Globally Consistent Quantitative Observations of Planktonic Ecosystems

Fabien Lombard1,2*, Emmanuel Boss3*, Anya M. Waite4, Meike Vogt5, Julia Uitz1, Lars Stemmann1, Heidi M. Sosik6, Jan Schulz7, Jean-Baptiste Romagnan8, Marc Picheral1, Jay Pearlman9, Mark D. Ohman10, Barbara Niehoff11, Klas O. Möller12, Patricia Miloslavich13,14, Ana Lara-Lopez13, Raphael Kudela15, Rubens M. Lopes16, Rainer Kiko17, Lee Karp-Boss3, Jules S. Jaffe18, Morten H. Iversen15,18, Jean-Olivier Irisson1, Katja Fennel19, Helena Hauss20, Lionel Guidi1, Gaby Gorsky1, Sarah L. C. Giering20, Peter Gaube21, Scott Gallager22, George Dubelaar22, Robert K. Cowen23, François Carlotti24, Christian Briserio-Avena25, Léo Berline24, Kelly Benoît-Biro25, Nicholas Bax13,26, Sonia Batter27, Sakina Dorotheé Ayata1,28, Luis Felipe Artigas29 and Ward Appeltans30

1 Sorbonne Université, CNRS, Laboratoire d’Océanographie de Villefranche, Villefranche-sur-Mer, France, 2 Institut Universitaire de France, Paris, France, 3 School of Marine Sciences, University of Maine, Orono, ME, United States, 4 Department of Oceanography, Ocean Frontier Institute, Dalhousie University, Halifax, NS, Canada, 5 Institute for Biogeochemistry and Pollutant Dynamics, ETH Zürich, Zurich, Switzerland, 6 Woods Hole Oceanographic Institution, Woods Hole, MA, United States, 7 Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany, 8 Ifremer Centre Atlantique, Unité Écologie et Modèles pour l’Haliutique (EMH), Nantes, France, 9 IEEE, Port Angeles, WA, United States, 10 Scripps Institution of Oceanography, UC San Diego, La Jolla, CA, United States, 11 Alfred-Wegener-Institut, Helmholtz Centre of Polar and Marine Research, Bremerhaven, Germany, 12 Institute of Coastal Research, Helmholtz-Zentrum Geesthacht, Geesthacht, Germany, 13 Institute for Marine and Antarctic Science, University of Tasmania, Hobart, TAS, Australia, 14 Departamento de Estudios Ambientales, Universidad Simon Bolivar, Caracas, Venezuela, 15 Department of Ocean Sciences, UC Santa Cruz, Santa Cruz, CA, United States, 16 Oceanographic Institute, University of São Paulo, São Paulo, Brazil, 17 GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany, 18 MARUM and Faculty of Geosciences, University of Bremen, Bremen, Germany, 19 Department of Oceanography, Dalhousie University, Halifax, NS, Canada, 20 Ocean Biogeochemistry and Ecosystems, National Oceanography Centre, Southampton, United Kingdom, 21 Applied Physics Laboratory, University of Washington, Seattle, WA, United States, 22 CytoBuoy b.v., Woerden, Netherlands, 23 Hatfield Marine Science Center, Oregon State University, Newport, OR, United States, 24 CNRS, IRD, MIO UM 110, Aix Marseille Utlav., Universite de Toulon, Marseille, France, 25 Monterey Bay Aquarium Research Institute, Moss Landing, CA, United States, 26 CSIRO Oceans and Atmosphere, Hobart, TAS, Australia, 27 The CPR Survey-MBA, Nanaimo, BC, Canada, 28 Institut de Systématique, Évolution, Biodiversité, Muséum National d’Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France, 29 Univ. Littoral Côte d’Opale, Univ. Lille, CNRS, UMR 8187, Laboratoire d’Océanologie et de Géosciences, Wimereux, France, 30 Intergovernmental Oceanographic Commission (IOC) of UNESCO, IOC Project office for IODE, Oostende, Belgium

In this paper we review the technologies available to make globally quantitative observations of particles in general—and plankton in particular—in the world oceans, and for sizes varying from sub-microns to centimeters. Some of these technologies have been available for years while others have only recently emerged. Use of these technologies is critical to improve understanding of the processes that control abundances, distributions and composition of plankton, provide data necessary to constrain and improve ecosystem and biogeochemical models, and forecast changes in marine ecosystems in light of climate change. In this paper we begin by providing the motivation for plankton observations, quantification and diversity qualification on a global scale. We then expand on the state-of-the-art, detailing a variety of relevant and (mostly) mature technologies and measurements, including bulk measurements of plankton, pigment composition, uses of genomic, optical and acoustical methods as well

Frontiers in Marine Science | www.frontiersin.org 1 April 2019 | Volume 6 | Article 196
1. CONTEXT AND RATIONALE WHY SHOULD WE OBSERVE PLANKTON AND PARTICLES IN THE OCEAN?

Plankton are organisms that have either null velocities or velocities significantly smaller than oceanic currents and thus are considered to travel with the water parcel they occupy (some can, however, through vertical swimming or change in buoyancy, move between water masses). Plankton are an extremely diverse group, spanning across several orders of magnitude in size (sub-microns to several a meter) and with representatives from all trophic levels (auto-, mixo- and hetero-trophic) and all domains of life, including bacteria, protists and small to large metazoa. Knowledge on distribution patterns and composition of planktonic communities is of great interest for many reasons: (1) Life is thought to have begun in the ocean, and planktonic organisms therefore provide important evolutionary insights on diversity, complex life forms and adaptations to highly variable environments. (2) Plankton form the foundation of most marine food webs, with phytoplankton accounting for approximately 50% of global primary production. The organic material produced by phytoplankton is consumed by herbivorous zooplankton that provide a crucial source of food for higher trophic levels, including important commercial species, marine mammals and sea birds. (3) Many important commercial species have early developmental stages that are planktonic. Time spent in the plankton is often considered a time of greatest vulnerability and hence a bottleneck for recruitment. (4) Planktonic organisms play important roles in global cycles of the majority of oceanic elements. (5) Because of their restricted environmental preferences and relatively short life spans, plankton abundance and composition react tightly to local and global environmental changes (e.g., Mackas and Beaugrand, 2010; Beaugrand et al., 2013, 2015; Edwards et al., 2013), and can serve as sentinel organisms of environment and water quality changes, and (6) some planktonic organisms can be toxic or cause disease and parasitism in animals, including commercial ones (Anderson et al., 2019). Given these important functions, the biomass and diversity of phytoplankton and zooplankton were identified as Essential Ocean Variables (EOVs) by the Global Ocean Observing System (Chiba et al., 2018; Miloslavich et al., 2018; Muller-Karger et al., 2018; Bax et al., 2019) as well as Essential Climate Variables (ECVs) under GCOS (Global Climate Observing System).

In addition, the upper ocean is teeming with organic and inorganic particles which can form larger aggregates (marine snow) very often colonized by micro-organisms. These aggregates can sink rapidly, creating one of the major fluxes of matter to the deep ocean (e.g., Kiko et al., 2017), forming the base of the food web in this dark ocean, and are an important vector in controlling nutrient distribution across the ocean. Recognizing this role, particulate matter was also identified as an EOV and ECV. In particular, the concentration of particulate matter provides quantitative information on spatial gradients and temporal variations in total particulate substrates, which are mostly comprised of organic material in the open ocean.

Understanding how planktonic communities change at regional and global scales, which is critical to address ocean health, food security, and biogeochemical cycles, requires a multidisciplinary approach to ecology. While planktonic organisms have been studied in great detail for decades, the relatively inconsistent and diverse use of various methods and techniques (taxonomic, isotopic, genomic, biogeochemical, etc.) makes it difficult to uncover global and/or long-term patterns and trends in abundances, diversity and composition of the plankton. As a result, global analyses are to date restricted to basic descriptors of plankton such as total abundance or biomass. For instance, phytoplankton distributions are most often described through chlorophyll a measurements and the concentrations of other pigments, which provide, through empirical relationships, limited taxonomic and size information on phytoplankton communities (Swan et al., 2016). Knowledge of the distribution of larger organisms often relies on the collation of numerous and heterogeneous observations from several decades and from diverse samplers, providing a crude description of zooplankton total biomass at a global scale (e.g., Moriarty and O’Brien, 2013; Moriarty et al., 2013; O’Brien et al., 2017) or through empirical relationships between several planktonic groups (Buitenhuis et al., 2013), with only little concern about diversity. Hence, our general understanding of plankton abundance and diversity is highly fragmented due to a paucity of data, heterogeneity among collection methods, analysis techniques, technology and

Keywords: plankton, imaging, OceanObs, autonomous platforms, global observing, EOVs, ECVs
scale of measurements. It is also fragmented because there is a lack of standardization of methods, data outputs and data curation.

Recently, new sensors, instruments, platforms (e.g., Argo floats see Roemmich et al., 2019, and gliders see Testor et al., 2019) and methods (imaging, acoustics, omics, etc.) have been developed and deployed to improve the spatiotemporal resolution of planktonic communities and particles (Powell and Ohman, 2015a; Brownlee et al., 2016; Hunter-Cevera et al., 2016; Ohman et al., 2018) even in hostile conditions (Grossmann et al., 2015). Here we suggest that with these technologies (as well as concurrent advances in software), we have the ability to collect and analyze significantly more global information on plankton distributions and diversity and at a finer scale (taxonomic, spatial, and temporal) than is currently done. Such a global plankton observation effort is not intended to replace fine, high-quality and highly precise local sampling and observations, usually done during specific oceanographic cruises focused on process studies or long-term observation sampling. Rather, the goal is to expand and upscale our observational capabilities at larger scales and with consistent data to provide a sustained observing system. To do so, we propose leveraging on existing international coordinated sampling programs and infrastructures (e.g., GO-SHIP Sloyan et al., 2019, OCEAN-SITES, Argo) by adding supplemental measurements that could be carried out without major changes to current sampling. In addition, integrating those supplementary measurements into existing long-term surveys would allow the description of key elements of the plankton system at both local and global scales.

This article is organized as follows; we first review the state of the art for plankton sampling in general and imaging in particular, focusing on commercially available instruments for operational reasons. We then detail what could be expected from a global plankton sampling program. We identify challenges associated with such a program and finally make recommendations on how to proceed over the next decade.

2. STATE OF THE ART

There are no global standard methodologies for the quantification of plankton across the different plankton groups (e.g., phyto-, zooplankton) or within each group. Few exceptions are analytical chemical methods, such as phytoplankton pigment analysis using HPLC and the use of WP2 nets from 0 to 200 m for zooplankton field sampling (UNESCO, 1968). This partly explains why existing plankton time series observations are difficult to compare except in their trends (O’Brien et al., 2017), since they are based on different methodologies and taxonomic resolution. Physical and biogeochemical marine research have systematic and sustained sampling programs that, at times, measure plankton-related variables, albeit at a low taxonomic and functional resolution (e.g., diel vertical migration pattern from acoustic Doppler current profilers, total in vivo fluorescence allowing the estimation of chlorophyll a concentrations). For decades physical and biogeochemical marine research has been tightly coordinated at the international level (e.g., GO-SHIP, GEOTRACES, OCEAN-SITES), which has not been the case for biological and ecological marine research.

One exception is the plankton-relevant data available from ocean color passive remote sensing, which provides global information on the distribution and concentration of chlorophyll a, organic carbon and phytoplankton carbon distributions in the surface ocean on a near-daily time scale (Siegel et al., 2013). Unfortunately, data are limited to the upper ocean and are only continuously available since 1996 (for a future perspective see Groom et al., 2019). Methods to obtain taxonomic and size information from remote sensing have been proposed (e.g., Alvain et al., 2008); however, these methods in general lack validation, and it is unclear if it is possible to use remote sensing to gain information about phytoplankton size and taxonomy beyond that which is globally correlated to chlorophyll (Chase et al., 2017). Recently, it has been shown that active remote sensing with atmospheric LIDAR (Light Detection and Ranging), not specifically tuned for ocean observations, can provide a proxy of concentration of upper ocean particles, even during the polar night and through thin cloud covering (Hostetler et al., 2018).

Another exception is the Continuous Plankton Recorder (CPR) surveys (Batten et al., 2003, 2019). CPR has the most extensive spatial coverage of any plankton sampling program and provides taxonomically resolved abundance data. Commercial ships or other ships-of-opportunity tow the CPR along their regular routes in near-surface waters (Batten et al., 2003). Lengthy and consistent time series exist for several locations such as the North Atlantic and North Sea (> 60 years), the Southern Ocean (> 25 years) and the North Pacific (19 years) as well as more recent regional surveys in the Mediterranean and around Australia, for example (Batten et al., 2018). It is, to date, the most cost-effective way to sample ocean basins over large distances. However, CPR sampling occurs at relatively low resolution (each sample represents 18.5 km of ocean), is confined to the surface ocean, and currently poorly covers the tropics and the sub-tropics (i.e., no global coverage). While the CPR is an internally consistent sampler, not all planktonic groups are well sampled or well resolved (very small, very large and/or fragile plankton especially), so abundance estimates are best considered as semi-quantitative, and integration with other data sources requires careful treatment (Richardson et al., 2006). Nevertheless, the existing lengthy CPR time series is invaluable and should be incorporated into continuing assessments of changing ecosystems.

3. MEASUREMENTS WITH INCREASING CAPABILITIES

While discrete water samples, net tows, CPR and ocean color provide important data for the study of oceanic plankton on basin scales, they are far from sufficient in their taxonomic, spatial and temporal resolution. There is a clear need for a systematic and sustained global sampling program that will take advantage of technological advances that provide significantly more taxonomic information. Such a program will likely build on the current standard measurements, expanding them to a wider
range of pelagic systems, to a finer spatial, temporal and also functional/taxonomic resolution by incorporating the growing field of -omics and the expanding suite of autonomous sensors and quantitative imaging technologies. In this section, we provide a list of existing sampling technologies for plankton. While all have achieved a technology readiness level (TRL) beyond the prototype stage, some of these technologies are more mature than others. Assessment of TRL depends, among other things, on method of deployment (e.g., AUV vs. R/V), but we avoid getting into these details here. Much about the TRL of a specific technology for a specific application can be assessed from the literature provided. Literature emanating from groups who have not developed the technology is particularly useful to assess the operational status of a given technology.

### 3.1. Analysis of Water Samples

Discrete sample analysis is performed on water samples taken from CTD-rosettes, surface buckets, water pumped from the surface into vessels (flow-through systems) and plankton nets. Such analyses often target the bulk properties of the underlying particle/plankton population, including chemical analysis of mass and elemental content, pigment content, size distribution and genetics.

#### 3.1.1. Bulk Mass and Elemental Composition

Total suspended mass in the upper open ocean is dominated by plankton-derived particles. The analysis of organic carbon, phosphorous, nitrogen and micro elements associated with bulk particulate samples retained on a filter provides essential descriptors of the dynamics of such particles. The associated methods have been determined and refined for decades (e.g., Hurd and Spencer, 1991; Cutter et al., 2017).

#### 3.1.2. Diagnostic Pigments

Information on phytoplankton diversity can be gained from High Performance Liquid Chromatography (HPLC) analysis of pigments present in bulk samples. Whereas, chlorophyll a is present in all phytoplankton (although in its divinyl form in prochlorophytes) and typically used as a proxy of phytoplankton biomass, accessory pigments vary with phytoplankton community composition, and some pigments can be used as biomarkers of specific taxa (Gieskes and Kraay, 1983; Jeffrey et al., 1997; Roy et al., 2011). Several pigment-based approaches have been proposed that allow estimating the relative contribution to chlorophyll a of different phytoplankton taxa (CHEMTAX algorithm Mackey et al., 1996) or taxonomic groupings or size classes (Claustre, 1994; Vidussi et al., 2001; Uitz et al., 2006; Aiken et al., 2008). Pigment-based methods have the advantage that they cover the whole phytoplankton assemblage in a single analysis and provide a quantitative assessment of phytoplankton community composition at the level of class or higher (Bax et al., 2001). However, lacking flow cytometry and microscopy validation, these methods can have large uncertainties linked to variability in accessory pigmentation within a given taxon or induced by environmental factors (Henriksen et al., 2002; Laviale and Neveux, 2011). It is also recognized that some biomarker pigments are not restricted to one single taxon, leading to ambiguity in the discrimination of some phytoplankton groups. Chlorophyll a can also be estimated from fluorometric and spectrophotometric analysis; however, such data can be significantly different from determination obtained with HPLC and are not recommended as a standard. However, they have significant value in conjunction with long-term time-series, where they have been used for years.

#### 3.1.3. Genomic and Next Generation Sequencing

Next generation sequencing (NGS, also called high throughput sequencing, HTS) provides relatively cost-effective and fast sequencing of DNA and RNA. Initially developed for microbial ecology, and now applied to any drop of water, it has been progressively applied to the whole marine ecosystem (e.g., metabarcoding, metagenomics, metatranscriptomics), encompassing plankton, nekton and even benthic organisms. Barcoding targets and amplifies a specific sequence that is highly conserved yet variable across taxa, such as the 18S ribosomal DNA or cytochrome oxidase (COI) genes for eukaryotes, in order to identify a specific organism. Metabarcoding (MetaB) is the application of barcoding at the scale of a whole water sample and allows the classification of organisms either expressed as in Operational Taxonomic Units (OTU) whose DNA is present in a given sample (Bucklin et al., 2016) or assigned to different taxa through the comparison with databases of named reference sequences. They represent the principal way of addressing biodiversity for prokaryotes and, to some extent, for pico- and nanoekaryotes as well. Similarly, metagenomics (MetaG) allows the study of all genes present in a given water sample, giving access to the reconstruction of full genomes (Metagenome-assembled genomes or MAGs) for non-model and uncultivated organisms (e.g., Delmont et al., 2018; Tully et al., 2018). RNA sequencing through transcriptomic and metatranscriptomic (MetaT) methods provide new insights into gene expression, and DNA or RNA sequences converted to amino acid sequences (e.g., proteins) allow the characterization of protein structures and phenotypic variations across planktonic communities (e.g., Carrade et al., 2018). These techniques can also be applied at the individual level of a single cell, giving new insight on the genomes and their expressions of uncultured species.

MetaB requires gene amplification using primers that attempt to be as universal as possible, but may still lead to differential amplification of some organisms, whereas MetaG could exhibit biases associated with the length of reads obtained through HTS (although long-read sequencing is currently being developed, it currently has a relatively high level of sequencing error). Due to these different steps, -omics methods provide results that, at present, should only be considered as relative, or semi-quantitative at best (e.g., Bucklin et al., 2016). Another limitation of these methods is the lack of reference sequences in databases, although the reconstruction of MAGs from MetaG data could partly overcome this issue.

MetaB and “environmental” DNA (eDNA) analyses, since they allow estimation of the diversity of samples with unprecedented taxonomic resolution and have become relatively affordable and automated (Ji et al., 2013), have the potential to be used in global
surveys to address the diversity of organisms. However, they may not yet be ready nor affordable for large-scale ecosystem surveys (Deiner et al., 2017), while the complexity of analyzing the resulting data remains a challenge. In fact, 10 years ago (Bowler et al., 2009) predicted that miniaturized ecogenomic sensors able to monitor almost all planktonic DNA and RNA would allow near real-time measurements of microbial activity in the ocean and associated biogeochemical processes such as carbon flux. However, such an achievement has not yet been reached, but the recent development of Oxford Nanopore MinION sequencing (Jain et al., 2016; Lu et al., 2016) and its first applications in the marine environment (Warwick-Dugdale et al., 2018) indicate that this possibility may become available soon.

3.2. Sensors
Sensors measuring bio-acoustical and bio-optical properties (see below) have been mounted on CTD-rosette frames, undulating vehicles, autonomous vehicles and floats, ship hulls (for bio-acoustics) and in-line flow-through systems (bio-optics). All sensors require calibrations to obtain absolute physical values comparable across instruments and proxy-calibrations to convert signals to biogeochemical parameters, as these measurements provide a proxy—not a direct—estimate of variables of interest. This is because while the forward problem is well defined (e.g., predict the acoustical or optical signal measured given a field of known organisms with specific optical or acoustical characteristics), the inverse problem is not (i.e., there are many possible configurations of organisms that could result in the observed signal; some, however, are more likely than others if we have additional knowledge). Hence, these methods are more useful when deployed in conjunction with other measurements.

3.2.1. Bio-Acoustics
Acoustic methods can reveal much about the spatial distribution and temporal dynamics of zooplankton. For example, echosounders led to the discoveries of the diel vertical migration of plankton and micronekton (Johnson, 1948) and their ubiquitous and dense but previously hidden aggregations (Cheriton et al., 2007). The ability of acoustic tools to simultaneously assess animals ranging in size from sub-mm to m allows ecological processes in the plankton to be examined (Kaartvedt, 2000; Ballón et al., 2011; Benoit-Bird and McManus, 2014; Powell and Ohman, 2015b) when appropriate frequencies are chosen. However, this ability also highlights a key challenge—separating animal types and accurately assessing the biomass of each. While these approaches have long been used for fish stock assessment and management of a number of species (MacLennan and Simmonds, 1992), in plankton, dramatic differences in body size, species composition, elastic properties of the animals and orientation markedly influence the acoustic reflectivity or target strength (Roberts and Jaffe, 2008; Briseño-Avena et al., 2018), coupled with the complexity of the community, making separation of taxa and assessment of biomass difficult. To address these challenges, research efforts have recently shifted from adding additional narrow-band signals (e.g., Holliday et al., 2009) to utilizing a continuum of frequencies to increase the amount of information available from acoustic returns (e.g., Jech et al., 2017). Acoustic measurements, however, will always have uncertainties; many of the greatest insights on zooplankton resulted from creative integration of multiple, complementary sampling devices including acoustics with nets, optics, imaging and animal tagging to take advantage of the different strengths and fill in the gaps of each approach (reviewed in Benoit-Bird and Lawson, 2016). Multi-sensor fusion efforts have the potential for wider application through the use of autonomous platforms, which resolves the limited range issue of high frequency acoustics. While bio-acoustic instruments have only begun to be deployed over long periods of time on autonomous platforms (Powell and Ohman, 2015a,b), we expect that strong development and wide use of these instruments will be seen in the next decade (Benoit-Bird et al., 2018).

3.2.2. Bio-Optics
Like acoustic, optical measurements are best used with complementary sampling approaches. Measurements of the optical characteristics of water (e.g., absorption, scattering, attenuation and fluorescence) in situ have been used for decades (e.g., Gardner et al., 2018) to characterize bulk properties associated with micrometer-size particles in general and phytoplankton in particular (near-forward scattering extends this range to a few 100 µm). Deployed on profiling floats, gliders, moorings, CTD rosette frames and in-line systems they are capable of providing high-resolution information on the spatial distribution of phytoplankton but with little specificity in terms of composition. They can also provide information on a few pigments beyond chlorophyll a (Chekalyuk and Hafez, 2008; Proctor and Roesler, 2010; Chase et al., 2013) which have been used to provide estimates of phytoplankton functional groups (e.g., MacIntyre et al., 2010; Houleiz et al., 2012). Data on basin scales from in-line systems and BGC-Argo floats have become available (Boss et al., 2013; Rembauville et al., 2017), which is useful, for example, to validate satellite-based algorithms (Werdell et al., 2013; Haëntjens et al., 2017). These techniques are also useful as proxies of particulate organic carbon in general and phytoplankton carbon in particular (Cetinić et al., 2012; Graff et al., 2015). Bio-optical sensors are more sensitive to fouling and require periodic cleaning or bio-shutters. Simple measurements, such as the Secchi disk, have been very useful to characterize the optical status of the upper ocean including long-term changes of plankton (Boyle et al., 2010).

3.3. Particle Size Distribution
Particle counters are designed to count particles and obtain information on their size. There are two main types of particle counters: (1) Electronic particle counters measure the change in impedance as an estimate of the physical volume occupied by a particle while it passes through a small aperture (Coulter et al., 1966; Graham, 2003). (2) Optical counters are based on shadow (HIAC counters and the LOPC, the latter no longer commercially available) or near-forward scattering of a laser beam as it encounters particles (Sequoia’s LISST). While many of these technologies have been around for decades, in recent years the LISST has been successfully deployed in open ocean environments on CTD rosette frames (Reynolds et al., 2010;
Barone et al., 2015; Leroux et al., 2018) and in flow-through systems (Boss et al., 2018). Note that optical counters are primarily sensitive to the cross-sectional area of particles while the resistance-based particle counter size is based on particle volume, hence the specific size associated with the particles, based on an equivalent sphere, is different in both cases. Additionally, handling and hydrodynamics associated with delivering the sample to the sizing instrument can affect the size determined, due to aggregation/disaggregation processes in the sample prior to sampling. Hence, some researchers intentionally disaggregate the sample prior to measurement (Milligan and Kranck, 1991).

3.4. Analysis of Individual Organisms and Particles

Imaging of individual organisms and particles, as long as the volume analyzed is well quantified, makes it possible to obtain simultaneously: (1) abundance of the different groups of plankton and their relative contribution to total abundance and biomass, (2) morphological or optical measurements on the organisms that can be used to obtain their biovolume as a proxy of their biomass, but also to derive size spectra of the imaged objects and (3) production of a digital archive of images and optical properties that can be shared or reprocessed if more information is needed. In addition, imaging systems can be operated on fresh samples on research vessels or in situ, as well as on concentrated or fixed samples. In some cases, images may reveal behavioral information (e.g., predator-prey interactions, parasitism, diurnal vertical migration) as well as physiological state and population conditions, from which rates may be inferred (e.g., lipid content and egg-production rates of ovigerous copepod species, Möller et al., 2015; growth rate of phytoplankton species, Dugenne et al., 2014).

Since the 1980s, a considerable amount of energy has been directed to produce prototypes of automated quantitative imaging devices [see reviews in Foote (2000), Wiebe and Benfield (2003), Benfield et al. (2007), Sieracki et al. (2010), Stemmann and Boss (2011)], among which some are now commercially available (Table 1). Some global patterns in large plankton communities and/or particle fluxes have emerged from intense use of these devices during ship surveys (Stemmann et al., 2008b; Bonato et al., 2015; Gündel et al., 2015; Thyssen et al., 2015; Biard et al., 2016; Waite et al., 2016) or on high frequency platforms (Thyssen et al., 2008, 2014). These instruments can be used in situ, in the laboratory and/or on a research vessel depending on their design, but they all share some common principles. Marine particles and plankton either pass by or are placed in a known volume illuminated by a specific light source. For optical devices, various optical measurements are made (e. g., fluorescence), while for imaging devices a picture is taken and measurements inferred from the picture (both could happen for the same object in the case of imaging flow cytometers). Images can be classified according to taxonomic or functional groups and living cells can be separated from aggregates and other non-living particles.

Imaging devices also provide common particle characteristics: each object’s size, shape and cross-sectional area can be determined as well as the intensity of light coming from each pixel of the particle, identified thanks to its optical or image characteristics, producing a large amount of raw data. Sometimes, these data are used to provide statistics for a given group (e.g., flow cytometry) or for given sizes. Each optical/imaging technique also comes with its own size range limitation (Table 1 and Figure 1). Small particles are often too small to be imaged efficiently (too few pixels, signal below threshold or near noise level), while larger organisms are too scarce to be sampled quantitatively (volume analyzed is too small), or too large to pass by the tubing of some devices, resulting in a narrow size range compared with the theoretical one (Figure 2). This “effective” range is often missing from technical documentation and needs to be determined experimentally. Additionally, to obtain taxonomic information from optical or imaging methods, there is a need for a computer-assisted human expert to classify organisms based on their optical properties (e.g., “gating” in flow cytometry) or on their image. While machine-learning methods are getting progressively more efficient (Luo et al., 2017), the final taxonomic resolution is often limited, and may include substantial errors (Culverhouse et al., 2003, 2006). The increased capabilities in automated recognition of images still needs to be complemented with taxonomic expertise. Training libraries associated with specific technologies are expanding and could be refined to be applied to new datasets as well as to revisit old datasets. This ensures that the taxonomic knowledge is preserved in the form of image libraries, in a fashion similar to plankton identification field guides (but not quite as rigorous as taxonomic keys). Finally, strong analytical and programming skills as well as computer resources are needed to sort and analyze these often very complex data, which combine organisms’ taxonomic, morphological and optical properties together with their concentration.

3.4.1. Flow Cytometers

Flow cytometry measures fluorescence and scattering signals from single virus, bacterial or protistan cells contained in a seawater sample. Flow cytometry fluidics are designed to orient individual particles through a capillary where they are illuminated sequentially by one or several laser beams. The scattering signals (forward and side scattering, FSC and SSC) accompanied with different fluorescence signal intensities are recorded for each individual object. Fluorescence is either natural to the cell (e.g., photosynthetic pigments) or originating from a specific stain (e.g., SYBR green stains, which bind to DNA). Therefore, flow cytometry can record simultaneously on the same particle several features representative of its size (FSC), granularity (SSC) or pigment quantity and composition (fluorescence). Note, however, that SSC depends on size, but also shape and refractive index of the analyzed cell (Green et al., 2003; Agagliate et al., 2018), and without a proper calibration against cells similar to the ones present in a sample (e.g., Laney and Sosik, 2014), the size of particles inferred from FSC could be significantly biased. Different populations of particles sharing similar characteristics are often lumped (a human-based “gating” step) to represent sub-populations of the underlying assembly.
<table>
<thead>
<tr>
<th>Instrument</th>
<th>total size range (ESD)</th>
<th>Size target for quantitative observations</th>
<th>Typical sample volume</th>
<th>Sampling</th>
<th>Sample</th>
<th>Condition of use</th>
<th>Max. operating range</th>
<th>Method</th>
<th>Approximate cost</th>
<th>Seller</th>
<th>Final results</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIAC</td>
<td>1.3–600 µm</td>
<td>tbe</td>
<td>25–100 ml/min</td>
<td>Water</td>
<td>Fresh (fixed) On board/ laboratory</td>
<td>na</td>
<td>Optical</td>
<td>26,000 €</td>
<td>Beckman-Coulter</td>
<td>Particles nb, size</td>
<td></td>
</tr>
<tr>
<td>Coulter Counter</td>
<td>0.4–1,600 µm</td>
<td>tbe</td>
<td>&gt; 8 ml/min</td>
<td>Water</td>
<td>Fresh (fixed) On board/ laboratory</td>
<td>na</td>
<td>Optical</td>
<td>45,000€</td>
<td>Beckman-Coulter</td>
<td>Particles nb, size</td>
<td></td>
</tr>
<tr>
<td>Flowcytometry</td>
<td>0.2–100 µm</td>
<td>0.2–20 µm</td>
<td>10 µL – 1 ml</td>
<td>Water</td>
<td>Fixed with Glutaraldehyde</td>
<td>Laboratory</td>
<td>na</td>
<td>Optical</td>
<td>-</td>
<td>several available</td>
<td></td>
</tr>
<tr>
<td>Imaging</td>
<td>5–150 µm</td>
<td>10–60 µm</td>
<td>5 ml (15 ml/h)</td>
<td>Water</td>
<td>Fresh (fixed) In situ / on board / laboratory</td>
<td>40 m</td>
<td>Optical/Imaging</td>
<td>$158,000</td>
<td>McLane research laboratories inc.</td>
<td>Particles nb, type, size</td>
<td></td>
</tr>
<tr>
<td>FlowCytobot</td>
<td>0.2–800 µm</td>
<td>(optical) 1–600 µm</td>
<td>0.5–5 ml (5–1,000 µ/min)</td>
<td>Water</td>
<td>Fresh (fixed) In situ / on board / laboratory</td>
<td>200 m</td>
<td>Optical (scanning flowcytometry) / imaging (optional)</td>
<td>90,000–130,000 depending on configuration</td>
<td>CytoBoy b. v.</td>
<td>Particles nb, type, size</td>
<td></td>
</tr>
<tr>
<td>FlowCam Nano</td>
<td>300 nm–30 µm</td>
<td>tbe</td>
<td>20 µL / min</td>
<td>Water</td>
<td>Fresh / Fixed On board/ laboratory</td>
<td>na</td>
<td>Imaging</td>
<td>$117,000</td>
<td>Fluid Imaging Technologies</td>
<td>Particles nb, type, size</td>
<td></td>
</tr>
<tr>
<td>FlowCam with different objectives:</td>
<td>- 2X 75–1,000 µm (water) 200 µm</td>
<td>- tbe</td>
<td>na 5 ml/min – 40 ml / min</td>
<td>Water + Plankton nets</td>
<td>Fresh/ Fixed On board/ laboratory</td>
<td>na</td>
<td>Imaging</td>
<td>$54,000</td>
<td>Fluid Imaging Technologies</td>
<td>Particles nb, type, size</td>
<td></td>
</tr>
<tr>
<td>FlowCam-macro</td>
<td>300–5,000 µm</td>
<td>tbe</td>
<td>100–900 ml/min (several m3 depending on net cast)</td>
<td>Plankton nets</td>
<td>Fresh / Fixed On board/ laboratory</td>
<td>na</td>
<td>Imaging</td>
<td>$54,000</td>
<td>Fluid Imaging Technologies</td>
<td>Particles nb, type, size</td>
<td></td>
</tr>
<tr>
<td>Zooscan</td>
<td>150–100,000 µm</td>
<td>240–5,000 µm (depending on nets)</td>
<td>several m3 depending on net cast</td>
<td>Plankton nets</td>
<td>Fresh / Fixed Laboratory</td>
<td>na</td>
<td>Imaging</td>
<td>$26,000€</td>
<td>Hydroptic</td>
<td>Particles nb, type, size</td>
<td></td>
</tr>
<tr>
<td>LISST-200</td>
<td>1–500 µm</td>
<td>tbe</td>
<td>-</td>
<td>Cast</td>
<td>Fresh</td>
<td>In situ</td>
<td>600 m</td>
<td>Optical</td>
<td>$37,500</td>
<td>Sequoia</td>
<td>Particles nb, size</td>
</tr>
<tr>
<td>LISST-Holo2</td>
<td>25–2,500 µm</td>
<td>tbe</td>
<td>30 ml/s</td>
<td>Cast</td>
<td>Fresh</td>
<td>In situ</td>
<td>3,000 m</td>
<td>Optical</td>
<td>$39,500</td>
<td>Sequoia</td>
<td>Particles nb, size</td>
</tr>
<tr>
<td>UVP-5</td>
<td>60–20,000 µm</td>
<td>60–2,000 µm (optical) 700–10,000 µm (imaging)</td>
<td>100–200 L per 5 m depth intervals</td>
<td>Cast</td>
<td>Fresh</td>
<td>In situ</td>
<td>6,000 m</td>
<td>Optical/Imaging</td>
<td>92,000 €</td>
<td>Hydroptic</td>
<td>Particles nb, type, size</td>
</tr>
<tr>
<td>UVP-6LP</td>
<td>60–20,000 µm</td>
<td>60–2,000 µm (optical) 500–10,000 µm (imaging)</td>
<td>-</td>
<td>Cast, gliders, Argo floats</td>
<td>Fresh</td>
<td>In situ</td>
<td>6,000 m</td>
<td>Optical/Imaging</td>
<td>15,000€</td>
<td>Hydroptic</td>
<td>Particles nb, type, size</td>
</tr>
</tbody>
</table>
| ISII-S     | 60 µm–130 mm          | 600 µm–130 mm (imaging)                 | 150 L/s at 5 kts     | Towed   | Fresh | In situ | 200 m | Focused shadowgraph | $100,000 | Bellamare-us.com | Particles nb, type, size |(Continued)
### TABLE 1 (Continued)

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Total size range (ESD)</th>
<th>Final results</th>
<th>Typical sample volume</th>
<th>Sampling</th>
<th>Condition of use</th>
<th>Max. operating range</th>
<th>Method Approximate cost</th>
<th>Seller Final results</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPICS</td>
<td>30 µm-20 mm</td>
<td>tbe</td>
<td>1 ml (10 m/s)</td>
<td>Fluids</td>
<td>In situ</td>
<td>1.000 – 5000 m</td>
<td>Imaging</td>
<td>Vision.com</td>
</tr>
<tr>
<td>VPR</td>
<td>30-50 µm</td>
<td>In situ</td>
<td>1.6-3.5 m</td>
<td>Towed</td>
<td></td>
<td>1,000 – 3,000 m</td>
<td>Imaging</td>
<td>ICO Systems</td>
</tr>
<tr>
<td>LOKI</td>
<td>25 µm-2 cm</td>
<td>In situ</td>
<td>1.6-3.5 m</td>
<td>Towed</td>
<td></td>
<td>1,000 or 3,000 m</td>
<td>Imaging</td>
<td>iSiTEC</td>
</tr>
</tbody>
</table>

3.4.2. Imaging Systems

#### 3.4.2.1. Imaging flow cytometry

Imaging flow cytometry (IFC) combines single-particle fluidics, optical characterization and the imaging of cells/colonies. The triggering of an image can be initiated based on different properties of the cell probed ahead of the camera (see Barteneva and Vorobjev, 2015 for a review). Four IFCs have been used routinely within aquatic research, the Imaging FlowCytobot® (Olson and Sosik, 2007; Sosik and Olson, 2007), the CytoSens/Cytobuoy® (Dubelaar et al., 1999), the FlowCam® (Flow Cytometer And Microscope; Sieracki et al., 1998 and the ZooCAM Colas et al., 2018, which differ in their approaches, outputs and size range (Table 1).

- The Imaging FlowCytobot records images for all particles above laser scattering and/or Chl a fluorescence trigger levels, with consistent image focus enabled by the hydrodynamic focusing principle of the conventional flow cytometric approach of sample injection into a sheath flow (the scattering and fluorescence data are preserved for further analysis). IFCB is a fully automated, submersible instrument with built-in design features (e.g., self-cleaning, onboard analysis of standard beads) that enable long (> 6 months) unattended deployments in the ocean. It is also routinely operated in flow-through systems during ship surveys.

- The CytoSense (and CytoSub which is a submersible version) measures FSC, SSC and multiple fluorescence signals. The scatter/fluorescence scans (pulse shape) show one-dimensional morphology and optical features of big cells, colonies, chains and filaments, converging to “normal” flow cytometry data (totals) for picoplankton. The optional camera makes bright field images of individual particles, hydrodynamically focused along their long axis by low shear acceleration in sheath fluid of particles. At high concentrations of particles which are detected at a fast rate (up to 10 k/s), all particles cannot be imaged and specific values of the optical scans are preselected to act as a trigger. Optical scattering and fluorescence output as well as imaging data are considered as the output of this instrument, while the combination of both properties will increase the classification efficiency.

- The FlowCam uses a similar imaging principle as the Imaging FlowCytobot (but lacks the hydrodynamic focusing provided by the sheath flow). Images are acquired either continuously (autotrigger mode) or after the detection of a fluorescent (Chl a) particle. FlowCam-nano and FlowCam-macro use only the autotrigger mode, respectively, to take pictures of smaller and larger objects and organisms.

- The ZooCAM uses an imaging principle similar to that of FlowCam-Macro.

Notably, depending on the trigger mode of these different IFC instruments, only some of the plankton may be characterized. When fluorescence triggering is used, organisms and particles with undetectable fluorescence will be missed (Reynolds et al., 2010). This does not apply if all organisms and particles are imaged by using scattering signals as an image trigger (e.g., “auto-trigger mode” for FlowCam, “unsupervised mode” for CytoSense).
3.4.2.2. Laboratory and in situ imaging systems

In order to obtain quantitative information on plankton > 100 µm, larger volumes of water need to be examined than is possible with IFCs. In situ imaging is non-destructive and can be combined with net sampling. However, there are numerous challenges. The most important criterion is the optimization of the trade-off between sensitivity, resolution, contrast and depth of field so that image quality allows taxonomic identification while the imaged volume is large enough for statistically relevant estimations of concentrations.

High magnification imaging at short distances results in a depth of field (DOF) of only a few millimeters and thus, in a high proportion of out-of-focus images not limited to the DOF (e.g., Schulz, 2013). To avoid motion blurring, short shutter speeds of a few microseconds are required (e.g., Davis et al., 1992; Schulz, 2013). Illumination adapted to the in situ and towing conditions should guarantee the image quality and high signal-to-noise ratio of the camera. Imaging systems can be adapted to illuminate a calibrated volume of water for more precise quantification of plankton and particles (Picheral et al., 2010). Due to the above-described trade-offs, in situ systems have a relatively restricted size range of operation and focus either on small size classes with a small volume imaged (e.g., CPICS, LOKI), where imaging of sizes less than a few millimeters with a high depth of field is close to the feasible border of the physical laws of optics (Schulz, 2013), or target larger fields of views with less details on organism morphology.

Another method to overcome the in situ constrains is to use imaging methods on net-collected plankton samples. In this case the conditions required for an optimal field of view and DOF may be met. However, net tows integrate the plankton over towing distances and can be intrusive, damaging some of the collected organisms. Plankton samples from nets may be imaged either by flow-through chambers (e.g., FlowCam-Macro, LOKI) or plankton scanners such as the ZooScan (Gorsky et al., 2010). Flow-through techniques limit the maximal size of organisms observable to the minimal diameter of its tubing. To avoid clogging, larger individuals are removed prior to analysis. Scanner-based approaches cannot be used in situ and are difficult to use at sea. Samples therefore have to be treated with fixatives that can modify the chemical composition of organisms, their color and, in some cases, their morphology. in situ imaging provides an alternative to study fragile taxa, such as gelatinous organisms, which may be damaged or destroyed by net tows (Remsen et al., 2004; Stemmann et al., 2008a).

In the following section, a selection of plankton imaging devices (other than the IFCs discussed above) that are commercially available, along with their imaging approach, are briefly introduced (more details are available in Table 1). As a general rule, information content on plankton increases with...
FIGURE 2 | Comparison of total planktonic organisms sampled with different collection methods and analyzed with different optical/imaging methods as a function of the size of organisms (expressed as equivalent spherical diameter; ESD). All sampling was done at the same station during Tara-Ocean Cruise at station 210 (Labrador Sea). Total organism biovolume per size classes were expressed as normalized biovolume size spectra (NBSS) by dividing the total biovolume within a size class by the biovolume interval of the considered size class. NBSS is representative of the number of organisms within a size class. Results were obtained by analyzing whole water with optical method (using an Accuri Flowcytometer) and imaging methods (Imaging FlowCyt Bot-IFC and Underwater Vision Profiler UVP-5) and plankton net samples of various mesh sizes using imaging methods (Flowcam with a 20X lens and Zooscan). All data are raw counts (removing only objects determined to be not living) and converted to biovolume using ellipsoidal calculations. The low count at the smaller size range of each observation corresponds to an underestimation of an object’s number due to both limited capabilities of each imaging device for small objects and net undersampling for small objects.

increasing volume and imaging frequency and with decreasing pixel-size. Additional desirable features are that the object be in focus with enhanced contrast and in color. Devices with a pixel size of > 50 µm are designed to study macro-plankton and large aggregates, while those with a pixel size of 5 µm can also resolve micro-plankton.

- The in situ Ichthyoplankton Imaging System (ISIIS; Cowen and Guigand, 2008) uses line scan cameras to produce shadowgraph images of plankton. The camera has a field of view of 13 × 13 cm with a depth-of-field of 50 cm and uses a 2,048 × 2,048 pixel sensor, resulting in a 63.5 µm pixel resolution. The system is equipped with environmental sensors, including CTD, fluorometer, dissolved oxygen, and PAR sensors. The volume resolution of the image captures a wide taxonomic range of mesozooplankton and with lower resolution, large protists and cyanobacteria (Luo et al., 2017). The data are transferred to an onboard computer via fiber optic in real time while the platform is being towed at 2.5 m s⁻¹, either undulating or at fixed-depth mode to a maximum depth of 150 m.

- The CPICS (Continuous Plankton Imaging and Classification Sensor, Grossmann et al., 2015) is based on a 6 Mpixel color camera and precisely aligned structured illumination to produce a darkfield image with high magnification (between 20X and 0.16X) that images respectively a small volume of seawater (between 2.6 × 10⁻⁵ µl and 34 ml depending on magnification) at high frequency [10 frames per second (fps)]. This instrument is deployed in situ (on rosette, ROV, AUV, glider or other autonomous vehicle, or buoy) with the possibility for real-time data link to vessel or shore.

- The Video Plankton Recorder II (VPR) is based on dark-field-illumination, operates a 1MPix (10 Bit b/w or color) camera (Davis et al., 2005) and images objects with 25 fps non-invasively and undisturbed in the water column. It images a small volume of seawater (app. 1 ml to 350 ml depending on calibration), and images are sent in real time onboard or shore via a fiber optic tow cable while the Digital Autonomous Video Plankton Recorder (DAVPR) is fully self contained.

- The Underwater Vision Profiler 5 HD (UVP, Picheral et al., 2010) operates a 4 MPix camera imaging a field of view of approximately 180 × 180 mm² about 200 mm in front of the camera. The UVP sizes marine snow, aggregates > 100µm and images plankton > 500 µm. It can be integrated on a CTD-Rosette system as a standard sensor delivering images indexed to the different environmental data collected at a rate of 20 images s⁻¹. Its depth range is 6,000 m. The UVP6-LP (low power) is a miniaturized and low power version which is designed to be deployed on ARGO float moorings, AUVs or gliders.
• The Lightframe On-sight Keyspecies Investigation (LOKI) system (Schulz et al., 2010) uses a flow-through chamber with an upstream plankton net. LOKI operates an industrial camera with up to 6 Mpxels at 15 µs shutter time, combined with a tailored high power LED flash unit to image a volume of approximately 20 × 20 × 5 mm³ in a flow-through chamber. Images of the LOKI system often allow the identification of fine morphological features as well as discrimination of developmental stages, sex and, in some cases, investigation of internal body structures (e.g., Schmid et al., 2018).

• In recent years two commercial systems have emerged which image particles using holography, which allows for 3-D reconstruction (see Sun et al., 2008) that can image microplankton. They take images at very short shutter times, scanning for a relatively large volume. The systems are the LISST-HOLO (now in its 2nd version, Sequoia Sci., 4.4 µm pixel size, 1,600 × 1,200 pixel frames, 20 fps) and the HoloSea (4Deep, 1.5 µm pixel size, 2,048 × 2,048 pixel frames, 22 fps). Given their novelty, only a small number of publications have been produced using them or their prototype systems (Bochdansky et al., 2017; Davies and Nepstad, 2017). Automated image reconstruction and recognition requires significant computational power, an issue that is likely to be alleviated in the near future.

• The ZooScan (Gorsky et al., 2010) is a plankton scanner which takes a high-resolution image of net- or bottle-collected plankton samples with a pixel size of 10.6 µm. It has built-in features making it possible to standardize the images of different ZooScans to remote control the image generation and to build a common image databases.

4. REQUIREMENTS FOR PLANKTON OBSERVING SYSTEMS

4.1. A Holistic View of Planktonic Communities

Our current vision of planktonic ecosystems is fragmented largely because of observational constraints. The concentration of organisms decreases with size, and therefore it is practically impossible to sample the full ecosystem, from sub-micrometer to meter-sized colonies of organisms, within a single sampling strategy or with a single methodology (the larger the organism, the larger the volume that needs to be sampled). Second, scientists often focus on a preferred taxonomic or functional range, leading to a kind of taxonomical blindness, with many details for some organisms and very few details for the ones not present in the taxonomical expertise of the specialist. The same is true for models (deYoung et al., 2004). We therefore recommend the adoption of a sustained sampling and analysis plan designed to cover the whole planktonic community, along with characterizing their environment, and incorporate ecosystem modeling in the design (e.g., Karsenti et al., 2011).

Few examples of such holistic studies exist (e.g., Waite et al., 2007; D’Alelio et al., 2015; Romagnan et al., 2015). Without such complete sampling of the planktonic community, the quantification of phytoplankton and zooplankton abundance or diversity, which are GOOS EOVs and GCOS ECV, could be biased.

4.2. Interfacing With Modeling Efforts

Marine plankton models are used to address a broad set of issues, ranging from basic science questions including those related to the dynamics of plankton blooms, plankton phenology and species succession (Bopp et al., 2005; Follows et al., 2007; Hashioka et al., 2013; Briseño-Avena et al., 2015; Kuhn et al., 2015), global nitrogen cycling (Jickells et al., 2017; Bianchi et al., 2018), zooplankton vertical migration (Bianchi et al., 2013), the generation of hypoxia (Fennel and Testa, 2019), the quantification of the biological carbon pump (DeVries et al., 2014; Laußkötter et al., 2016), climate change projections (Bopp et al., 2013) and effects of plankton dynamics on ecosystem services related to fisheries (Megrey et al., 2007; Rose et al., 2007; Lefort et al., 2015) to very applied purposes such as the provision of forecasts and reanalyses of oceanic physico-chemical and biogeochemical-relevant variables (see Fennel et al., 2019). However, most current global models represent only limited aspects of the functional diversity of plankton via so-called plankton functional types (Le Quéré et al., 2005) to simulate primary production within the euphotic zone, the transfer of biomass to higher trophic levels, and the pathways of biological carbon sequestration in the deep ocean.

Biogeochemical models typically lack detailed descriptions of zooplankton because of the lack of relevant observations (Buitenhuis et al., 2006), and also because the prevailing paradigm within the modeling community is that global biogeochemical cycles are primarily driven by bottom-up rather than top-down controls. This paradigm is increasingly being refuted by experimental evidence (Lima-Mendez et al., 2015; Guidi et al., 2016) and model studies, which emphasize the important role of zooplankton grazing on nutrient cycling and the biological carbon pump (e.g., Le Quéré et al., 2016), phytoplankton phenology (Hashioka et al., 2013; Ward et al., 2013a) and phytoplankton diversity (Prowe et al., 2012; Vallina et al., 2014). Complex population-structured zooplankton models exist (Hofmann and Ambler, 1988; Carlotti and Wolf, 1998), and models with multiple zooplankton functional groups are increasingly being developed (e.g., Ward, 2018). However, these models have not been widely applied because of the large number of poorly constrained parameters required to describe multiple functional groups and their life stages, physiological traits and ecological interactions (Anderson, 2005; Barton et al., 2013), and because modeling numerous species/stages is computationally expensive. Trait-based modeling frameworks that can simulate an unlimited number of plankton functional types by using allometric relationships between physiological and/or morphological traits (Ward et al., 2012; Ward, 2018) are promising in this regard, but zooplankton functional groups are still underrepresented because of the complexity and heterogeneity of this group, and because the small number of observational data sets available for model development and calibration remains problematic (Carlotti and Poggiale, 2010; Barton et al., 2013; Benedetti et al., 2015; Brun et al., 2016). Most global models now represent up to three size classes...
of zooplankton, albeit with fixed metabolic rates (e.g., Kishi et al., 2007; Saille et al., 2013). A further increase in model complexity may be warranted (Benedetti et al., 2018) since taxon- or functional group-specific trophic interactions with phytoplankton and diel vertical migration patterns (Bianchi et al., 2013) may be important to understanding ecosystem dynamics and global biogeochemical cycling (Guidi et al., 2016).

Furthermore, global models generally treat the mesopelagic and deep ocean as a black box because of a lack of ecological understanding of its microbial ecosystems and limited observations (Arístegui et al., 2009). In plankton models, organic matter is routed from phyto- and zooplankton to a few explicit particulate and dissolved detrital pools (Bopp et al., 2013; DeVries et al., 2014; Laukötter et al., 2016), sometimes accounting for their variable reactivity (Aumont et al., 2017; Bianchi et al., 2018), which then transport particulate organic matter from the surface to the deep ocean via sinking. However, an accurate simulation of nutrient and plankton dynamics in the euphotic zone does not necessarily imply an accurate representation of observed export fluxes (Bagniewski et al., 2011) or consensus on its major pathways within the marine ecosystem (Laukötter et al., 2016). Marine particle fluxes are known to display strong regional and temporal variability in response to different production regimes and their seasonality, or to the presence of OMZ (Haake et al., 1992; Van Mooy et al., 2002; Guidi et al., 2015), but this variability is not yet well represented in models. Better representations of particle dynamics (Aumont et al., 2017) and particle-plankton interactions may improve simulation of the important mechanisms governing global biogeochemical cycles of carbon and other essential elements, and may thus lead to improved model projections of ocean carbon cycling (Kriest and Evans, 2000; Gehlen et al., 2006), but carbon pathways in models still depend strongly on the data sets used for optimization (Bisson et al., 2018) and the parameterization of their ecosystem modules (Laukötter et al., 2016).

Robust marine ecosystem/ocean biogeochemistry models that relate climate change to fish production or relating changes in the strength of the biological pump to changes in plankton community structure require an adequate description of phytoplankton, zooplankton and particle compartments in the upper kilometers of the ocean and the mesopelagic, as well as consensus on the major pathways of organic matter transport and transformation, i.e., the fluxes between them. The optimal level of complexity of global models for each biogeochemical or ecological application remains to be determined, and may vary according to the specific target application (Ward et al., 2013b). Therefore, the acquisition of quality-controlled, standardized, global in situ data is essential for the development and validation of mechanistic end-to-end models that optimize the balance between fidelity and simplicity for the continuum of the plankton food web, and within the entire water column.

In order to be useful for marine plankton models targeted for biogeochemical and fisheries applications, observation development should provide robust, global information on major taxa (concentration as well as relevant rates), the size distribution of the major plankton groups, as well as the particle size distribution at a biologically relevant spatiotemporal resolution (Capotondi et al., 2019; Fennel et al., 2019). Imaging and other techniques detailed above provide such information even for the deep ocean based on data from in situ devices. Yet many observational data sets still lack robust quantification of their uncertainty range, which is essential for model calibration, and some observations are still of limited use for model applications due to the lack of a common set of comparable standard products. A close collaboration between modelers and experimentalists can guarantee the usefulness of data products for this important user group, and thus a better quantification of the present and future biogeochemical functioning of the ocean and health of its ecosystem (Siegel et al., 2014).

### 4.3. Achieving a Globally Consistent and Holistic Plankton Observing System:

Given the TRL of the technologies for the observation of plankton biomass and diversity reviewed above, these types of technologies are ready to be included in a global observation network (Miloslavich et al., 2018).

As a first step, integration of plankton measurements into existing global observing systems (e.g., GO-SHIP Sloyan et al., 2019, OCEAN-SITES, BGC-ARGO Roemmich et al., 2019) could serve as the vehicle to ensure wider integration in the future. Observing systems are international and distributed (i.e., do not depend on a single nation), therefore requiring a high level of international coordination and standardization. Such efforts could have a direct benefit for smaller scale observing systems if the lessons learned and data infrastructure were shared with them.

Ocean sites where time series observations are conducted (Benway et al., 2019) provide natural foci for experimental process studies. If possible, observational technologies should be augmented with measurements of functional interactions and rate processes (e.g., primary and net community production, growth and grazing rates, predator-prey interactions, export fluxes) since understanding such processes is critical to linking ocean observations with functional models of biogeochemistry and ocean food webs.

The observations should follow basic procedures in order to guarantee the best comparable outputs for present and future studies:

1. **Methods** should be standardized across the whole observing system and should be validated outside of the lab which developed them. Where necessary, protocols of best practice should be produced and made widely available (e.g., OceanBestPractices, 2018). Standards for each measurement should be defined so that all related measurements can be quantitatively compared with respect to the standard. Standardization of methods should include standardizing the methodology for quality control and of data curation.

2. **Plankton samples** should be associated with environmental variables acquired simultaneously (e.g., from remote and local sensors). More information about the environment as well as the plankton (e.g., using different sensing and sampling...
3. Sampling should include as complete a spectrum of plankton as possible, spanning size and function, across a variety of places, depths and times. De-correlation scales in time and space should be taken into consideration to maximize the use of resources. Sustainable observations at specific locations and along transects enable the quantification of temporal trends.

4. Methods should be cross-compared (e.g., nets, CPR, imaging systems) and inter-calibrated to ensure that uncertainties and potential bias are known. When using proxies for particles and plankton properties (e.g., bio-optics and bio-acoustics), periodic ground-truth ensures biases are constrained. Measurements of related/complementary variables (e.g., pigment concentration, POC and phytoplankton volume from FCMs and IFCs of surface samples, bio-acoustic and quantitative imaging) can be used to point out anomalous data for flagging if the different measurements are not consistent with established relationships.

5. Protocols for adopting new technology should be in place (e.g., how long should there be side-by-side deployment and inter-comparison before replacement of old with new). These documents should identify the advantages and limitations of the measurements to provide realistic quantified uncertainties. There are a few existing inter-comparisons between existing sampling devices (notably between nets, e.g., Stehle et al., 2007 and references therein) which have been conducted, with few exceptions, years after the introduction of the new technology. Similarly, imaging techniques have seldom been inter-compared and calibrated with field-specific standards (e.g., Colas et al., 2018), and the inter-comparison has often taken place long after the introduction of a new device (e.g., Schultes and Lopes, 2009; Reynolds et al., 2010; García-Comas et al., 2011; Forest et al., 2012; Thyssen et al., 2014; Le Bourg et al., 2015).

6. Because of the potential improvement of techniques and changes in scientific interests, it is very important that physical plankton samples obtained with nets or water samples be properly archived for future reanalysis. Similarly, as image analysis methods improve, images (or optical properties) could be reanalyzed at later dates.

5. CHALLENGES

5.1. Deciding on What to Measure With a Finite Budget

There is a need to develop a strategy regarding the methods and measurements to be done. Prioritization should be based on: (1) cost (e.g., human capital needed, cost of analysis and/or instrument, ability to take advantage of existing already-funded efforts), (2) central variables in the context of global ecosystems (EOVs) and biogeochemical processes and (3) associated ecological information content. Emphasis should be placed on a holistic sampling program with sufficient redundancy to ensure success and reduce uncertainties. It follows that scientists versed in modeling, observations, data mining and marine resource management should be consulted to optimize this strategy, which will likely involve both in situ observations as well as samples for laboratory analysis.

5.2. Sampling Design and Constraints

Counting statistics indicate that, to treat a population as continuous, about 400 individual particles need to be enumerated in the volume analyzed (Siegol, 1998). This requirement means that while the distribution of bacteria can be assessed with samples as small as a milliliter, to sample macro-zooplankton, hundreds of liters may be necessary. Spatial heterogeneity further increases the number of required samples. This constraint also means that the sampling volume will increase dramatically with the desire to obtain more resolved taxonomic information, particularly for larger organisms. Plankton nets provide means to sample large volumes (and hence large number of individuals) while selecting for specific groups that are well sampled within nets and integrating vertical and horizontal gradients. Imaging systems, while better suited for targeted spatial sampling and fragile organisms, need to be deployed in such a way as to obtain sufficient numbers of individuals and to minimize avoidance of in situ instruments by the targeted organisms (e.g., minimal physical perturbation to the environment, appropriate lighting to avoid biasing observations especially at night or at deep depth).

5.3. Managing and Integrating the Massive Data Flow Originating From a Network of Different Instruments

An integrated observation system will be useful and relevant only if the data collected are available and meet user needs. Data dissemination platforms should comply with common standards of Findability, Accessibility, Interoperability, and Reusability (FAIR; Wilkinson et al., 2016). At the moment plankton-relevant oceanographic data are scattered within different databases (e.g., OBIS, COPEPOD, PANGAEA, IGMETS, SeaBASS, BCO-DMO, EMODNET), which are data repositories of variable ease of use. Each of these data portals offers only a partial view of the ecosystem, focusing on very specific features (e.g., species presence for OBIS, zoo- and phyto-plankton biomass for COPEPOD, pigments and optical parameters in SeaBASS) and often lacking links to contextual data originating from the same sampling event or cruise. For a modeler interested in constructing a holistic view of the ecosystem during a particular field effort, they have to invest significant effort to find and access to relevant data in all the different repositories. Linking between them could significantly facilitate more complete exploitation of the data (Benway et al., 2019; Tanhua et al., 2019b). Indeed, OBIS has recently adopted the Event Core format of Darwin Core and developed the Extended Measurement or Fact Extension, enabling linking sampling facts including environmental measurements to an event hierarchy and biotic measurements (e.g., biomass, absence/presence, fatty acids, pigments) to the occurrence records (De Pooter et al., 2017). Another important limitation is the lack of uncertainties associated with the data in most data repositories as well as the lack of defined quality control annotations. Hence, the users...
are often left with the need to develop their own uncertainty estimates, for example to propagate in their calculations or to test model sensitivity. Beyond the necessity to integrate the data sets, users of such databases should be regularly consulted to ensure they are fit for purpose.

5.4. Curating, Validating and Distributing Imaging Data

Imaging systems deployed for oceanographic research worldwide already collect millions of images a year. Large-scale infrastructure is therefore needed to host and distribute this wealth of data. Such infrastructure should provide a collaborative way to visualize and classify images, perform quality control on these identifications, and share the resulting data in an open-access manner.

A prototype infrastructure for quantitative imaging datasets (EcoTaxa) has been developed. EcoTaxa is a web application (http://ecotaxa.obs-vlfr.fr) that allows users to store images of individual organisms and associated metadata in its database, efficiently classify these objects within a universal taxonomic reference, and export the resulting data for further analyses. It uses machine learning, combining classical approaches and Convolutional Neural Networks in a user-friendly way, to help ecologists, even those with no computer-science background, classify large numbers of images (typically >10,000/operator/day). It is meant to be collaborative, allowing an unlimited number of users to interact on the same dataset (all using the same taxonomic tree). Their work can be iterative (by correcting each other’s mistakes) while always retaining the full history of identification for each object, and machine learning models can be built based on all classified images in the database.

Currently (as of Jan 25th 2019), EcoTaxa hosts >72 million images of plankton (42% of which have had their identification validated by a human operator), collected with more than seven different instruments, over the world's oceans. In 3 years, it attracted ~550 registered users from ~160 institutions. While functional, this prototype needs to be deployed more widely, on servers backed by academic institutions, and improved in various ways. For example, the database may not perform as well with billions rather than the current millions of images. The taxonomic back-end will benefit from the unification of phylogenies allowed by nucleic acid sequencing (e.g., https://unieuk.org). Finally, the machine learning back-end should benefit from various ongoing efforts to infuse more computer science knowledge into ocean sciences. The available database of images has already proved to be a good tool to foster such initiatives by providing a wealth of classified data (Elineau et al., 2018). Efforts are partly funded and underway to address the limitations outlined above.

EcoTaxa also includes a 'Particle module' (http://ecotaxa.obs-vlfr.fr/part/) intended to store, visualize, and export data originating from in situ instruments that quantify marine snow (UVP, LISST, LOPC). In this application, all data originating from the same sampling event (CTD data, marine snow abundance, and plankton identified from images) are gathered in the same dataset and can be downloaded by visitors. Data from any device counting and measuring particles can be easily integrated into this module.

5.5. Near Real-Time Data Processing

For certain applications (e.g., adaptive sampling of deployed assets), near-real-time data processing and quality control is critical. Much effort has been invested in BGC-Argo to process the biogeochemical data in near-real-time (including flagging data that are likely problematic, Roemmich et al., 2019). Similar efforts have also been invested on developing such strategies for fixed infrastructures (e.g., IOOS, 2018). Global coordination of such efforts and the addition into them of more plankton-relevant parameters will ensure optimal use of these data (e.g., Tanhua et al., 2019a). In addition, delayed-mode processing is often needed for plankton variables (e.g., once water samples needed to calibrate proxies are available).

5.6. Enhancing Capacity and Knowledge

The community with the knowledge necessary to collect, quality control, analyze and interpret plankton data is small, and the current assets to sample the world's oceans are extremely limited given the task at hand (to monitor the distribution of plankton-relevant variables in time and space throughout the world's oceans). In addition, taxonomic expertise has declined, which severely limits many types of investigations. Significant efforts in education, collaboration (sharing of expertise, computer codes, instruments, annotated image databases, etc.) and invention of sensors to be deployed for longer periods of time on autonomous platforms can help mitigate this limitation. In addition, the creation and adoption of best practices can be an important element in training to increase the uniformity and interoperability of measurements (Pearlman et al., 2018). Finally, plankton ecology needs to develop further to take full advantage of the measurements collected.

5.7. Maximizing Return on Investment

To ensure sustained and continuing investment in plankton monitoring, it is critical to enlarge the community of active users of such data. An approach could include the generation of simple indicators summarizing the complex ecological data (simple and user friendly but not simplistic) and will require the education of users through virtual tools, workshops, summer schools, and capacity building to include potential user communities and stakeholders (e.g., from modeling, resource management, environmental agencies and private companies).

6. RECOMMENDATIONS ON HOW TO PROCEED DURING THE NEXT TEN YEARS

Building a global observation system to describe planktonic key variables and predict the functioning of the pelagic ecosystems requires a stepwise approach with regional-scale experiments as pilot projects while engaging with existing global programs and infrastructures to increase their sampling capacity. Such pilot studies, combining in situ sensors deployed on long-endurance platforms with satellite sensors, ship cruises and in conjunction with data-assimilating biogeochemical-ecological
models, will provide a scalable template from which to grow and improve upon. Acquisition of biological variables should be performed in a consistent framework to facilitate inter-comparison between methods and projects. To ensure quality of data, adherence to agreed upon best practice protocols, including specific actions such as inter-calibration, is necessary. We suggest that the following itemized strategies should be used to prioritize investments and provide below each some example activities (some may contribute to more than one priority):

6.1. Priority I: Make the Best of Existing Data, Share Publicly, Inter-calibrate/inter-compare Existing Observations and Technologies, Work on Common Protocols, Improve Accessibility of Existing Databases and Searchability of Data at Multiple Levels of Organization

Promote free data and information sharing by using open access publication strategies of both articles and source data and following a "FAIR" principle (Wilkinson et al., 2016). Build robust distributed networks for collection, distribution and curation of data (like Argo and EcoTaxa) that do not depend on one country's funding and that serve the full scientific community worldwide (see Tanhua et al., 2019b). Channel the data to global and consistent public databases such as OBIS or other existing platforms in consultation with modelers. Make sure funding agencies that require data curation from funded PI’s are consulted and in agreement with adopted approach.

Ensure there are experts available to assist the larger community with quality assessment and control (QA/QC) of plankton-relevant data collected by non-specialists. For imaging in particular, ensure there is ready access to taxonomic information/expertise, and find ways to reward people who serve the community (e.g., encourage data publication with citable DOI for databases). Make lists and databases of taxa at different organization levels (for example, one with 5 taxa that are well identified and one with few tens of taxa when automatic sorting is followed by detailed human annotation).

Work on best practice documents to ensure worldwide methodology is consistent (Pearlman et al., 2018) and interchangeable. Encourage the contributions of documented best practices to a global scale repository to facilitate consensus and adaption of common methodologies.

6.2. Priority II: Generate Novel Data in a Reasonable Way, i.e., Using Common Global Standards for Data Generation, Taxonomic Identification, Quantification of Uncertainty, Comparing Against Standards

Invest in efforts to homogenize plankton-relevant variables between national and international programs, and to inter-compare and inter-calibrate methods to ensure measurements can be integrated into a global observing system and modeling framework.

Invest in the integration of tried-and-true technologies for plankton measurements on globally coordinated programs collecting time-series and transects which currently do not measure these systematically (Weller et al., 2019). Proceed with this integration incrementally - e.g., start with one GO-SHIP line that is occupied annually and use it as an example, and as the basis for writing the appropriate protocols and the expansion to other transects and cruises. Invest in the integration of mature and calibrated sensors to AUVs to expand coverage in space and time. Take into account in planning that the synergistic value of colocated measurements, providing significantly more information than the sum of the individual measurements.

6.3. Priority III: Think About Collaboration, Co-funding and Joint Projects Between Different User Groups

Closely interact with modelers to derive and produce outputs/indices relevant to the models. Organize workshops with modelers to explain the added value of the new measurements, and interact closely with modelers to identify their needs (e.g., to ensure measurements target what models are sensitive to) and make them aware of observation efforts. Apply for co-funding for experimental and modeling studies to encourage co-involvement of both fields in the initial planning and final use of the results (e.g., Tara expeditions).

Organize summer schools and workshops, dedicated to students, early career scientists, senior scientists, but also stakeholders and policy makers, to ensure transfer of theoretical (plankton ecology, diversity and taxonomy) and technical knowledge, consistency in processing (e.g., image annotation), dissemination, and increased use of data collected. Document the material provided through the summer school in a repository (e.g., videos of classes, PDFs of materials, etc.) to support long-term opportunities for training.

Promote the exchange of expertise between northern and southern hemisphere countries by disseminating the existing expertise on plankton imaging/optical measurements to scientists and students from less-developed countries, particularly in the southern hemisphere, where the largest gaps in our knowledge still persist, but where sometimes taxonomic expertise is still strong when it has been progressively lost in northern countries.

6.4. Priority IV: Innovate and Develop New and Better Technologies for the Development of Augmented Observations

Keep innovating! Keep investing in the development of new sensor/sensing, of new analytic tools and pipelines on the use of automated optical/imaging techniques for building new biodiversity indicators.

Take advantage of developments from other fields (e.g., computer science, optical engineering, etc.) to improve our current sensing systems. Look for opportunities to reduce the costs of measurements without decreasing their quality (e.g., Wang et al., 2019), in such a way that future equipment could be
embarked with as standard tools on cruises and/or autonomous platforms (floats, gliders).

**AUTHOR CONTRIBUTIONS**

FL and EB coordinated and contributed equally to the writing of the manuscript. AMW, MV, JU, LS, HMS, JS, J-BR, MP, JP, MDO, BN, KOM, PM, AL-L, RKi, RML, RKu, LK-B, JSJ, MHI, J-OI, HH, LG, GG, SLG, PG, SG, KE, GD, RKC, FC, CB-A, LB, KB-B, NB, SB, SDA, LFA, and WA provided input on the full manuscript.

**ACKNOWLEDGMENTS**

 Much of this manuscript flows from discussions of the authors with the members of SCOR working groups 150 (TOMCAT) and 154 (P-OBS) as well as discussions with the greater community in various GOOS workshops. We also thank Mike

**REFERENCES**


Lombard et al. Quantitative Observations of Plankton Ecosystem


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.