



Inter-comparison of phytoplankton functional type phenology metrics derived from ocean color algorithms and Earth System Models



Tihomir S. Kostadinov ^{a,*}, Anna Cabré ^{b,1}, Harish Vedantham ^{c,2}, Irina Marinov ^b, Astrid Bracher ^{d,l}, Robert J.W. Brewin ^{e,m}, Annick Bricaud ^f, Takafumi Hirata ^{g,n}, Toru Hirawake ^h, Nick J. Hardman-Mountford ⁱ, Colleen Mouw ^{j,3}, Shovonlal Roy ^k, Julia Uitz ^f

^a Department of Geography and the Environment, Univ. of Richmond, 28 Westhampton Way, Richmond, VA 23173, USA

^b Department of Earth & Environmental Science, Univ. of Pennsylvania, Philadelphia, PA, USA

^c Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, The Netherlands

^d Alfred-Wegener-Institute for Polar and Marine Research, Buesestr. 24, 27570 Bremerhaven, Germany

^e Plymouth Marine Laboratory (PML), Prospect Place, The Hoe, Plymouth PL1 3DH, UK

^f Sorbonne Universités, UPMC Univ Paris 06, CNRS, Laboratoire d'Océanographie de Villefranche, 181 chemin du Lazaret, 06230 Villefranche-sur-Mer, France

^g Faculty of Environmental Earth Science, Hokkaido University, N10W5, Kita-Ku, Sapporo, 060-0810, Hokkaido, Japan

^h Faculty of Fisheries Sciences, Hokkaido University, 3-1-1 Minato-cho, Hakodate 041-8611, Japan

ⁱ CSIRO Oceans and Atmosphere, Indian Ocean Marine Research Centre, Crawley, Western Australia, Australia

^j Department of Geological and Mining Engineering and Sciences, Michigan Technological University, Houghton, MI 49931, USA

^k Department of Geography and Environmental Science, School of Agriculture, Policy and Development, University of Reading, Whiteknights, Reading RG6 6AB, UK

^l Institute of Environmental Physics, University Bremen, Otto Hahn Allee 1, 28359 Bremen, Germany

^m National Centre for Earth Observation, PML, Plymouth PL1 3DH, UK

ⁿ CREST, Japan Science and Technology Agency

ARTICLE INFO

Article history:

Received 16 December 2015

Received in revised form 9 November 2016

Accepted 17 November 2016

Available online 28 December 2016

Keywords:

Phytoplankton bloom

Phenology

Phytoplankton functional types

Microplankton

Ocean color algorithms

Inter-comparison

CMIP5 Earth System Models

Discrete Fourier Transform

ABSTRACT

Ocean color remote sensing of chlorophyll concentration has revolutionized our understanding of the biology of the oceans. However, a comprehensive understanding of the structure and function of oceanic ecosystems requires the characterization of the spatio-temporal variability of various phytoplankton functional types (PFTs), which have differing biogeochemical roles. Thus, recent bio-optical algorithm developments have focused on retrieval of various PFTs. It is important to validate and inter-compare the existing PFT algorithms; however direct comparison of retrieved variables is non-trivial because in those algorithms PFTs are defined differently. Thus, it is more plausible and potentially more informative to focus on emergent properties of PFTs, such as phenology. Furthermore, ocean color satellite PFT data sets can play a pivotal role in informing and/or validating the biogeochemical routines of Earth System Models. Here, the phenological characteristics of 10 PFT satellite algorithms and 7 latest-generation climate models from the Coupled Model Inter-comparison Project (CMIP5) are inter-compared as part of the International Satellite PFT Algorithm Inter-comparison Project. The comparison is based on monthly satellite data (mostly SeaWiFS) for the 2003–2007 period. The phenological analysis is based on the fraction of microplankton or a similar variable for the satellite algorithms and on the carbon biomass due to diatoms for the climate models. The seasonal cycle is estimated on a per-pixel basis as a sum of sinusoidal harmonics, derived from the Discrete Fourier Transform of the variable time series. Peak analysis is then applied to the estimated seasonal signal and the following phenological parameters are quantified for each satellite algorithm and climate model: seasonal amplitude, percent seasonal variance, month of maximum, and bloom duration. Secondary/double blooms occur in many areas and are also quantified. The algorithms and the models are quantitatively compared based on these emergent phenological parameters. Results indicate that while algorithms agree to a first order on a global scale, large differences among them exist; differences are analyzed in detail for two Longhurst regions in the North Atlantic: North Atlantic Drift Region (NADR) and North Atlantic

* Corresponding author.

E-mail address: tkostadi@richmond.edu (T.S. Kostadinov).

¹ Now at Dept. of Physical and Technological Oceanography, Marine Science Institute, CSIC, Barcelona, Spain.

² Now at Cahill Center for Astronomy and Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA.

³ Now at University of Rhode Island, Graduate School of Oceanography, Narragansett, RI 02882, USA.

Subtropical Gyre West (NASW). Seasonal cycles explain the most variance in zonal bands in the seasonally-stratified subtropics at about 30° latitude in the satellite PFT data. The CMIP5 models do not reproduce this pattern, exhibiting higher seasonality in mid and high-latitudes and generally much more spatially homogeneous patterns in phenological indices compared to satellite data. Satellite data indicate a complex structure of double blooms in the Equatorial region and mid-latitudes, and single blooms on the poleward edges of the subtropical gyres. In contrast, the CMIP5 models show single annual blooms over most of the ocean except for the Equatorial band and Arabian Sea.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Marine phytoplankton play an important role in the global carbon cycle via oxygenic photosynthesis and the biological pump (Field et al., 1998; Eppley and Peterson, 1979; Falkowski et al., 1998; IPCC, 2013; Siegel et al., 2014). Since the late 1990's, ocean color remote sensing has enhanced our understanding of oceanic ecosystems via continuous global estimates of total chlorophyll *a* concentration ([Chl], mg m⁻³, henceforth referred to simply as Chl), interpreted as a proxy for phytoplankton biomass (e.g. McClain, 2009; Siegel et al., 2013). However, total Chl does not provide a full description of the ecosystem. Phytoplankton have different morphological (size and shape) and physiological (growth and mortality rates, response to nutrient, temperature and light conditions) characteristics and different resulting biogeochemical and ecological roles (e.g. silica or iron requirements, calcification, sinking rates, feeding characteristics) and are thus grouped accordingly into phytoplankton functional types (PFTs, e.g. IOCCG, 2014). Phytoplankton community structure influences many fundamental components of the marine biogeochemical cycle, including: phytoplankton physiology; nutrient uptake; nutrient cycling; growth rates; metabolic rates; deep-ocean carbon export; and the transfer of energy through the marine food web (IOCCG, 2014). Therefore, detailed characterization of PFTs, and not only total Chl, is required to develop predictive understanding of the ocean's role in climate on various time scales (e.g. Le Quéré et al., 2005; Hood et al., 2006; Stock et al., 2014) and inform climate models. One of the primary distinguishing characteristics of the different PFTs is cell size, which is considered to be a master trait (Marañón, 2015) and is correlated to first order with biogeochemical function (e.g. Le Quéré et al., 2005). Size partitioning has been used as a first-order proxy for PFT classification (e.g. Vidussi et al., 2001; Le Quéré et al., 2005; Uitz et al., 2006; Kostadinov et al., 2010).

Satellite remote sensing provides a comprehensive observation method to characterize the global spatio-temporal distribution of PFTs (e.g. McClain, 2009; Siegel et al., 2013). Space-borne platforms can provide continuous sampling at the required resolution in time and space in order to facilitate the development of more complex "dynamic green ocean models" (Le Quéré et al., 2005) that include multiple functional types and resolve important biogeochemical processes (IOCCG, 2014, Ch. 1, Sect. 1.5). Multiple satellite bio-optical algorithms for the retrievals of various PFTs have been developed in the last decade as a result. One class of algorithms is based on total abundance and the premise that smaller cells are associated with oligotrophic conditions whereas larger cells are associated with eutrophic conditions (Chisholm, 1992) – such algorithms are described by Brewin et al. (2010), Hirata et al. (2011) and Uitz et al. (2006). Another class of algorithms relies on various spectral features. The PHYSAT algorithm exploits second-order anomalies of reflectance spectra (Alvain et al., 2005; Alvain et al., 2008), whereas several other algorithms are based on either absorption (Bracher et al., 2009; Ciotti and Bricaud, 2006; Mouw and Yoder, 2010; Roy et al., 2011; Roy et al., 2013), or backscattering (Kostadinov et al., 2009; Kostadinov et al., 2010; Kostadinov et al., 2016), or a hybrid of absorption and backscattering (Fujiwara et al., 2011).

Brewin et al. (2011) conducted the first systematic inter-comparison of PFT algorithms designed to identify "dominant" PFTs in the oceans. With the increasing publication of new PFT algorithms (IOCCG, 2014),

an international team of PFT algorithm developers and scientists was tasked to perform a follow-up inter-comparison exercise (Hirata et al., 2012; Hirata, 2015); this study reports results from a component of this inter-comparison project. A summary of the available algorithms and their technical basis can be found in Table 1 (also see IOCCG, 2014). The various algorithms use different PFT definitions and retrieve different variables that are based on various sets of assumptions, and hence are not necessarily directly comparable. Some retrieve several taxonomic groups, others – size fractions based on Chl or volume (Table 1; IOCCG, 2014). PFT algorithms often aim to quantify the size structure of the phytoplankton population by defining three phytoplankton size classes (PSCs) – picoplankton (<2 µm), nanoplankton (2–20 µm), and microplankton (>20 µm) (Sieburth et al., 1978). This is justified because size is considered a master trait (e.g. Marañón, 2015), but we caution that differences exist between PFTs and PSCs, even though this terminology is often used interchangeably.

Here we compare the algorithm outputs in terms of a key emergent property: phytoplankton phenology. Since seasonal cycles are a key property of ecosystems, it is important to assess to what degree different algorithms agree in terms of phenology, i.e. how consistently they capture the annual progression of phytoplankton blooming and subsequent senescence. If the timing of a bloom were slightly shifted between two data sets, direct comparison of the variables at each time step would yield disparate and meaningless results, whereas phenological analysis will identify the offset in timing (Platt et al., 2009). The amount of algorithms spread about an ensemble mean can be indicative of our confidence in retrieving a certain phenological parameter (e.g. timing of annual bloom), and overall results of the comparison can guide further algorithm improvements.

We use the Discrete Fourier Transform (DFT) to first model the seasonal cycle as a summation of sinusoids derived from the annual frequency band and its harmonics (integer multiples). We then quantify phenological parameters of interest using the modeled seasonal cycle. The phenology inter-comparison is based on global ocean color data (SeaWiFS and SCIAMACHY) for the period 2003–2007, using microplankton fraction or the most similar available variable from each participating PFT/PSC algorithm (Table 1). Increases in the absolute or fractional amount of large phytoplankton or diatoms were considered here to define a bloom, which is consistent with the established ecological idea that higher chlorophyll concentrations are associated with eutrophy and a relatively higher dominance of large phytoplankton (e.g. Chisholm, 1992; Loisel et al., 2006; Kostadinov et al., 2010; Marañón, 2015). We quantify the timing, amplitude and duration of blooms, as well as the fraction of variance explained by the modeled seasonal cycle. We compare these phenological parameters among the PFT algorithms. The same phenological parameters are also compared for the NASA chlorophyll product (OC4v6 Chl), as well as contemporary diatom carbon biomass provided by seven CMIP5 Earth System Models (ESMs). Our goal is not to rank the satellite algorithms and CMIP5 models in terms of quality; rather we strive to identify spatial patterns of agreement and disagreement among the algorithms in an effort to guide future improvements. Additionally, the comparison to the ESM ensemble is aimed at guiding future improvements in biogeochemical and climate modeling, a key goal of the Earth system science community (IPCC, 2013).

Table 1

Overview of the PFT/PSC algorithms used and the relevant variable(s) from which phenological parameters were derived. SW10 refers to SeaWiFS monthly mapped 9 km global $R_{rs}(\lambda)$ data for the 2003–2007 period (Section 2.1). Monthly data for 2003–2007 from SCIAMACHY on ENVISAT was only used for PhytoDOAS and has 1/2 degree spatial resolution. The variables provided by most algorithm are dimensionless, i.e. fractions of a total, most commonly – chlorophyll-a concentration (Chl). This is indicated by a dash in the table. If *in situ* data were used in algorithm development, the region from which the data came is indicated. N/A means no *in situ* data were directly used in the algorithm development (not including validation) (see references for details).

Algorithm publication(s)	Acronym	Variables analyzed	Units	Input data	Algorithm class/basis	Variables retrieved	Region of development
Alvain et al. (2005, 2008)	PHYSAT	Frequency of detection of diatoms	% of days	SW10	$R_{rs}(\lambda)$ second-order anomalies (radiance-based)	Multiple taxonomic PFTs	North Atlantic; Equatorial & Tropical South Pacific; Southern Ocean
Bracher et al. (2009), Sadeghi et al. (2012)	PhytoDOAS	Diatoms Chl	mg m^{-3}	SCIAMACHY	Differential absorption from hyperspectral data (absorption-based)	Multiple taxonomic PFTs	Uses PFT-specific $a_{ph}(\lambda)$
Brewin et al. (2010)	BR10	Microplankton – fraction of Chl	–	SW10	Abundance-based	Size structure	Atlantic Ocean
Ciotti and Bricaud (2006), Bricaud et al. (2012)	CB06	$1 - S_f$, where S_f = fraction of small phytoplankton	–	SW10	Absorption-based	Size structure	Global (<i>in situ</i> data used for picoplankton basis vector)
Fujiwara et al. (2011)	FUJI11	Microplankton – fraction of Chl	–	SW10	Absorption- and backscattering-based	Size structure	Arctic-North Pacific
Hirata et al. (2011)	OC-PFT	Microplankton – fraction of Chl	–	SW10	Abundance-based	Size structure	Global (coastal and shelf waters excluded)
Kostadinov et al. (2009, 2010)	KSM09	Microplankton - volume fraction	–	SW10	Backscattering-based	Size structure	N/A
Mouw and Yoder (2010)	MY10	S_{fm} , fraction of large phytoplankton	–	SW10	Absorption-based	Size structure	Global
Roy et al. (2011, 2013)	ROY13	Microplankton – fraction of Chl	–	SW10	Absorption-based	Size structure	Global
Uitz et al. (2006)	UITZ06	Microplankton – fraction of Chl	–	SW10	Abundance-based	Size structure	Global (case-2 waters excluded)
O'Reilly et al. (1998, 2000)	OC4v6	Chl	mg m^{-3}	SW10	Band-ratio $R_{rs}(\lambda)$ based (radiance-based)	Chl	Global

2. Data and methods

2.1. Input satellite data

All algorithms with the exception of PhytoDOAS use monthly global 9 km Level 3 mapped SeaWiFS remote-sensing reflectance, ($R_{rs}(\lambda)$, reprocessing R2010.0) from January 2003 to December 2007 as input (60 monthly maps total). These data as well as the corresponding monthly OC4v6 Chl data (O'Reilly et al., 2000) and monthly composites of daily averaged photosynthetically available radiation (PAR, mol photons m^{-2} day $^{-1}$) from the same reprocessing were downloaded from the NASA Ocean Biology Distributed Active Archive Center (OB.DAAC) maintained by the Ocean Biology Processing Group (OBPG) (<http://oceandata.sci.gsfc.nasa.gov/>) (NASA Goddard Spaceflight Center, 2010). $R_{rs}(\lambda)$ data were processed by the individual algorithm providers. The hyperspectral PhytoDOAS algorithm is based on Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) level-1 top-of-atmosphere radiance data. SCIAMACHY was a satellite sensor with a native pixel size is 30 km by 60 km which operated from 2002 to 2012 on the ENVISAT satellite. These processing details were agreed upon by the International PFT Inter-comparison Project Team (Hirata et al., 2012). SeaWiFS Chl data were analyzed in the same way as the PFT algorithms data for comparison purposes. PAR data were used for verification of the DFT phenological algorithm (Supplement Part 1).

2.2. PFT/PSC algorithm output pre-processing

The PFT/PSC algorithms were used to derive phenological parameters using the variable most closely corresponding to either microplankton Chl [mg m^{-3}] or microplankton/large phytoplankton fraction [% of total Chl]. The rationale behind this choice is 1) blooming/more eutrophic conditions are on average characterized by an increase in total

and fractional large phytoplankton biomass; and 2) this is the most common variable among all available algorithms. Table 1 summarizes the respective variables used in the phenological analysis, indicates the acronym used here for each algorithm and provides additional relevant information. For additional algorithm methodologies details, see the references in Table 1 and IOC/CG (2014).

Monthly data from all algorithms were down-sampled to 1 degree resolution using two-dimensional convolution with a 12×12 top hat averaging kernel (2×2 in the case of PhytoDOAS due to its different resolution). Missing data in the original resolution were ignored in the averaging; however, if <50% of the pixels being averaged were valid data, the pixel in the down-sampled image was assigned a missing data value. For PhytoDOAS, even a single valid pixel of the four being averaged produced a valid pixel in the down-sampled image. The registration of PhytoDOAS images was changed from grid/node to cell/pixel (NCEI, 2015) in order to match all other down-sampled imagery. All variables were down-sampled in linear space with the exception of Chl and the PhytoDOAS data, which were down-sampled in log10 space, since Chl values tend to vary lognormally spatially (Campbell, 1995). The log space spatial average was weighted appropriately for any present zeros, which cannot participate in a log average (Habib, 2012). Note that taking an arithmetic average in log space approximates the median of the data in linear space (Campbell, 1995).

2.3. Phenological parameters via Discrete Fourier Transform (DFT); metrics of algorithm and model inter-comparison

An increase in the absolute or fractional amount of large phytoplankton or diatoms is considered a bloom, and the bloom peak is considered the maximum of these values, respectively. A time series of each algorithm's relevant microplankton or diatom variable (Table 1) was constructed at each pixel at 1-degree resolution. Data were gap-filled temporally by linear interpolation (no extrapolation was applied).

If >45% of the data points were missing or if there was a continuous run of missing data longer than 8 months anywhere in the time series, data for that pixel was not used in the analysis. The mean was subtracted from each time series. Interannual variability is not explicitly considered in this study; however, possible secular trends are removed by detrending (by subtracting a least-squares line fit to the data), and other interannual variability in the study period of 2003–2007 is taken into account implicitly because the DFT is computed over the entire time series. If data were missing at the edges of the time series, they were filled with zeros. The DFT was then used to transform the time series to the frequency domain. The Fourier coefficients at a frequency of one cycle per year ($f = 1 \text{ yr}^{-1}$) and all its available harmonics (integer multiples) were used to model the seasonal cycle as a sum of sinusoids of varying phases and amplitudes. Supplement Part 1 describes the details of the DFT analysis and seasonal cycle modeling, and Supplement Fig. S1 illustrates an example modeled seasonal cycle.

Peak analysis was performed on the modeled annual cycle signal using the MATLAB® (R2014b) Signal Processing Toolbox® routine `findpeaks` in order to determine the timing of the local maxima, the signal height (the value of the signal at the peak) and the width of the signal at half-height. The modeled signal minimum value was subtracted from the signal before peak analysis to ensure correct height determinations. Signal edge effects were taken into account. In order to avoid detection of small secondary peaks (many of which can be artifacts of the modeling), only peaks whose prominence was >10% of the signal range (maximum minus minimum value) and which were at least 2.5 months apart from each other were detected. Prominence here is equivalent to topographic prominence and can be thought of as the intrinsic height of the peak relative to other nearby peaks. The same phenological analysis was applied to diatom carbon output from 7 CMIP5 models – details of the methodology and model information are provided in Supplement Part 2 and in Cabré et al. (2015). The following phenological parameters and inter-comparison metrics were derived from the peak analysis for both the PFT algorithms and the CMIP5 models:

1. **Seasonal amplitude** of the primary bloom, determined as half the height of the most prominent (highest) peak (Supplement Fig. S1), was inter-compared qualitatively because variables are on different scales (even among algorithms that have the same units, the methodologies are different).
2. The **month of maximum** of the primary bloom, determined as the month where the DFT-modeled seasonal cycle is maximum. The ensemble mean for the algorithms and CMIP5 models (calculated separately for each ensemble) was used to quantitatively compare the month of maximum of each algorithm to the ensemble mean month of maximum for all algorithms. We also compared the Chl and the ensemble mean CMIP5 model months of maxima. Variances in month of maxima were quantified for the algorithms and the CMIP5 models separately. The month of maximum was averaged across algorithms and models using circular statistical methods to ensure a properly estimated mean and variance (Supplement Part 3).
3. **Duration** of the primary bloom (in days), determined as the width of the most prominent peak at the peak's half-height level. Ensemble mean and standard deviation of bloom duration were calculated for the CMIP5 models and algorithms, and individual algorithm durations were qualitatively compared.
4. **Percent seasonal variance**, i.e. fraction of the data variance explained by the modeled seasonal cycle as opposed to other processes (e.g. one-time events, multiannual processes, and inter-annual variability due to climate oscillation modes like ENSO) and noise. An area in which this fraction is very high is characterized by a very clean seasonal signal i.e. little variance contribution by other processes. It is calculated as the sum of power at $f = 1 \text{ yr}^{-1}$ and its harmonic frequencies (Supplement Part 1), divided by total variance of the input data. The ensemble mean among the algorithms and models was calculated.

At least three algorithms (or two CMIP5 models) were required to participate in the ensemble means of month of maximum, percent seasonal variance and bloom duration for the ensemble statistics to be considered valid. If present, the second most prominent peak, representing a possible secondary bloom, was also characterized by using the above phenological parameters and the following was also derived: difference in months between the primary and secondary bloom, and relative prominence of the secondary blooms (ratio of the prominence of the secondary peak to the prominence of the primary peak). The fraction of satellite algorithms or CMIP5 models exhibiting two annual peaks was mapped, indicating in which regions the majority of algorithms or models agree that there is a secondary bloom. The difference in months between the maxima of the primary and secondary blooms was compared qualitatively among the algorithms. Finally, regionally binned analysis was performed for the following Longhurst (1998) provinces 1) Westerlies – North Atlantic Drift – NADR and 2) North Atlantic Subtropical Gyre West – NASW. Available data in these regions were spatially averaged and the resulting single time series per region and algorithm/model were analyzed. Methodological details of this analysis are provided in Supplement Part 4.

3. Results and discussion

3.1. Seasonal amplitude

The seasonal amplitude (Fig. 1) of the relevant PFT variables (Table 1) quantifies the strength of the seasonal cycle at a particular location. The tropical ocean and the oligotrophic subtropical gyres (defined here as the regions delineated by the climatological SeaWiFS Chl = 0.08 mg m⁻³ isoline and having Chl values less than this value) were generally characterized by low seasonal amplitudes of Chl and microplankton across all the algorithms. However, according to most algorithms, the southern edge of the South Pacific gyre, at around 30°S, was characterized by a band of higher seasonal amplitude, particularly evident in the UITZ06 and ROY13 data. Notably, in BR10 and UITZ06, some of the areas of lowest amplitude for microplankton are just equatorward of the Chl isoline delineating the gyre. Similarly, in ROY13 the gyres are not actually the places with the lowest amplitudes; rather they occur just equatorward of them in the Pacific. According to most algorithms, the highest seasonal amplitudes occur at the temperate and subpolar latitudes and in coastal zones, particularly the North Atlantic, the Northwestern Pacific, and in the monsoon-driven upwelling region off of the Arabian Peninsula. Circumpolar bands of high and low seasonal amplitude characterize the Southern Ocean, but the spatial details of the bands differ across algorithms. Overall, the large-scale spatial patterns of seasonal amplitude are similar, across algorithms of the same type. The CB06 algorithm retrieved almost no valid data over the gyres, and the PHYSAT data sparsity precludes meaningful DFT analysis over most areas; thus no PHYSAT map was included in Fig. 1 and subsequent maps (see Section 3.6 for regionally binned analyses of PHYSAT data).

3.2. Percent seasonal variance

The ensemble-mean percent seasonal variance for all 10 PFT algorithms (Fig. 2A) reveals several oceanic zones where the seasonal cycle is particularly clean/reproducible. This is especially prominent at the poleward boundaries of the subtropical gyres, around 30–40° latitude in both hemispheres, where 70–80% of the signal variability is explained by the seasonal cycle represented by the DFT. These highly-seasonally variable regions correspond to sharp transitions in surface Chl, as observed in *in situ* observations and satellite data (e.g. Glover et al., 1994). In the Pacific this feature is known as the transition zone chlorophyll front (TZCF), which migrates from 30–35°N in winter to 40–45°N in summer; its migration is due primarily to wind-driven seasonal variations in Ekman pumping and Ekman advection of nutrients

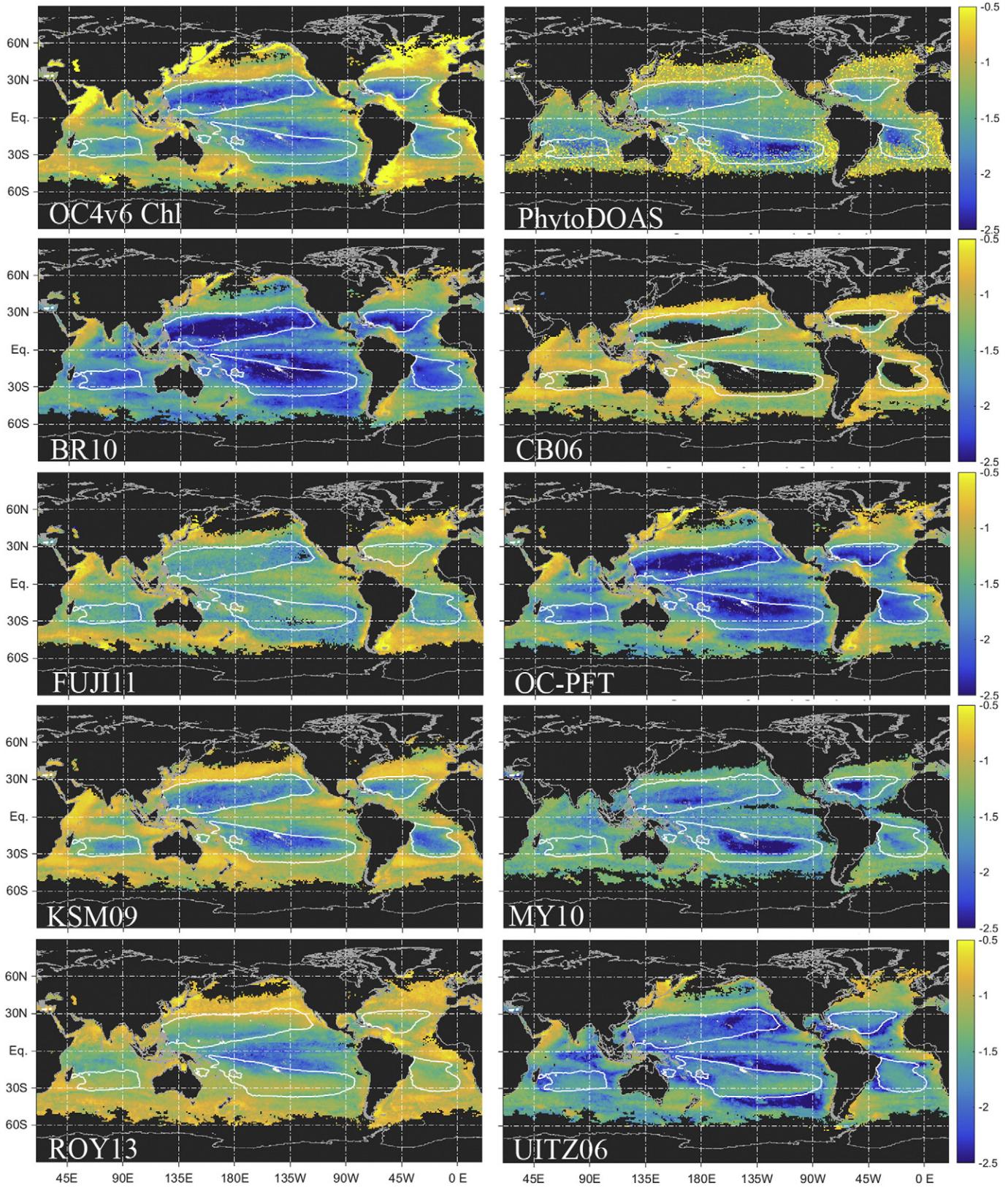


Fig. 1. Seasonal amplitude of Chl and the large phytoplankton/diatoms variables of 9 PFT satellite ocean color algorithms (Table 1). PHYSAT is not shown due to data sparsity. The same logarithmic color scale applies to all maps, and units are as indicated in Table 1. The isoline of climatological Chl = 0.08 mg m^{-3} is plotted as a solid white contour. All pixels where valid phenological analysis was performed are mapped.

(e.g., Bograd et al., 2004). The equivalent North Atlantic seasonally-stratified subtropics coincide with the mid-latitude biome of Levy et al. (2005) and are described as a nutrient-limited regime.

Equivalent seasonally-stratified, nutrient-limited subtropics with high seasonal variability are present in the Southern Ocean band around 30°S.

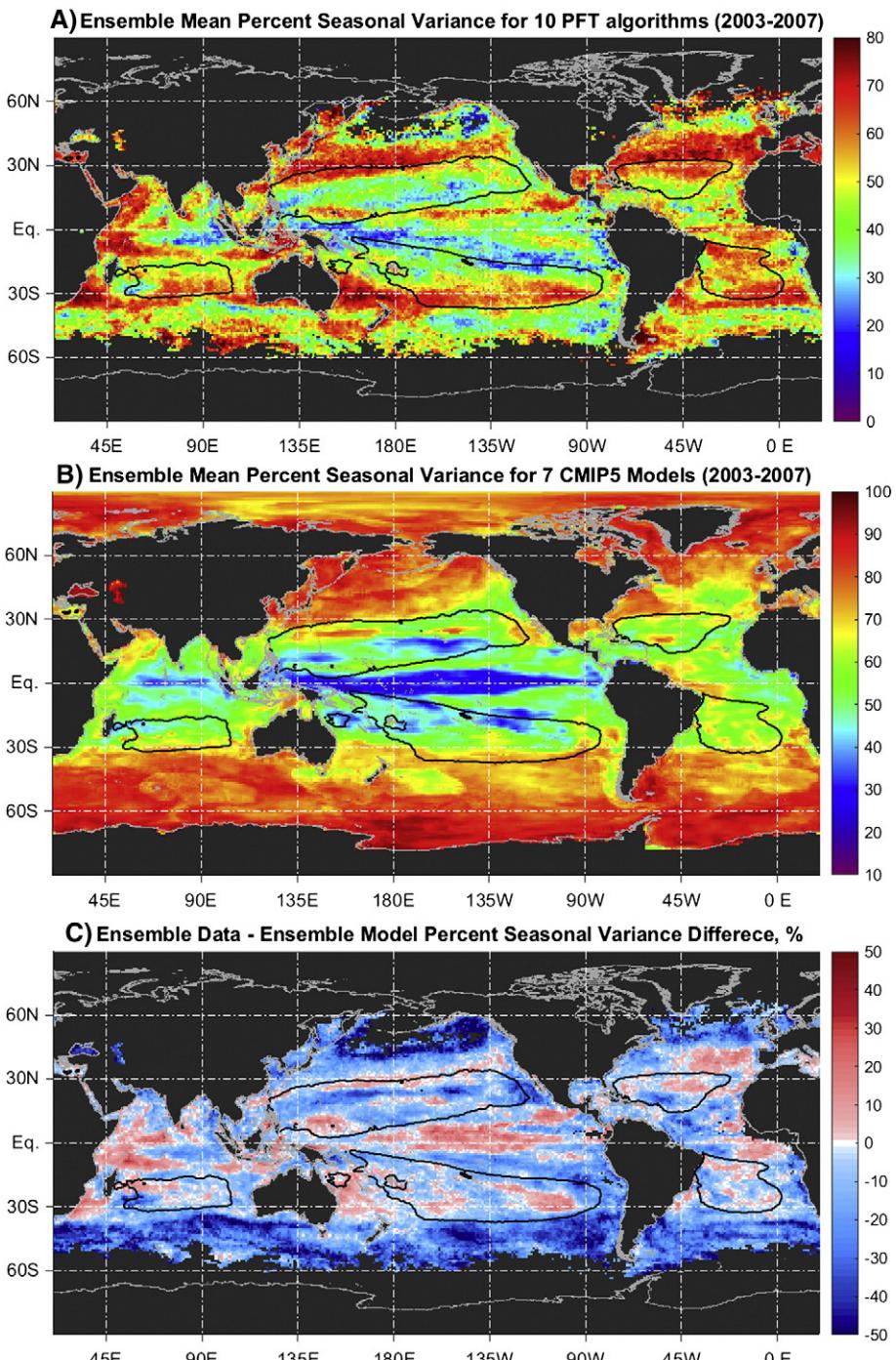


Fig. 2. Ensemble mean percent seasonal variance for the 2003–2007 period for A) the 10 PFT algorithms (Table 1) and B) the 7 CMIP5 models (Table S1). C) The difference in percent seasonal variance between the satellite data and the models (positive difference means satellite data percent seasonal variance is larger than the model value). The isoline of climatological Chl = 0.08 mg m⁻³ is plotted as a solid black contour.

Detailed analysis of the KSM09-based carbon biomass from SeaWiFS (Kostadinov et al., 2016) and CMIP5 model output shows that the regions of high percent seasonal variance of Fig. 2A exhibit strong and reproducible seasonality and are mainly dominated by nano- and microphytoplankton during the bloom months, and by picophytoplankton during the low-biomass summer months (Cabré et al., 2016). These ensemble mean PFT-based results are consistent with the Chl-based analysis of Sapiano et al. (2012) who also determined that the poleward fringes of the subtropical gyres have the best seasonality statistical fits.

In contrast to the satellite data, CMIP5 model diatom biomass exhibits smoother spatial variability of percent variance explained by the seasonal cycle (Fig. 2B, model ensemble mean; Fig. 2C, model-data difference map), with much broader regions characterized by >60% of variance due to the annual cycle. The percent variance explained by the seasonal cycle is a much stronger function of latitude (and hence the seasonality of insolation) in models. Poleward of 40°S/40°N, models show stronger seasonal variance compared to satellite data. In contrast, the Equatorial regions, especially in the Pacific and Indian Oceans, have lower percent annual variance as compared to the satellite data.

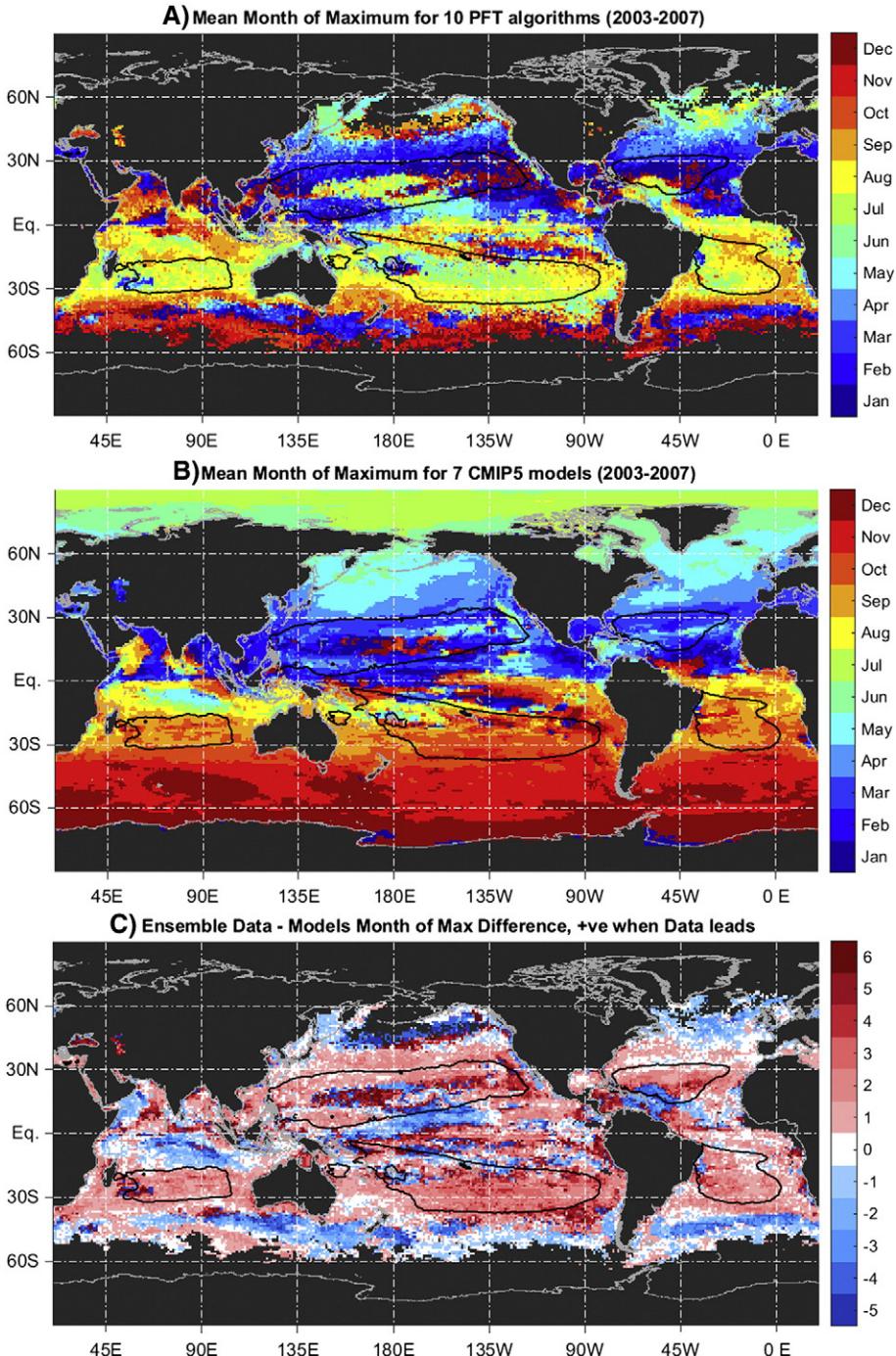


Fig. 3. Ensemble mean month of maximum of A) fraction large phytoplankton/diatoms (or Chl) among 10 PFT satellite algorithms (Table 1), and B) diatom biomass among 7 CMIP5 models. C) The difference (in months) between the ensemble means of the satellite data and the models. A positive difference means the satellite data is leading the models, i.e. the bloom peak occurs earlier in the data than in the model. The isoline of climatological Chl = 0.08 mg m⁻³ is shown (black solid contour). The means and the respective difference should be treated with caution in areas where a considerable number of algorithms or models exhibit low fraction of variance explained by the seasonal cycle (see Supplement Fig. S13).

3.3. Month of maximum of the primary bloom

The ensemble-mean month of maximum of the primary peak (Fig. 3A) varies as a function of latitude to first order, due to the seasonality of insolation, which is a primary physical driver of oceanic ecosystems, controlling both light and nutrient availability (via heating rates, wind patterns and mixing). This latitudinal dependence is most pronounced in the North Atlantic, where the subtropics experience a late winter-early spring bloom, northern temperate latitudes experience maximum blooms in May and June, and

subarctic regions – as late as August. Equivalently, for much of the Southern Hemisphere subtropical seas, the maximum of the bloom occurs in late austral winter – July and August (Fig. 3A). The Southern Ocean blooms later during austral late spring and summer, mostly in November through February. However, the Southern Ocean exhibits an interesting banded structure where large phytoplankton (microplankton, diatoms) bloom earlier (November and December) in a zonal band around 50°S, as compared to a nearly continuous band just to the north, at about 45°S, which tends to bloom later in January and February.

The CMIP5 ensemble mean month of maximum of diatom carbon biomass (Fig. 3B) exhibits a similar spatial pattern with latitude and is spatially less noisy than the satellite PFT estimates. In general the models place the blooms later in time over most of the ocean, as indicated by the algorithm-model difference map (Fig. 3C, red colors indicate the data peak leads the model peak in time). Notable exceptions are some areas in or near the subtropical gyres, the Equatorial Upwelling, and the higher latitudes (e.g. the models do not reproduce the aforementioned banded structure in the Southern Ocean), where models place the blooms earlier in time. The algorithm-model difference is about one month over much of the ocean area (Fig. 3C, pale red or blue). This difference is not randomly distributed and exhibits definite spatial patterns, pointing to latitudinal biases in processes and understanding of seasonality in models.

The differences between the PFT algorithm's ensemble mean month of maximum and the month of maximum for Chl are small (Supplement Fig. S2, top left panel), indicating that the PFT algorithm ensemble mean month of maximum for microplankton (or similar variable, Table 1) appears to be representative of that for total Chl. Several algorithms (BR10, OC-PFT, UITZ06) are abundance-based (Table 1), i.e. their PFT retrievals are a strong function of Chl; thus it is not surprising that their individual differences with the Chl month of maximum are relatively small. So the ensemble mean month of maximum may be driven by the abundance-based PFT algorithms. Two of the spectral-based models (CB06 and MY10) also exhibit generally small differences with the Chl results. Differences among other algorithms with respect to the month of maximum can be larger; while for most of the ocean and for most algorithms the differences are not very large, considerable discrepancies persist in significant ocean areas where month of maxima difference can reach up to 5 or 6 months. More details, including possible reasons for the observed differences, are discussed in Supplement Part 5. An alternative way to quantify the level of agreement among the satellite algorithms or the CMIP5 models is the circular variance of the month of maxima (Supplement Fig. S3). Note that in areas where percent seasonal variance is low (Fig. 2A and Supplement Fig. S13), the concept of month of maximum for the seasonal cycle breaks down and results in these areas should be interpreted with caution. See Section 3.7 and this Supplement Part 9 for details.

We note that from a resource management standpoint, a difference of a month can be very significant, especially with respect to the effect on higher trophic levels. For example, Platt et al. (2003) conclude that differences of three weeks in the timing of the spring algal bloom can have large influences on the survival index of fish larvae in the Northwest Atlantic. Koeller et al. (2009) discuss the coupling of the phenologies of phytoplankton and shrimp in the North Atlantic. In general, whether a difference of one or two months among the algorithms and CMIP5 models is significant for practical applications will depend on the specific application. Which PFT algorithm or algorithm ensemble may be best suited to inform a certain decision will also depend on the issue at hand, as the algorithms have different theoretical bases. However, in general, practical applications would best be addressed by using daily or 8-day data, rather than the monthly data used here for the global inter-comparison.

3.4. Primary bloom duration

The ensemble mean of the duration of the primary annual bloom among the 10 PFT algorithms (Fig. 4A) indicates that over much of the ocean the bloom duration is about 100–120 days, i.e. about 3–4 months. Maximum durations tend to occur at the poleward and to a lesser extend, the equatorward fringes of the subtropical gyres. These maximum bands are most prominent in the Pacific. Large portions of the interior of the southern hemisphere subtropical gyres also exhibit long bloom duration. These results are consistent to first order with the SeaWiFS Chl-based phenological analysis of Racault et al. (2012) and Sapiano et al. (2012), who use different methodologies. Sapiano et al. (2012) observe

longer bloom durations than the analysis here. They note that their results are indeed longer than most previous studies and also caution that in areas of double blooms, their duration indicates the combined duration of the blooms in some cases. Additionally, our analysis may exhibit shorter durations if there is a taxonomical succession, as our analysis indicates the bloom of only microplankton/diatoms. Sapiano et al. (2012) note that bloom durations do not tend to exhibit a simple pattern of decrease with higher latitudes, which is consistent with our observations (Fig. 4A), and different from the result of Racault et al. (2012). The PFT ensemble mean exhibits high spatial frequency noise, and there is no clear pattern of decreasing bloom duration with increasing latitude. The same is noted by Sapiano et al. (2012) and is also apparent in the analysis of Racault et al. (2012) to some degree, but note that they use a coarser spatial smoothing. Maps of primary bloom duration for Chl and the individual PFT algorithms are shown in Supplement Fig. S4 and agreement among the algorithms and CMIP5 models is quantified by the variance in primary bloom duration (Supplement Fig. S5); additional details are discussed in Supplement Part 6.

The CMIP5 models, in contrast to the PFT algorithm data, exhibit much smoother bloom duration spatially (Fig. 4B), and the most prominent duration maxima occur at the equatorward fringes and inside of the subtropical gyres. Compared to the satellite data, the higher latitudes exhibit a much more obvious progression towards shorter bloom durations, dropping to below 2 months for polar latitudes. Models fail to capture secondary peaks, which are especially important at high latitudes, as explained in the following section. The lack of secondary peaks might contribute to a shorter and cleaner definition of bloom duration in models when compared to data. The difference in bloom duration between the models and the data (Fig. 4C) confirms that in general, models exhibit longer blooms in the gyres and shorter bloom at latitudes higher than ~30°.

3.5. Secondary blooms

The presence of secondary blooms can be detected, because several harmonics were used in the DFT analysis (Section 2.3 and Supplement Section S1). In the mid-latitudes, the second bloom is usually a secondary bloom of smaller amplitude in the respective hemisphere's autumn (e.g. Sapiano et al., 2012). To summarize the PFT algorithm and model consensus about where secondary blooms occur, the fraction of algorithms that exhibit a single annual peak (Fig. 5A) versus a double annual peak (Fig. 5B) is employed. Both maps exhibit well-defined latitudinal banding, where most algorithms exhibit a single peak at the poleward fringes of the subtropical gyres (~30° latitude), a double peak around 40–45° in both hemispheres, and again a single peak at higher subpolar latitudes of about 60°. Previous studies have identified and studied this pattern, using Chl data (Sapiano et al., 2012) together with ecosystem modeling (Platt et al., 2009). Cushing (1959) qualitatively described a single peak at higher latitudes and a double peak at lower temperate latitudes, which is consistent with the PFT observations summarized here, as well as the ecosystem model of Platt et al. (2009). Chl time series from SeaWiFS in the North Atlantic analyzed in Platt et al. (2009) are also generally consistent with these observations, as is the analysis of Cabré et al. (2016). In general the zonal bands of single vs. double peak run slightly SW to NE in the Northern Hemisphere, which is most pronounced in the Pacific and is apparent in both the analyses here (Fig. 5A) and the maps of Sapiano et al. (2012). Importantly, caution should be employed when interpreting results from areas with low seasonal variance (Section 3.7, Fig. 2A and Supplement Fig. S13A). More details on the secondary blooms in the PFT data sets, including phase difference with the respective primary bloom and fractional prominence analysis, are provided in Supplement Part 7.

The CMIP5 models exhibit a very different pattern of single (Fig. 5C) vs. double peaks (Fig. 5D), as compared to the PFT and Chl satellite data. Double peaks are predominantly found only along the Equator and in the Arabian Sea and the Bay of Bengal. In these areas the models are

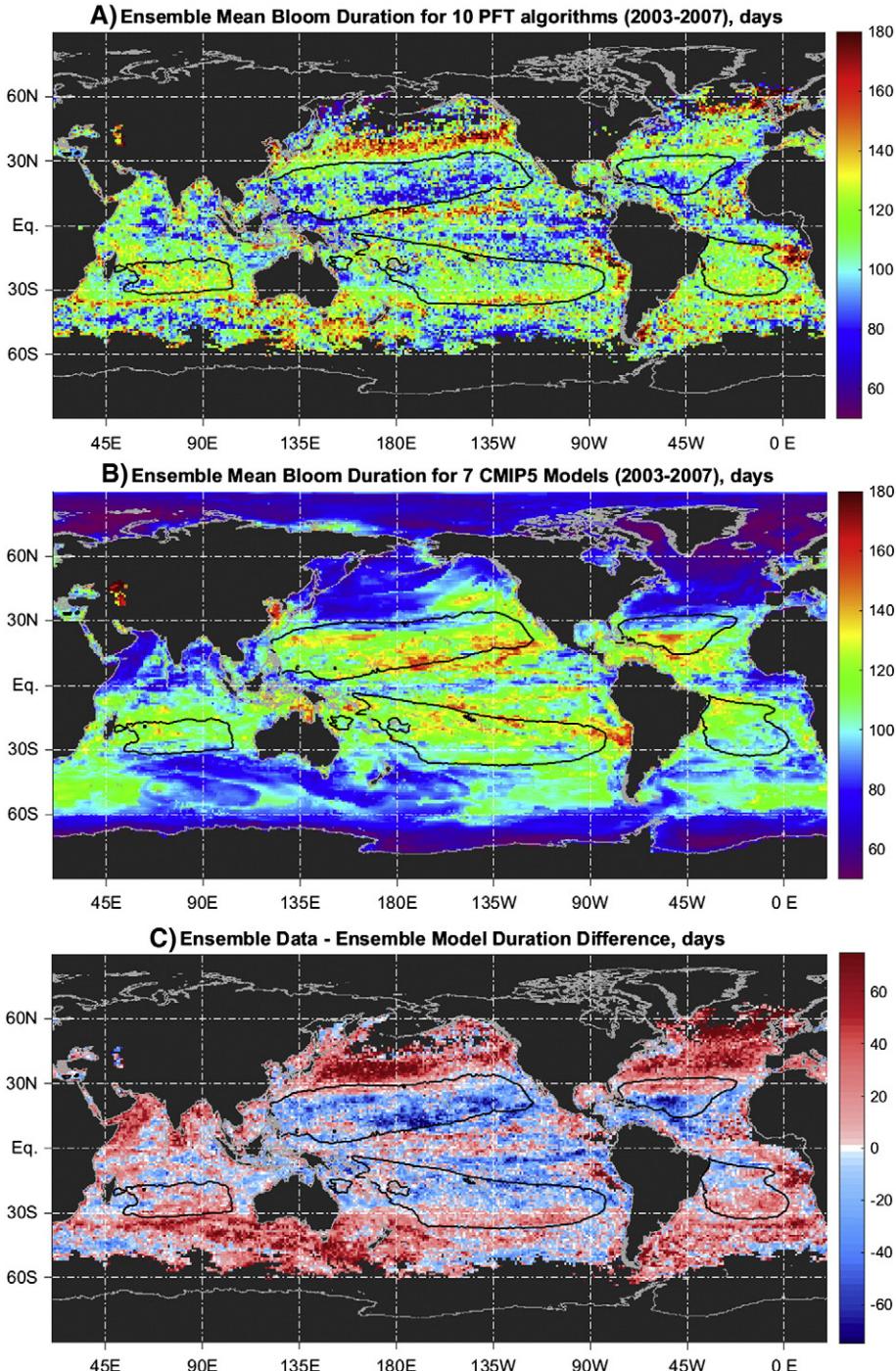


Fig. 4. Ensemble mean primary bloom duration (in days) for A) large phytoplankton/diatoms among 10 PFT satellite algorithms, and B) diatom biomass among 7 CMIP5 models. C) Difference in bloom duration between the satellite data and the models (positive when data bloom duration is larger). Duration is defined as the width of the modeled seasonal signal at half the bloom peak height. The isoline of climatological Chl = 0.08 mg m⁻³ is shown (black solid contour).

in agreement with the satellite data (cf. Fig. 5A and B), although the data are noisier. Models fail to reproduce the secondary peak occurring at mid-latitudes around 40–45° in satellite data.

3.6. North Atlantic regionally binned analysis

The analyses of seasonal variance (Fig. 2) and that of number of peaks (Fig. 5) suggest the presence of the following North Atlantic biomes with significant seasonal amplitude: (a) a *subpolar* regime north of about 50°N, with a single (June to August) light-limited

biomass peak; and (b) a *transitional*, subpolar–subtropical regime between 35°–50°N with two annual peaks, resulting from an alternation of light and nutrient limited conditions (e.g. Evans and Parslow, 1985), and (c) a seasonally varying *Northern subtropics* regime centered around 30°N, characterized by a single annual peak in winter or early spring and high seasonal variability. Monthly averages from regionally-binned satellite PFT time series for two Longhurst marine biogeographic provinces, the North Atlantic Drift Region (NADR) and the North Atlantic Subtropical gyre – West (NASW) (Supplement Fig. S9), are exhibited in Fig. 6A and B, respectively.

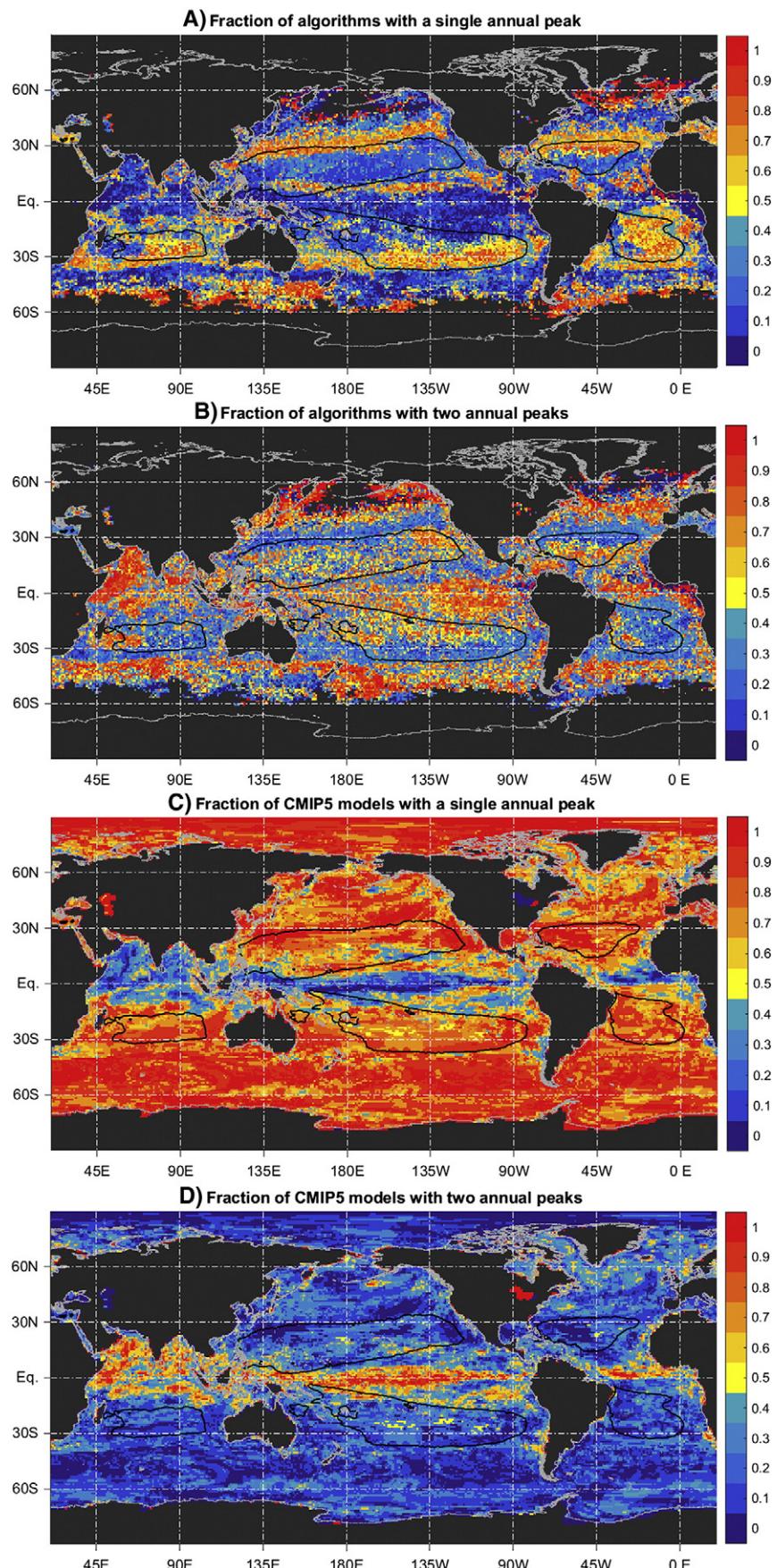


Fig. 5. Fraction of PFT algorithms exhibiting (A) a single annual peak or (B) two peaks in one annual cycle. Fraction of CMIP5 algorithms that exhibit (C) a single annual peak and (D) two peaks in one annual cycle. The fraction is calculated from all algorithms (or models, respectively) that have valid phenology metrics calculated at each pixel (Supplement Fig. S6). The isoline of climatological Chl = 0.08 mg m^{-3} (black solid contour) is shown on all panels.

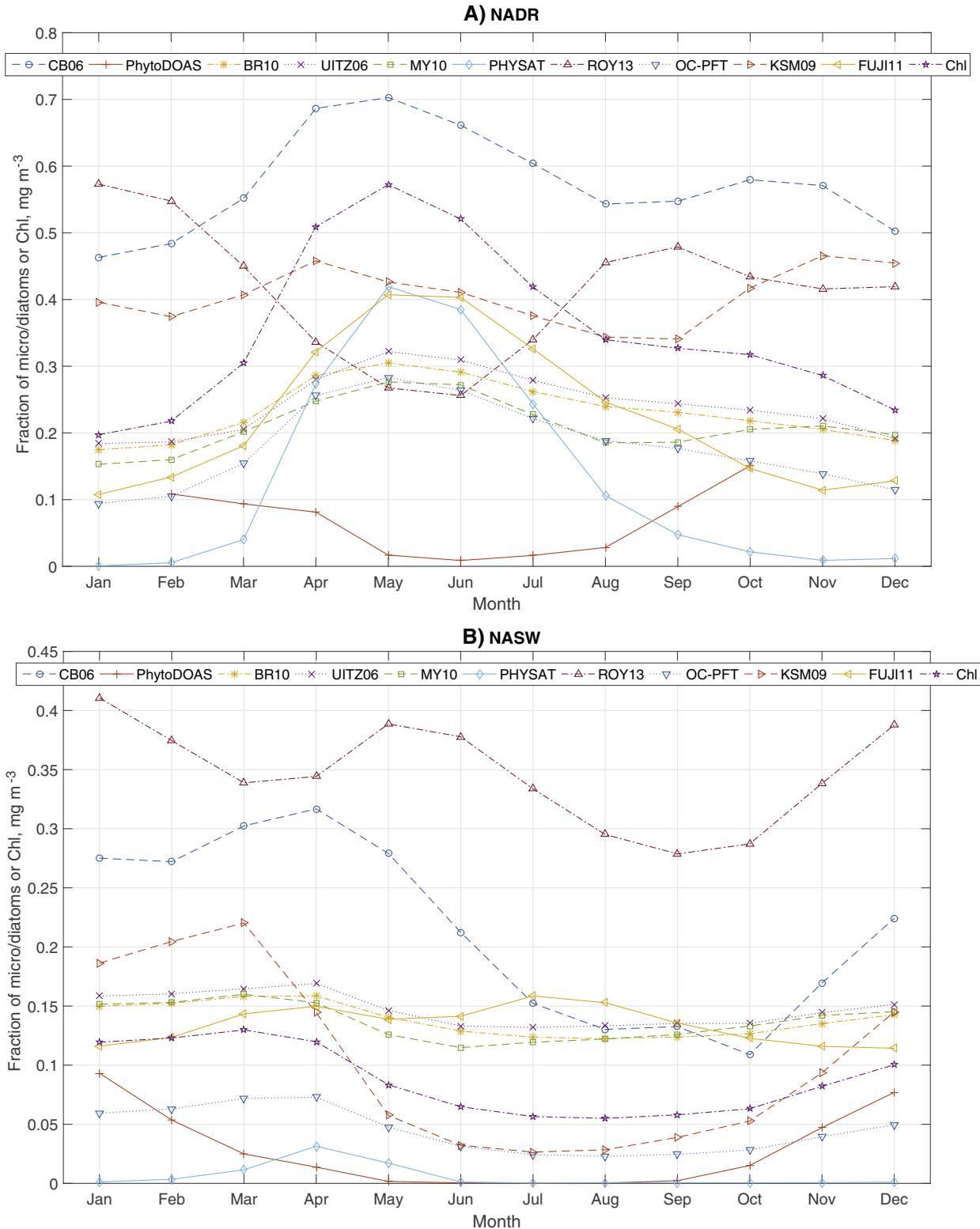


Fig. 6. Time series plots of the monthly climatologies of the PFT algorithm variables and OC4v6 Chl (units given in Table 1) and for two example Longhurst (1998) biogeographic provinces as follows: A) North Atlantic Drift Region (NADR); B) The Western North Atlantic Subtropical Gyral Province (NASW). Y-axis scales are different between the two panels. See Section 3.6 and Supplement Parts 4 and 8 for more information and Supplement Fig. S9 for a map of the provinces.

The NADR province straddles the subpolar and the transitional regimes (representing mostly the transitional biome), whereas NASW represents the Northern subtropics regime. The corresponding month of maxima are illustrated for the same Longhurst provinces for the 10 PFT algorithms and Chl (Fig. 7A, B) and for the 7

CMIP5 models (Fig. 7C, D). The time series in both regions (Fig. 6) illustrate that the PFT algorithms exhibit different absolute values and amplitudes/ranges of their variables (Table 1), even if they are mostly referred to as large/micro phytoplankton. This is expected since the algorithms have differing theoretical bases. The spectra of the

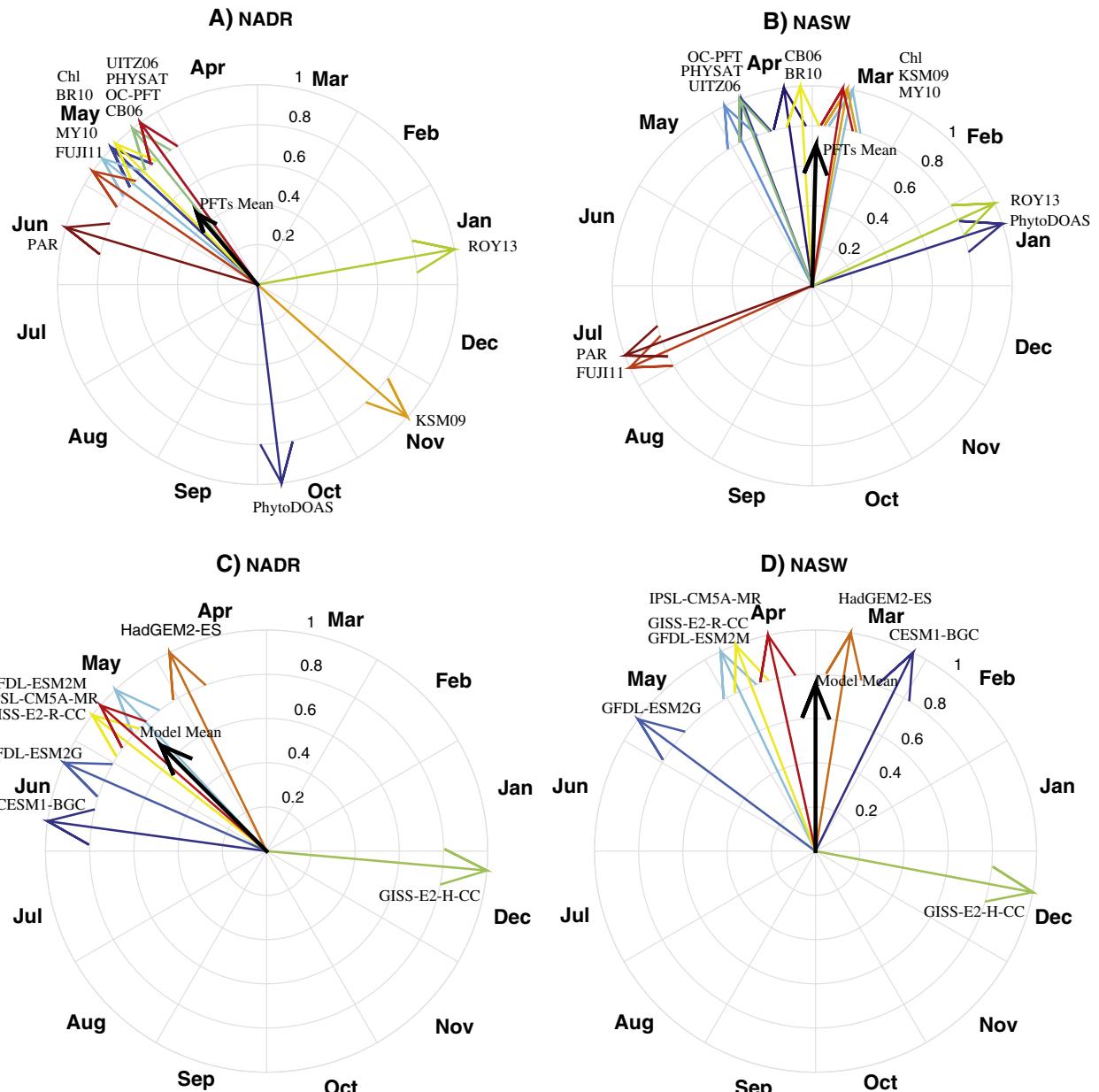


Fig. 7. Months of maxima of the regionally binned PFT algorithm variables (Table 1) (top panels – A and B) and CMIP5 models' diatom carbon biomass (bottom panels – C and D). The following Longhurst (1998) provinces are displayed as examples: A and C) North Atlantic Drift Region (NADR); B and D) The Western North Atlantic Subtropical Gyral Province (NASW). See Supplement Fig. S9 for a map of the provinces. The black arrow's direction indicates the PFT algorithms' or CMIP5 models' ensemble mean month of maximum (circular), and its length indicates one minus the circular variance in month of maxima among the algorithms (an arrow length of one indicates zero variance, and an arrow length of zero – maximal variance of one). The beginning of each month is marked by the abbreviated month name on the polar plots. See Section 2.3 and Supplement Part 4 for methodology details. Supplement Table S1 lists the CMIP5 models used with their acronyms as they appear here.

corresponding complete 5-year time series (Supplement Fig. S10) have their strongest peak at $f = 1 \text{ yr}^{-1}$, indicating that the annual seasonal cycle is a first order source of variability. 2nd and 3rd harmonics often represent additional notable peaks.

Most algorithms in Longhurst's NADR province (Supplement Fig. S9) agree that there is a well pronounced annual maximum in May (Figs. 6A, 7A); however ROY13 and PhytoDOAS indicate minima then instead, exhibiting very different phasing of the seasonal cycle. Also KSM09 exhibits relatively small range of the seasonal cycle there compared to other algorithms, as well as a double peak in April and November (Fig. 7A). The reasons for the lack of complete agreement among the PFT algorithms regarding the month of maximum could not be resolved by this study and require further investigation; this suggests that more

in situ validation and algorithm development data is required. Three of the seven CMIP5 models also place the maximum in May (Fig. 7C), but for some it is in April or June, and December for the GISS-E2-H-CC model, resulting in a higher variance of month of maximum for the models as compared to the satellite data for NADR (cf. length of black arrows in Fig. 7A and C). Some inter-model and data-model differences could be due to the relatively small size of the Longhurst provinces with respect to the coarse model resolutions and the fact that some models may place the equivalent biome in a different location due to different model physics. The NADR province straddles regions where most satellite algorithms indicate double peaks (Fig. 5A and B); while CMIP5 models show single annual peaks (Fig. 5C). The fall peak in satellite data is generally weaker than the spring one (Fig. 6A); CB06, KSM09

and MY10 exhibit more noticeable fall blooms. Note that these fall peaks may or may not be detected by the DFT analysis here based on prominence criteria (Section 2.3). The presence of double peaks is reflected in the spectra having a pronounced peak at $f = 2 \text{ yr}^{-1}$ (Supplement Fig. S10) that is almost as high as the primary peak at $f = 1 \text{ yr}^{-1}$. The NADR spectra indeed exhibit higher overall variance (more power) than the NASW spectra, and also more even distribution of power between the annual and semi-annual peaks, indicating the NASW area is characterized by a single annual peak, and NADR has a fall secondary peak. The strength of this peak in relation to the primary one may depend on the unit (fractional vs. absolute) used in the time series analysis (Section 3.7 and Supplement Part 7). The high variance of NADR is expected, as this area is known for its spectacular North Atlantic blooms in the spring (e.g. Siegel et al., 2002; Behrenfeld, 2010). Since NADR straddles two different regimes with respect to single vs. double peaks (cf. Fig. 5A and B and Supplement Fig. S9), this analysis illustrates the limitations of a regionally binned approach using classically defined biogeographic provinces.

Further to the south, the NASW province straddles the $\text{Chl} = 0.08 \text{ mg m}^{-3}$ climatological isoline used here to delineate the gyre (Supplement Fig. S9) and is a typical northern-subtropical region. It is mostly characterized by a single peak in most algorithms and exhibits some of the cleanest seasonal cycles globally (Figs. 2A and 5A). It is considerably more oligotrophic and hence has lower fraction of microplankton (and total Chl) than NADR, year-round (cf. y-axis scales of Fig. 6A vs. B). While some algorithms indicate a strong winter-spring peak (KSM09, CB06), the annual range of others is a lot smaller and some even exhibit double peaks (ROY13, FUJI11). Seven algorithms agree on a maximum for the primary bloom in March or April (Fig. 7B), but ROY13 and PhytoDOAS indicate a January peak, and FUJI11 indicates a July peak for the primary bloom. Similarly, most CMIP5 models agree on a March or April bloom peak (Fig. 7D). As stated above, the DFT spectra of the satellite data (Supplement Fig. S10B) indicate that overall variance is lower than NADR, and the first harmonic contains proportionately more power than the second harmonic, indicating a single annual peak. A comparative analysis for time-series at the Bermuda Atlantic Time Series (BATS) station, representative of the NASW province and the Northern subtropics regime, is provided in Supplement Part 8 (Supplement Fig. S11), including discussion of some complementary *in situ* biogeochemical data. Supplement Part 8 also provides details on mechanisms in the regions of interest discussed here.

In conclusion, regional binning of the satellite data sets reduces noise, and allows for inspection of a limited number of actual time series from various regions. It also allows for analysis in regions or algorithms that suffer from data sparsity (particularly true of the PHYSAT algorithm) that precludes the DFT analysis on a per-pixel basis. On the other hand, if the chosen regions are spatially heterogeneous, results may be misleading or meaningless, and regions may have an arbitrary definition not necessarily relevant to the phenology at question here. Future monitoring of biology and biogeochemistry at multiple specific representative locations such as BATS is critically important for validation and inter-comparison of satellite algorithms.

3.7. Sources of uncertainty

There are multiple sources of uncertainty that can affect the DFT phenology analysis presented, both related to intrinsic ecosystem characteristics and features and limitations of the DFT technique. Here we discuss some important considerations and we provide more details in Supplement Part 9. The advantages of the DFT method include relative insensitivity to noise and the ability to isolate the variance in the frequencies of interest for phenological studies (de Beurs and Henebry, 2010). In addition, multi-year time series are summarized with a single metric from the modeled signal and the DFT approach eliminates the

need to consider a “sliding season” to ensure the annual cycle is properly described everywhere (Racault et al., 2012). The DFT is a statistical method to fit data to a sum of sines and cosines of fixed frequencies and varying phases. As such, one disadvantage is that representation of seasonal cycles that have non-sinusoidal waveforms requires artificial placement of variance in high frequencies where most noise resides. For example, Wilson and Qiu (2008) observe sharp spikes in seasonal Chl blooms in some areas of the oligotrophic gyres. If higher frequencies are present in the data, significant aliasing can occur, confounding the analysis by placing spurious energy in lower frequencies. The aliasing problem can be addressed by analyzing weekly or daily data to test for the significance of higher order harmonics in the data. That said, Fourier analysis is among the best available techniques to analyze cyclical phenomena and partition variance in frequency components of interest, so it is strongly suited for analyzing seasonal cycles.

The correct retrieval of the phenological parameters of interest using the DFT technique was verified in several ways: by examining an example time series (Supplement Fig. S1); by validation against direct maxima finding via peak analysis of the time series of the monthly climatologies of the respective PFT satellite data, and relating the results to percent seasonal variance (Supplement Fig. S12); and by using the SeaWiFS PAR data set (as discussed in Supplement Section S1). The overall assessment is that the DFT technique correctly identifies phenological parameters if percent variance explained by the seasonal harmonics is 30% or more. Phase derived by the DFT techniques in frequency bands where power density is low can be stochastic and meaningless; therefore month of maxima determinations can be unreliable or random and should be treated with caution in areas of low percent seasonal variance (Fig. 2; Supplement Fig. S13). Supplement Fig. S12A illustrates the fraction of pixels for each algorithm that exhibit a given percent seasonal variance. Comparison of direct maxima finding vs. the DFT determinations of the month of maxima indicate that in places where percent seasonal variance drops below 30%, differences between the two methods can be >2 months for over 10% of the pixels for most algorithms (Supplement Fig. S12B). It was therefore determined that the DFT technique results should be interpreted with caution or not used in analyses in places where percent seasonal variance drops below 30%. A map of the number of algorithms exhibiting percent seasonal variance <30% is shown in Supplement Fig. S13A, and the analogous map for the CMIP5 models is shown in Supplement Fig. S13B. It is worth noting that analyzing phenology in places with low percent seasonality becomes intrinsically meaningless, regardless of what technique is used. Additionally, ensemble metrics for the PFT algorithms should be interpreted with caution where few algorithms contribute to the mean (Supplement Fig. S6). Finally, results should also be interpreted with caution at high latitudes where data can be sparse, especially in the respective winter months. This is discussed in Supplement Section S1.

It is important to note that it is not exactly equivalent to analyze absolute biomass indicators (Chl, carbon) and fractions (as done here for most PFT algorithms), e.g. there are indications that the secondary bloom is more dominated by larger sizes than the primary peak (Cabré et al., 2016) (see also Supplement Part 5 and Part 7). For example, a secondary peak can look about as high as the primary one in terms of percent, but can be much weaker in terms of Chl. Mathematically, percent microplankton can increase without an accompanying increase (or even with a decrease) in total or microplankton absolute biomass or Chl. However, such situations are likely to be atypical on a global scale according to modern ecosystem understanding. Absolute carbon biomass of diatoms was analyzed for the CMIP5 models, which is not exactly equivalent to the algorithm variables. Further analysis should focus on comparing phenology calculated from fractions vs. absolute biomass indicators. Finally, physiological adaptation affects the Chl to carbon ratio and can decouple Chl and carbon variability, especially in lower-latitude oceans (e.g. Behrenfeld et al., 2005; Siegel et al., 2013); it is best to express size-fractionated or PFT-specific biomass

in terms of carbon units, as done by Kostadinov et al., 2016 who re-cast the PFTs in terms of carbon using the KSM09 particle size distribution algorithm and allometric relationships (Menden-Deuer and Lessard, 2000). In spite of the above considerations, we stress that the primary purpose of this work is PFT algorithm inter-comparison, and all PFT algorithms are analyzed in an equivalent way here.

The phenological parameters described here were derived only from the large phytoplankton/microplankton variable (Table 1) or for diatoms in the case of the CMIP5 models and the PhytoDOAS and PHYSAT algorithms. While in much of the temperate, high latitude and upwelling regions it is indeed larger phytoplankton that dominate the bloom, in the more oligotrophic subtropics and tropics nanoplankton can dominate the seasonal maximum. Also, diatoms can be found within the nanoplankton fraction or the microplankton can be dominated by other large phytoplankton (e.g. dinoflagellates). Various PFT algorithms provide nanoplankton and other PFT variables that need to be analyzed and compared in future work. The analysis presented here is global; however, some algorithms were developed/parameterized with data sets of specific limited geographic coverage. Those would not be expected to necessarily perform well outside of their area of development. Notably, FUJI11 was developed for the Arctic. Not surprisingly, this algorithm differs more substantially from other algorithms in the tropics and subtropics. The Southern Ocean presents atypical bio-optical characteristics (e.g. Uitz et al., 2006) some implications of which are discussed in Supplement Part 9.

4. Concluding remarks

We used the Discrete Fourier Transform (DFT) to derive and inter-compare phenological parameters for the 2003–2007 period among 1) the fraction of Chl corresponding to microplankton (or a closely related variable) from 10 satellite ocean color algorithms, 2) satellite determinations of chlorophyll concentration, and 3) diatom biomass from 7 CMIP5 climate models. The phenological parameters derived were amplitude, month of maximum, percent variance explained by the seasonal cycle, bloom duration, and secondary bloom characteristics. Results indicate that PFT algorithms agree only to first order globally. Enough qualitative and quantitative differences between the algorithms are detected (e.g. Fig. 6) to make a further comprehensive, global validation exercise a high priority. While validation is outside the scope of this work, a separate working group has been formed within the PFT Inter-comparison Project to perform a comprehensive validation exercise (Bracher et al., 2015). Validation itself is challenging (Brewin et al., 2011), as *in situ* HPLC-derived PFTs have their own limitations and do not necessarily correspond to the way non-HPLC-based algorithms define their variables. To allow for these subtle differences in the variables retrieved, it may be best to test all algorithms against a comprehensive *in situ* data set incorporating co-located radiometric, bio-optical (pigment concentration, IOPs) and derived biological quantities (phytoplankton Chl, size structure, etc.) in order to resolve whether differences are due to algorithm uncertainties or to actual biogeophysical differences between, for example, the timing of maximum diatom biomass vs. maximum of large cells biomass (Bracher et al., 2015).

Comparison of phenological parameters in CMIP5 model output (diatom carbon biomass) to those of satellite data suggests that the month of maximum is fairly well represented in models (albeit with a systematic, latitude-dependent bias), while other phenological characteristics show a number of important biases in CMIP5 models: 1) more pronounced seasonal variability in the models, e.g. a smoother latitudinal progression and less local spatial variability in phenological indices such as month of maximum and bloom duration, 2) while seasonality in the satellite data is cleanest (i.e. percent variance explained is highest) along a zonal band at 30° latitude in both, this feature is not reflected in the CMIP5 models, and 3) models exhibit a single annual biomass peak over most of the ocean, except for the Equatorial band,

whereas secondary blooms tend to occur in zonal bands in temperate latitudes in the satellite data but are not well captured in the climate models. These biases are probably due to over-simplification of processes in models and a lower response to interannual variability than in reality as also discussed in Cabré et al. (2016). Additionally, the coarse 1° resolution does not allow a proper representation of coastal processes and some frontal dynamics in models. We note that for many modelers, the PFT products derived from satellite algorithms are considered as observations (not algorithm products per se) and used as a reference for validating model outputs (e.g. Le Quéré et al., 2005; Bopp et al., 2005; Stock et al., 2014). To improve the utility of satellite algorithms for this purpose, better uncertainty characterization should be considered high priority.

A fruitful way forward would be to use algorithms of different theoretical bases together to increase the degrees of freedom and solve for more variables. Advent of hyperspectral sensors in the near future is expected to improve our ability to discern small spectral differences arising from the different PFTs, as indicated, for example, by the development of the PhytoDOAS algorithm. Therefore, future direction of efforts towards development of hyperspectral algorithms is desirable, keeping in mind that there is a fundamental limit on the additional degrees of freedom available (Lee et al., 2007). While chlorophyll is certainly a useful variable, it is carbon biomass in the living phytoplankton that is the variable of most direct relevance to carbon cycle and biogeochemical studies; it is also the unit of PFT accounting in climate models (Supplement Table S1). The carbon-based algorithm of Kostadinov et al. (2016) could be used in conjunction with an algorithm partitioning Chl (e.g. BR10) in order to assess physiological status and productivity by size class (Behrenfeld et al., 2005; Uitz et al., 2010).

Importantly, PFT algorithms and bio-optical algorithms in general could improve by moving towards analytical approaches based more on first principles rather than empirical relationships, i.e. being mechanistic in nature. Most of the existing PFT algorithms contain a high degree of empiricism. Empirical algorithms rely on statistical relationships derived during a certain environmental state and are thus not predictive in nature. Should the underlying relationship change, the algorithm uncertainties will increase. Mechanistic models should remain more robust under changing environmental conditions of the future, e.g. due to climate change.

Acknowledgments

This work was performed with funding from NASA Ocean Biology and Biogeochemistry Program (grant #NNX13AC92G to Irina Marinov and Tihomir S. Kostadinov). The contribution of A. Bracher, R. Brewin and A. Bricaud was partly funded via the ESA SEOM SY-4Sci Synergy project SynSenPFT. We thank Tilman Dinter and Bernard Gentili for help with the PhytoDOAS and CB06 algorithm processing, respectively, Amane Fujiwara for leading development of the FUJI11 algorithm, and Aurea Ciotti for leading the development of the CB06 algorithm. We thank David Shields (specifically for producing Supplement Fig. S11C and processing BATS data, and for work on the Chl gyre contour lines), Svetlana Milutinović, and Danica Fine for providing help in the completion of this work. We thank Libe Washburn for FFT processing advice. We thank Jordan Rosenthal for his compass plot labeling script (used here with modifications in Fig. 7). All data processing, analysis and plotting for the phenological analysis in this work was done in MATLAB®. We additionally acknowledge all the satellite algorithm providers (and their funding agencies) for their support and providing their data. The coastlines displayed on maps were extracted with the NOAA/NGDC GEODAS-NG software using the L1 layer of the GSHHG v2.2.3 (Wessel and Smith, 1996) coastline data set. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table S1 of this paper) for producing and making available their model output. For CMIP the U.S. Department of

Energy's Program for Climate Model Diagnosis and Inter-comparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We are grateful to three anonymous reviewers and the editor whose comments and suggestions improved the quality of this manuscript.

Appendix A. Supplementary Material

Supplementary material to this article can be found online at <http://dx.doi.org/10.1016/j.rse.2016.11.014>.

References

- Alvain, S., Moulin, C., Dandonneau, Y., Bréon, F.M., 2005. Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep-Sea Res. I Oceanogr. Res. Pap.* 52 (11), 1989–2004.
- Alvain, S., Moulin, C., Dandonneau, Y., Loisel, H., 2008. Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view. *Glob. Biogeochem. Cycles* 22:GB3001. <http://dx.doi.org/10.1029/2007GB003154>.
- Behrenfeld, M.J., Boss, E., Siegel, D.A., Shea, D.M., 2005. Carbon-based ocean productivity and phytoplankton physiology from space. *Glob. Biogeochem. Cycles* 19:GB1006. <http://dx.doi.org/10.1029/2004GB002299>.
- Behrenfeld, M.J., 2010. Abandoning Sverdrup's critical depth hypothesis on phytoplankton blooms. *Ecology* 91 (4), 977–989.
- de Beurs, K.M., Henebry, G.M., 2010. Spatio-temporal statistical methods for modelling land surface phenology. *Phenological research*, Springer, Netherlands, pp. 177–208.
- Bograd, S.J., Foley, D.G., Schwinger, F.B., Wilson, C., Laurs, R.M., Polovina, J.J., Howell, E.A., Brainard, R.E., 2004. On the seasonal and interannual migrations of the transition zone chlorophyll front. *Geophys. Res. Lett.* 31:L17204. <http://dx.doi.org/10.1029/2004GL020637>.
- Bopp, L., Aumont, O., Cadule, P., Alvain, S., Gehlen, M., 2005. Response of diatoms distribution to global warming and potential implications: a global model study. *Geophys. Res. Lett.* 32 (19):L19606. <http://dx.doi.org/10.1029/2005GL023653>.
- Bracher, A., Vountas, M., Dinter, T., Burrows, J.P., Röttgers, R., Peeken, I., 2009. Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences* 6:751–764. <http://dx.doi.org/10.5194/bg-6-751-2009>.
- Bracher, N., Hardman-Mountford, N., Hirata, T., Bernard, S., Brewin, R., Bricaud, A., Brotas, V., Chase, A., Ciotti, A., Choi, J.-K., Clementson, L., Devred, E., Di Giacomo, P., Dupouy, C., Kim, W., Kostadinov, T., Kwiatkowska, E., Lavender, S., Moisan, T., Mouw, C., Son, S., Sosik, H., Uitz, J., Werdell, J., Zheng, G., 2015. Report on IOC-CCG workshop "Phytoplankton composition from space: towards a validation strategy for satellite algorithms" 25–26 Oct 2014, Portland, ME, USA. <http://www.iocccg.org/groups/report-PFTworkshopOct2014.pdf>; as NASA technical Memorandum 217528_01-22-15. http://www.iocccg.org/groups/PFT-TM_2015-217528_01-22-15.pdf.
- Brewin, R.J., Sathyendranath, S., Hirata, T., Lavender, S.J., Barciela, R.M., Hardman-Mountford, N.J., 2010. A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Model.* 221 (11), 1472–1483.
- Brewin, R.J., Hardman-Mountford, N.J., Lavender, S.J., Raitsos, D.E., Hirata, T., Uitz, J., Devred, E., Bricaud, A., Ciotti, A., Gentili, B., 2011. An inter-comparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sens. Environ.* 115 (2), 325–339.
- Bricaud, A., Ciotti, A.M., Gentili, B., 2012. Spatial-temporal variations in phytoplankton size and colored detrital matter absorption at global and regional scales, as derived from twelve years of SeaWiFS data (1998–2009). *Global Biogeochem. Cycles* 26:GB1010. <http://dx.doi.org/10.1029/2010GB003952>.
- Cabré, A., Marinov, I., Leung, S., 2015. Consistent global responses of marine ecosystems to future climate change across the IPCC AR5 earth system models. *Clim. Dyn.* 45:1253. <http://dx.doi.org/10.1007/s00382-014-2374-3>.
- Cabré, A., Shields, D., Marinov, I., Kostadinov, T.S., 2016. Phenology of size-partitioned phytoplankton carbon-biomass from ocean color remote sensing and CMIP5 models. *Front. Mar. Sci.* 3:39. <http://dx.doi.org/10.3389/fmars.2016.00039>.
- Campbell, J.W., 1995. The lognormal distribution as a model for bio-optical variability in the sea. *J. Geophys. Res. Oceans* 100 (C7), 13237–13254.
- Chisholm, S.W., 1992. Phytoplankton size. In: Falkowski, P.G., Woodhead, A.D. (Eds.), *Primary Productivity and Biogeochemical Cycles in the Sea*. Plenum Press, New York, pp. 213–237.
- Ciotti, A.M., Bricaud, A., 2006. Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. *Limnol. Oceanogr. Methods* 4, 237–253.
- Cushing, D.H., 1959. The seasonal variation in oceanic production as a problem in population dynamics. *J. Conseil.* 24 (3), 455–464.
- Eppley, R.W., Peterson, B.J., 1979. Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* 282, 677.
- Evans, G.T., Parslow, J.S., 1985. A model of annual plankton cycles. *Biol. Oceanogr.* 3 (3), 327–347.
- Falkowski, P.G., Barber, R.T., Smetacek, V., 1998. Biogeochemical controls and feedbacks on ocean primary production. *Science* 281, 200–206.
- Field, C.B., Behrenfeld, M.J., Randerson, J.T., Falkowski, P., 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281, 237–240.
- Fujiwara, A., Hirawake, T., Suzuki, K., Saitoh, S.-I., 2011. Remote sensing of size structure of phytoplankton communities using optical properties of the Chukchi and Bering Sea shelf region. *Biogeosciences* 8:3567–3580. <http://dx.doi.org/10.5194/bg-8-3567-2011>.
- Glover, D.M., Wroblewski, J.S., McClain, C.R., 1994. Dynamics of the transition zone in coastal zone color scanner-sensed ocean color in the North Pacific during oceanographic spring. *J. Geophys. Res.* 99, 7501–7511.
- Habib, E.A., 2012. Geometric mean for negative and zero values. *Int. J. Res. Rev. Appl. Sci.* 11, 419–432.
- Hirata, T., Hardman-Mountford, N.J., Brewin, R.J.W., Aiken, J., Barlow, R., Suzuki, K., Isada, T., Howell, E., Hashioka, T., Noguchi-Aita, M., Yamanaka, Y., 2011. Synoptic relationships between surface chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* 8:311–327. <http://dx.doi.org/10.5194/bg-8-311-2011>.
- Hirata, T., Hardman-Mountford, N., Brewin, R.J.W., 2012. Comparing satellite-based phytoplankton classification methods. *EOS Trans. Am. Geophys. Union* 93 (6), 59.
- Hirata, T., 2015. Satellite phytoplankton functional type algorithm inter-comparison project. <http://pft.ees.hokudai.ac.jp/satellite/index.shtml> (last access: Oct. 16, 2015).
- Hood, R.R., Laws, E.A., Armstrong, R.A., Bates, N.R., Brown, C.W., Carlson, C.A., Chai, F., Doney, S.C., Falkowski, P.G., Feely, R.A., Friedrichs, M.A.M., Landry, M.R., Moore, J.K., Nelson, D.M., Richardson, T.L., Salihoglu, B., Schartau, M., Toole, D.A., Wiggert, J.D., 2006. Pelagic functional group modeling: progress, challenges and prospects. *Deep-Sea Res. II* 53, 459–512.
- IOC-CCG, 2014. Phytoplankton functional types from space. In: Sathyendranath, S. (Ed.), *Reports of the International Ocean-Colour Coordinating Group*, No. 15. IOC-CCG, Dartmouth, Canada.
- IPCC, 2013. Climate change 2013: the physical science basis. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA (1535 pp.).
- Koeller, P., Fuentes-Yaco, C., Platt, T., Sathyendranath, S., Richards, A., Ouellet, P., Aschan, M., 2009. Basin-scale coherence in phenology of shrimps and phytoplankton in the North Atlantic Ocean. *Science* 324 (5928), 791–793.
- Kostadinov, T.S., Siegel, D.A., Maritorena, S., 2009. Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res.* <http://dx.doi.org/10.1029/2009JC005303>.
- Kostadinov, T.S., Siegel, D.A., Maritorena, S., 2010. Global variability of phytoplankton functional types from space: assessment via the particle size distribution. *Biogeosciences* 7 (10), 3239–3257.
- Kostadinov, T.S., Milutinović, S., Marinov, I., Cabré, A., 2016. Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. *Ocean Sci.* 12:561–575. <http://dx.doi.org/10.5194/os-12-561-2016>.
- Le Quéré, C., Harrison, S.P., Prentice, I.C., Buitenhuis, E.T., Aumont, O., Bopp, L., Claustre, H., Cotrim Da Cunha, L., Geider, R., Giraud, X., Klaas, C., Kohfeld, K.E., Legendre, L., Manizza, M., Platt, T., Rivkin, R.B., Sathyendranath, S., Uitz, J., Watson, A.J., Wolf-Gladrow, D., 2005. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob. Chang. Biol.* 11, 2016–2040.
- Lee, Z.P., Carder, K., Arnone, R., He, M.X., 2007. Determination of primary spectral bands for remote sensing of aquatic environments. *Sensors* 2007 (7), 3428–3441.
- Levy, M., Lehahn, Y., Andre, J.M., Memery, L., Loisel, H., Heifetz, E., 2005. Production regimes in the northeast Atlantic: a study based on Sea-viewing Wide Field-of-view Sensor (SeaWiFS) chlorophyll and ocean general circulation model mixed layer depth. *J. Geophys. Res.* 110. <http://dx.doi.org/10.1029/2004jc002771>.
- Loisel, H., Nicolas, J.-M., Sciandra, A., Stramski, D., Poteau, A., 2006. Spectral dependency of optical backscattering by marine particles from satellite remote sensing of the global ocean. *J. Geophys. Res.* 111:C09024. <http://dx.doi.org/10.1029/2005JC003367>.
- Longhurst, A.R., 1998. *Ecological Geography of the Sea*. Academic Press, San Diego (397 pp.).
- McClain, C.R., 2009. A decade of satellite ocean color observations. *Annu. Rev. Mar. Sci.* 1, 19–42.
- Marañón, E., 2015. Cell size as a key determinant of phytoplankton metabolism and community structure. *Annu. Rev. Mar. Sci.* 7 (1). <http://dx.doi.org/10.1146/annurev-marine-010814-015955>.
- Menden-Deuer, S., Lessard, E.J., 2000. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.* 45, 569–579.
- Mouw, C.B., Yoder, J.A., 2010. Optical determination of phytoplankton size composition from global SeaWiFS imagery. *J. Geophys. Res.* 115:C12018. <http://dx.doi.org/10.1029/2010JC006337>.
- NASA Goddard Space Flight Center, Ocean Biology Distributed Active Archive Center, 2010e. Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Ocean Color Data, NASA OB.DAAC, Greenbelt, MD, USA. Reprocessing R2010.0. Maintained by NASA Ocean Biology Distributed Active Archive Center (OB.DAAC). Goddard Space Flight Center, Greenbelt MD.
- NCEI: National Centers for Environmental Information (NOAA), 2015. Registration of structured square-cell grids. URL. <https://www.ngdc.noaa.gov/mgg/global/gridregistration.html> (last accessed: Dec. 8, 2015).
- O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., McClain, C., 1998. Ocean chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.* 103 (C11), 24, 937–24, 953.
- O'Reilly, J.E., Maritorena, S., Siegel, D.A., O'Brien, M.C., Toole, D., Mitchell, B.G., et al., 2000. Ocean color chlorophyll-a algorithms for SeaWiFS, OC2, and OC4: version 4, SeaWiFS post-launch calibration and validation analyses, part 3. NASA/TM 206892 vol. 11, pp. 9–23.

- Platt, T., Fuentes-Yaco, C., Frank, K.T., 2003. Marine ecology: spring algal bloom and larval fish survival. *Nature* 423 (6938), 398–399.
- Platt, T., White III, G.N., Zhai, L., Sathyendranath, S., Roy, S., 2009. The phenology of phytoplankton blooms: ecosystem indicators from remote sensing. *Ecol. Model.* 220 (21), 3057–3069.
- Racault, M.F., Le Quéré, C., Buitenhuis, E., Sathyendranath, S., Platt, T., 2012. Phytoplankton phenology in the global ocean. *Ecol. Indic.* 14 (1), 152–163.
- Roy, S., Sathyendranath, S., Platt, T., 2011. Retrieval of phytoplankton size from bio-optical measurements: theory and applications. *J. R. Soc. Interface* 8 (58), 650–660.
- Roy, S., Sathyendranath, S., Bouman, H., Platt, T., 2013. The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sens. Environ.* 139, 185–197.
- Sadeghi, A., Dinter, T., Vountas, M., Taylor, B., Peeken, I., Altenburg Soppa, M., Bracher, A., 2012. Improvements to the PhytoDOAS method for identification of coccolithophores using hyper-spectral satellite data. *Ocean Sci.* 8, 1055–1070.
- Sapiro, M.R.P., Brown, C.W., Schollaert Uz, S., Vargas, M., 2012. Establishing a global climatology of marine phytoplankton phenological characteristics. *J. Geophys. Res.* 117: C08026. <http://dx.doi.org/10.1029/2012JC007958>.
- Sieburth, J.M., Smetacek, V., Lenz, J., 1978. Pelagic ecosystem structure: heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnol. Oceanogr.* 23, 1256–1263.
- Siegel, D.A., Doney, S.C., Yoder, J.A., 2002. The North Atlantic spring phytoplankton bloom and Sverdrup's critical depth hypothesis. *Science* 296 (5568), 730–733.
- Siegel, D.A., Behrenfeld, M.J., Maritorena, S., McClain, C.R., Antoine, D., Bailey, S.W., Bontempi, P.S., Boss, E.S., Dierssen, H.M., Doney, S.C., Eplee Jr., R.E., Evans, R.H., Feldman, G.C., Fields, E., Franz, B.A., Kuring, N.A., Mengelt, C., Nelson, N.B., Patt, F.S., Robinson, W.D., Sarmiento, J.L., Swan, C.M., Werdell, P.J., Westberry, T.K., Wilding, J.G., Yoder, J.A., 2013. Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission. *Remote Sens. Environ.* 135 (0), 77–91.
- Siegel, D.A., Buesseler, K.O., Doney, S.C., Sailley, S.F., Behrenfeld, M.J., Boyd, P.W., 2014. Global assessment of ocean carbon export by combining satellite observations and food-web models. *Glob. Biogeochem. Cycles* 28:181–196. <http://dx.doi.org/10.1002/2013GB004743>.
- Stock, C.A., Dunne, J.P., John, J.G., 2014. Global-scale carbon and energy flows through the marine planktonic food web: an analysis with a coupled physical-biological model. *Prog. Oceanogr.* 120:1–28. <http://dx.doi.org/10.1016/j.pocean.2013.07.001>.
- Uitz, J., Claustre, H., Morel, A., Hooker, S.B., 2006. Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.* 111:C08005. <http://dx.doi.org/10.1029/2005JC003207>.
- Uitz, J., Claustre, H., Gentili, B., Stramski, D., 2010. Phytoplankton class-specific primary production in the world's oceans: seasonal and interannual variability from satellite observations. *Global Biogeochem. Cycles* 24:GB3016. <http://dx.doi.org/10.1029/2009GB003680>.
- Vidussi, F., Claustre, H., Manca, B.B., Luchetta, A., Marty, J.C., 2001. Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *J. Geophys. Res.* 106 (C9), 19, 939–19, 956.
- Wessel, P., Smith, W.H.F., 1996. Global self-consistent, hierarchical, high-resolution shoreline database. *J. Geophys. Res.* 101 (#B4), 8741–8743.
- Wilson, C., Qiu, X., 2008. Global distribution of summer chlorophyll blooms in the oligotrophic gyres. *Prog. Oceanogr.* 78 (2), 107–134.