Sensing the ocean biological carbon pump from space: A review of capabilities, concepts, research gaps and future developments

Robert J.W. Brewin a,b,*, Shubha Sathyendranath b,c, Trevor Platt b, Heather Bouman d, Stefano Ciavatta b,c, Giorgio Dall’Olmo b,c, James Dingle b, Steve Groom b, Bror Jönsson b, Tihomir S. Kostadinov e, Gemma Kulk b, Marko Laine h, Victor Martínez-Vicente h, Stella Psarra g, Dionyssios E. Raitsos h, Katherine Richardson i, Marie-Hélène Rio j, Cécile S. Rousseaux k,l, Joe Salisbury m, Jamie D. Shutler a, Peter Walker b

a Centre for Geography and Environmental Science, College of Life and Environmental Sciences, University of Exeter, Penryn, Cornwall, United Kingdom
b Plymouth Marine Laboratory, Plymouth, Devon, United Kingdom
c National Centre for Earth Observation, Plymouth Marine Laboratory, Plymouth, Devon, United Kingdom
d Department of Earth Sciences, University of Oxford, Oxford, United Kingdom
e Department of Liberal Studies, California State University San Marcos, San Marcos, CA, USA
f Finnish Meteorological Institute, Helsinki, Finland
g Hellenic Centre for Marine Research, Crete, Greece
h Department of Biology, National and Kapodistrian University of Athens, Athens, Greece
i Globe Institute, University of Copenhagen, Denmark
j European Space Agency, European Space Research Institute (ESRIN), Frascati, Italy
k Universities Space Research Association, Columbia, MD, USA
l Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA
m College of Engineering and Physical Sciences, University of New Hampshire, Durham, NH, USA

ARTICLE INFO

Keywords:
Ocean
Carbon cycle
Satellite
Biology

ABSTRACT

The element carbon plays a central role in climate and life on Earth. It is capable of moving among the geosphere, cryosphere, atmosphere, biosphere and hydrosphere. This flow of carbon is referred to as the Earth’s carbon cycle. It is also intimately linked to the cycling of other elements and compounds. The ocean plays a fundamental role in Earth’s carbon cycle, helping to regulate atmospheric CO2 concentration. The ocean biological carbon pump (OBCP), defined as a set of processes that transfer organic carbon from the surface to the deep ocean, is at the heart of the ocean carbon cycle. Monitoring the OBCP is critical to understanding how the Earth’s carbon cycle is changing. At present, satellite remote sensing is the only tool available for viewing the entire surface ocean at high temporal and spatial scales. In this paper, we review methods for monitoring the OBCP with a focus on satellites. We begin by providing an overview of the OBCP, defining and describing the pools of carbon in the ocean, and the processes controlling fluxes of carbon between the pools, from the surface to the deep ocean, and among ocean, land and atmosphere. We then examine how field measurements, from ship and autonomous platforms, complement satellite observations, provide validation points for satellite products and lead to a more complete view of the OBCP than would be possible from satellite observations alone. A thorough analysis is then provided on methods used for monitoring the OBCP from satellite platforms, covering current capabilities, concepts and gaps, and the requirement for uncertainties in satellite products. We finish by discussing the potential for producing a satellite-based carbon budget for the oceans, the advantages of integrating satellite-based observations with ecosystem models and field measurements, and future opportunities in space, all with a view towards bringing satellite observations into the limelight of ocean carbon research.

* Corresponding author at: Centre for Geography and Environmental Science, College of Life and Environmental Sciences, University of Exeter, Penryn, Cornwall, United Kingdom.
E-mail address: r.brewin@exeter.ac.uk (R.J.W. Brewin).

https://doi.org/10.1016/j.earscirev.2021.103604
Received 21 July 2020; Received in revised form 16 February 2021; Accepted 13 March 2021
Available online 19 March 2021
0012-8252/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Introduction

The ocean biological carbon pump (OBCP) can be defined as a suite of biological, physical, and chemical processes that contributes to, and controls, the transfer of organic carbon (in dissolved and particulate forms), and calcium carbonate, from the surface layer to the deep ocean (Volk and Hoffert, 1985). The magnitude of this pump has been estimated to be between 4 and 12 Pg C year\(^{-1}\) (Laws et al., 2000; Henson et al., 2011; DeVries and Weber, 2017). The OBCP has a profound impact on the functioning of the ocean and the planet. It helps modulate the CO\(_2\) concentration and pH of the ocean, which together with physical processes, impacts the transfer of CO\(_2\) between the ocean and atmosphere. If the OBCP were to be switched off, it has been estimated that atmospheric CO\(_2\) concentration would be 50% higher than it is today (Parekh et al., 2006), and consequently, we would be living in a very different world. The carbon exported to the deep ocean through the OBCP supports meso-, bathy- and abysso-pelagic food webs, and controls the vertical and horizontal distribution of elements in the ocean. Studying and monitoring the OBCP is thus of paramount importance for understanding and predicting environmental and ecological changes at the planetary scale.

There are a variety of techniques we use to monitor the OBCP, which can be deployed from various platforms such as ships, autonomous vehicles, and satellites. In this paper, we review these tools with a focus on satellite-based methods. We highlight current capabilities in space-based OBCP monitoring, identifying the components of the OBCP not amenable to observations from space, and how we can fill these gaps. We discuss the future of satellite-based OBCP monitoring, and how satellites, monitoring electromagnetic radiation at different frequencies and resolutions, can be used in synergy with other monitoring tools (e.g. ships-based and autonomous-based) and ecosystem models, to advance our understanding of the OBCP and ultimately the global carbon cycle.

2. Overview of the Ocean Carbon Cycle

2.1. The Ocean Biological Carbon Pump (OBCP)

The OBCP depends on a number of key pools, components and processes that influence its functioning (Fig. 1). There are four main pools of carbon in the ocean.

- **Dissolved Inorganic Carbon (DIC)** is the largest pool. It constitutes around 38,000 Pg C (Hedges, 1992) and includes: dissolved carbon dioxide (CO\(_2\)); bicarbonate (HCO\(_3\)-); carbonate (CO\(_3^{2-}\)); and carbonic acid (H\(_2\)CO\(_3\)). The equilibrium between carbonic acid and carbonic acid determines the pH of the seawater. Carbon dioxide dissolves easily in water and its solubility is inversely related to temperature. Dissolved CO\(_2\) is taken up in the process of photosynthesis, and can reduce the partial pressure of CO\(_2\) in the seawater, favouring drawdown from the atmosphere. The reverse process respiration, releases CO\(_2\) back into the water, can increase partial pressure of CO\(_2\) in the seawater, favouring release back to the atmosphere. The formation of calcium carbonate by organisms such as coccolithophores has the effect of releasing CO\(_2\) into the water (Zeebe and Wolf-Gladrow, 2001; Rost and Riebesell, 2004; Zeebe, 2012).

- **Dissolved Organic Carbon (DOC)** is the next largest pool at around 662 Pg C (Hansell and Carlson, 2013). DOC can be classified according to its reactivity as refractory, semi-labile or labile. The labile pool constitutes around 0.2 Pg C, is bioavailable, and has a high production rate (~15–25 Pg C year\(^{-1}\); Hansell, 2013). The refractory component is the biggest pool (~642 Pg C ± 32; Hansell and Carlson, 2013), but has a very low turnover rate (0.043 Pg C year\(^{-1}\); Hansell, 2013). The turnover time for refractory DOC is thought to be greater than 1000 years (Williams and Druehl, 1987; Druffel et al., 2016).

- **Particulate Organic Carbon (POC)** constitutes around 2.3 Pg C (Stramska and Gieyszor, 2015; CEOS, 2014) and is relatively small compared with DIC and DOC. Though small in size, this pool is highly dynamic, having the highest turnover rate of any organic carbon pool on the planet (Sarmiento and Gruber, 2006). Driven by primary...
production, it produces around 50 Pg C y$^{-1}$ globally (Longhurst et al., 1995; Sathyendranath et al., 2019b; Kulk et al., 2020). It can be separated into living (e.g. phytoplankton, zooplankton, bacteria) and non-living (e.g. detritus) material. Of these, the phytoplankton carbon is particularly important, because of its role in marine primary production, and also because it serves as the food resource for all the larger organisms in the pelagic ecosystem.

### Particulate Inorganic Carbon (PIC) is the smallest of the pools at around 0.03 Pg C (Hopkins et al., 2019). It is present in the form of calcium carbonate (CaCO$_3$) in particulate form, and impacts the carbonate system and pH of the seawater. Estimates for PIC production are in the region of 0.8–1.4 Pg C y$^{-1}$, with at least 65% of it being dissolved in the upper water column, the rest contributing to deep sediments (Feely et al., 2004). Coccolithophores and foraminifera are estimated to be the dominant sources of PIC in the open ocean (Schiebel, 2002; Feely et al., 2004). The PIC pool is of particular importance due to its role in the ocean carbonate system, and in facilitating the export of carbon to the deep ocean through the carbonate pump, whereby PIC is exported out of the photic zone and deposited in the bottom sediments (Riebesell et al., 2000).

There are a series of key processes that determine the fluxes of carbon between these pools (Fig. 1).

- **Primary production** converts DIC, the form of CO$_2$, to POC, in the form of phytoplankton tissue, through the processes of photosynthesis. Light and nutrients are required for photosynthesis by autotrophic phytoplankton. Primary production is the engine behind the OBCP and it occurs within the euphotic zone (the region of the surface ocean that receives enough light for photosynthesis to occur). Gross primary production represents total organic carbon production irrespective of respiration. Net primary production is defined as gross primary production minus autotrophic respiration. In the ocean net primary production is equivalent to that in the terrestrial environment (~50 Pg C y$^{-1}$; Field et al., 1998).

- **Grazing** by heterotrophic organisms. The carbon produced by primary production can be transferred from phytoplankton to other POC components (heterotrophic organisms, such as zooplankton and nekton), through the process of grazing. This can occur on various time scales, for example, in the surface ocean soon after primary production, or at deeper depths long after primary production, following the sinking of phytoplankton. For simplicity, in this review we consider consumption (which occurs when all organisms take up carbon) under the same category as grazing. The fraction of organic carbon from primary production consumed by organisms becomes smaller as you move up the food chain, from herbivores (secondary production) through to carnivorous organisms (tertiary production), such as fish and even humans.

- **Secondary production** refers to the growth rate of herbivores. Only a very small fraction of the carbon grazed by herbivores is used for growth. The ratio of secondary production to primary production is used as an index of the efficiency with which organic carbon is transferred up the marine food chain (Sigman and Hain, 2012).

- **Respiration** is the conversion of POC, in the form of bacteria, phytoplankton, zooplankton and other higher trophic levels (or labile DOC), to DIC, in the form of CO$_2$. It is the principal mechanism of remineralisation in the ocean. This process is the reverse of photosynthesis, results in the consumption of oxygen and can occur at various depths.

- **Excretion** is the release of metabolic waste by marine organisms (Wiebe and Smith, 1977). The majority of carbon released by excretion is in the form of DOC.

- **Aggregation** is the joining together of different components into a cluster. It is controlled in part by coagulation and flocculation and influences the sinking rate of particles, which is critical to the export of material from surface to deep waters (Riley, 1963; Burd and Jackson, 2009).

- **Fragmentation** is the breaking apart of larger particles (or aggregates) into smaller components. This can occur through physical (e.g. shear stress) or biological (e.g. grazing, sloppy feeding) processes. It has recently been demonstrated that fragmentation is the primary process controlling the downhill flux of POC through the mesopelagic zone (Briggs et al., 2020).

- **Non-predatory mortality** is death of phytoplankton and zooplankton through non-predatory means. This can be caused, for example, by cell lysis or viral infection (Kirchman, 1999). Other causes include senescence, temperature change, light exposure, physical and chemical stresses, parasitism and food-related death (Tang et al., 2014).

- **Solubility**. The solubility of CO$_2$ in the ocean is primarily dependent on seawater temperature (Takahashi et al., 2002). Cooler, deeper water can store larger amounts of dissolved CO$_2$ in the form of DIC than warmer surface water.

- **Calcification**. Marine organisms, notably coccolithophores, and to a lesser extent foraminifera and pteropods, produce calcium-carbonate particles, which can sink out the upper ocean and settle at the bottom of the sea. The chemical processes involved in the formation of calcium carbonate release CO$_2$ into the water (Zeebe, 2012).

The OBCP by definition is concerned primarily with the transfer of organic carbon from the surface layer of the ocean (epipelagic or euphotic zone) to the deep ocean (mesopelagic and below). This can be achieved through three main routes, or pathways (Fig. 1, also see Boyd et al. (2019) and Le Moigne (2019) for recent reviews on the topic).

- **Physical transport**. Carbon can be transported vertically through physical processes. These may include: subduction by large-scale ocean circulation (e.g. see Lévy et al., 2013; eddy-driven subduction (Lévy et al., 2001; Omand et al., 2015; Llort et al., 2018; Respland et al., 2019); and the detrainment of carbon due to fluctuations in the mixed-layer depth (Stramska, 2010; Dall’Olmo et al., 2016; Lacour et al., 2019).

- **Sinking of particles** (gravitational pump). Particles can sink from the epipelagic to the mesopelagic zone (Sanders et al., 2014; Cael and Bisson, 2018; Bisson et al., 2020). The gravitational sinking of POC has been closely related to the spring phytoplankton bloom at high latitudes (Martin et al., 2011). The particles can sink in the form of dead phytoplankton cells, in the form of aggregates of particles, and in the form of faecal pellets. The sinking rates of the particles are dependent on their size, density and morphology (Stemmann and Boss, 2012).

- **Migration**. Organisms capable of vertical migration (e.g. zooplankton and carnivorous organisms) can transfer carbon from the epipelagic to the mesopelagic (Longhurst et al., 1990; Steinberg and Landry, 2017). Diel zooplankton migration is believed to be a large component of this flux. During the night, zooplankton migrate from the mesopelagic to the epipelagic to graze on phytoplankton. During the day, they return back to the mesopelagic (partly to avoid predation) and synthesise this carbon, resulting in a net downward flux of organic carbon. The vertical transport of carbon from migration can also occur on longer time-scales, for example, through the hibernation of copepods (and in some cases mortality at the hibernation depth) in winter in the mesopelagic zone, referred to as the seasonal-lipid pump (Jonasdóttir et al., 2015).

Once the carbon reaches the mesopelagic zone, its fate is dependent on which pool it belongs to, and on the extent to which the various processes control it (Fig. 1). Part of the carbon exported to the mesopelagic zone continues to sink and can contribute to sequestration, where it is consequently stored in the deep ocean for long timescales. There are three production terms that often feature in descriptions of the
2. Interactions with the atmosphere and land

The ocean surface is constantly exchanging carbon dioxide (CO₂) with the atmosphere. It has been estimated that, since the start of the industrial revolution, approximately a quarter of anthropogenically emitted CO₂ has been taken up by the surface ocean (Friedlingstein et al., 2019). This uptake varies spatially, seasonally, and on interannual scales (Watson et al., 2009; Wanninkhof et al., 2013; Landschützer et al., 2016; Gruber et al., 2019). The exchange of carbon between the atmosphere and ocean is dependent on concentrations in the atmosphere and ocean and turbulent exchange between the two (Woolf et al., 2016). The concentrations (combination of temperature and partial pressure of CO₂ in the water and atmosphere dictate the gradient of transfer, and any CO₂ gas transfer across the interface varies with environmental conditions and surface water constituents. The in-water CO₂ concentration can then be modulated by biological uptake or respiration. Once within the surface ocean, the gaseous CO₂ reacts with the seawater and becomes part of the buffered inorganic carbonate system comprised of aqueous CO₂, carbonic acid, bicarbonate and carbonate, and the sum of these different species is the total inorganic carbon, often called the dissolved inorganic carbon (DIC). The combination of species that determine the capacity of the water to resist changes in pH is referred to as the total alkalinity (AT). Satellite data are critical for quantifying synoptic-scale ocean gas fluxes, and for a recent review of these capabilities along with the use of satellite data for studying the surface inorganic carbonate system, the reader is referred to the work of Shutler et al. (2019).

There are also exchanges in carbon between the ocean and the land. Riverine outflow can transfer carbon (DIC, DOC, POC and PIC) produced or stored on land to the ocean (Tranvik et al., 2009). This outflow also contains nutrients that can support new production in coastal and open ocean waters (Seitzinger et al., 2010). Current estimates suggest that 0.455 to 0.78 Pg C enters the ocean each year due to rivers (Jacobson et al., 2020). It is often used as a measure of the net amount of carbon removed from the atmosphere through the OBCP. Note that, at certain time scales at which steady state may be assumed to prevail, and for appropriate depth scales, new production and export production may be equivalent to each other.

3. Field observations of carbon pools and fluxes in the ocean

While this review focuses primarily on satellite-based techniques for monitoring the OBCP, their development is highly dependent on data collected in the ocean. Satellite products depend heavily on precise in situ measurements to test algorithms and validate products. Field measurements are required to fill gaps in satellite products, whether it be through observations of pools and fluxes that are not currently accessible to remote sensing, or through observations at depth. This section provides an overview of key tools and methods used in the ocean for monitoring components of the OBCP.

3.1. Ship and laboratory-based measurements

A wide range of laboratory and field techniques have been developed to measure pools and fluxes of carbon in the ocean. In the following section (see also Table 1) we outline some of the key techniques and tools used in the field and laboratory. Table 2 lists examples of in situ databases, synthesis, data portals and observational programmes available to the community.

3.1.1. DIC and components of the carbonate system

The complete carbonate system can be quantified in situ by measurement of two or more core variables within a water sample (see Section “Two Out of Six” in Zeebe, 2012, for details) along with temperature and pressure: pH, carbon dioxide fugacity (fCO₂), total alkalinity (TA), and total dissolved inorganic carbon (DIC). Shipboard and in situ analyses of these components often rely on dissimilar instrumentation and methodologies, including: ion-sensitive field effect transistors; infrared detection; and titrimetric, colourmetric, spectrophotometric, and potentiometric approaches (Liu et al., 2013; Byrne, 2014; Martz et al., 2015).

3.1.2. DOC

DOC is typically separated from POC by filtration through a 0.2 μm filter. The DOC is then separated from DIC by acidification and sparging with CO₂-free gas (Sharp, 2002). This is typically based on the oxidation of organic compounds to CO₂ (typically by high-temperature combustion or UV, historically using persulfate methods) and is determined using either a non-dispersive infrared CO₂ analyser or through colourmetric measurements of CO₂ (Sharp, 2002). Colourless Dissolved Organic Matter (CDOM), that can be used as a tracer for DOC in coastal regions (see Section 4.1.2), is typically measured from discrete samples using long-path-length capillary waveguides with single-beam spectrophotometers (Bricaud et al., 2010; Nelson et al., 2007), though higher-frequency ship-based measurements are possible using underway spectrophotometry (see Dall’Olmo et al., 2017).

3.1.3. POC

Particulate organic carbon is typically defined as all the organic carbon that is retained by GF/F filters (nominal pore size of 0.7 μm). Samples are typically collected on pre-combusted (450 °C) GF/F filters and dried overnight at 65 °C before analysis. To remove particulate inorganic carbon, filters are acidified either by adding low-carbon HCl directly or by overnight exposure to the fumes of a concentrated HCl solution in a desiccator. Filters are then dried, packed in pre-combusted tin capsules, combusted in an elemental analyser to convert the organic carbon in CO₂ (Sharp, 1974). The liberated CO₂ is finally detected by thermal conductivity. Acetanilide is used as a standard. For further details on measuring POC, the reader is referred to protocols recently published by the IOCCG (Chaves et al., 2020).

3.1.4. Phytoplankton carbon

Phytoplankton carbon is notoriously difficult to quantify and no standard method exists. Current techniques include: measuring chlorophyll-α concentration then converting to carbon by making assumptions on the carbon to chlorophyll-α ratio (e.g. Sathyendranath et al., 2009); measuring cell counts (e.g. from light microscopy and flow cytometry) for different communities of phytoplankton, then making...
### Table 1

Key OBCP tools and techniques used in the laboratory and field. References are examples (see main text for a comprehensive list).

<table>
<thead>
<tr>
<th>Pools and fluxes</th>
<th>Common techniques</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pools</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOC</td>
<td>Samples separated from POC by 0.2μm filtration. DOC separated from DIC by acidification and sparging. DOC quantified through oxidation of organic compounds to CO₂, and measured using an infrared CO₂ analyser or by colourmetric methods.</td>
<td>Sharp (2002)</td>
</tr>
<tr>
<td>Total POC</td>
<td>Organic carbon retained on pre-combusted GF/F filter, dried then acidified to remove PIC. Filters combusted in an elemental analyser to convert the organic carbon in CO₂, which is then detected by thermal conductivity.</td>
<td>Sharp (1974) Chaves et al. (2020)</td>
</tr>
<tr>
<td><strong>Phytoplankton Carbon</strong></td>
<td>Proposed techniques include: measuring chlorophyll-a concentration then converting to carbon; measuring cell counts then converting bioluminescence and carbon; calibrating flow cytometer scattering and video-based imaging; flow cytometric sorted samples; and x-ray microanalysis.</td>
<td>Olson et al. (2003), Graff et al. (2015) &amp; Martínez-Vicente et al. (2017)</td>
</tr>
<tr>
<td><strong>Zooplankton Carbon</strong></td>
<td>Zooplankton cell counts (from nets, optical counters or video-based cameras) and bioluminescence and carbon conversions. Acoustic signals have also been related to zooplankton biomass.</td>
<td>Gallienne et al. (1996)</td>
</tr>
<tr>
<td><strong>Bacterial Carbon</strong></td>
<td>Techniques include: cell counts, bioluminescence estimates and carbon conversions; and separation of bacteria by tangential flow filtration and use of high-temperature catalytic oxidation.</td>
<td>Fukuda et al. (1998) &amp; Martínez-Vicente et al. (2016)</td>
</tr>
<tr>
<td><strong>PIC</strong></td>
<td>Inorganic carbon retained on pre-combusted GF/F filter, rinsed with 0.2 μm filtered seawater, extracted in a bath using 50% HCl, and quantified by atomic absorption spectrometry.</td>
<td>Mitchell et al. (2017)</td>
</tr>
<tr>
<td><strong>Fluxes</strong></td>
<td>Use of in-vivo techniques (e.g. ¹³C assimilation and O₂ evolution), triple isotope methods, O₂/Ar measurements, measurements of bulk properties, optical methods and proxies, and quantifying upper and lower limits.</td>
<td>Haney (1971), Daro (1978), Landry and Hassett (1982), White and Roman (1991) &amp; Kaiser and Benner (2012)</td>
</tr>
</tbody>
</table>

### Table 1 (continued)

<table>
<thead>
<tr>
<th>Pools and fluxes</th>
<th>Common techniques</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grazing</td>
<td>¹⁴C method, dilution techniques, radiotransfer methods, and measuring ingestion of [methyl-⁷H] methyamine hydrochloride (⁷H-MeA) labelled particles. Consumption has been estimated from the composition of major bioelements and biochemicals in the ocean.</td>
<td>Fuhrman and Azam (1990) &amp; Poulet et al. (1995)</td>
</tr>
<tr>
<td>Secondary production</td>
<td>Zooplankton production has been estimated from measurements of zooplankton biomass, hatching success and fecundity, and bacterial production from measurement of change in bacterial abundance in filtered seawater, and through bacterial deoxyribonucleic acid synthesis.</td>
<td>Robinson and Williams (2005)</td>
</tr>
<tr>
<td>Respiration</td>
<td>Inverse modeling of the community composition and activity, measuring the rate of production and consumption of a product or a reactant, the assay of an appropriate respiratory enzyme or enzyme system, and biomass-based predictions.</td>
<td>Sharp (1979) and Lancelot (1979)</td>
</tr>
<tr>
<td>Excretion</td>
<td>Radiocarbon methods or kinetic-based methods.</td>
<td>Burd and Jackson (2009)</td>
</tr>
<tr>
<td>Fragmentation</td>
<td>Monitoring abundances and size distributions of marine particles as a function of time, laboratory video-based analysis, and using optical measurements.</td>
<td>Tang et al. (2014), Agusti et al. (1996) &amp; Bale et al. (2011)</td>
</tr>
<tr>
<td>Solubility</td>
<td>Solubility of CO₂ is dependent on temperature and salinity.</td>
<td>Daniels et al. (2018)</td>
</tr>
<tr>
<td>Caleification</td>
<td>Two techniques are commonly used. The first involves filtering ¹³C-labelled samples through two filters, one is funed with acid to remove</td>
<td>(continued on next page)</td>
</tr>
</tbody>
</table>
3.1.5. Zooplankton and bacterial carbon

Techniques used to estimate zooplankton carbon typically involve counting microzooplankton (e.g. using nets, optical counters or video-based cameras), and using biovolume estimates (or length-weight relationships) and carbon conversions (Gallienne et al., 1996; Alcaraz et al., 2003). Efforts have also been made to relate acoustic signals to zooplankton biomass (Powell and Ohman, 2012). Techniques used to estimate bacterial carbon typically involve cell counts, biovolume estimates and carbon conversions. Other techniques include separation of bacteria from phytoplankton and detritus by tangential flow filtration, then use of high-temperature catalytic oxidation (HTCO) to determine carbon and nitrogen content (Fukuda et al., 1998; Martínez-Vicente et al., 2012, 2016).

3.1.6. PIC

Particulate inorganic carbon is typically defined as the inorganic carbon that is retained on a GF/F filter following filtration. Samples are collected on pre-combusted GF/F filters, rinsed with 0.2 μm filtered seawater, and extracted in a 40 °C bath by means of 50% HCl. Particulate calcium is then quantified by atomic absorption spectrometry and converted to PIC (Mitchell et al., 2017). It can also be estimated through the acid-labile backscattering technique and through inductively coupled plasma optical emission spectrometry (Balch et al., 2005).

3.1.7. Primary production, new production and export production

There are a wide range of methods used to measure primary production in the ocean. For a full review on the topic, the reader is referred to the recent works of Regaudie-de-Gioux et al. (2014), Sathyendranath et al. (2019b) and Church et al. (2019). Common methods used include: in-vivo techniques (14C and 13C assimilation; O2 evolution; 15NO3 assimilation; 15NH4 assimilation; 18O2 evolution); triple isotope methods; O2/Ar measurements; bulk properties (NO3 flux to the photic zone, O2 utilisation rate, net O2 accumulation in photic zone, 234Th/238U, 3H/3He); optical methods (double-flash fluorescence; passive fluorescence); and monitoring upper and lower limits (sedimentation rate below the photic zone, optimal energy conversion of photons absorbed, depletion of winter accumulation of NO3), typically through the use of sediment traps (Buesseler et al., 2007).

Table 2

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Data Syntheses</td>
<td>Synthesis of in-situ bio-optical observations of spectral remote-sensing reflectances, concentrations of chlorophyll-a, spectral inherent optical properties, spectral diffuse attenuation coefficients and total suspended matter.</td>
<td>Valente et al. (2019a, 2019b)</td>
</tr>
<tr>
<td>OC-CCI in-situ data</td>
<td>Synthesis of POC measurements from the ESA POCO project, including data from: PANGAEA; NASA SeaBASS (Werdell and Bailey, 2005); data compiled by Martiny et al. (2014); data from BCO-DMO; data from AMT (Rasse et al., 2017); data from cruises in the Southern Ocean (Thomalla et al., 2017).</td>
<td>Evers-King et al. (2017)</td>
</tr>
<tr>
<td>POC flux data</td>
<td>Synthesis of POC flux estimates from POC concentration observations derived from sediment traps and 233Th, compiled across the global ocean.</td>
<td>Mouv et al. (2016a, 2016b)</td>
</tr>
<tr>
<td>Picoplankton carbon data</td>
<td>Synthesis of more than 12,000 global observations of picophytoplankton abundance converted into carbon concentrations from the ESA POCO project.</td>
<td>Martínez-Vicente et al. (2017)</td>
</tr>
<tr>
<td>Cocolithophore calcification rates</td>
<td>Synthesis of 2765 CaCO3 production measurements, the majority of which were measured using 12 to 24 h incorporation of radioactive carbon (14C) into acid-labile inorganic carbon (CaCO3).</td>
<td>Daniels et al. (2018)</td>
</tr>
<tr>
<td>Photosynthesis Irradiance parameters</td>
<td>Global synthesis of photosynthesis-irradiance parameters from a range of oceanographic regimes as an aid to examining the basin-scale variability in the photophysiological response of marine phytoplankton.</td>
<td>Bouman et al. (2018)</td>
</tr>
<tr>
<td>MAREDAT</td>
<td>Synthesis of carbon biomass for 11 plankton functional types (picophytoplankton, diatoms, coccolithophores, Phaeocystis, pteropods, foraminifers, mesozooplankton, pteropods and macrozooplankton) plus phytoplankton pigment data, compiled as part of the MArine Ecosystem biomass DAtA (MAREDAT) initiative (<a href="http://www.pangaea.de/search?q=maredat">http://www.pangaea.de/search?q=maredat</a>).</td>
<td>Buitenhuis et al. (2013)</td>
</tr>
<tr>
<td>GOCAD</td>
<td>The Global Ocean Carbon Algorithm Database (GOCAD) is a global, carbon-focused synthesis of optical (apparent and inherent optical properties, hyperspectral and extending to the UV), physical, biogeochemical and carbon data (e.g. DOC and POC), for use in satellite algorithm development.</td>
<td>Aurin et al. (2018)</td>
</tr>
<tr>
<td>LDEO pCO2 data</td>
<td>Synthesis of 13.5 million measurements of surface water</td>
<td>Takahashi et al. (2019)</td>
</tr>
<tr>
<td>Name</td>
<td>Description</td>
<td>Reference</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>SOCAT fCO₂ data</td>
<td>Synthesis of surface ocean fCO₂ (fugacity of carbon dioxide) observations. The latest SOCAT version (v2019) has 25.7 million global observations from 1957 to 2019 (<a href="https://www.socat.info">https://www.socat.info</a>).</td>
<td>Bakker et al. (2016)</td>
</tr>
<tr>
<td>General Databases and Portals</td>
<td>Copernicus In Situ Component: Includes in-situ ocean data (ship, drifters, floats and buoys) for use in combination with space-based data. It includes contributions of the member states to the Copernicus programme, and benefits from international efforts to collect and share data.</td>
<td><a href="https://insitu.coperni">https://insitu.coperni</a> eu</td>
</tr>
<tr>
<td></td>
<td>EMODataNet: The European Marine Observation and Data Network (EMODnet) has contributions from 150 organisations assembling marine data, products and metadata to integrate these resources and make them more publicly-available, using quality-assured, standardised and harmonised methods.</td>
<td><a href="https://www.emodnet.eu">https://www.emodnet.eu</a></td>
</tr>
<tr>
<td></td>
<td>OceanSites: OceanSites is a global network of long-term, reference stations measuring many variables and at the full depth of the water column.</td>
<td><a href="http://www.oceansites.org">http://www.oceansites.org</a></td>
</tr>
<tr>
<td></td>
<td>Argo Programme: Profiles of physical variables (ocean temperature, salinity and pressure) from the global network array of Argo floats.</td>
<td><a href="http://wwwargo.ucsd.edu">http://wwwargo.ucsd.edu</a></td>
</tr>
<tr>
<td></td>
<td>Biogeochemical-Argo Programme: An extension of the Argo float array containing profiles of biogeochemical variables (in addition to physical variables), including: pH, oxygen, nitrate, chlorophyll, suspended particles, and downwelling irradiance.</td>
<td><a href="https://biogeochemical-argo.org">https://biogeochemical-argo.org</a></td>
</tr>
<tr>
<td></td>
<td>SeaBASS: The SeaWiFS Bio-optical Archive (SeaBASS) is a publicly shared archive of in-situ oceanographic and atmospheric data maintained by the NASA Ocean Biology Processing Group.</td>
<td><a href="https://seabass.gsfc.nasa.gov">https://seabass.gsfc.nasa.gov</a></td>
</tr>
<tr>
<td></td>
<td>BCO-DMS: The Biological and Chemical Oceanography Data Management Office (BCO-DMS) is a facility where marine biogeochemical and ecological data and information can easily be disseminated, protected, and stored on short and intermediate time-frames. Created to serve principal investigators funded by the US National Science Foundation (NSF).</td>
<td><a href="https://www.bco-dmo.org">https://www.bco-dmo.org</a></td>
</tr>
<tr>
<td></td>
<td>PANGAEA: Open Access library aimed at archiving, publishing and distributing georeferenced data from earth system research.</td>
<td><a href="https://www.pangaea.de/">https://www.pangaea.de/</a></td>
</tr>
<tr>
<td></td>
<td>OCADS: The Ocean Carbon Data System (OCADS) is a data management project located within the NOAA National Centers for Environmental Information.</td>
<td><a href="https://www.noc.noaa.gov/ocads/">https://www.noc.noaa.gov/ocads/</a></td>
</tr>
<tr>
<td></td>
<td>Simons CMAP: Simons Collaborative Marine Atlas Project (Simons CMAP) is an open-source data portal interconnecting data sets across oceanography disciplines.</td>
<td><a href="http://simonsmap.com">http://simonsmap.com</a></td>
</tr>
<tr>
<td></td>
<td>COPEPOD: Coastal and Oceanic Plankton Ecology, Production, and Observation Database (COPEPOD). A global plankton database that contains over 400,000 observations of copepods, along with other zooplankton, phytoplankton, and microbial plankton taxa.</td>
<td><a href="https://www.st.nmfs.noaa.gov/copepod/">https://www.st.nmfs.noaa.gov/copepod/</a></td>
</tr>
</tbody>
</table>

Observational programmes:
- HOT: The Hawaii Ocean Time-series (HOT) program has been making repeated observations of the hydrography, chemistry and biology of the water column at a station north of Oahu, Hawaii since 1988. [https://www.hahana.soest.hawaii.edu/hot/](https://www.hahana.soest.hawaii.edu/hot/)
- MOBY: The NOAA funded Marine Optical BuoY (MOBY) provides vicarious calibration of ocean colour satellites. It is an autonomous optical buoy which is moored off the island of Lanai in Hawaii. [http://bats.bios.edu](http://bats.bios.edu)
- BATS: The Bermuda Atlantic Time-series Study (BATS) has been making repeated observations on the physical, biological, and chemical properties of the ocean every month since 1988 in an area of the western Atlantic Ocean (~32.17°N, 64.5°N). The programme is supplemented by biweekly Hydrostation-S cruises to a neighbouring location that began in 1954. [http://www.obs-vlfr.fr/Boussole/html/projet/introduction.php](http://www.obs-vlfr.fr/Boussole/html/projet/introduction.php)
- CARAICO: The Carbon Retention In A Coloured Ocean (CARAICO) is an ocean time series program situated in the Cariaco Basin with the objective of studying the relationship between surface primary production, regional hydrography, physical forcing variables, and the settling flux of POC. [http://imars.marine.usf.edu/caraco](http://imars.marine.usf.edu/caraco)
- WCO: The Western Channel Observatory (WCO) is an oceanographic time-series (dating back to 1903) and marine biodiversity reference site in the Western English Channel. In-situ measurements are undertaken weekly at coastal station L4 and fortnightly at open shelf station E1. [https://www.westerncanalexobservatory.org.uk](https://www.westerncanalexobservatory.org.uk)
- AMT: The Atlantic Meridional Transect (AMT) is a multidisciplinary programme which undertakes biological, chemical and physical oceanographic research during an annual voyage between the UK and destinations in the South Atlantic. [https://www.amt-uk.org](https://www.amt-uk.org)
- CPR: The Continuous Plankton Recorder (CPR) Survey has been collecting data in the North Atlantic and the North Sea on the biogeography and ecology of plankton since 1931, and more recently expanded to other regions around the globe. The CPR is an instrument designed...
3.1.8. Grazing and secondary production

Common methods for estimating grazing rates include: the $^{14}$C method (Daro, 1978); dilution techniques (Landry and Hassett, 1982); the radioisotope method using $^{32}$P-labelled yeast (Rhodotorula) (Haney, 1971); and measuring ingestion of [methyl-$^{3}$H]methylamine hydrochloride ($^{3}$H-MeA) labelled particles (White and Roman, 1991). Consumption has been estimated from vertical measurements of the composition of major bioelements and biochemicals in the ocean (Kaiser and Benner, 2012). Secondary production by zooplankton has been estimated using information on zooplankton biomass, hatching success and fecundity, derived from laboratory mesocosm experiments or ship-based incubations (Poulet et al., 1995). Secondary production by bacteria has been estimated from measurements of increasing bacterial abundance with time in filtered seawater, and through bacterial deoxyribonucleic acid synthesis by tritiated thymidine incorporation in seawater (Fuhrman and Azam, 1980).

3.1.9. Respiration and excretion

Planktonic respiration rates can be derived through various means (see Robinson and Williams, 2005, for a review on the topic). Common methods include: inverse modeling of the community composition and activity; measuring the rate of production and consumption of a product or a reactant; the assay of an appropriate respiratory enzyme or enzyme system; and biomass-based predictions (Robinson and Williams, 2005). Excretion rates are typically determined using either the radiocarbon method indirectly (Fogg, 1952; Watt and Fogg, 1966; Sharp, 1979) or directly using kinetic-based methods (Lancelot, 1979).

3.1.10. Aggregation and fragmentation

Aggregation has been measured in laboratory-based experiments using floculators (Waite et al., 1997), and using mesocosm based experiments (Passow and Aldredge, 1995), or in the field using photographic observations (Suzuki and Kato, 1953) and high resolution video analysis (Gorsky et al., 2000). For a full review on the topic of aggregation the reader is referred to the work of Burd and Jackson (2009). Fragmentation has been measured through comparisons of abundances and size distributions of marine particles in time (Dilling and Aldredge, 2000), laboratory video-based analysis (Goldthwaite et al., 2004), and using optical measurements to track changes in small and large particles (Briggs et al., 2020).

3.1.11. Non-predatory mortality

Non-predatory plankton mortality has been estimated using microscopic methods. These typically involve differentiating between living and dead plankton using staining based methods or morphological characteristics (Tang et al., 2014). Phytoplankton cell lysis rates have been estimated by using spectrofluorometric methods to quantify dissolved esterase activity (Agusti et al., 1998), flow cytometric staining has been used for monitoring phytoplankton vitality (Peperzak and Brussaard, 2011), and algal senescence has been estimated through pigment-based methods (Bale et al., 2011).

3.1.12. Solubility

The solubility of CO$_2$ is dependent on temperature and salinity and various equations have been proposed relating CO$_2$ solubility to these physical variables (Murray and Riley, 1971; Weiss, 1974).

3.1.13. Calcification

Rates of calcification, or CaCO$_3$ production, are typically measured using $^{14}$C. This was first proposed by Paasche (1962, 1963), who demonstrated in the laboratory that radiocaive $^{14}$C could be used to infer the production of both organic (photosynthesis) and inorganic carbon (calcification) by coccolithophores (Daniels et al., 2018). At present two techniques are commonly used: 1) following incubation, $^{14}$C-labelled samples are filtered through two filters, one is used to give an estimate of total particulate production, the other is fumed with acid to remove Ca$^{14}$CO$_3$ giving an estimate of particulate organic production, with calcification representing the difference between filters; 2) using the micro-diffusion technique, that captures $^{14}$CO$_2$ liberated from Ca$^{14}$CO$_3$, which gives a measure of calcification. Additional techniques have been developed, for example, using $^{45}$Ca as a tracer (Van der Wal et al., 1995). For further details on these techniques, and data on calcification rates, the reader is referred to Daniels et al. (2018).

3.1.14. Air-sea CO$_2$ transfer

In-situ methods used to measure or determine the exchange of CO$_2$ between the ocean and atmosphere include: tracer studies using ambient gases ($^{13}$CO$_2$) that integrate the flux over years; deliberately introduced tracers that integrate the flux over days; and direct eddy-covariance flux measurements on hourly timescales (Landwehr et al., 2018). A recent review of the history and current capabilities of these approaches can be found in Shulter et al. (2019).

3.2. Autonomous platforms

Over the past few decades there has been a rapid increase in the use of autonomous platforms for ocean sampling (Bushinsky et al., 2019; Chai et al., 2020). While not the main focus of this review, this section discusses the types of autonomous platforms currently in use for monitoring the OBCP and directs the reader to additional reviews on the topic. Autonomous platforms have the advantage over ship-based methods in that they are capable of operating in hostile environments, at temporal and spatial scales that challenge research ship operation, and operating below the eyes of the satellite. They are also cheap to

Table 2 (continued)

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTER The Long-Term Ecological Research (LTER) programme, supported by the US National Science Foundation, consists of 28 sites throughout the United States, Puerto Rico and Tahiti. Key marine sites include the Palmer LTER, California Current Ecosystem LTER, and Northeast U.S. Shelf.</td>
<td><a href="https://lternet.edu">https://lternet.edu</a></td>
<td></td>
</tr>
<tr>
<td>Station Stonica Located in the central Adriatic Sea, near the island of Vis, monthly phytoplankton primary production measurements have been collected at Station Stonica since April 1962, providing one of the longest time-series of primary production measurements to date.</td>
<td>Kovac et al. (2018a)</td>
<td></td>
</tr>
<tr>
<td>EXPORTS EExport Processes in the Ocean from Remote Sensing (EXPORTS) is a NASA funded programme aiming to develop a predictive understanding of the export and fate of global ocean net primary production and its implications for present and future climates.</td>
<td><a href="https://oceannexports.org">https://oceannexports.org</a></td>
<td></td>
</tr>
<tr>
<td>NAAMES The North Atlantic Aerosols and Marine Ecosystems Study (NAAMES) is a NASA funded programme designed to improve understanding of Earth’s ocean ecosystem-aerosol-cloud system.</td>
<td>Behrenfeld et al. (2019b)</td>
<td></td>
</tr>
</tbody>
</table>
Examples of autonomous ocean platforms include:

- **Argo and biogeochemical Argo (BGC-Argo) floats.** These are floating devices that profile the vertical structure of the water column, measuring key physical (temperature, salinity and pressure) and biogeochemical and related variables (pH, oxygen, nitrate, chlorophyll, backscattering and irradiance) usually down to 1000 or 2000 m, transmitting the data via satellite communication. The network of Argo and BGC-Argo floats is transforming our understanding of OBCP processes and fluxes (e.g. Johnson et al., 2009; Boss and Behrenfeld, 2010; Dall Olmo and Mork, 2014; Dall’Olmo et al., 2016; Bishop et al., 2016; Lacour et al., 2017; Mignot et al., 2018; Ardyña et al., 2019; Estapa et al., 2019; Gittings et al., 2019; Briggs et al., 2020), and there is huge potential for integrating vertical data from BGC-Argo floats with surface satellite observations, for improving the monitoring of carbon pools and fluxes like primary production (Richardson and Bendtsen, 2019). For a recent review of the topic the reader is referred to the work of Claustre et al. (2020).

- **Underwater gliders** are autonomous underwater vehicles that, unlike floats, which float passively as a preprogrammed pressure level and predetermined time (Bork et al., 2008), can profile vertically (typically down to around 1000 m, depending on design) and move horizontally by changing their buoyancy and using wings (Stommel, 1989; Rudnick, 2016). Equipped with physical, biological and chemical sensors, they can be programmed to operate in hostile or inaccessible environments (e.g. under sea ice), can track physical, chemical and biological features, and can operate for long durations, measuring key components of the OBCP (e.g. Hemsley et al., 2015). They are particularly useful for complementing ship-cruises, for instance, by putting discrete observations collected by ship in the context of larger-scale oceanic processes. Towed undulated gliders, equipped with sensors, have also been used to monitor at depth (Rudnick and Cole, 2011). Similar to BGC-Argo floats, they have been transforming our understanding of the OBCP (e.g. Allkire et al., 2014; Omand et al., 2015; Bol et al., 2018). For a recent review of the topic the reader is referred to the work of Rudnick (2016).

- **Unmanned surface vehicles,** including wave gliders (Daniel et al., 2011), sea gliders (Eriksen et al., 2001) and sail drones (Gentemann et al., 2020), are small (typically 2-7 m long) surface platforms powered by wind, wave and solar energy. These platforms are capable of operating for long duration, covering vast distances (100 km per day), and when equipped with environmental sensors, can sample the biological, chemical and physical properties of the surface ocean, as well as the air-sea interface and lower atmosphere (e.g. wind and surface irradiance), and transmit the data via satellite communication. They are increasingly being used to complement satellite observations (Villareal and Wilson, 2014; Vazquez-Cuervo et al., 2019; Scott et al., 2020), and reveal ocean and atmosphere dynamics in remote regions (Mitarai and McWilliams, 2016; Thomson and Girton, 2017) that are relevant to ocean carbon cycling.

- **Next generation autonomous buoys and drifters.** New autonomous buoys (stationary platforms) and drifters (Lagrangian platforms) are being developed, with satellite tracking and transmission, that are capable of hosting physical, biological and chemical sensors for monitoring the surface ocean, and for use in satellite-algorithm validation (e.g. Brown et al., 2019; Le Menn et al., 2019). These systems provide the capability of being produced at low cost and deployed en masse, and consequently, increasing surface ocean observations.

- **Unmanned aerial vehicles.** Unmanned aerial vehicles (UAVs) – in particular portable and affordable, aerial drones – are now being used in physical and biological oceanography. Systems have been developed for measuring: SST and surface salinity (Lee et al., 2016; McIntyre and Gasiewski, 2007); for collecting water samples (Ore et al., 2015; Terada et al., 2018); measuring aerosols over the ocean (Corrigan et al., 2008, useful for atmospherically-correcting ocean-colour data); for measuring chlorophyll, macroalgae and harmful algal blooms (Su, 2017; Xu et al., 2018; Luy et al., 2017); and monitoring the ecology of large ocean animals (Durban et al., 2016). For a recent review of the topic the reader is referred to the work of Johnston (2019).

- **Ship-based autonomy.** There is a long history of marine scientists using ships of opportunity for autonomous ocean sampling. The Continuous Plankton Recorder (CPR) survey uses an instrument towed from the stern of volunteer ships designed to capture plankton samples, that has been successfully operating since 1931. In more recent years, autonomous systems have been integrated into commercial ships for monitoring sea-surface and atmospheric properties relevant to the marine carbon cycle, including: partial pressure of CO₂ (pCO₂), phytoplankton concentration, biogeochemical properties, above-water and in water marine optical and thermal properties, ocean physics and atmospheric gases (e.g. Donlon et al., 2008; Pierrot et al., 2009; Slade et al., 2010; Giliuzza et al., 2013; Simis and Olsson, 2013; Petersen, 2014).

- **Marine animal platforms.** Marine animals are used as platforms for oceanographic sampling, particularly in the polar seas (Fedak, 2004, 2013; Harcourt et al., 2019). Seals (Fedak, 2004; Meredith et al., 2011), turtles (Mahon et al., 2005), whales (Lydersen et al., 2002), penguins (Xavier et al., 2018), sea birds (Yoda et al., 2014; Yonehara et al., 2016), even humans (Brewin et al., 2015a, 2017b; Bresnahan et al., 2016, 2017b; Wright et al., 2016), have all been proposed as oceanographic platforms. Data from such platforms have been used for evaluating satellite observations (Brewin et al., 2017c; Keates et al., 2020) and integrated with other autonomous observations to gain new insight into biological oceanography in remote regions (Carranza et al., 2018).

The capacity of autonomous platforms for hosting physical, chemical and biological sensors varies with each platform. Common physical sensors hosted on in-water platforms are temperature, salinity and pressure sensors (CTDs), useful for understanding physical controls on ocean carbon cycling, for example, the retrieval of ocean surface currents from drifting buoys, which play a key role in advecting carbon pools. The most common technology used to derive information on carbon pools and fluxes are optical sensors. Frequently-used optical sensors include fluorometers and backscattering meters (single photodetectors), which are typically used to estimate chlorophyll concentration and POCl, respectively, through establishing relationships between the optical signals and concentrations of the biogeochemical variable of interest. These sensors are primarily used to derive information on small particles, though analysis of optical spikes offer a way to access information on larger particles, including aggregates and zooplankton (Briggs et al., 2011, 2013, 2020; Burt and Tortell, 2018). Light sensors are also a common feature on autonomous platforms, useful for quantifying light availability for photosynthesis and light attenuation. Absorption and attenuation meters have been deployed on larger autonomous platforms (e.g. Cunningham et al., 2003), for measuring beam attenuation and absorption properties of water constituents. Fluorometers can also be tuned to measure fluorescence by CDOM, offering a route into monitoring the DOC pool in some regions.

Advancement in microelectronics, battery technology and wireless communications have presented opportunities to integrate new optical sensors onto autonomous platforms, including: transmissometers (Bishop et al., 2004; Bishop and Wood, 2009); photodetector arrays (Miles et al., 2018); holographic systems (Anderson et al., 2018), and photographic systems (Høienjens et al., 2020). Data from these sensors are capable of providing insight into pools and fluxes of carbon, for example, the particle size distribution, plankton composition, particle biolume, inorganic and organic content of particles, sinking velocities, particle fluxes, sedimentation rates, and zooplankton migration. For a full review on the topic see the recent work of Giering et al. (2020).
Oxygen and pH probes are increasingly being integrated into autonomous platforms, with potential for estimating carbonate chemistry (Bresnahan et al., 2016) and surface ocean pCO₂ (Williams et al., 2017), as are acoustic sensors, for monitoring zooplankton, currents (ADCPs), bathymetry, marine mammals, wind and rain (McCollum and Krajewski, 1997; Baumgartner et al., 2013; Guileen et al., 2014; Todd et al., 2017; Cauchy et al., 2018). Nutrient sensors can be used on BGC-Argo floats and underwater gliders (Johnson et al., 2013), and there are increasing efforts to integrate miniature lab-on-a-chip style chemical sensors (Beaton et al., 2011; Vincent et al., 2018) as well as water samplers (Bresnahan et al., 2017a) on autonomous platforms. As microelectronic and battery technologies develop, there will be potential to miniaturise other oceanographic technologies for use on autonomous platforms.

Despite the huge potential, there are challenges in the operation and maintenance of sensors (e.g. bio-fouling) on autonomous platforms, and at present in maintaining their global distribution. It is critical that data quality assurances are in place. This can be done through rigorous, community-assessed protocols, sensor calibrations, monitoring of sensor drift and stability, cross-checking with existing datasets, validation, database management, and data filtering (e.g. Xing et al., 2011, 2012; Organelli et al., 2016; Claustre et al., 2020). Efforts also need to be placed on quantifying uncertainties (Williams et al., 2017). For sustained observations of the OBCP, an integrative approach has been suggested, combining autonomous observations with other field data, models and satellite remote sensing (Claustre et al., 2010).

4. Monitoring carbon pools and fluxes from space

Satellite remote sensing is the only method capable of viewing our entire surface ocean synoptically, and at a relatively high temporal frequency (daily) and spatial resolution (0.3 to 25 km). While providing powerful horizontal (global) and temporal coverage, satellite ocean observations have limitations, for example:

- Satellite ocean observations are limited to the surface layer of the ocean. The satellite signals in the visible portion of the electromagnetic spectrum are representative of about one optical depth (Zaneveld et al., 2005) \(1/K_d\), where \(K_d\) (m\(^{-1}\)) represent the diffuse attenuation coefficient of the water column (up to 50 m depending on water clarity and the wavelength used). Other parts of the electromagnetic spectrum utilised for ocean observations from satellite (e.g., thermal and microwave regions) are representative of less than a millimetre at the surface.
- Passive satellite visible and thermal remote sensing is limited to cloud-free conditions, and for visible sensing, low to moderate sunzenith angles.
- Satellite observations often have limitations and difficulties in coastal regions. For example, for visible remote sensing, the optical complexity of coastal waters makes it challenging to derive some pools and fluxes of carbon (IOCCG, 2000; Kostadinov et al., 2007).
- Atmospheric correction of satellite data is often more complex in coastal regions (Hu et al., 2000).
- Satellite ocean observations are limited to information that can be communicated by electromagnetic radiation. The primary observables (e.g., remote-sensing reflectance, brightness temperature and backscatter) need to be related to surrogate (geophysical/biogeochemical/ecological) variables used to compute the carbon pools and fluxes. These relationships can be complex and prone to large uncertainties.
- To resolve some very fine spatial and temporal scale physical oceanographic processes at the surface, complementary information to satellite data is often required.

Therefore, like other technologies, satellite data often needs to be integrated with other observations and models, for algorithm development, for extrapolating the surface signal down to depth, for filling gaps in the data, and for resolving fine scale surface processes. Techniques used to monitor pools and fluxes of ocean carbon using satellite data lie on a scale between direct and indirect techniques, with some using aspects of both.

- **Direct techniques** (i.e. detection methods) are dependent on whether the carbon pool in question modifies (absorbs and/or scatters) visible electromagnetic radiation to the extent its presence and its concentration leaves an imprint on the signal that can be seen from space.
- **Indirect techniques** involve relating the pool or flux of interest to some variable (or variables) that can be retrieved with confidence from space (e.g. through ocean-colour, thermal radiometry, microwave radiometry and satellite altimetry). This approach essentially uses satellite remote-sensing as a method of extrapolation. Proposed relationships between satellite variables and pools (or fluxes) can be mechanistic, empirical or statistical.

In this section, we provide an overview of the satellite techniques used to derive information on pools and fluxes (see also Table 3).

4.1. Pools of carbon from space

4.1.1. DIC from space

DIC cannot be directly observed from space. However, indirect approaches have been proposed. Land et al. (2015) identified the potential of these indirect approaches and Land et al. (2019) evaluated the precision and accuracy of these approaches for regional and global waters. These approaches can use a combination of satellite sea-surface temperature (SST), the major controller of the solubility of CO₂, and satellite retrievals of sea surface salinity (SSS), and sometimes use climatology nutrient data or satellite ocean-colour observations, for interpreting the biological modulation of DIC. DIC has been estimated using these indirect approaches in the Equatorial Pacific (Loukos et al., 2000), the Arctic (Arrigo et al., 2010) and the Arabian Sea (Sarma, 2003), the Caribbean, the Bay of Bengal, the Amazon outflow, and globally (Land et al., 2019). As discussed by Land et al. (2015), such approaches require tuning for different regions of the ocean, are based on in-situ data-derived empirical relationships between DIC variables and the oceanic variables, and so may require tuning for a particular region and season (e.g., Sarma, 2003; Sarma et al., 2006).

4.1.2. DOC from space

From a remote-sensing perspective, the ensemble of all DOC pools does not have a strong optical signature that can be used to quantify its concentration in surface waters. A small component (~10%) of the DOC pool is chromophoric (coloured), and thus can be directly monitored by ocean colour. Satellite retrievals of spectral absorption by CDOM, and combined spectral absorption by CDOM and coloured detrital matter (CDM), are routinely produced by space agencies, and have been shown to perform with reasonable accuracy in satellite validation exercises (Siegel et al., 2013; Mannino et al., 2014; Loisel et al., 2014; Brewin et al., 2015c). Though representative of “matter” rather than “carbon”, and acknowledging difficulties in separating CDOM and detrital absorption owing to their similar spectral signatures, there have been increasing efforts to use satellite CDOM and CDM products for monitoring DOC.

In coastal and shelf regions, strong correlations between CDOM absorption (magnitude and spectral slope) and DOC have been observed (Ferrari et al., 1996; Vodaček et al., 1997; Bowers et al., 2004; Fichot and Benner, 2012; Tehran et al., 2013). Mannino et al. (2008) analysed a large dataset of CDOM absorption (m\(^{-1}\)) and DOC concentration (μmol C L\(^{-1}\)) in the Mid Atlantic Bight and established empirical relationships between the two variables that changed with season. Application of these algorithms to satellite CDOM absorption estimates indicated a
Table 3
Techniques for monitoring the OBCP from satellites. (For Maturity, N = Not currently possible, I = In it’s infancy; Y = Possible). References are examples (see main text for more).

<table>
<thead>
<tr>
<th>Pools and fluxes</th>
<th>Maturity</th>
<th>Common techniques</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pools DIC</td>
<td>I</td>
<td>Indirect approaches that integrate space-based sea surface salinity, temperature and ocean colour observations.</td>
<td>Land et al. (2015) &amp; Shutler et al. (2019)</td>
</tr>
<tr>
<td>DOC</td>
<td>I</td>
<td>Using empirical relationships between CDOM absorption (retrievable from satellite ocean colour) and DOC. These have been found to work well in coastal waters and the Arctic Ocean, but not in the open ocean where the relationship breaks down. There is potential for developing indirect approaches in the open-ocean using multiple satellite products.</td>
<td>Mannino et al. (2008) &amp; Matsuoka et al. (2017)</td>
</tr>
<tr>
<td>POC Total POC</td>
<td>Y</td>
<td>Methods include: empirical band ratio algorithms; empirical band-difference algorithms; backscattering-based algorithms; algorithms using satellite estimates of backscattering and chlorophyll; and algorithms using satellite estimates of diffuse attenuation, and a two-step relationship between diffuse attenuation and beam attenuation.</td>
<td>Stramski et al. (2008), Evers-King et al. (2017) &amp; Le et al. (2018)</td>
</tr>
<tr>
<td>Phytoplankton carbon</td>
<td>I</td>
<td>Methods include: relationships between phytoplankton carbon and chlorophyll; reflectance-based; absorption-based; and backscattering-based. Algorithms are available for partitioning phytoplankton</td>
<td>Martinez-Vicente et al. (2017), Kostadinov et al. (2016), Roy et al. (2017) &amp; Sathyendranath et al. (2020b)</td>
</tr>
<tr>
<td>Fluxes Primary production</td>
<td>Y</td>
<td>All approaches relate elevated water-leaving radiance (or reflectance), a result of elevated backscattering due to the presence of coccoliths, to PIC concentrations. Methods proposed include: band-ratio approaches; band-difference approaches; optical water type methods; and semi-analytical inversions and look-up-tables.</td>
<td>Gordon et al. (2001), Balch et al. (2005), Mitchell et al. (2017) &amp; Kondrik et al. (2019)</td>
</tr>
<tr>
<td>Zoo-plankton carbon</td>
<td>I</td>
<td>Methods include: indirect approaches, through estimates of the slope of the particle size spectrum, or biophysical models driven by visible and thermal ocean remote sensing products; and some emerging direct approaches, for large congregations of certain species, or through active satellite ocean colour (lidar) methods.</td>
<td>Strömberg et al. (2009), Basedow et al. (2019) &amp; Behrenfeld et al. (2019a)</td>
</tr>
<tr>
<td>Bacteria carbon</td>
<td>I</td>
<td>Indirect methods have been proposed for mapping bacterial abundance for some species, typically using multiple satellite products.</td>
<td>Grimes et al. (2014), Racault et al. (2019) &amp; Larsen et al. (2015)</td>
</tr>
</tbody>
</table>

(continued on next page)
capacity to retrieve DOC within 10% uncertainty (mean absolute percentage difference). Mannino et al. (2016) updated these relationships using a larger dataset, and developed a neural network algorithm to model the vertical profile of DOC from the surface concentrations (retrievable from satellite) and physical model data. They then combined satellite estimates of water column DOC with physical circulation model simulations to quantify the shelf boundary fluxes of DOC in the Mid Atlantic Bight.

Relationships between CDOM absorption and DOC have also been established for Arctic regions. Using the method of Fichot and Benner (2012), that relates terrigenous DOC concentration in coastal waters to the exponential slope of the absorption coefficient of CDOM in the 275 to 295 nm range (S275–295), Fichot et al. (2013) estimated the first pan-Arctic distribution of terrigenous dissolved organic matter and continental runoff in the surface Arctic Ocean, by estimating S275–295 from satellite remote-sensing reflectance data. In a further study, Matsuoka et al. (2017) demonstrated that satellite estimates of CDOM absorption across the Arctic region can be retrieved with an uncertainty of 12%, and showed a tight correlation between CDOM absorption and DOC (Matsuoka et al., 2012), demonstrating that DOC in river-influenced coastal waters can be retrieved from satellite ocean colour with an average uncertainty of 28%. Le Fouest et al. (2018) showed good agreement between the Matsuoka et al. (2017) satellite estimates of DOC and modelled DOC, and used the data to assess the concentration of terrestri
d DOC in surface waters in the western Arctic Ocean.

Rivers are a major source of allochthonous DOC in the ocean. Assuming DOC behaves as a conservative property, then it will change with salinity as the freshwater gets mixed with seawater. It has in fact been shown that the absorption by CDOM in coastal waters under the influence of freshwater influx is closely related to salinity. For example, Bowers et al. (2000) have proposed a remote-sensing algorithm for estimating the concentration of CDOM in the Clyde Sea, which could then be used to estimate salinity in waters affected by freshwater. Today, with the advent of salinity sensors in space, the potential exists for using salinity as a tracer for DOC influx. But there is also emerging evidence that terrigenous DOC may be susceptible to microbial degradation, based on a study of DOC from the Congo River (Spencer et al., 2009). They suggested that the CDOM absorption coefficient could potentially be used for studying photochemical degradation of terrigenous DOC. It would therefore be desirable to compare the timescales involved in the microbial degradation with those of mixing between freshwater and seawater, to ensure the degradation did not introduce significant errors into the calculation of DOC from salinity, or, if necessary, to correct for such processes. The algorithms noted above have been developed for coastal waters, and are regional in their implementation. The question remains whether an algorithm that would be globally applicable in all coastal waters might be established, for detection of riverine DOC inputs into the ocean. Efforts towards establishing the data required to answer this question are underway (Aurin et al., 2018). Many of these algo
rithms are designed to study the dynamics of allochthonous DOC (e.g. from river runoff), as opposed to biogenic, autochthonous DOC pro
duced in the ocean through microbial cycling.

In the open ocean, Nelson and Siegel (2013), in a major study involving over a thousand samples from a variety of locations and depths, concluded that no relationship existed between CDOM absorp
tion and DOC concentration, hence demonstrating major challenges in using satellite CDOM absorption products to quantify open-ocean DOC. Considering the different components of DOC (labile and semi-labile), their timescales, vertical structure, processes such as photobleaching, and the biotic and abiotic factors controlling DOC processes (Hansell, 2013), indirect satellite methods for open ocean DOC estimation could be developed, using conceptual, empirical or statistical models, driven by multiple satellite products (e.g. chlorophyll, plankton type, light, salinity and SST). In a recent study, Roshan and DeVries (2017) com
bined an artificial neural network estimate of the global DOC distribu
tion, developed using input data on chemical (e.g. nutrients), physical (e.g. depth, salinity and temperature), and biological variables (depth of the euphotic zone and chlorophyll concentration), with a data-constrained ocean circulation model, to produce the first observationally based global-scale assessment of DOC production and export. Many of these physical and biological products are available through satellite, suggesting the potential to develop similar satellite-driven approaches.

### Table 3 (continued)

<table>
<thead>
<tr>
<th>Pools and fluxes</th>
<th>Maturity</th>
<th>Common techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Export production</td>
<td>Y</td>
<td>Previous attempts have focused on empirical relationships between export production, SST and primary production. Satellite-data driven food-web export flux models have also been proposed.</td>
</tr>
<tr>
<td>Grazing</td>
<td>N</td>
<td>Not currently retrievable from space (but see Section 4.2.3).</td>
</tr>
<tr>
<td>Secondary production</td>
<td>N</td>
<td>Not currently retrievable from space (but see Section 4.1.3).</td>
</tr>
<tr>
<td>Respiration</td>
<td>N</td>
<td>Not currently retrievable from space (but see Section 4.2.3).</td>
</tr>
<tr>
<td>Excretion</td>
<td>N</td>
<td>Not currently retrievable from space (but see Section 4.2.3).</td>
</tr>
<tr>
<td>Aggregation</td>
<td>N</td>
<td>Not currently retrievable from space.</td>
</tr>
<tr>
<td>Fragmentation</td>
<td>N</td>
<td>Not currently retrievable from space (but see Section 4.2.3).</td>
</tr>
<tr>
<td>Non-predatory mortality</td>
<td>N</td>
<td>Not currently retrievable from space.</td>
</tr>
<tr>
<td>Solubility</td>
<td>I</td>
<td>Feasible at the surface with estimates of SST and SSS.</td>
</tr>
<tr>
<td>Calcification</td>
<td>N</td>
<td>Not currently retrievable from space (but see Section 4.1.4).</td>
</tr>
<tr>
<td>Air-sea CO₂ transfer</td>
<td>I</td>
<td>Regional, indirect approaches have been proposed. These involve forcing empirical equations and statistical models with satellite retrievals of SST, SSS and chlorophyll-a, individually or in combination.</td>
</tr>
</tbody>
</table>

---

1. Laws et al. (2000, 2011) & Siegel et al. (2014)
2. Mannino et al. (2016)
4. Raphael et al. (2014)
5. Bowers et al. (2000)
7. Fichot et al. (2013)
8. Mannino et al. (2016)
10. Matsuoka et al. (2017)
11. Matsuoka et al. (2012)
12. Fichot et al. (2013)
4.1.3. POC from space

4.1.3.1. Total POC from space. POC is accessible through satellite remote-sensing of ocean colour because particles suspended in seawater that make up the POC influence the bulk seawater inherent optical properties, namely the particle beam attenuation, scattering and backscattering coefficients, which are all sensitive to particle abundance, to the composition of particles (complex refractive index), their size distribution, shape and their internal structure (e.g. Stramski and Kiefer, 1991; Organelli et al., 2018). Empirical relationships between POC and particle beam attenuation measured from optical instruments (transmissometers) have been quantified, and used to monitor POC optically (Gardner et al., 1993; Bishop, 1999; Claustre et al., 1999; Stramska and Stramski, 2005). In addition, relationships between POC and particle backscattering and absorption have also been quantified (e.g. Stramski et al., 2005; Allison et al., 2010; Rasse et al., 2017), the latter two directly retrievable from satellite ocean colour inversion algorithms (acknowledging indirect estimates of beam attenuation have also been attempted, see Roesler and Boss, 2003).

The first attempts to use satellite ocean colour to estimate POC synthetically were made by Stramski et al. (1999). Using satellite remote sensing reflectance data, and empirical relationships between reflectance and particle backscattering on the one hand, and POC and particle backscattering on the other hand, basin-scale POC concentrations were quantified in the Southern Ocean. Global and regional satellite POC estimates quickly followed (e.g. Loisel et al., 2001, 2002; Mishonov et al., 2003; Gardner et al., 2006). At present, there are various algorithms designed to estimate POC from satellite ocean-colour remote sensing (Evers-King et al., 2017), including: empirical band ratio algorithms (Stramski et al., 2008); empirical band-difference algorithms (Le et al., 2018); backscattering-based algorithms (Stramski et al., 2008), using satellite estimates of backscattering from inversion algorithms, and empirical relationships between POC and backscattering (Stramski et al., 1999, 2008); algorithms using satellite estimates of backscattering and chlorophyll (Loisel et al., 2002); algorithms using satellite estimates of diffuse attenuation, and a two-step relationship between diffuse attenuation and beam attenuation, and the beam attenuation and POC (Gardner et al., 2006); and algorithms based on magnitude and spectral shape of the backscattering coefficient, and their relationships with the particle size distribution and POC (Kostadinov et al., 2009, 2016). Global validation has indicated some satellite POC algorithms perform with similar precision and accuracy to satellite ocean-colour chlorophyll-a algorithms (Swirgin and Stramski, 2015), highlighting POC as one of the carbon pools retrievable from satellite with high confidence. Intercomparison and validation exercises have demonstrated good performance by empirical band ratio algorithms, band-difference algorithms and algorithms using satellite estimates of backscattering and chlorophyll (Evers-King et al., 2017; Le et al., 2018; Tran et al., 2019), but that algorithm performance varies regionally (Allison et al., 2010), and that blending of region-specific algorithms may provide the best way forward for generating global POC products (Evers-King et al., 2017).

In a study by Evers-King et al. (2017), uncertainties in POC products were mapped globally for a range of POC algorithms, using fuzzy-logic optical water type mapping (Moore et al., 2009; Jackson et al., 2017a). They found that for some of the better performing algorithms, the majority of satellite pixels at global scale fell within an error range that is widely accepted by the ocean colour community (< 30%, Evers-King et al., 2017). Many of the approaches described above are designed for surface open-ocean waters, that are less optically complex when compared with surface coastal waters. However, increasing efforts are being made to develop satellite POC algorithms for surface coastal waters (Le et al., 2017; Tran et al., 2019). Approaches have been developed for extrapolating surface POC estimates in the open-ocean vertically, to quantify depth-integrated POC from space (Duforet-Gaurier et al., 2010).

4.1.3.2. Phytoplankton carbon from space. Owing to the central role of phytoplankton carbon in the OBCP, there are increasing efforts to develop and refine algorithms for deriving phytoplankton carbon from satellite data (Martinez-Vicente et al., 2017). Methods include those based on backscattering at a single wavelength, empirical relationships based on chlorophyll concentration, and methods based on allometric considerations combined with either phytoplankton absorption characteristics (Roy et al., 2017) or the spectral slope of backscattering (Kostadinov et al., 2016). Behrenfeld et al. (2005) produced estimates of phytoplankton carbon as a linear function of satellite estimates of particle backscattering at 443 nm (Martorena et al., 2002), following subtraction of a fixed background contribution to backscattering from non-algal particles. Further refinements to the approach, either to the linear relationship or by including a variable background contribution, have been proposed (Martinez-Vicente et al., 2013; Bellacci et al., 2016, 2018). Sathyendranath et al. (2009) developed a simple conceptual model to infer phytoplankton carbon as a function of chlorophyll-a concentration. The model was tuned to a large dataset of POC and chlorophyll-a concentration using quantile regression. Similar relationships have been proposed using other datasets (Buck et al., 1996; Marañón et al., 2014), but this was the first application of the model to satellite observations, for mapping phytoplankton carbon at large scales.

The last couple of decades have shown considerable progress and proliferation in satellite algorithms developed to identify various aspects of phytoplankton community structure (size structure, functional types, diversity, dominant phytoplankton types) from satellite data. The progress achieved and the prospects for further developments are outlined in a number of recent reports, reviews and scientific roadmaps (i.e. IOC, 2014; Bracher et al., 2017; Mouw et al., 2017). Synergy between ocean-colour products and SCIAMACHY data have been exploited (Losa et al., 2017), as well as synergy with other satellite products (Raitos et al., 2008; Palacz et al., 2013; Ward, 2015; Brewin et al., 2017a; Lange et al., 2018; Sun et al., 2019; Moore and Brown, 2020). An active grass-roots level initiative of interested scientists has assembled validation data, and carried out inter-comparison exercises (Brewin et al., 2011; Hirata et al., 2012; Kostadinov et al., 2017; Mouw et al., 2017). However, much of the work carried out in this domain, so far, has focussed on partitioning the phytoplankton community according to their contributions to the total chlorophyll concentration, rather than carbon concentration.

Only a small number of attempts have addressed the problem of partitioning phytoplankton carbon into fractions residing in various size classes or functional types. These include the work of Kostadinov et al. (2016) and of Roy et al. (2017), both of which are based on allometric considerations. The method of Kostadinov et al. (2016) uses the satellite-derived particle-backscattering signal as the starting point, and makes assumptions about how the total particle pool in the pelagic ocean is partitioned between phytoplankton and other particulate matter. The method of Roy et al. (2017) is based on phytoplankton absorption coefficient in the red part of the spectrum, and its modification by the size structure and concentration of phytoplankton. Both methods assume that the size spectrum of phytoplankton follows a power-law distribution, with cell numbers increasing with decreasing cell size.

Recently, Sathyendranath et al. (2020b) proposed an alternative approach to deriving phytoplankton carbon and size fractions. In their approach, they estimate the chlorophyll-to-carbon ratio of a phytoplankton population using a physiological model (Jackson et al., 2017b) driven by light (photosynthetically available radiation, PAR), modified to account for spectral distribution of light and spectral variations in light absorption by phytoplankton. The model then estimates the maximum chlorophyll-to-carbon ratio of a phytoplankton population attained as light approaches zero, which is allowed to vary as a function of the phytoplankton community structure (Brewin et al., 2015b), and
used estimates of the photosynthesis-irradiance parameters (Melin and Høegh-Krogstrup, 2011). Once the chlorophyll-to-carbon ratio is estimated, it becomes possible to transform a chlorophyll-based size-class model to one that is based on carbon. The advantage of this method is the consistency it shows between chlorophyll-based biomass estimates and carbon-based biomass estimates as well as with primary production, bringing us towards unification of photo-acclimation models of phytoplankton (which deal with changes in chlorophyll-to-carbon ratio) and primary production models. Such consistency is paramount in work towards closing the budget of the biologically-mediated carbon cycle in the ocean. Key to implementing such a model through remote sensing is a method to map photosynthesis-irradiance parameters, especially the photo-adaptation parameter, at space and time scales that are consistent with satellite-derived information on chlorophyll-a and PAR (Kovac et al., 2018b; Kulk et al., 2020). The method is also useful for exploring the physiological responses of phytoplankton to changes in the marine environment in the context of climate change, and consequent changes in pools and fluxes of carbon in the ocean.

To date, validation efforts of satellite phytoplankton carbon products (Martinez-Vicente et al., 2017; Burt et al., 2018) suggest higher uncertainties (e.g. >35%) than some standard ocean colour products (e.g. chlorophyll-a concentration) and some other carbon pools (e.g. POC, Evers-King et al., 2017). Part of the reason for this is that phytoplankton carbon is difficult to quantify in situ and consequently in situ phytoplankton carbon measurements have high uncertainties (see section 3.1.4).

4.1.3.3. Zooplankton carbon from space. It has traditionally been thought that surface-dwelling zooplankton cannot be observed directly from space, considering they do not require pigments for capturing light energy like phytoplankton do, owing to their preference for transparency, as a means to avoid predation, and their smaller abundance when compared with phytoplankton (less likely to influence bulk inherent optical properties). Until recently, only indirect approaches have been suggested for mapping zooplankton biomass from satellite. Strömberg et al. (2009) developed a model that related the flow of energy from primary production (derived using satellite data) to zooplankton biomass. The model was parameterised with primary production estimated from the SeaWiFS sensor and a subset of a global dataset of zooplankton biomass. The model was then validated with the remaining zooplankton data and used to produce global maps of zooplankton biomass. Similar models, based on the metabolic theory of ecology (Platt and Denman, 1977, 1978), driven by remotely-sensed variables (e.g. primary production and SST), have also been used for predicting biomass in higher trophic levels (e.g. Jennings et al., 2008). Satellite products of phytoplankton community structure also have the potential for predicting energy flow to higher trophic levels (Siegel et al., 2014). For example, in the Benguela ecosystem, it has been suggested that flagellates favour the growth of sardines, whereas diatoms favour the growth of anchovy (Cury et al., 2008). Bio-physical models to estimate zooplankton biomass, using satellite chlorophyll-a and SST as input, have been developed for regions of the Indian Ocean (e.g. Solanki et al., 2015). Mahesh et al. (2018) developed a regional model of mesozooplankton biomass in the Bay of Bengal, as a function of SST and chlorophyll-a concentration, that was subsequently applied to MODIS data to estimate mesozooplankton biomass. Other novel ways of deriving zooplankton biomass using satellite observations are also being pursued, for example, the use of productivity fronts (Draou et al., 2019). It has recently been demonstrated that certain zooplankton species, such as the red coloured copepod Calanus finmarchicus, that can congregate in large surface patches, are capable of modifying the colour of the ocean such that it is detectable directly from satellite (Bashedov et al., 2019). Such direct methods of zooplankton remote-sensing have the potential to provide new understanding on the distribution and behaviour of the species, with implications for ocean carbon cycling.

Furthermore, it has recently been shown that satellite active ocean colour (lidar) observations are capable of detecting zooplankton diel vertical migration patterns (Behrenfeld et al., 2019a), through comparison of day-night particle backscattering differences, with implications for monitoring zooplankton biomass. Direct zooplankton monitoring from space appears to be an emerging field of research.

4.1.3.4. Bacterial carbon from space. Submicron photosynthetic cyanobacteria are typically treated as part of the phytoplankton pool. Whereas algorithms have been developed to detect cyanobacteria from space exploiting their absorption characteristics or their ecological niches (IOC, 2014), we are not aware of satellite algorithms developed to detect other pigmented bacteria. Stomp et al. (2007) have provided information on the absorption spectra and major pigments of some pigmented bacteria, such as green sulphur bacteria, purple sulphur bacteria and purple non-sulphur bacteria, and Grossart et al. (2009) have isolated a blue pigment Glaucoldiato from two strains of marine Prostobacteria. No doubt algorithm development for such pigmented bacteria are hampered by the low magnitude of the signal, the difficulty with differentiating them from some phytoplankton groups, and moreover because their habitats are not typically targeted through remote sensing. For example, purple sulphur bacteria require anoxic, but illuminated locations where hydrogen sulphide accumulates.

Potential exists for exploitation of the backscattering signal from bacteria. Theoretical work based on homogeneous sphere models has indicated bacteria can contribute significantly to particle backscattering in the open ocean (Ulloa et al., 1994; Stramski et al., 2004), implying a route to direct detection. Belluccio et al. (2018) have estimated the backscattering coefficient for non-algal particles (inclusive of heterotrophic bacteria, detritus and viruses) using satellite ocean colour observations (see also Belluccio et al., 2019, for application of a similar approach to BGC-Argo float data), and observed large regional variations. However, the contribution of bacteria to backscattering, predicted theoretically using homogeneous sphere models, may underestimate the contribution by larger particles as indicated by coated sphere modelling (Meyer, 1979; Quijantes and Bernard, 2004; Dall’Omo et al., 2009; Organelli et al., 2018). Lack of comprehensive knowledge of the particular assemblage characteristics and size distribution for sub-micron particles is a significant impediment to identifying the sources of backscattering in the ocean (Stramski et al., 2004). To date, and with the exception of the night-time detection of bioluminescent bacteria such as Vibrio harveyi (Lapota et al., 1988) using satellite observations (Miller et al., 2005), there have not been attempts (to our knowledge) to directly estimate heterotrophic bacteria from space.

Indirect methods have been suggested and Grimes et al. (2014) provide an overview of some of these methods. Indirect methods typically focus on monitoring cyanobacteria (considered in this review under phytoplankton carbon) and vibrios, the latter tied with health and water quality, for instance, associated with outbreaks of cholera. Approaches have been proposed to map Vibrio cholerae using satellite visible and thermal radiometry (chlorophyll-a and SST), altimetry (sea surface height), and precipitation (e.g. Lobitz et al., 2006; Ford et al., 2009; de Magny et al., 2011). For a recent review on the topic of Vibrio cholerae detection using satellite data, the reader is referred to Racault et al. (2019) and Sathyendranath et al. (2020a).

Driving advanced statistical tools with multiple satellite ocean products could offer a route to mapping heterotrophic bacteria at global scale. For example, in a regional study, Larsen et al. (2015) extrapolated satellite products could offer a route to mapping heterotrophic bacteria at global scale. For example, in a regional study, Larsen et al. (2015) extrapolated satellite products could offer a route to mapping heterotrophic bacteria at global scale. For example, in a regional study, Larsen et al. (2015) extrapolated satellite products could offer a route to mapping heterotrophic bacteria at global scale. For example, in a regional study, Larsen et al. (2015) extrapolated
communities relies on the large-scale structure in ecological provinces of the ocean (Longhurst, 2007) that can be detected from space. Gomez-Pereira et al. (2010) showed that the composition of flavobacterial communities varied with ecological provinces detected from space.

4.1.4. PIC from space

Detection of PIC using satellite ocean-colour data primarily exploits the impact of calcium carbonate (CaCO$_3$) from coccolithophores (a group of phytoplankton in the nano-size range, typically around 5 μm in diameter) on the visible light field. Coccoliths are individual plates of CaCO$_3$ formed by coccolithophores that aggregate to form a coccosphere surrounding the cell. Because coccoliths are composed of inorganic CaCO$_3$ which has higher real index of refraction than organic cellular material, they are characterised by strong backscattering. In large densities, the coccoliths turn the colour of the water a pale turquoise, visible to the naked eye (Birkenes and Braarud, 1952; Berge, 1962).

The first observations of the impact of coccolithophore blooms on satellite ocean-colour imagery were made using Landsat MSS4 images (Le Fevre et al., 1983) and using Coastal Zone Colour Scanner (CZCS) images (Holligan et al., 1983). In the seminal paper by Holligan et al. (1983), patches of strong reflectance in CZCS imagery thought to be caused by coccolithophore blooms were confirmed with ship-based surveys. Similar combined satellite and ship-based studies followed using satellite images from the Advanced Very High Resolution Radiometer (AVHRR) (GREMPA, 1988). The first global distribution of coccolithophore blooms using CZCS satellite data were produced by Brown and Yoder (1994), revealing coverage of around 1.4 × 10$^8$ km$^2$, with 71% of this area located in subpolar waters. They identified blooms and classified these features using normalised water leaving radiance at 440 nm and 550 nm and using ratios between 440/520 nm, 440/550 nm and 520/550 nm. This work identified difficulties associated with misclassification of bright waters affected by sediment concentrations, and under-detection of CaCO$_3$ compared with in-situ measurements. Follow-on algorithms were developed for the SeaWiFS sensor, either adapting the Brown and Yoder (1994) technique (Iglesias-Rodríguez et al., 2002) or making use of near infrared bands (Gordon et al., 2001).

To bridge the gap between the CZCS and more current ocean colour satellites, and building on work of Groom and Holligan (1987), Smyth et al. (2004) used data from AVHRR and a semi-analytical algorithm to detect coccolithophores, again based on their high reflectances. Recently, Loveday and Smyth (2018) have produced a consistently calibrated 40-year-long data set of coccolithophore bloom occurrence derived from AVHRR.

Balch et al. (2005) developed a technique for use with MODIS sensors, using normalised water leaving radiance at 440 and 550 nm and a look-up table generated to calculate coccolithophore concentrations. The look-up table was based on the inversion of a semi-analytical model to separate and derive pigment concentration and coccolithophore concentration. The underlying relationships in this semi-analytical model capture the effects of coccolithophores on backscattering. This study highlighted the error in estimation of chlorophyll-a concentration when signal in the 440 and 550 nm bands was attributed to pigment concentration rather than backscattering by coccolithophores. Challenges in estimating of PIC using the Balch et al. (2005) approach include the use of the absolute magnitude of water-leaving radiance, which can be sensitive to errors from atmospheric correction. Similarly, elevated water-leaving radiance, a result of elevated backscattering, may be due to the presence of other constituents which increase scattering, such as diatom silica, or sediments. Because of the effect of sediments, these methods are difficult to employ in optically-complex (case 2) waters, where there are substantial non-algal scattering components, or in cases of blooms of high biomass or of species with complex intracellular structure, where it may be inappropriate to characterise phytoplankton scattering with the formulations of Balch et al. (2005) (Chami et al., 2006; Robertson Lain et al., 2014). Further techniques have been developed to reduce these false positive classifications in coastal zones, for example, by subtraction of stationary (background) structures to allow improved spectral detection of the coccolithophores (Shutler et al., 2010). Recent work has also added coccolithophore bloom spectra to optical water type classifiers (Moore et al., 2012). Hyperspectral sensors have also been used to detect coccolithophores, for example, based on differential absorption spectroscopy, applied to SCIAMACHY data (Sadeghi et al., 2012). For examples of each approach and a detailed history the author is referred to IOCCC (2014).

Many of these satellite-based methods for detecting coccolithophores from space have had the objective of mapping the aerial extent of the blooms. Some others (e.g. Sadeghi et al., 2012) focus on estimating the chlorophyll-a associated with coccolithophores, which is not necessarily indicative of the coccoliths (PIC) produced by these organisms. Only a few (e.g. Gordon et al., 2001; Balch et al., 2005; Mitchell et al., 2017; Kondrik et al., 2019) have attempted to take it a step further, towards estimation of PIC. This typically involves quantifying the number of coccoliths in surface waters, and making assumptions on per-coccolith PIC concentrations. Recent advancements in satellite-based PIC methods include: the development of band-difference PIC approaches (Mitchell et al., 2017), which have been seen to perform with better accuracy than standard methods (Balch et al., 2005) and found to be more resilient to errors in atmospheric-correction; methods that relate surface PIC to water-column integrated PIC (down to 100 m or to the euphotic depth, Balch et al., 2018), allowing quantification of PIC below the depths seen by the satellite (Höpkins et al., 2019); and quantification of the influence of coccolithophores on surface CO$_2$ concentrations and gas fluxes (Shutler et al., 2013; Kondrik et al., 2019).

4.2. Fluxes of carbon from space

4.2.1. Primary production from space

Following the development of the first satellite ocean-colour algorithms for deriving phytoplankton biomass, during the CZCS era, it became desirable to convert these fields of biomass into fields of primary production (Platt and Herman, 1983; Robinson, 1983; Perry, 1986; Sathyendranath et al., 1999b). In the 1980’s, there were some regional and basin-scale attempts to convert these fields directly (Smith et al., 1982; Brown et al., 1985; Eppley et al., 1985; Lorenzen et al., 1988). However, accurate conversion requires knowledge of, in addition to phytoplankton biomass (as indexed through satellite estimates of chlorophyll-a concentration), information on light availability (available through satellite PAR products), and the response of the phytoplankton to the available light, through parameters that describe the photosynthesis-irradiance curve (assimilation number (maximum photosynthetic rate) and initial slope of the curve; Platt et al., 1980). Some of these components can vary over the day, with depth, species composition, physiological state, and with the wavelength and angular distribution of light (Platt and Herman, 1983; Platt et al., 1990; Sathyendranath and Platt, 1989). Platt (1986) and Platt et al. (1988) described the application of this approach in the context of satellite remote-sensing, variants of which have had subsequent success in quantifying primary production at local and regional scales (e.g. Platt and Sathyendranath, 1988; Platt et al., 1991; Sathyendranath et al., 1989, 1991, 1995; Kuring et al., 1990; Morel, 1991; Morel and André, 1991). Platt and Sathyendranath (1993) provide a detailed user guide on the types of models used in computing marine primary production from remotely-sensed data on ocean colour, which illustrates, supported by more recent work (e.g. Sathyendranath and Platt, 2007; Sathyendranath et al., 2020a), that model variants conform to the same basic formulation, with the same set of parameters. Targeting the model parameters appears to be the key to accurate estimates of primary production from space (Platt et al., 1992).

The first global estimates of satellite primary production using CZCS were provided by Longhurst et al. (1995), who estimated global net oceanic primary production to be 45 to 50 Gt C yr$^{-1}$. Other global estimates using CZCS quickly followed (e.g. Antoine et al., 1996; Behnfeldt
and Falkowski, 1997), with most techniques converging at around 50 Gt C y\(^{-1}\) (see also Carr et al., 2006; Sathyendranath et al., 2019b), equivalent to primary production in the terrestrial biosphere (Field et al., 1998). Following the CZCS era, a variety of primary production models (including available light, absorbed light, inherent optical property based, and carbon based) have been applied to subsequent ocean-colour missions (SeaWiFS, MODIS and MERIS) for mapping primary production (e.g. Bosc et al., 2004; Behrenfeld et al., 2005; Smyth et al., 2005; Platt et al., 2008; Westberry et al., 2008; Tilstone et al., 2015; Arrigo and van Dijken, 2015), demonstrating regulation of global primary production by climate variability (Behrenfeld et al., 2001, 2006; Racault et al., 2017; Taboada et al., 2019; Kulk et al., 2020). Kulk et al. (2020) provide global primary production estimates (recently updated), using a 20-year ocean-colour data record (Sathyendranath et al., 2019a) together with a global database on photosynthesis-irradiance curve parameters (Bouman et al., 2018) and vertical biomass parameters (Platt et al., 1988; Longhurst et al., 1995) partitioned into ecological provinces and varied seasonally, and quantified global primary production to be around 50 Gt C y\(^{-1}\) over the period of 1998 to 2018.

Considerable resources have been invested by space agencies in round-robin comparisons and validation of products for estimating ocean primary production, for example, the NASA Primary Production Algorithm Round Robin series (PPARR; Campbell et al., 2002; Carr et al., 2006; Friedrichs et al., 2009; Saba et al., 2010, 2011; Lee et al., 2015a). These comparisons have emphasised that model performance varies with complexity and trophic state, and with region and season. The challenges in deriving trends in satellite primary production, the importance of minimising the uncertainties in model inputs, model parameters, and in the in-situ measurements used for validation, were also highlighted in the PPARR series. In general, satellite, depth-integrated primary production models have an uncertainty (root mean square deviation) of around 0.2 to 0.5 in log\(_{10}\)\(^{\text{EP}}\) space, when matched to and compared with in-situ data (e.g. Campbell et al., 2002; Friedrichs et al., 2009; Tilstone et al., 2009; Brewin et al., 2017d). Based on a formal error analysis, Platt et al. (1995) reported that uncertainties (standard error) in primary production were about 50% on primary-production estimates at a single pixel.

In addition to estimates of total primary production, there have been efforts to map primary production in different phytoplankton size classes from satellite data (Mouw and Yoder, 2005; Silio-Calzada et al., 2008; Uitz et al., 2009, 2010, 2012; Hirata et al., 2009; Brewin et al., 2017d; Curran et al., 2018). This is useful considering cell size influences many key processes in biogeochemistry and marine ecology (Chisholm, 1992; Marañón, 2009, 2015) and many marine biogeochemistry models use a size-based partitioning for phytoplankton (e.g. Aumont et al., 2003; Ward et al., 2012; Butenschon et al., 2016). Other satellite algorithms have been developed for computing new production (e.g. Sathyendranath et al., 1991; Goes et al., 2000; Coles et al., 2004; Tilstone et al., 2015), export production (see Section 4.2.2), and to account for the the inhibition of photosynthesis by ultraviolet radiation (Cullen et al., 2012).

One of the principal difficulties in modeling primary production from space, considering that many models conform to the same formulation (Sathyendranath and Platt, 2007; Sathyendranath et al., 2020b), is in the assignment of the photosynthetic parameters on a per-pixel basis. Various efforts have been made in this direction, including: direct use of in-situ data, by averaging data seasonally and regionally, using biogeochemical properties (Longhurst et al., 1995; Sathyendranath et al., 1995; Kulk et al., 2020), or interrogating the data using statistical methods such as nearest-neighbour techniques (Platt et al., 2008); and indirect methods that tie parameters to one (or more) environmental variable retrievable from satellite data, such as SST, light, chlorophyll and phytoplankton size structure (Eppley, 1972; Behrenfeld and Falkowski, 1997; Bouman et al., 2005; Claustre et al., 2005; Uitz et al., 2008; Sathyendranath et al., 2009; Saux Picart et al., 2014; Brewin et al., 2017d; Tilstone et al., 2017). Future advances in this direction rely on the continued production and development of in-situ datasets (Richardson et al., 2016; Bouman et al., 2018; Kulk et al., 2020), and improvements in satellite PAR PROPARs (Frouin et al., 2018). The use of ecosystem models run in data assimilation mode is another avenue to infer model parameters (Roy et al., 2012; DeVries and Weber, 2017). Improving the vertical parametrisations of satellite models through integration of data from autonomous platforms (e.g. BGC-Argo floats) is another exciting avenue of work. For example, Richardson and Bendtsen (2019) have demonstrated that the depth of the nutricline, that can be measured by attaching nutrient sensors to BGC-Argo floats, is useful for understanding vertical variations in primary production. For further discussions on future strategies to improving satellite-based ocean primary production models, the reader is referred to Lee et al. (2015b).

4.2.2. Export production from space

Estimating Export Production (EP) from satellite-derived properties requires that we establish how much of primary production is exported out of the surface layer. Most of the carbon generated by primary production is consumed by heterotrophs. The remaining carbon (net community production, NCP, defined as primary production minus community respiration) is, if aggregated over sufficiently large temporal and spatial scales, equivalent to EP, which is the amount of organic matter produced in the ocean by primary production that is not recycled (remineralised) before it sinks into the aphotic zone and which serves as the upper bound for carbon sequestration (Platt et al., 1989; Siegel et al., 2016). Currently, NCP and EP are not derived from physiological relationships, as is the case for primary production, since no such relationships are available for community respiration. Also, NCP and EP depend explicitly on the temporal and spatial domains over which they are defined. For example, in-situ measurements of mixed layer NCP cannot be converted directly to EP since the newly-produced biomass might be consumed before it can be exported to the aphotic zone. It is therefore not yet possible to define mechanistic relationships between EP and properties derived from remote sensing.

Instead, a common approach used by the scientific community to estimate EP from satellite derived products, is to rely on empirical correlations of in-situ measurements of vertical carbon fluxes, with properties that can be derived from satellite. These approaches can be traced back to early work where proxies of export (e.g. the ratio of new production to total production, defined as the f-ratio, or particle fluxes from sediment traps) were empirically related to primary production, chlorophyll-a concentration and depth (Eppley and Peterson, 1979; Suess, 1980; Betzer et al., 1984; Pace et al., 1987; Wassman, 1990), or to nutrient concentrations (Platt and Harrison, 1985; Harrison et al., 1987). One of the first techniques, that directly used satellite data, was proposed by Laws et al. (2000). In this approach the ratio of export production to primary production (denoted as the ef-ratio) is approximated based on empirical relationships with SST. Once the ef-ratio is computed, it can be used with satellite primary production to estimate EP. The study estimated global EP at between 11.1 and 20.9 Pg C y\(^{-1}\). Their empirical model was tuned using a food-web ecosystem model, rather than in-situ data. However, Laws et al. (2011) presented updates to the approach, where the ef-ratio was approximated using an empirical equation driven by SST and primary production, and tuned to in-situ data. Application of the model to satellite data suggested global EP of around 9–13 Gt C y\(^{-1}\). Other satellite-based empirical algorithms have been presented, for example, Prieur et al. (2005) modelled EP as an empirical function of SST, chlorophyll-a and primary production, and Henson et al. (2011) as an empirical function of SST and primary production. Using these empirical approaches, satellite-based global estimates of EP range between 4 and 12 Pg C y\(^{-1}\). Though proven useful, these empirical approaches are, by their very nature, limited in their ability to predict EP (Stukel et al., 2015; Palevsky et al., 2016).

Simple food-web based models, driven by satellite observations, that are more mechanistic in nature than the aforementioned empirical
approaches, have also been proposed as a means for estimating EP from space. The link between size and export is relatively well established, with large cells thought to contribute more to EP than smaller phytoplankton. For example, Tremblay et al. (1997) demonstrated the fraction of large phytoplankton (>5 μm) to total biomass to be linearly correlated with the f-ratio. These approaches have gained momentum owing partly to the development of satellite algorithms for deriving phytoplankton size structure (IOC, 2014). Siegel et al. (2014) used a food-web model, driven by satellite estimates of primary production (Westberry et al., 2008), eutrophic depth (Morel et al., 2007), and the slope of the particle size spectrum (Kostadinov et al., 2009), to predict the production of sinking zooplankton feces and algal aggregates, and consequently EP. Their model estimates global EP at 6 Pg C yr⁻¹. In an analysis of this approach, Stukel et al. (2015) found it to perform as well or better than purely empirical approaches, and suggest food-web relationships and grazing dynamics are crucial to improving the accuracy of export predictions made from satellite-derived products. Li et al. (2018) also found the Siegel et al. (2014) model to work reasonably well in the northern South China Sea. Despite this progress, considerable unexplained variance in export remains and needs to be explored to improve satellite EP estimates (Stukel et al., 2015). Moving towards more mechanistic relationships (e.g. DeVries et al., 2014) between satellite-derived products and EP is a major goal in the NASA EXPORTS programme (Siegel et al., 2016).

As more in-situ data become available, improvements in both empirical and food-web based satellite EP models are expected. One solution is to tune the models on a more regional basis (Sathyendranath et al., 1991; Li et al., 2018). This could be done through the use of biogeochemical provinces, for example, Longhurst regions (e.g. Reygondeau et al., 2013), with sufficient data richness to perform the tuning. Other avenues of investigation include: studying the phytoplankton seasonal cycles (phenology) from space (Platt et al., 2009; Racault et al., 2012), critical in food-web dynamics (Platt et al., 2003), over long-time periods, which may offer insight into inter-annual variability in EP (Racault, 2009); and consideration of the time-scales of variability, since ecosystems that vary on shorter timescales and exhibit more episodic production events are often expected to export a greater proportion of their primary production (Platt et al., 1989; Bueseler, 1998; Lutz et al., 2007; Henson et al., 2012; Brown et al., 2014).

Of the three export pathways, the majority of studies estimating EP from satellite data focus on gravitational sinking, only implicitly including, or in some cases neglecting, the physical pathway and the migration pathway. Recent evidence that these other pathways contribute significantly to the OBCP (Boyd et al., 2019), has led to efforts to incorporate them explicitly into satellite estimates, for example, by including the contribution of zooplankton vertical migration (Archibald et al., 2019; Behrenfeld et al., 2019a) and the contribution to carbon export from fluctuations in the mixed-layer (Stramska, 2016; Dall’Olmo et al., 2016). As we learn more about these pathways, perhaps exploring and exploiting the use of ocean physical datasets, we can seek to improve their representation and consequently the space-based estimates of EP.

4.2.3. Loss rates from space

Individually, the loss rates (grazing, respiration, sinking, excretion, mortality) cannot be retrieved directly from remote-sensing. However, methods have been proposed to derive the total loss rate (sum of the losses) using remote-sensing methods. These techniques typically involve monitoring changes in phytoplankton properties at a given location, over time. Zhai et al. (2008) implemented a method, proposed by Platt and Sathyendranath (2008), that involves taking a series of images of chlorophyll concentration that allow the calculation, by difference at successive time steps, of the rate of chlorophyll increase (difference between growth and total loss). The difference in the rate of chlorophyll increase and the photosynthetic rate (also estimated from satellite data), gives the total loss rate. The recovered loss rates showed clear seasonal cycles and abrupt shifts during the spring bloom in the North West Atlantic. Using a similar approach, Behrenfeld (2010) computed the rate of carbon increase (referred to as the net specific growth rate, the specific growth rate minus loss rate) from successive time steps in phytoplankton carbon concentrations retrieved from satellite data. The sign in net specific growth rate was subsequently used to monitor the balance between specific growth and loss rates, and used for interpreting phytoplankton blooms (Behrenfeld, 2014).

To make best use of this sequential time-series approach, one needs to correct for the effect of physical processes (advection and mixing; Zhai et al., 2008). This requires additional information on velocity components, and horizontal and vertical eddy diffusivity. The magnitude and importance of this correction is likely to vary regionally. One solution to addressing this is through Lagrangian-based analysis of satellite-data, to track the water mass, which can be achieved through combining satellite biological observations (ocean colour) with ocean circulation velocity fields derived from an operational model (Jonsson et al., 2009) or derived through physical satellite ocean observations (Lehahn et al., 2018; Sencioni et al., 2018). Other exciting avenues for determining information on loss rates from ocean colour include harnessing recent, and soon to be launched, geostationary ocean-colour satellites (see Section 7) together with diel optical techniques (Dall’Olmo et al., 2011; Kheireddine and Antoine, 2014).

4.2.4. Air-sea CO₂ transfer-related variables from space

Regional, indirect approaches have been proposed for estimating the exchange of CO₂ between the ocean and atmosphere from satellite data (see Section 4.1.1 and Land et al., 2015). As discussed in Section 4.1.1, these approaches involve forcing empirical equations or statistical models, typically tuned with in-situ observations, with satellite retrievals of SST, SSS and chlorophyll-a, individually or in combination (Ono et al., 2004; Sarma et al., 2006; Gledhill et al., 2008; Friedrich and Oschlies, 2009). For further reviews on the topic the reader is referred to Land et al. (2015), Land et al. (2019) and Shuter et al. (2019).

4.3. Uncertainties in space-based carbon estimates

4.3.1. Uncertainties in individual satellite products

To enable responsible and informed uptake of satellite carbon products by the user community, it is essential that uncertainties in individual carbon products be quantified, ideally on a per-pixel basis (Moore et al., 2009; Jackson et al., 2017a; Sathyendranath et al., 2017, 2019a). This helps algorithm providers to establish whether the products meet user requirements, and to identify areas where further development is needed. Two common approaches to quantify uncertainties are through validation of satellite products with independent in-situ data or through error propagation methods.

Determining uncertainties through validation of satellite products typically requires matching the satellite data in time and space with the in-situ observations, having characterised the uncertainties in the in-situ observations. Such comparisons will never be perfect, owing to the differences in the spatial scales between satellite and in-situ observations. Once a match-up dataset is produced, the two data streams can be compared quantitatively, using a common set of statistical tests (Brewin et al., 2015c; 2016; Seegers et al., 2018), for example, the root mean square deviation. With a sufficiently large dataset, the match-ups can be partitioned into trophic categories, regions, or by using other discretisation methods. One approach that has been gaining momentum is the use of optical-class-based, fuzzy-logic methods (Moore et al., 2009; Jackson et al., 2017a). In this method, metrics are calculated for all the match-up data in each optical class. By deriving the fuzzy membership probabilities of a given pixel’s spectrum to the set of optical classes, per-pixel metrics can be computed by weighing the metric in each class by the memberships of the pixel. This method of computing per-pixel uncertainties has been used for mapping uncertainties in satellite-based estimates of POC, phytoplankton carbon and phytoplankton
chlorophyll in different size classes or in different functional types (Brewin et al., 2017a; Evers-King et al., 2017; Martínez-Vicente et al., 2017). Further refinements to this approach could be made by incorporating functions that relate uncertainties to the known factors that might influence the quality of a satellite product (e.g. Land et al., 2018).

Determining uncertainties though error propagation techniques involves quantifying uncertainty in all components of the measurement equation, starting with instrument calibration, and propagating the errors through to the derived carbon products, using either the principles of formal analytical error propagation, or using statistical methods like Monte-Carlo techniques. Standardised frameworks exist (e.g. BIPM, 2008) and should ideally be used consistently across products. Error propagation techniques have been used previously in satellite primary production models (e.g. Sathyendranath et al., 1991; Platt et al., 1995; Brewin et al., 2017d) and phytoplankton carbon models (e.g. Kostadinov et al., 2009, 2016). For a comprehensive review on uncertainties in satellite ocean-colour products, the reader is referred to the recent IOCCG report on the topic (IOCCG, 2019).

### 4.3.2. Uncertainties in multiple satellite products

In the context of quantifying the ocean carbon budget, in which all the pools and fluxes have to fit in a mutually consistent way, it is important to not only consider the uncertainties in individual products, but to analyse multiple products to identify discrepancies and associated uncertainties with a view to close the ocean carbon budget in a satisfactory manner. This requires that we analyse each of the products in relation to all the other products, and see whether they hold together in a coherent fashion. Similarity, bio-optical closure among input data products is equally important. In the case of carbon products, we can illustrate this point with a couple of examples taken from the current status of the field. The bio-available, labile DOC pool is currently estimated to be about 0.2 Pg with annual production rate of 15–25 Pg C yr$^{-1}$ (Hansell, 2013). This yields a turnover rate of 100 times per year, or roughly once in every 3–4 days. The source of this labile pool is phytoplankton (Hansell and Carlson, 2013). According to Falkowski et al. (1998), the size of the phytoplankton pool is 1 Pg. With global primary production estimated to be about 50 Pg C yr$^{-1}$, we get a turnover rate of 50 per year for phytoplankton. We immediately note a discrepancy of a factor of two between the two turnover rates, which merits further investigation. How can phytoplankton, turning over once every week, generate a labile DOC pool that is turning over twice as fast? Clearly, this is an indication of uncertainty in the overall budget, that can be reduced only if the two turnover rates can be brought closer to each other, or if we find another explanation for the discrepancy (e.g. microbial activity at the interface between labile and semi-labile DOC enhancing the DOC rates). This example demonstrates that, to reduce uncertainties in the overall carbon budget, future efforts are needed to compare individual products against others, to check for consistency, and to reduce discrepancies.

### 5. Integrating satellite observations with models

As highlighted in Table 3, many components of the OBCP are not directly observable from space. Furthermore, the components that are observable are limited to conditions (e.g. a cloud-free atmosphere) that allow for satellite remote-sensing. One avenue to infer the hidden pools and fluxes, and to fill gaps in data, is through assimilation of satellite-based observations into ecosystem models. There are increasing efforts to integrate radiative transfer models into ecosystem models, such that the ecosystem models themselves become capable of simulating ocean colour data (e.g. Jones et al., 2016; Gregg et al., 2017; Dukiewicz et al., 2018, 2019). Integrating satellite observations with models can be used to improve model parameterisation, improve hindcasts and forecasts of biogeochemical pools and fluxes, and identify processes poorly represented by models, that can be subsequently improved in future model design (Fennel et al., 2019).

Using satellite data assimilation to improve model parameters is a powerful method of inferring rates not accessible from space. For example, Roy et al. (2012) assimilated satellite-based chlorophyll data into a nutrient-phytoplankton-zooplankton (NPZ) model to infer specific growth rates and mortality rates of phytoplankton. DeVries and Weber (2017) developed a global model of the OBCP that assimilates satellite primary production and particle size data with oceanographic tracer observations to constrain rates and patterns of organic matter production, export, and remineralization in the ocean. Their OBCP model combines an ecosystem model, an ocean circulation model and an organic matter transport and remineralisation model. Model parameters are adjusted by an iterative process to achieve optimal consistency with observed global-scale data sets of sinking POC fluxes, DOC concentrations and dissolved O$_2$ distributions. The model outputs an export flux of 9 Pg C yr$^{-1}$, and offers insight into the controls and distributions of various pools and fluxes not seen from space. Major challenges in model parameter optimisation using data-assimilation techniques are the often large number of parameters used in models, that some parameters are correlated, and that errors in some parameters can be compensated by errors in others. The success of these approaches can depend somewhat on the complexity of the model (Friedrichs et al., 2006).

Satellite data assimilation can also be used directly to improve hindcasts and forecasts of carbon pools and fluxes. The NASA Ocean Biogeochemical Model (NOBM), a coupled, atmosphere, ecosystem, radiative transfer and circulation model, has a history of being used for such purposes (Gregg, 2001, 2008). Decadal trends in phytoplankton biomass, primary production and phytoplankton composition have been estimated using the NOBM with ocean-colour data assimilation (Roussaux and Gregg, 2015; Gregg et al., 2017; Gregg and Rousseaux, 2019). The model has also been used to estimate pCO$_2$ and carbon dioxide fluxes (Gregg et al., 2014; Ott et al., 2015). The assimilation of satellite chlorophyll data into various ecosystem models has demonstrated improvements in carbon cycle variables (e.g. Ford and Barciela, 2017; Pradhan et al., 2019). Ciavatta et al. (2018) assimilated satellite-based chlorophyll data for four different phytoplankton groups (using the model of Brewin et al., 2017c), into the European Regional Seas Ecosystem Model (ERSEM) and showed improvements in the ability of the model to hindcast the air-sea flux of CO$_2$. Similar results were also found in a forecast simulation with a (pre)operational model (Skakala et al., 2018). A prerequisite for successful data assimilation is accurate assignment of uncertainties in the satellite observations and the model simulations. A point to note is that increased performance in carbon fluxes and pools using data assimilation might be accompanied by reduced performance in other model compartments (Ciavatta et al., 2018). For a comprehensive review on synergy between satellite and ecosystem model data, the reader is referred to the recent IOCCG report on the topic (IOCCG, 2020).

### 6. Bringing satellite observations into the limelight

As this review demonstrates, recent years have seen great progress in the use of satellite observations to detect or infer various pools and fluxes of carbon in the ocean as well as air-sea fluxes of carbon. Much of the information necessary to establish the OBCP is already accessible to remote sensing, and the potential exists for improving the products and for generating additional products (Table 3). Satellite observations meet the requirements for global coverage and high spatial and temporal resolution. With ocean-colour observations, the creation of long-term, climate-quality products has been a challenge, because of the finite life span of sensors in space, in combination with the differences in the characteristics of each sensor launched into space (e.g. wavebands), making it difficult to eliminate inter-sensor biases. But that situation is changing now, with bias-corrected, multi-sensor, error-characterised products being generated (Sathyendranath et al., 2019a), and the length of the continuous time series now exceeding two decades. The prospects for further improvements in the quality of the products and in
the length of the uninterrupted time series is excellent, with initiatives such as the ESA’s Sentinel and NASA’s PACE missions (Donlon et al., 2012; Remer et al., 2019). The access to many of the key components of the OBCP implies that we are now in a position to examine not only the influence of the ocean biota on the air-sea flux of carbon dioxide, but also to evaluate any potential impact of climate change on the integrity of marine ecosystems. When combined with in situ observing systems (Tables 1 and 2), we are in a position to close the biological carbon budget of the ocean with unprecedented rigour, and to develop a predictive, mechanistic understanding of the processes involved, rather than solely rely on empirical relationships.

Notwithstanding the benefits of integrating satellite remote sensing with ecosystem and climate models, as well as with autonomous observations and with other field datasets, as discussed above, producing a purely satellite-based carbon budget for the oceans, for the pools and fluxes that can be estimated, could be a useful way of independently verifying ecosystem and climate models. The international Global Carbon Project was established to produce a common and mutually agreed knowledge base to support policy debate and action to slow down and ultimately stop the increase of greenhouse gases in the atmosphere (http://www.globalcarbonproject.org/about/index.htm). Every year, a picture of the global carbon cycle is produced, including both its physical and human dimensions, together with the interactions and feedbacks between them, based on a combination of models and observations. The latest report (Friedlingstein et al., 2019) suggests the oceans currently absorb 2.5 Gt C year{\(^{-1}\)} (±0.6 Gt C year{\(^{-1}\}), close to the figure for the terrestrial biosphere (3.2 ±0.6 Gt C year{\(^{-1}\}), and accounting for the absorption of around 24% of the fossil CO\(_2\) emissions. The results also suggest in recent periods (2009–2018), the ocean CO\(_2\) sink appears to have intensified. This work emphasises the importance of the oceans in the global carbon cycle. Yet, large uncertainties still remain in how the oceans will respond to increased CO\(_2\) emissions in the future. The conclusions drawn from the latest Global Carbon project are from models which, at best, include satellite-observations implicitly in their simulations (e.g. through validation). One conclusion from the Friedlingstein et al. (2019) work is that the ocean models underestimate CO\(_2\) variability outside the tropics. DeVries et al. (2019) have also noted that ocean biogeochemistry models tend to underestimate inter-annual variability in ocean uptake of atmospheric carbon, compared with in situ observation-based methods. Similar comparisons with satellite-based observations are essential, to improve our understanding of the current state of the OBCP as well as model-based forecasts. Here, we argue that it is time for our community to bring satellite-based ocean carbon observations into the limelight. Our understanding of the ocean carbon cycle can improve with new opportunities in space, and it is our duty to use satellite observations to help strive towards a characterisation and understanding of the ocean carbon cycle, and its role in the Earth’s carbon cycle, that is better than anything that has been achieved before. This is an objective of the ESA Biological Pump and Carbon Exchange (BICEP) project.

7. New opportunities in space and future recommendations

In this section, we highlight a series of opportunities and make recommendations that can facilitate the advancement of satellite-based ocean carbon observations.

- **Satellite data records.** Over recent years, there has been large international investment in the production of satellite data records (Cazenave et al., 2019; Groom et al., 2019; O’Carroll et al., 2019; Vinogradova et al., 2019). These efforts typically involve stitching together satellite data from different platforms and sensors, correcting for differences in wavelengths and for systematic biases, and producing a seamless data record, suitable for monitoring long-term (climate) change. The European Space Agency (ESA) Climate Change Initiative (CCI) is one such example (Hollmann et al., 2013), where ocean data records have been produced for many of the input datasets needed for producing carbon-based products, such as ocean colour (OC-CCI; Sathyendranath et al., 2019a), SST (SST-CCI; Merchant et al., 2014, 2019), and sea-level (SST-CCI; Legeais et al., 2018). NASA and NOAA have also developed similar initiatives (e.g. NASA’s MEaSUREs programme and NOAA’s Climate Data Record Programme; Maritorena et al., 2010; Bates et al., 2016). Many of these programmes are developed with consideration of commitments made by space agencies and intergovernmental organisations, for continued support of operational ocean satellites in the coming decades (e.g. Sentinel Copernicus programme). These records provide the ideal platform for exploring long-term change in carbon pools and fluxes (e.g. Mélin, 2016; Mélin et al., 2017; Kulk et al., 2020).

- **Increased spectral resolution of ocean colour.** At present, the majority of operational ocean-colour satellite sensors are multi-spectral by nature, with select wavebands (typically 6–10) spread out over the visible spectral range, with each band having a spectral response function (band-width) in the order of 10–20 nm. By increasing the spectral resolution of the sensors, it is feasible to acquire additional information on the properties of phytoplankton and other water constituents, such as phytoplankton composition, by monitoring spectral features (e.g. from accessory pigment absorption) associated with different phytoplankton types (IOCCG, 2014; Cael et al., 2020). Direct measurements of phytoplankton composition would help improve OBCP analysis, owing to the differing roles of phytoplankton in the cycling of carbon (IOCCG, 2014). Additionally, increased spectral resolution in the UV and short wave infrared regions (as well as satellite polarimetry) are expected to improve atmospheric-correction (Frouin et al., 2019). Planned hyperspectral ocean-colour sensors like NASA’s PACE (Werdell et al., 2019; Remer et al., 2019) and ESA’s CHIME (one of the 6 High Priority Candidate missions for the Copernicus Expansion) will facilitate these developments, though challenges remain, for example, in treating the signal to noise ratio for narrower bands, and that the degrees of freedom in hyperspectral data are likely a lot less than the number of bands (Cael et al., 2020).

- **Increased spatial and temporal coverage.** New satellite platforms are opening the door to improved spatial and temporal coverage. Geostationary ocean colour sensors like the Korean GOCI satellite, have the capability to monitor diel cycles in phytoplankton properties, and dramatically enhance daily coverage of ocean colour (Choi et al., 2012). Under the most recent NASA Earth Venture Instrument solicitation, a geostationary, hyperspectral, ocean-colour instrument was selected. The instrument, Geostationary Littoral Imaging and Monitoring Radiometer (GLIMR) gathers data at hourly frequencies and is expected to launch in the 2026 time frame. Our community is well-positioned to take advantage of developments in micro- and nano- satellites (CubeSats; Vanhellemont, 2019). These small platforms can be launched cheaply, in large swarms, and, when cross-calibrated, have the potential to monitor the ocean with unprecedented spatial and temporal coverage. There is already an example of a successful launch of an ocean-colour CubeSat (e.g. SeaHawk; Schuler and Holmes, 2016). Tapping into high resolution satellites (e.g. Landsat and Sentinel-2) for improved monitoring of the carbon cycle in coastal waters and lakes is also an exciting avenue of research (Pahlevan et al., 2020).

- **New satellite approaches.** There have been successful examples of research using satellites originally designed for other purposes (e.g. atmospheric monitoring) for monitoring ocean properties relevant to the OBCP. For instance, the SCIAMACHY instrument onboard ESA ENVISAT, developed originally for monitoring atmospheric constituents, has found use for monitoring phytoplankton composition in the ocean (Bracher et al., 2009; Lusa et al., 2017). Similarly, the on-going ESA S5p + project is investigating the use of the atmospheric TROPOMI sensor onboard Sentinel-5p to retrieve new ocean-colour products such as phytoplankton functional types,
sun-induced marine fluorescence and light attenuation. Satellite lidar sensors, like those onboard the NASA and CNES CALIPSO satellite, designed for monitoring aerosol composition and clouds, have found use in ocean ecosystem monitoring (see reviews by Neukermans et al., 2018; Jamet et al., 2019). Similar applications could emerge from the Doppler wind lidar on-board ESA’s Earth Explorer mission Aeolus. With a long history in monitoring ocean properties from aircraft (Churnside, 2014), attaching these active ocean-colour sensors to satellite platforms can reveal parts of the ocean not viewable with passive systems (e.g. polar ocean in winter), monitor at night and see further in the depth dimension. This has led to new insight relevant to the OBCP, for example, resolving the full seasonal cycles of phytoplankton in the polar ocean (Behrenfeld et al., 2017), and viewing the diel migration of zooplankton (Behrenfeld et al., 2019a).

- **New statistical tools.** Driven by developments in big data (extremely large data sets), rapid advances in cloud-based computing (e.g. Google Earth Engine) and statistical techniques, such as artificial intelligence, are underway. These tools are well-placed to help address some of the challenges we face in monitoring the OBCP from space: for instance, developing techniques that use multiple satellite products to target pools and fluxes of carbon not currently feasible to monitor from space, or for integrating satellite data with field data, models and autonomous observations. Advanced statistical and machine learning techniques can be useful for highly non-linear inversion algorithm development (e.g. neural networks) and model data assimilation.

- **Open Science.** Open Science is the way forward for accelerating the research process and maximising impact. It is essential that our community follows an approach of data sharing and knowledge transfer. This involves promoting: open access datasets, in community standard data formats, available in open-access repositories; open sharing of methods, for reproducibility and knowledge transfer, with code documented in open software repositories; open source publications; and openly available education resources, promoting the importance of the work to younger generations of scientists, students and the general public.

- **In-situ observations.** Considering the development, evaluation, uncertainty computation, and calibration of satellite remote-sensing algorithms for carbon pools and fluxes are highly-dependent on in situ observations, increasing resources should be assigned to improving in-situ methods for measuring carbon pools and fluxes, embracing new technologies, autonomous platforms, and developing community protocols for in situ data gathering, with high emphasis on uncertainty characterisation. Additionally, and considering the importance of system vicarious calibration of ocean colour sensors and the validation of core satellite platforms used as input to carbon algorithms (e.g., remote sensing reflectance), sustained support should be placed on funding campaigns to obtain Fiducial Reference Measurements (Banks et al., 2020).

- **Bridging scales in observations and models.** Given the need to integrate satellite, in-situ and model-based datasets, for improving our understanding of carbon pools and fluxes, increasing emphasis should be placed of methods that can address the challenge in bridging the contrasting temporal and spatial scales applicable to these datasets.

- **Harmonising carbon products across different planetary domains.** Moving beyond the ocean carbon cycle towards the Earth’s carbon budget, increasing emphasis should be placed on harmonising satellite carbon products across different planetary domains (ocean, land, ice and air). This harmonisation should consider consistency in data product format and algorithm approach. For example, the fundamental equation of photosynthesis is the same for the ocean and land, offering scope to develop unified algorithms of primary production. This harmonisation and consistency will ultimately improve Earth’s carbon budget and estimates of cross-domain carbon fluxes.

- **Improved data distribution of satellite carbon products.** Significant improvements have been made in recent years by space agencies and operational satellite services to provide satellite data to end users in a variety of formats, and with a faster turn-around from data capture to delivery (Groom et al., 2019). Increasing emphasis should be placed on integrating satellite-based carbon products into these data streams, to ensure easy access to end users.

### 8. Summary

The ocean biological carbon pump (OBCP) transfers organic carbon from the surface to the deep ocean, helping to regulate atmospheric CO₂ concentrations. Monitoring the OBCP is vital to predicting changes in the Earth’s carbon cycle, and consequently changes in life and climate. In this paper, we have defined the pools and fluxes of carbon that make up the OBCP and reviewed in-situ (sampling from ships and autonomous platforms) and remote methods for monitoring them, with a major focus on satellite remote sensing. We discussed the advantages of producing a satellite-based carbon budget for the oceans, as an independent means for evaluating ecosystem models, but also advantages of integrating satellite-based observations with ecosystems models, to access the pools not currently retrievable from space, and also to integrate datasets (in situ and satellite) together in one platform and provide global coverage. We finished by touching on future opportunities in space, with the goal of bringing satellite observations to the attention of Earth System carbon research.

### Author Contributions

Robert J. W. Brewin wrote the original draft with significant input from Shubha Sathyendranath & Trevor Platt. All authors contributed to the writing and editing of subsequent versions.

### Funding

This work was funded through a European Space Agency (ESA) project “Biological Pump and Carbon Exchange Processes (BICEP)” and by the Simons Foundation Project “Collaboration on Computational Biogeochemical Modeling of Marine Ecosystems (CBIOMES)” (549947, SS). It was also supported by the UK National Centre for Earth Observation (NCEO). Additional support from the Ocean Colour Component of the Climate Change Initiative of the European Space Agency (ESA) is gratefully acknowledged. Tihó S. Kostadinov acknowledges support from NASA grant #80NSSC19K0297 and California State University San Marcos. Cécile S. Rousseaux acknowledges support from NASA grant NNH15ZDA01N-OB.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

During the writing of this review, we very sadly lost one of our co-authors Professor Trevor Platt FRS. Trevor’s input to this review was substantial. The topic of the ocean carbon cycle, and the use of optical measurements for monitoring and understanding carbon cycling, has been at the centre of Trevor’s research over the past five decades. His contributions to the topic have been pioneering. He has left an enduring mark on our field and his presence will continue in those he has worked with and alongside, and through his vast body of publications. He will be dearly missed. We dedicate this paper to the memory of Trevor.


Hansell, D.A., Carlson, C., 2013. Localised refractory dissolved organic carbon sinks in


Hess, J., 2008. Elemental and functional group composition of plankton from various


GREMPA, 1988. Satellite (AVHRR/NOAA-9) and ship studies of a coccolithophorid


R.J.W. Brewin et al.