M., & Wood, K. (2021). Sea Ice. In T.A. Moon, M.L. Druckenmiller, & R.L. Thoman (Eds.), *Arctic Report Card 2021*

Menden-Deuer, S., & Lessard, E.J. (2000). Carbon to volume relationships for dinoflagellates, diatoms and other protist plankton. *Limnology and Oceanography, 45*, 569-579

Michalak, A.M. (2016). Study role of climate change in extreme threats to water quality. *Nature, 535*, 349-350

1452    Millie, D.F., Schofield, O.M., Kirkpatrick, G.J., Johnsen, G., Tester, P.A., & Vinyard, B.T. (1997). Detection
1453    of harmful algal blooms using photopigments and absorption signatures: A case study of the Florida red
1454    tide dinoflagellate, Gymnodinium breve. *Limnology and Oceanography, 42*, 1240-1251
1455    Mitra, A., Flynn, K.J., Tillmann, U., Raven, J.A., Caron, D., Stoecker, D.K., Not, F., Hansen, P.J., Hallegraeff,
1456    G., Sanders, R., Wilken, S., McManus, G., Johnson, M., Pitta, P., Våge, S., Berge, T., Calbet, A., Thingstad,
1457    F., Jeong, H.J., Burkholder, J., Glibert, P.M., Granéli, E., & Lundgren, V. (2016). Defining Planktonic Protist
1458    Functional Groups on Mechanisms for Energy and Nutrient Acquisition: Incorporation of Diverse
1459    Mixotrophic Strategies. Protist, 167, 106-120
1460    Mobley, C. (Ed.) (2022). The Oceanic Optics Book. Dartmouth, NS, Canada: International Ocean Colour
1461    Coordinating Group (IOCCG)
1462    Moore, T.S., & Brown, C.W. (2020). Incorporating environmental data in abundance-based algorithms for
1463    deriving phytoplankton size classes in the Atlantic Ocean. *Remote Sensing of Environment, 240*, 111689
1464    Moore, T.S., Dowell, M.D., & Franz, B.A. (2012). Detection of coccolithophore blooms in ocean color
1465    satellite imagery: A generalized approach for use with multiple sensors. *Remote Sensing of Environment,*
1466    *117*, 249-263
1467    Morel, A., & Berthon, J. F. (1989). Surface pigments, algal biomass profiles, and potential production of
1468    the euphotic layer: Relationships reinvestigated in view of remote-sensing applications. *Limnology and*
1469    *oceanography*, 34, 1545-1562.
1470    Morel, A., & Prieur, L. (1977). Analysis of variation in ocean color. *Limnology and Oceanography, 22*, 709-
1471    722
1472    Mouw, C.B., Hardman-Mountford, N.J., Alvain, S., Bracher, A., Brewin, R.J.W., Bricaud, A., Ciotti, A.M.,
1473    Devred, E., Fujiwara, A., Hirata, T., Hirawake, T., Kostadinov, T.S., Roy, S., & Uitz, J. (2017). A Consumer's
1474    Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean. *Frontiers in*
1475    *Marine Science, 4*
1476    Mouw, C.B., & Yoder, J.A. (2010). Optical determination of phytoplankton size composition from global
1477    SeaWiFS imagery. *Journal of Geophysical Research: Oceans, 115*
1478    Muller-Karger, F.E., Hestir, E., Ade, C., Turpie, K., Roberts, D.A., … & Jetz, W. (2018). Satellite sensor
1479    requirements for monitoring essential biodiversity variables of coastal ecosystems. *Ecological*
1480    *Applications, 28*, 749-760
1481    Muller-Karger, F. E., Miloslavich, P., Bax, N. J., Simmons, S., Costello, M. J., Sousa Pinto, I., ... & Geller, G.
1482    (2018). Advancing marine biological observations and data requirements of the complementary essential
1483    ocean variables (EOVs) and essential biodiversity variables (EBVs) frameworks. Frontiers in Marine
1484    Science, 211.
1485    Nardelli, S.C., Gray, P.C., Stammerjohn, S.E., & Schofield, O. (2023). Characterizing coastal phytoplankton
1486    seasonal succession patterns on the West Antarctic Peninsula. *Limnology and Oceanography*.
1487    NASA (2010). Responding to the Challenge of Climate and Environmental Change: NASA's Plan for a
1488    Climate-Centric Architecture for Earth Observations and Applications from Space. (p. 48)
1489    National Academies of Sciences, Engineering, and Medicine (2022). A research strategy for ocean-based
1490    carbon dioxide removal and sequestration. In. Washington, DC The National Academies Press.
1491    National Research Council (2007). Earth science and applications from space: National imperatives for
1492    the next decade and beyond. *The National Academies Press, Washington, DC, 10*, 11820
1493    Neeley, A., Beaulieu, S., Proctor, C., Cetinic, I., Futrelle, J., Soto Ramos, I., Sosik, H., Devred, E., Karp-Boss,
1494    L., & Picheral, M. (2021). Standards and practices for reporting plankton and other particle observations
1495    from images. In, *Ocean Carbon and Biogeochemistry Publications*: Ocean Carbon and Biogeochemistry
1496    Neeley, A., Harris, L., & Frey, K. (2018). Unraveling Phytoplankton Community Dynamics in the Northern
1497    Chukchi Sea Under Sea-Ice-Covered and Sea-Ice-Free Conditions. *Geophysical Research Letters, 45*, 7663-
1498    7671

1499    Neeley, A.R., Lomas, M.W., Mannino, A., Thomas, C., & Vandermeulen, R. (2022). Impact of Growth
1500    Phase, Pigment Adaptation, and Climate Change Conditions on the Cellular Pigment and Carbon Content
1501    of Fifty-One Phytoplankton Isolates. *Journal of Phycology, 58*, 669-690
1502    Neeley, A.R., Soto-Ramos, I., & Proctor, C. (2023). Standards and Best Practices for Reporting Flow
1503    Cytometry Observations: a technical manual, Version 1.1. Greenbelt, MD., NASA Goddard Space Flight
1504    Center, 31pp.
1505    O'Shea, R.E., Pahlevan, N., Smith, B., Bresciani, M., Egerton, T., Giardino, C., Li, L., Moore, T., Ruiz-Verdu,
1506    A., Ruberg, S., Simis, S.G.H., Stumpf, R., & Vaičiūtė, D. (2021). Advancing cyanobacteria biomass
1507    estimation from hyperspectral observations: Demonstrations with HICO and PRISMA imagery. *Remote
1508    Sensing of Environment, 266*, 112693
1509    Olson, R.J., Chisholm, S.W., Zettler, E.R., & Armbrust, E.V. (1988). Analysis of Synechococcus pigment
1510    types in the sea using single and dual beam flow cytometry. *Deep Sea Research Part A. Oceanographic
1511    Research Papers, 35*, 425-440
1512    Olson, R.J., & Sosik, H.M. (2007). A submersible imaging-in-flow instrument to analyze nano-and
1513    microplankton: Imaging FlowCytobot. *Limnology and Oceanography: Methods, 5*, 195-203
1514    O'Reilly, John E., et al. "Ocean color chlorophyll algorithms for SeaWiFS." *Journal of Geophysical
1515    Research: Oceans* 103.C11 (1998): 24937-24953.
1516    Orenstein, E.C., Ayata, S.-D., Maps, F., Becker, É.C., Benedetti, F., Biard, T., …& Irisson, J.-O. (2022).
1517    Machine learning techniques to characterize functional traits of plankton from image data. *Limnology
1518    and Oceanography, 67*, 1647-1669
1519    Organelli, E., Dall'Olmo, G., Brewin, R.J.W., Tarran, G.A., Boss, E., & Bricaud, A. (2018). The open-ocean
1520    missing backscattering is in the structural complexity of particles. *Nature Communications, 9*, 5439
1521    PACE mission (2020). Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission PACE Science Data
1522    Product Validation Plan. NASA GSFC, Greenbelt.
1523    Pahlevan, N., Smith, B., Binding, C., Gurlin, D., Li, L., Bresciani, M., & Giardino, C. (2021). Hyperspectral
1524    retrievals of phytoplankton absorption and chlorophyll-a in inland and nearshore coastal waters. *Remote
1525    Sensing of Environment, 253*, 112200
1526    Palacios, S.L., Kudela, R.M., Guild, L.S., Negrey, K.H., Torres-Perez, J., & Broughton, J. (2015). Remote
1527    sensing of phytoplankton functional types in the coastal ocean from the HyspIRI Preparatory Flight
1528    Campaign. *Remote Sensing of Environment, 167*, 269-280
1529    Palacz, A., St. John, M., Brevin, R., Hirata, T., & Gregg, W. (2013). Distribution of phytoplankton functional
1530    types in high-nitrate low-chlorophyll waters in a new diagnostic ecological indicator model.
1531    *Biogeosciences, 10*, 7553-7574
1532    Pan, X., Mannino, A., Russ, M.E., Hooker, S.B., & Harding Jr, L.W. (2010). Remote sensing of
1533    phytoplankton pigment distribution in the United States northeast coast. *Remote Sensing of
1534    Environment, 114*, 2403-2416
1535    Parada, A.E., Needham, D.M., & Fuhrman, J.A. (2016). Every base matters: assessing small subunit rRNA
1536    primers for marine microbiomes with mock communities, time series and global field samples.
1537    *Environmental microbiology, 18*, 1403-1414
1538    Picheral, M., Guidi, L., Stemmann, L., Karl, D.M., Iddaoud, G., & Gorsky, G. (2010). The Underwater Vision
1539    Profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and
1540    zooplankton. *Limnology and Oceanography: Methods, 8*, 462-473
1541    Poulin, C., Zhang, X., Yang, P., & Huot, Y. (2018). Diel variations of the attenuation, backscattering and
1542    absorption coefficients of four phytoplankton species and comparison with spherical, coated spherical
1543    and hexahedral particle optical models. *Journal of Quantitative Spectroscopy and Radiative Transfer,
1544    217*, 288-304
1545    Racault, M.-F., Sathyendranath, S., Brewin, R.J.W., Raitsos, D.E., Jackson, T., & Platt, T. (2017). Impact of
1546    El Niño Variability on Oceanic Phytoplankton. *Frontiers in Marine Science, 4*

Raitsos, D.E., Lavender, S.J., Maravelias, C.D., Haralabous, J., Richardson, A.J., & Reid, P.C. (2008). Identifying four phytoplankton functional types from space: An ecological approach. *Limnology and Oceanography, 53*, 605-613

Rasconi, S., Gall, A., Winter, K., & Kainz, M.J. (2015). Increasing Water Temperature Triggers Dominance of Small Freshwater Plankton. *PLoS ONE, 10*, e0140449

Remer, L.A., Knobelspiesse, K., Zhai, P.-W., Xu, F., Kalashnikova, O.V., Chowdhary, J., Hasekamp, O., Dubovik, O., Wu, L., Ahmad, Z., Boss, E., Cairns, B., Coddington, O., Davis, A.B., Dierssen, H.M., Diner, D.J., Franz, B., Frouin, R., Gao, B.-C., Ibrahim, A., Levy, R.C., Martins, J.V., Omar, A.H., & Torres, O. (2019). Retrieving Aerosol Characteristics From the PACE Mission, Part 2: Multi-Angle and Polarimetry. *Frontiers in Environmental Science*, 7

Rêve-Lamarche, A.-H., Alvain, S., Racault, M.-F., Dessailly, D., Guiselin, N., Jamet, C., Vantrepotte, V., & Beaugrand, G. (2017). Estimation of the potential detection of diatom assemblages based on ocean color radiance anomalies in the North Sea. *Frontiers in Marine Science, 4*, 408

Rivero-Calle, S., Gnanadesikan, A., Del Castillo, C.E., Balch, W.M., & Guikema, S.D. (2015). Multidecadal increase in North Atlantic coccolithophores and the potential role of rising $CO_2$. *Science, 350*, 1533-1537

Roelke, D., Kennedy, C., & Weidemann, A. (1999). Use of discriminant and fourth-derivative analyses with high-resolution absorption spectra for phytoplankton research: Limitations at varied signal-to-noise ratio and spectral resolution. *Gulf of Mexico Science, 17*, 75-86

Roesler, C.S., & Boss, E. (2003). Spectral beam attenuation coefficient retrieved from ocean color inversion. *Geophysical Research Letters, 30,* 1468

Rose, J.M., Caron, D.A., Sieracki, M.E., & Poulton, N. (2004). Counting heterotrophic nanoplanktonic protists in cultures and aquatic communities by flow cytometry. *Aquatic Microbial Ecology, 34*, 263-277

Rousseaux, C.S., & Gregg, W.W. (2012). Climate variability and phytoplankton composition in the Pacific Ocean. *Journal of Geophysical Research: Oceans, 117*, n/a-n/a

Roy, S., Sathyendranath, S., Bouman, H., & Platt, T. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sensing of Environment, 139*, 185-197

Ryan, J.P., Chavez, F.P., & Bellingham, J.G. (2005). Physical-biological coupling in Monterey Bay, California: topographic influences on phytoplankton ecology. *Marine Ecology-Progress Series, 287*, 23-32

Ryan, J.P., Davis, C.O., Tufillaro, N.B., Kudela, R.M., & Gao, B.-C. (2014). Application of the Hyperspectral Imager for the Coastal Ocean to Phytoplankton Ecology Studies in Monterey Bay, CA, USA. *Remote Sensing, 6*, 1007-1025

Sadeghi, A., Dinter, T., Vountas, M., Taylor, B., Altenburg-Soppa, M., & Bracher, A. (2012). Remote sensing of coccolithophore blooms in selected oceanic regions using the PhytoDOAS method applied to hyper-spectral satellite data. *Biogeosciences, 9*, 2127-2143

Sathyendranath, S., Watts, L., Devred, E., Platt, T., Caverhill, C., & Maass, H. (2004). Discrimination of diatoms from other phytoplankton using ocean-colour data. *Marine Ecology-Progress Series, 272*, 59-68

Satinsky, B.M., Gifford, S.M., Crump, B.C., & Moran, M.A. (2013). Chapter Twelve - Use of Internal Standards for Quantitative Metatranscriptome and Metagenome Analysis. In E.F. DeLong (Ed.), *Methods in Enzymology* (pp. 237-250): Academic Press

Sato, M., Takeda, S., & Furuya, K. (2006). Effects of long-term sample preservation on flow cytometric analysis of natural populations of pico-and nanophytoplankton. *Journal of Oceanography, 62*, 903-908

Sayer, A.M., Lelli, L., Cairns, B., van Diedenhoven, B., Ibrahim, A., Knobelspiesse, K., Korkin, S., & Werdell, P.J. (2022). The CHROMA cloud top pressure retrieval algorithm for the Plankton, Aerosol, Cloud, ocean Ecosytem (PACE) satellite mission. *Atmos. Meas. Tech. Discuss*., 2022, 1-40

Schaeffer, B.A., Loftin, K.A., Stumpf, R.P., & Werdell, P.J. (2015). Agencies collaborate, develop a cyanobacteria assessment network. *Eos, Earth and Space Science News*, 96

1595    Schoening, T., Durden, J. M., Faber, C., Felden, J., Heger, K., Hoving, H. J. T., … & Zurowietz, M. (2022).
1596    Making marine image data FAIR. *Scientific data, 9*, 414
1597    Schofield, O., Brown, M., Kohut, J., Nardelli, S., Saba, G., Waite, N., & Ducklow, H. (2018). Changes in the
1598    upper ocean mixed layer and phytoplankton productivity along the West Antarctic Peninsula.
1599    *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376*,
1600    20170173
1601    Shaju, S.S., Minu, P., Srikanth, A.S., Ashraf, P.M., Vijayan, A.K., & Meenakumari, B. (2015). Decomposition
1602    study of in vivo phytoplankton absorption spectra aimed at identifying the pigments and the
1603    phytoplankton group in complex case 2 coastal waters of the Arabian Sea. *Oceanological and
1604    Hydrobiological Studies, 44*, 282-293
1605    Shulman, I., Frolov, S., Anderson, S., Penta, B., Gould, R., Sakalaukus, P., & Ladner, S. (2013). Impact of
1606    bio-optical data assimilation on short-term coupled physical, bio-optical model predictions. *Journal of
1607    Geophysical Research: Oceans, 118*, 2215-2230
1608    Siegel, D.A., Behrenfeld, M.J., Maritorena, S., McClain, C.R., Antoine, D., …& Yoder, J.A. (2013). Regional
1609    to global assessments of phytoplankton dynamics from the SeaWiFS mission. *Remote Sensing of
1610    Environment, 135*, 77-91
1611    Siegel, D.A., DeVries, T., Cetinić, I., & Bisson, K.M. (2023). Quantifying the Ocean's Biological Pump and Its
1612    Carbon Cycle Impacts on Global Scales. *Annual Review of Marine Science, 15*, null
1613    Sieracki, C.K., Sieracki, M.E., & Yentsch, C.S. (1998). An imaging-in-flow system for automated analysis of
1614    marine microplankton. *Marine Ecology Progress Series*, 285-296
1615    Skákala, J., Ford, D., Brewin, R.J.W., McEwan, R., Kay, S., Taylor, B., de Mora, L., & Ciavatta, S. (2018). The
1616    Assimilation of Phytoplankton Functional Types for Operational Forecasting in the Northwest European
1617    Shelf. *Journal of Geophysical Research: Oceans, 123*, 5230-5247
1618    Smith, M.E., & Bernard, S. (2020). Satellite Ocean Color Based Harmful Algal Bloom Indicators for
1619    Aquaculture Decision Support in the Southern Benguela. *Frontiers in Marine Science, 7*
1620    Snyder, J., Boss, E., Weatherbee, R., Thomas, A.C., Brady, D., & Newell, C. (2017). Oyster Aquaculture Site
1621    Selection Using Landsat 8-Derived Sea Surface Temperature, Turbidity, and Chlorophyll a. *Frontiers in
1622    Marine Science, 4*
1623    Soppa, M.A., Hirata, T., Silva, B., Dinter, T., Peeken, I., Wiegmann, S., & Bracher, A. (2014). Global
1624    Retrieval of Diatom Abundance Based on Phytoplankton Pigments and Satellite Data. *Remote Sensing, 6*,
1625    10089-10106
1626    Sosik, H.M., Olson, R.J., Neubert, M.G., Shalapyonok, A., & Solow, A.R. (2003). Growth rates of coastal
1627    phytoplankton from time-series measurements with a submersible flow cytometer. *Limnology and
1628    Oceanography, 48*, 1756-1765
1629    Soto, I.M., Cannizzaro, J., Muller-Karger, F.E., Hu, C., Wolny, J., & Goldgof, D. (2015). Evaluation and
1630    optimization of remote sensing techniques for detection of Karenia brevis blooms on the West Florida
1631    Shelf. *Remote Sensing of Environment, 170*, 239-254
1632    Stramski, D., Bricaud, A., & Morel, A. (2001). Modeling the inherent optical properties of the ocean based
1633    on the detailed composition of planktonic community. *Applied Optics* 40, 2929–2945.
1634    Stramski, D. & Kiefer, D. A. (1991). Light scattering by microorganisms in the open ocean. *Progress in
1635    Oceanogaphy* 28, 343–383
1636    Stramski, D. & Reynolds, R.A. (1993). Diel variations in the optical properties of a marine diatom.
1637    *Limnology and Oceanography*, 38, 1347–1364
1638    Stramski, D., Shalapyonok, A.& R. A. Reynolds (1995). Optical characterization of the oceanic unicellular
1639    cyanobacterium Synechococcus grown under a day-night cycle in natural irradiance. *Journal of
1640    Geophysical Research*, 100, 13295–13307
1641    Stramska, M., & Stramski, D. (2005). Effects of a nonuniform vertical profile of chlorophyll concentration
1642    on remote-sensing reflectance of the ocean. *Applied Optics, 44,* 1735-1747

1643    Subramaniam, A., Brown, C.W., Hood, R.R., Carpenter, E.J., & Capone, D.G. (2001). Detecting
1644    Trichodesmium blooms in SeaWiFS imagery. *Deep Sea Research Part II: Topical Studies in Oceanography,*
1645    *49*, 107-121

1646    Subramaniam, A., & Carpenter, E.J. (1994). An empirically derived protocol for the detection of blooms
1647    of the marine cyanobacterium Trichodesmium using CZCS imagery. *International Journal of Remote*
1648    *Sensing, 15*, 1559-1569

1649    Sun, X., Shen, F., Brewin, R.J., Li, M., & Zhu, Q. (2022). Light absorption spectra of naturally mixed
1650    phytoplankton assemblages for retrieval of phytoplankton group composition in coastal oceans.
1651    *Limnology and Oceanography, 67*, 946-961

1652    Swalwell, J.E., Ribalet, F., & Armbrust, E.V. (2011). SeaFlow: A novel underway flow-cytometer for
1653    continuous observations of phytoplankton in the ocean. *Limnology and Oceanography: Methods, 9*, 466-
1654    477

1655    Sydor, M., Gould, R.W., Arnone, R.A., Haltrin, V.I., & Goode, W. (2004). Uniqueness in remote sensing of
1656    the inherent optical properties of ocean water. *Applied Optics, 43*, 2156-2162

1657    Taylor, B.B., Taylor, M.H., Dinter, T., & Bracher, A. (2013). Estimation of relative phycoerythrin
1658    concentrations from hyperspectral underwater radiance measurements—A statistical approach. *Journal*
1659    *of Geophysical Research: Oceans, 118*, 2948-2960

1660    Thompson, P.A., & Carstensen, J. (2022). Global observing for phytoplankton? A perspective. *Journal of*
1661    *Plankton Research*

1662    Thyssen, M., Alvain, S., Lefèbvre, A., Dessailly, D., Rijkeboer, M., Guiselin, N., Creach, V., & Artigas, L.F.
1663    (2015). High-resolution analysis of a North Sea phytoplankton community structure based on in situ flow
1664    cytometry observations and potential implication for remote sensing. *Biogeosciences, 12*, 4051-4066

1665    Torrecilla, E., Stramski, D., Reynolds, R.A., Millán-Núñez, E., & Piera, J. (2011). Cluster analysis of
1666    hyperspectral optical data for discriminating phytoplankton pigment assemblages in the open ocean.
1667    *Remote Sensing of Environment, 115*, 2578-2593

1668    Turner, K. J., Mouw, C. B., Hyde, K. J., Morse, R., & Ciochetto, A. B. (2021). Optimization and assessment
1669    of phytoplankton size class algorithms for ocean color data on the Northeast US continental shelf.
1670    *Remote Sensing of Environment*, 267, 112729.

1671    Uitz, J., Claustre, H., Morel, A., & Hooker, S.B. (2006). Vertical distribution of phytoplankton communities
1672    in open ocean: An assessment based on surface chlorophyll. *Journal of Geophysical Research, 111*,
1673    C08005

1674    Uitz, J., Stramski, D., Reynolds, R.A., & Dubranna, J. (2015). Assessing phytoplankton community
1675    composition from hyperspectral measurements of phytoplankton absorption coefficient and remote-
1676    sensing reflectance in open-ocean environments. *Remote Sensing of Environment, 171*, 58-74

1677    van Heukelem, L., & Hooker, S.B. (2011). The importance of a quality assurance plan for method
1678    validation and minimizing uncertainties in the HPLC analysis of phytoplankton pigments. In C.A.
1679    Llewellyn, E.S. Egeland, G. Johnsen, & S. Roy (Eds.), *Phytoplankton Pigments: Characterization,*
1680    *Chemotaxonomy and Applications in Oceanography* (pp. 195-256). Cambridge: Cambridge University
1681    Press

1682    Vandermeulen, R.A., Mannino, A., Neeley, A., Werdell, J., & Arnone, R. (2017). Determining the optimal
1683    spectral sampling frequency and uncertainty thresholds for hyperspectral remote sensing of ocean color.
1684    *Optics Express, 25*, A785-A797

1685    Vaulot, D., Courties, C., & Partensky, F. (1989). A simple method to preserve oceanic phytoplankton for
1686    flow cytometric analyses. *Cytometry, 10*, 629-635

1687    Verbeek, L., Gall, A., Hillebrand, H., & Striebel, M. (2018). Warming and oligotrophication cause shifts in
1688    freshwater phytoplankton communities. *Global Change Biology, 24*, 4532-4543

Vidussi, F., Claustre, H., Manca, B.B., Luchetta, A., & Marty, J.-C. (2001). Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *Journal of Geophysical Research, 106*, 19939-19956

Vuorio, K., Lepistoe, L., & Holopainen, A.L. (2007). Intercalibrations of freshwater phytoplankton analyses. *Boreal Environment Research, 12*, 561-569

Wang, L., Ou, L., Huang, K., Chai, C., Wang, Z., Wang, X., & Jiang, T. (2018). Determination of the spatial and temporal variability of phytoplankton community structure in Daya Bay via HPLC-CHEMTAX pigment analysis. *Journal of Oceanology and Limnology, 36*, 750-760

Wang, S., Xiao, C., Ishizaka, J., Qiu, Z., Sun, D., Xu, Q., Zhu, Y., Huan, Y., & Watanabe, Y. (2016). Statistical approach for the retrieval of phytoplankton community structures from in situ fluorescence measurements. *Optics Express, 24*, 23635-23653

Ward, B.A., Dutkiewicz, S., Jahn, O., & Follows, M.J. (2012). A size-structured food-web model for the global ocean. *Limnology and Oceanography, 57*, 1877-1891

Ward, B.A., & Follows, M.J. (2016). Marine mixotrophy increases trophic transfer efficiency, mean organism size, and vertical carbon flux. *Proceedings of the National Academy of Sciences, 113*, 2958-2963

Werdell, P.J., Behrenfeld, M.J., Bontempi, P.S., Boss, E., Cairns, B., Davis, G.T., Franz, B.A., Gliese, U.B., Gorman, E.T., Hasekamp, O., Knobelspiesse, K.D., Mannino, A., Martins, J.V., McClain, C.R., Meister, G., & Remer, L.A. (2019). The Plankton, Aerosol, Cloud, Ocean Ecosystem Mission: Status, Science, Advances. *Bulletin of the American Meteorological Society, 100*, 1775-1794

Werdell, P.J., Franz, B.A., Bailey, S.W., Feldman, G.C., Boss, E., Brando, V.E., Dowell, M., Hirata, T., Lavender, S.J., Lee, Z.P., Loisel, H., Maritorena, S., Melin, F., Moore, T.S., Smyth, T.J., Antoine, D., Devred, E., d'Andon, O.H.F., & Mangin, A. (2013). Generalized ocean color inversion model for retrieving marine inherent optical properties. *Applied Optics, 52*, 2019-2037

Werdell, P.J., Roesler, C.S., & Goes, J.I. (2014). Discrimination of phytoplankton functional groups using an ocean reflectance inversion model. *Applied Optics, 53*, 4833-4849

Westberry, T., Siegel, D.A., & Subramaniam, A. (2005). An improved bio-optical model for the remote sensing of Trichodesmium spp. blooms. *Journal of Geophysical Research: Oceans, 110*

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., … & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*, 160018

Willén, E. (1976). A simplified method of phytoplankton counting. *British Phycological Journal, 11*, 265-278

Wolanin, A., Soppa, M.A., & Bracher, A. (2016). Investigation of spectral band requirements for improving retrievals of phytoplankton functional types. *Remote Sensing, 8*, 871

Wolny, J.L., Tomlinson, M.C., Schollaert Uz, S., Egerton, T.A., McKay, J.R., Meredith, A., Reece, K.S., Scott, G.P., & Stumpf, R.P. (2020). Current and Future Remote Sensing of Harmful Algal Blooms in the Chesapeake Bay to Support the Shellfish Industry. *Frontiers in Marine Science, 7*

Worden, A.Z., Nolan, J.K., & Palenik, B. (2004). Assessing the dynamics and ecology of marine picophytoplankton: The importance of the eukaryotic component. *Limnology and Oceanography, 49*, 168-179

Wynne, T.T., R.P. Stumpf, M.C. Tomlinson, R.A. Warner, P.A. Tester, J. Dyble & Fahnenstiel, G.L. (2008). Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes. *International Journal of Remote Sensing, 29*, 3665-3672

Xi, H., Hieronymi, M., Krasemann, H., & Röttgers, R. (2017). Phytoplankton Group Identification Using Simulated and In situ Hyperspectral Remote Sensing Reflectance. *Frontiers in Marine Science, 4*

Xi, H., Hieronymi, M., Röttgers, R., Krasemann, H., & Qiu, Z. (2015). Hyperspectral Differentiation of Phytoplankton Taxonomic Groups: A Comparison between Using Remote Sensing Reflectance and Absorption Spectra. *Remote Sensing, 7*, 14781

1737 Xi, H., Losa, S.N., Mangin, A., Garnesson, P., Bretagnon, M., Demaria, J., Soppa, M.A., Hembise Fanton
1738 d'Andon, O., & Bracher, A. (2021). Global Chlorophyll a Concentrations of Phytoplankton Functional
1739 Types With Detailed Uncertainty Assessment Using Multisensor Ocean Color and Sea Surface
1740 Temperature Satellite Products. *Journal of Geophysical Research: Oceans, 126*, e2020JC017127
1741 Xi, H., Losa, S.N., Mangin, A., Soppa, M.A., Garnesson, P., Demaria, J., Liu, Y., d'Andon, O.H.F., & Bracher,
1742 A. (2020). Global retrieval of phytoplankton functional types based on empirical orthogonal functions
1743 using CMEMS GlobColour merged products and further extension to OLCI data. *Remote Sensing of
1744 Environment, 240*, 111704
1745 Xiao, Y., & Friedrichs, M. (2014). The assimilation of satellite-derived data into a one-dimensional lower
1746 trophic level marine ecosystem model. *Journal of Geophysical Research: Oceans, 119*, n/a-n/a
1747 Ye, H., Zhang, B., Liao, X., Li, T., Shen, Q., Zhang, F., Zhu, J., & Li, J. (2019). Gaussian decomposition and
1748 component pigment spectral analysis of phytoplankton absorption spectra. *Journal of Oceanology and
1749 Limnology, 37*, 1542-1554
1750 Yeh, Y.C., McNichol, J., Needham, D.M., Fichot, E.B., Berdjeb, L., & Fuhrman, J.A. (2021). Comprehensive
1751 single-PCR 16S and 18S rRNA community analysis validated with mock communities, and estimation of
1752 sequencing bias against 18S. *Environmental Microbiology, 23*, 3240-3250
1753 Zapata, M., Jeffrey, S.W., Wright, S.W., Rodríguez, F., Garrido, J.L., & Clementson, L. (2004).
1754 Photosynthetic pigments in 37 species (65 strains) of Haptophyta: implications for oceanography and
1755 chemotaxonomy. *Marine Ecology Progress Series, 270*, 83-102
1756 Zibordi, G., Voss, K.J., Johnson, B.C., & Mueller, J. (2019). Protocols for Satellite Ocean Colour Sensor
1757 Validation: In Situ Optical Radiometry. In, IOCCG Protocols Series. Dartmouth, NS, Canada: IOCCG
1758 Zhang, M., Ibrahim, A., Franz, B.A., Ahmad, Z., & Sayer, A.M. (2022). Estimating pixel-level uncertainty in
1759 ocean color retrievals from MODIS. *Optics Express, 30*, 31415-31438
1760 Zhu, F., Massana, R., Not, F., Marie, D., & Vaulot, D. (2005). Mapping of picoeucaryotes in marine
1761 ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiology Ecology, 52*, 79-92
1762 Zhu, Q., Shen, F., Shang, P., Pan, Y., & Li, M. (2019). Hyperspectral Remote Sensing of Phytoplankton
1763 Species Composition Based on Transfer Learning. *Remote Sensing, 11*, 2001
1764

1765    Figures and Tables.



1766

1767    *Figure 1. Artist's rendition of NASA's Plankton, Aerosol, Cloud, ocean Ecosystem satellite flying*

1768    *over cloud-covered Earth's Ocean. The Ocean Color Instrument is visible on the front of the*

1769    *platform, while two polarimeters (located on the bottom side of the platform) are not visible from*

1770    *this view. Credit: NASA Scientific Visualization Studio.*
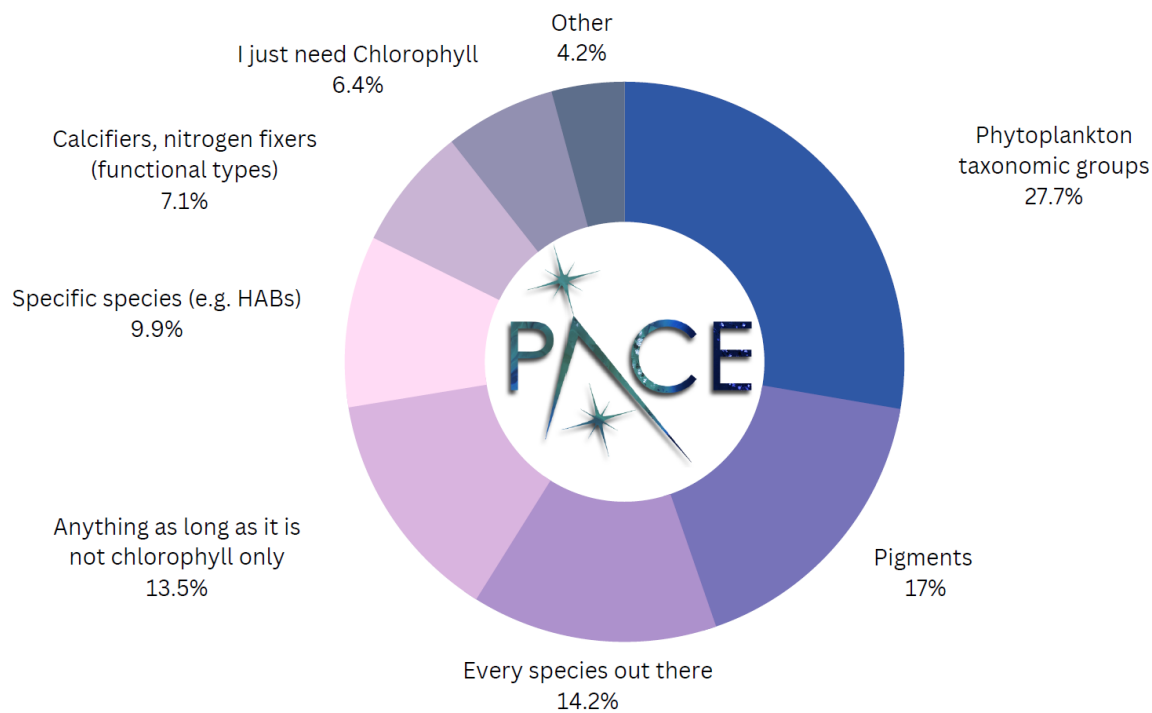
1771

1772

1773
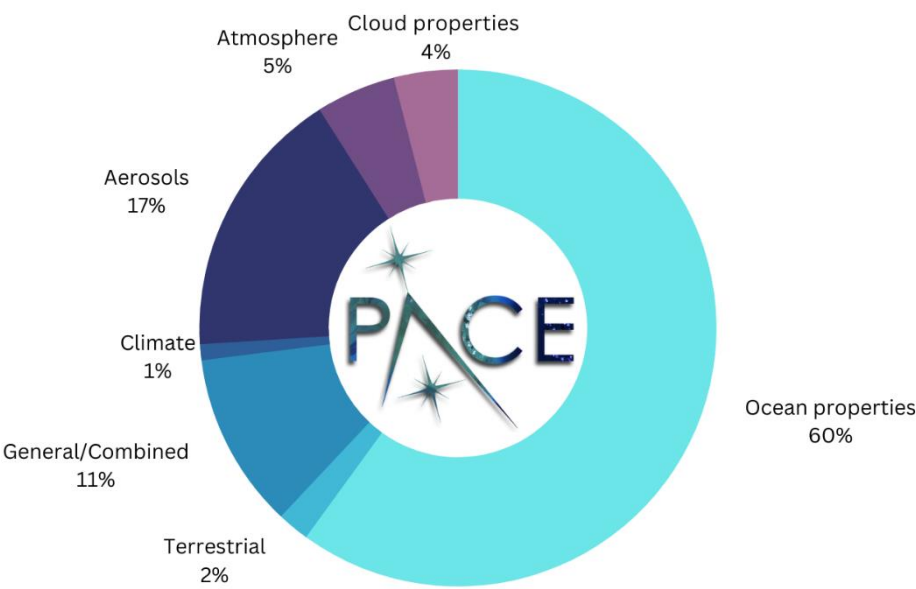


1774

1775 *Figure 2. Self-identified research focus areas grouped into thematic areas, as dictated by PACE*

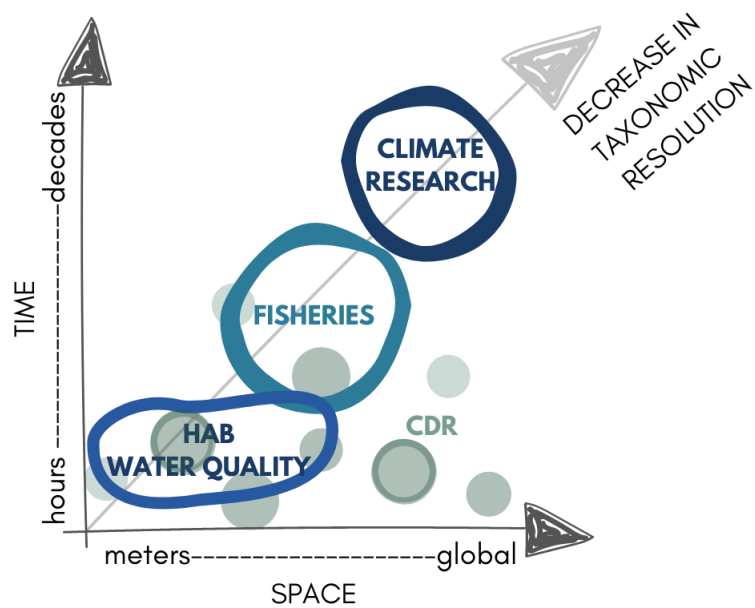1776 *Community of Practice members.*

1777

Figure 3. Needs of the PACE Community of Practice, based on the responses collected from a questionnaire with predetermined answers or other (optionally including a write-in answer).

1783

1784   Figure 4. PACE will provide data to answer scientific questions about phytoplankton community

1785   composition (PCC) at local to global scales. The user community, already actively engaged in

1786   developing frameworks and pipelines that use PACE data or support mission objectives, comes

1787   from different sectors, whose scientific objectives are sometimes distinct and sometimes

1788   overlapping. This illustration provides example research interests at different scales and from

1789   different sectors.

1790

1791

1792
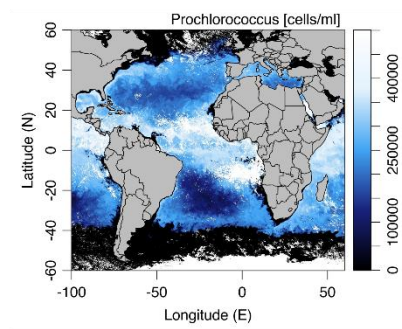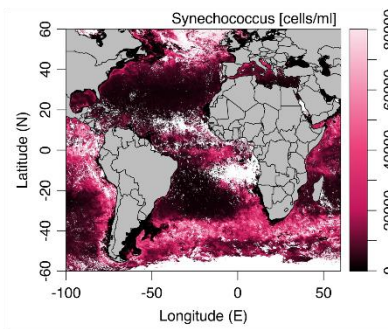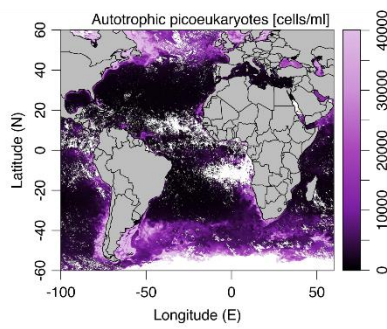
1793

*Figure 5. Spatial distribution of the autotrophic picoeucaryotes, Synechococcus, and Prochlorococcus,*

*stated as concentration of cells per volume, derived from MODIS Aqua (modified after Lange et al. 2020).*

1796

1797
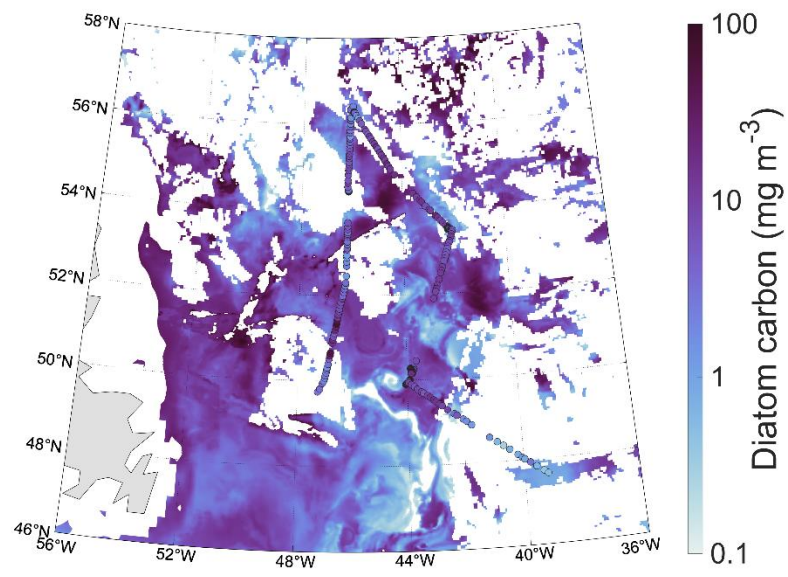
*Figure 6. Diatom Carbon concentration derived from three component neural network model, following*

*Chase et al. 2022.*

1798

1799

1800

1801
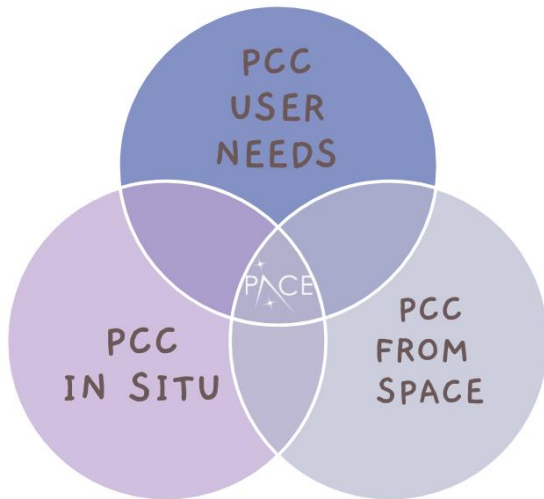
1802

1803 *Figure 7. PCC from PACE is an intersection of user's needs, facilitated and constrained by the*

1804 *spaceborne technology aboard the satellite, and our capability to measure and define the*

1805 *phytoplankton community structure in-situ.*

1806

1807

1808    *Table 1 Instrument specifications for OCI, HARP2 and SPEXone (modified from Table 2 from*

1809    *Werdell et al. (2019))*

|  | OCI | HARP2 | SPEXone |
|---|---|---|---|
| UV-NIR range (bandwidth) | Continuous from 340 to 890nm* in 2.5-nm steps (5) | 441 (16), 549 (10), 665 (11), and 866 (40) nm (nominal) | Continuous from 385 to 770 nm in 2-4nm steps |
| SWIR channels (bandwidth) | 940 (45), 1,038 (75), 1,250 (30), 1,378 (15), 1,615 (75), 2,130 (50) and 2,260 (75) nm | None | None |
| Polarized bands | None | All | Continuous from 385 to 770 nm in 15-45nm steps |
| Number of viewing angles | One, with fore-aft instrument tilt of ±20˚ to avoid sun glint | 10 for 440, 550 and 870 nm and 60 for 670 nm (spaced over 114˚) | 5 (-57˚, -20˚, 0˚, 20˚, 57˚) |
| Swath width | ±56.6˚ (2,663 km at 20˚ tilt) | ±47˚ (1,556km at nadir) | ±4˚ (100 km at nadir) |
| Global coverage | 1-2+ days | 2 days | ~30 days |
| Ground pixel | 1 km at nadir | 3 km | 2.5 km |
| Institution | GSFC | UMBC | SRON/Airbus |

1810

1811

1812 *Table 2. Comparison of different methods, capabilities, and their connection to expected PACE*

1813 *products (from perspective of algorithm development and validation)*

1814

| Method | Volume sampled (approx.) | Size range detected (approx.) | Taxonomic resolution | What is actually measured? | PACE-relevant examples |
|---|---|---|---|---|---|
| Light microscopy | 1-200 mL | 10-200 µm (nominal) | To species level | Cell concentrations and/or biovolume (can estimate cellular carbon) | Brewin et al. (2011a); Soto et al. (2015); Wolny et al. (2020) |
| HPLC pigments | 1-10 L | >0.3 µm (nominal; 0.7 µm for non-combusted GF/F) | Group level | Pigment concentrations | Bracher et al. (2015); Chase et al. (2017); Kramer et al. (2022); Uitz et al. (2015) |
| Flow cytometry | 0.1-2 mL | 0.2-50 µm (instrument and volume dependent) | *Prochlorococcus, Synechococcus,* pico- and nano-eukaryotes | Cell concentrations and optical properties (can estimate cellular carbon) | Lange et al. (2020), Kramer et al. (2020) |
| Imaging-in-flow cytometry | 2-5 mL | 6-150 µm (nominal) | To species level | Cell concentrations + biovolumes (can estimate cellular carbon) | Chase et al. (2022) |
| DNA metabarcoding | 0.5-10 L | >0.2 µm (filter dependent, nominal) | To species level | Relative sequence abundances and/or barcode gene concentrations | Catlett et al. (2021) |

1815
1816

1817

1818

1819    *Table 3. Compilation of published algorithms to assess phytoplankton community composition. Algorithms are considered global if they*

1820    *are designed for/applied to more than one major ocean.*

1821

| Application | PCC product(s) | Algorithm validation data | Remote sensing approaches | Hyperspectral (or polarization?) in situ approaches |
|---|---|---|---|---|
| Global | Taxonomic group(s) | Direct cell observation (cultures and/or field microscopy) | Subramaniam et al. (2001); Westberry et al. (2005) Subramaniam and Carpenter (1994) | |
| | | Pigment concentrations | Alvain et al. (2005); Alvain et al. (2008); Ben Mustapha et al. (2014); Bracher et al. (2009); Hirata et al. (2011); Losa et al. (2017); Moore et al. (2012); Palacz et al. (2013); Sadeghi et al. (2012); Soppa et al. (2014); Xi et al. (2020) | Torrecilla et al. (2011) |
| | | Spectral signatures | Brown and Yoder (1994) | |
| | Size classes, size index, or PSD | Pigment concentrations | Brewin et al. (2010); Brewin et al. (2015); Devred et al. (2006); Devred et al. (2011); Fujiwara et al. (2011); Hirata et al. (2008); Hirata et al. (2011); Kostadinov et al. (2010); Li et al. (2013); Moore and Brown (2020); Mouw and Yoder (2010); Roy et al. (2013, spectral a_ph also used in development); Uitz et al. (2006) | |
| | | Mie modeling, Coated Spheres model | Kostadinov et al. (2009); Kostadinov et al. (2022) | |
| | | Spectral signatures | Bricaud et al. (2012) | |
| | Accessory pigments | Pigment concentrations | O'Shea et al. (2021); Wang et al. (2018) | Bracher et al. (2015); Chase et al. (2013); (Chase et al. 2017); Kramer et al. (2022); Taylor et al. (2013); Uitz et al. (2015) |
| Regional /Local | Taxonomic group(s) | Direct cell observation (microscopy of cultures and/or field data or imaging-in-flow cytometry) | Chase et al. (2022); Raitsos et al. (2008) Rêve-Lamarche et al. (2017) | Kirkpatrick et al. (2000); Lubac et al. (2008); Millie et al. (1997); Xi et al. (2017); Xi et al. (2015) |
| | | Pigment concentrations | Di Cicco et al. (2017); Kramer et al. (2018); Palacios et al. (2015); Sathyendranath et al. (2004); Werdell et al. (2014) | Catlett and Siegel (2018); Isada et al. (2015); Shaju et al. (2015) |
| | | Spectral signatures | | Craig et al. (2006); Wynne et al. (2008) |
| | Size classes, size index, or PSD | Pigment concentrations | Gittings et al. (2019) | |
| | | Spectral signatures | Ciotti and Bricaud (2006) | |
| | Accessory pigments | Pigment concentrations | Bracher et al. (2015); Pan et al. (2010); Sun et al. (2022) | Aguirre-Gómez et al. (2001); Hoepffner and Sathyendranath (1991); Hoepffner and Sathyendranath (1993); Liu et al. (2019); Lohrenz et al. (2003); Wang et al. (2016); Ye et al. (2019) |

1822
1823

1824    *Table 4. Acronyms and symbols used in the text.*

| | |
|---|---|
| CDOM | Colored Dissolved Organic Matter |
| CDR | Carbon Dioxide Removal |
| CHEMTAX | CHEMical TAXonomy |
| CHIME3 | Copernicus Hyperspectral Imaging Mission for the Environment |
| *Chl* a | Chlorophyll a |
| CMIP | Climate Research Programme Coupled Model Intercomparison Project |
| CoP | Community of Practice |
| DNA | Deoxyribonucleic Acid |
| EOSDIS | NASA's Earth Observing System Data and Information System |
| EOV | Essential Ocean Variable |
| ESM | Earth System Models |
| FAIR | Findable, Accessible, Interoperable and Reusable |
| FCM | Flow cytometry |
| GIOP | Generalized IOP algorithm |
| GLIMR | Geosynchronous Littoral Imaging and Monitoring Radiometer |
| GSFC | Goddard Space Flight Center |
| HABs | Harmful Algal Blooms |
| HARP2 | Hyper Angular Research Polarimeter |
| HICO | Hyperspectral Imager for the Coastal Ocean |
| HNLC | High Nutrient Low Chlorophyll |
| HPLC | High performance liquid chromatography |
| IFCB | Imaging FlowCytobot |
| IOCCG | *International Ocean-Colour Coordinating Group* |
| IPCC | Intergovernmental Panel on Climate Change |
| GCOM-C | Global Change Observation Mission – Climate satellite |
| JAXA | Japan Aerospace Exploration Agency |
| LEO | Low Earth Orbit |
| MAAs | Mycosporine-like Amino Acids |
| MERIS | Medium Resolution Imaging Spectrometer |
| MODIS | Moderate Resolution Imaging Spectroradiometer |
| NASA | National Aeronautics and Space Administration |
| NASA NOMAD | NASA bio-Optical Marine Algorithm Dataset |
| NIR | Near-Infrared |
| OB.DAAC | NASA's Ocean Biology Distributed Active Archive Center |
| OCI | Ocean Color Instrument |
| PACE | NASA's Plankton, Aerosol, Cloud, ocean Ecology |
| PAR | Photosynthetically Active Radiation |
| PCC | Phytoplankton Community Composition |
| PhytoDOAS | Extension of the Differential Optical Absorption Spectroscopy (DOAS, a method for detection of atmospheric trace gases), developed for remote identification of oceanic phytoplankton groups |
| POLDER | Polarization and Directionality of the Earth's Reflectance |
| PSC | Phytoplankton Size Class |
| $R_{rs}(\lambda)$ | Remote Sensing Reflectance |

| | |
|---|---|
| SBG | Surface Biology and Geology mission |
| SCIAMACHY | Scanning Imaging Absorption Spectrometer for Atmospheric Chartography |
| SeaBASS | SeaWiFS Bio-optical Archive and Storage System |
| SeaWiFS | Sea-viewing Wide Field-of-view Sensor |
| SMD | Science Mission Directorate |
| SPEXone | Spectro-polarimeter for Planetary Exploration |
| SPM | Suspended Particulate Matter |
| SRON | Netherlands Institute for Space Research |
| SST | Sea Surface Temperature |
| SWIR | Short-wave Infrared |
| TOA | Top-of-atmosphere |
| UMBC | University of Maryland, Baltimore County |
| UV | Ultraviolet |
| UVP | Underwater Vision Profiler |
| VIIRS | Visible Infrared Imaging Radiometer Suite |

1825