



An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing

Robert J.W. Brewin^{a,b,*}, Nick J. Hardman-Mountford^{b,c}, Samantha J. Lavender^{a,d}, Dionysios E. Raitsos^{e,f}, Takafumi Hirata^{b,c,1}, Julia Uitz^g, Emmanuel Devred^{h,i}, Annick Bricaud^j, Aurea Ciotti^k, Bernard Gentili^j

^a School of Marine Science and Engineering, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, UK

^b National Centre for Earth Observation, PML, Plymouth, UK

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

^d ARGANS Ltd, Unit 3, Drake Building, Tamar Science Park, Derriford, Plymouth, PL6 8BY, UK

^e Hellenic Centre for Marine Research (HCMR), 46.7 km Athens-Sounio, PO Box 712, 190 13 Anavissos, Attica, Greece

^f Sir Alister Hardy Foundation for Ocean Science (SAHFOS), The Laboratory, Citadel Hill, Plymouth, PL1 2PB, UK

^g Marine Physical Laboratory (MPL), Scripps Institution of Oceanography, University of California at San Diego, 9500 Gilman Drive, LaJolla, CA 92093-0238, USA

^h Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, B3H 4J1, Canada

ⁱ Ocean Science Division, Bedford Institute of Oceanography, Box 1006, Dartmouth, Nova Scotia, B2Y 4A2, Canada

^j Laboratoire d'Océanographie de Villefranche, Université Pierre et Marie Curie-Paris 6, CNRS, Villefranche-sur-mer, France

^k UNESP-Campus do Litoral Paulista, Praça Infante Dom Henrique S/N, São Vicente, São Paulo CEP 11220-900, Brazil

ARTICLE INFO

Article history:

Received 26 August 2009

Received in revised form 26 August 2010

Accepted 12 September 2010

Keywords:

Phytoplankton

Size

Ocean colour

Remote sensing

Pigment

Chlorophyll-a

SeaWiFS

Absorption

ABSTRACT

Satellite remote sensing of ocean colour is the only method currently available for synoptically measuring wide-area properties of ocean ecosystems, such as phytoplankton chlorophyll biomass. Recently, a variety of bio-optical and ecological methods have been established that use satellite data to identify and differentiate between either phytoplankton functional types (PFTs) or phytoplankton size classes (PSCs). In this study, several of these techniques were evaluated against *in situ* observations to determine their ability to detect dominant phytoplankton size classes (micro-, nano- and picoplankton). The techniques are applied to a 10-year ocean-colour data series from the SeaWiFS satellite sensor and compared with *in situ* data (6504 samples) from a variety of locations in the global ocean. Results show that spectral-response, ecological and abundance-based approaches can all perform with similar accuracy. Detection of microplankton and picoplankton were generally better than detection of nanoplankton. Abundance-based approaches were shown to provide better spatial retrieval of PSCs. Individual model performance varied according to PSC, input satellite data sources and *in situ* validation data types. Uncertainty in the comparison procedure and data sources was considered. Improved availability of *in situ* observations would aid ongoing research in this field.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Phytoplankton functional types (PFTs) refer to phytoplankton that have a specific function with regard to the scientific question being addressed (Le Quééré et al., 2005; Nair et al., 2008). For instance, diatoms are responsible for ~20% of global carbon fixation (Nelson et al., 1995) and are major contributors to the biogeochemical cycling of silicon (Falcioro et al., 2000). In terms of primary production and the global carbon cycle, cell size, referred to here as phytoplankton size class (PSC), has previously been adopted to classify the functional groups (Sieburth

et al., 1978). According to the conceptual model of Sieburth et al. (1978), the autotrophic pool is split into picoplankton (<2 µm), nanoplankton (2–20 µm) and microplankton (>20 µm) contributions. While from a biogeochemical perspective the cell size functional classification may not be fully satisfactory (see Nair et al. (2008)) many ecological and biogeochemical processes are related to cell size. These processes include light absorption, as influenced by the cellular pigment composition and packaging effect (Bricaud et al., 1995, 2004; Duysens, 1956; Kirk, 1975; Morel & Bricaud, 1981; Prieur & Sathyendranath, 1981), nutrient uptake (Probyn, 1985; Sunda & Huntsman, 1997), sinking rate and export (Boyd & Newton, 1999; Laws et al., 2000; Michaels & Silver, 1988).

In recent years, a variety of bio-optical methods have been established that use satellite data to identify and differentiate between either PFTs (e.g. Alvain et al., 2005, 2008; Aiken et al., 2007; Raitsos et al., 2008) or PSCs (e.g. Ciotti & Bricaud, 2006; Devred et al., 2006; Uitz et al., 2006; Hirata et al., 2008). Platt et al. (2006) describe the detection of

* Corresponding author. A410 Portland Square, Drake Circus, University of Plymouth, Plymouth, PL4 8AA, UK. Tel.: +44 1752 584883.

E-mail address: robert.brewin@plymouth.ac.uk (R.J.W. Brewin).

¹ Present address: Faculty of Environmental Earth Science, Hokkaido University N10W5, Kita-Ku, Sapporo, 060-0810, Japan.

different phytoplankton communities from satellite as a major challenge in ocean optics, which is further complicated by the sparseness of *in situ* data required to validate these algorithms.

PSC measurements from space have been incorporated into primary production Earth Observation models, resulting in a greater understanding of the contribution of different PSCs to global ocean primary production (Claustre et al., 2005; Hirata et al., 2009; Mouw & Yoder, 2005; Silió-Calzada et al., 2008; Uitz et al., 2009, 2008, 2010). PFT measurements could potentially be used to verify PFT ecological models, as the lack of *in situ* PFT data is regarded as a major problem in PFT ecosystem modelling (Anderson, 2005). However, in order to confidently use these satellite products, as with any satellite-derived geophysical or biogeochemical product, validation exercises need to be conducted to ascertain accuracy and limitations. This is of particular importance when a field of research, such as the detection of PFTs from satellite data, is in its early stages of development, referred to as the research mode (Platt et al., 2008). In such cases, validation exercises may be used to raise questions that can guide future efforts in the field.

Previous attempts to validate geophysical or biogeochemical algorithms have resulted in intercomparison efforts. These include the SeaWiFS Bio-optical Algorithm Mini-workshop (SeaBAM; O'Reilly et al., 1998), designed to identify chlorophyll-a algorithms suitable for operational use by SeaWiFS; the primary production algorithm round robin series (PPARR) (Campbell et al., 2002; Carr et al., 2006; Friedrichs et al., 2009), designed to evaluate and compare models which estimate primary production from ocean colour; and the Inherent Optical Properties (IOP) Algorithm Workshop (Werdell, 2009), designed to provide an international forum, data, and the processing framework within which a new generation of global IOP products can be developed and evaluated.

The PPARR series was supported by NASA for over a decade and structured itself around three main tasks: (i) to compare participating bio-optical models against globally representative *in situ* data (PPARR1 and PPARR2; Campbell et al., 2002); (ii) to see how the global output of the participating bio-optical models compare with each other (PPARR3; Carr et al., 2006); and (iii) to focus on particular geographical areas where extensive *in situ* data is available, and hence the phytoplankton ecology is known, and quantitative testing of the participating bio-optical models can be made (PPARR3 and PPARR4; Friedrichs et al., 2009; Saba et al., 2010).

Following the structure of the PPARR series, this study aims to take the first step towards the validation and comparison of different PFT and PSC satellite algorithms by comparing current approaches to six different sources of *in situ* data, spanning from 1997 to 2007, in order to assess their ability at detecting size classes. By applying these satellite techniques to the 10-year SeaWiFS ocean-colour data series and by comparing the results with *in situ* data, a better understanding of the performance of these algorithms can be gained and issues can be raised which may influence future efforts in the field.

2. Review of current approaches for detecting PFTs and PSCs from satellite

In this section we review the current approaches for identifying and detecting multiple PFTs and PSCs from satellite data. Such approaches can be categorised into three groups that: (i) use the spectral response of optical properties to distinguish between different phytoplankton groups; (ii) rely on phytoplankton abundance to infer information on the size structure; or (iii) rely on other information in addition to ocean-colour data to distinguish between different phytoplankton groups.

2.1. Spectral-response-based approaches

Spectral-response algorithms utilise differences in the optical signatures of specific phytoplankton groups to distinguish among them. Alvain et al. (2005) (extended in Alvain et al. (2008)) developed

an apparent optical property-based (AOP) method that used a large set of *in situ* pigment inventories with coincident ocean-colour spectral measurements. The method is designed to detect satellite pixels that are dominated by nanoeukaryotes (and separately *Phaeocystis* and coccolithophores), two types of picoplankton (*Prochlorococcus* and *Synechococcus*-like cyanobacteria) and diatoms. The method involved producing a look-up table (LUT) of the mean normalised water leaving radiances $L_{wn}(\lambda)$, for a given chlorophyll-a concentration $L_{wn}(\lambda, [\text{chl}])$ (the chlorophyll-a concentration is hereafter denoted [chl]). Then the satellite $L_{wn}(\lambda)$ was normalised by $L_{wn}(\lambda, [\text{chl}])$ in order to develop empirical correlations between spectral characteristics and HPLC-based diagnostic pigments. The technique was developed using $L_{wn}(\lambda, [\text{chl}])$ measured in case 1 open ocean waters where phytoplankton may be considered responsible for most of the variations in the marine optical properties (Morel & Prieur, 1977), therefore, it is only applicable to such an area.

Ciotti et al. (2002) assessed the dominant cell size of phytoplankton and their absorption spectra for a wide variety of surface waters. The phytoplankton were characterised according to their dominant cell size and taxonomic group, and the relationship between this classification and the spectral shape of the phytoplankton absorption coefficient ($a_{ph}(\lambda)$) for the whole assemblage was described. Using a two-component mathematical model, a dimensionless “size factor”, varying between 0 (100% microplankton) and 1 (100% picoplankton), was adopted to specify the complementary contribution to the normalised $a_{ph}(\lambda)$ of the smallest and largest cells in the dataset. It was found that, by classifying the cell size of the dominant organism into pico- (<2 μm), ultra- (2–5 μm), nano- (5–20 μm) or microplankton (>20 μm), more than 80% of the variability in spectral shape of $a_{ph}(\lambda)$ from 400 to 700 nm could be explained. Ciotti and Bricaud (2006) used the inherent optical property (IOP) model of Loisel and Stramski (2000) to retrieve total absorption from satellite data (i.e. SeaWiFS reflectances). A non-linear optimisation method was used to decompose the total absorption into phytoplankton and coloured detrital matter (CDM) absorption. The non-linear optimisation method also retrieved the phytoplankton size factor and the slope of the exponential decrease of the absorption coefficient of CDM with wavelength.

Spectral-response approaches rely on the covariation between some spectral features of optical properties and the dominant PFTs or PSCs. Accurately exploiting the spectral characteristics of different phytoplankton groups to identify and distinguish among them may not always be successful. Previous research into the variability of particulate absorption spectra has suggested that differences in the shape of the phytoplankton absorption spectra for different communities of phytoplankton are too small to detect from satellite (Garver et al., 1994). Problems with spectral-response algorithms can also occur when distinguishing among different phytoplankton groups with the same or similar optical signatures. Furthermore, there are difficulties dealing with variations in the spectral characteristics of the same phytoplankton group or species due to growth conditions, nutrient availability and light regimes (Nair et al., 2008). However, unlike abundance-based approaches, spectral-response approaches can detect different phytoplankton groups with the same [chl] biomass providing they have contrasting optical signatures. For example, coccolithophores produce calcite plates, or coccoliths which are highly reflective (Holligan et al., 1983) and algorithms have been proposed to identify them from other phytoplankton with similar biomass based on their spectral characteristics (Ackleson et al., 1994; Brown & Podesta, 1997; Brown & Yoder, 1994). Other satellite spectral-response algorithms have also been proposed for distinct phytoplankton groups, for instance, the cyanobacterium *Trichodesmium* (Subramaniam et al., 1999) and diatoms (Sathyendranath et al., 2004).

2.2. Abundance-based approaches

Abundance-based approaches rely on typically-observed relationships between the trophic status of the environment and the type of phytoplankton (Margalef, 1967, 1978). The trophic status can be

linked directly to biogeochemical parameters such as [chl] or related to variables such as the magnitude of $a_{ph}(\lambda)$.

Devred et al. (2006) extended the two-population absorption model of Sathyendranath et al. (2001) to retrieve $a_{ph}(\lambda)$ from [chl], assuming the assemblages of the two populations vary as the phytoplankton biomass changes. The model was applied to *in situ* data collected from six regions during 34 cruises. Significant seasonal and regional changes in the spectral form and magnitude of the specific absorption coefficients of small- and large-celled populations were identified, which were then related to changes in species composition (micro- and combined nano-picoplankton). The model was also found to be consistent with pigment analysis performed on the same datasets. The parameters of the model provide direct bio-optical and biological interpretation, furthermore, the model is flexible, in that it can also function as a spectral-response algorithm (see Section 4.1.5).

Uitz et al. (2006) examined the potential for using near-surface [chl] to infer the column-integrated phytoplankton biomass, its vertical distribution and its community composition. They analysed an extensive set of High-Performance Liquid Chromatography (HPLC) determined pigment data collected in open ocean waters. Using the detailed pigment composition and specific diagnostic pigments, the chlorophyll contribution and vertical distribution of three PSCs (micro-, nano- and picoplankton) were assessed. The results lead to an empirical parameterisation enabling vertical [chl] profiles of each PSC to be inferred from the knowledge of satellite-based [chl], the euphotic depth (Morel & Maritorena, 2001) and the mixed-layer depth (de Boyer Montégut et al., 2004).

Aiken et al. (2007) used ranges in the absorption coefficient and [chl] to classify phytoplankton into three different size classes in the Benguela upwelling. They used the backscattering characteristics to sub-divide the size classes into functional types. The distributions of the dominant phytoplankton types compared well with the observations from station HPLC data. Diatom and dinoflagellate populations were located in shallow, recently upwelled water (cold, with high nutrients) while flagellates (and prokaryotes) occurred in nutrient-poor offshore water. The validation demonstrated that an empirical analysis of remotely sensed data can be used to determine the distributions of phytoplankton functional types.

Hirata et al. (2008) used HPLC data to explore the relationship between the direct optical properties of phytoplankton and PSC. Phytoplankton were classified into their dominant taxonomic size classes of pico-, nano- and microplankton using diagnostic pigment analysis (Uitz et al., 2006; Vidussi et al., 2001) extended to account for picoeukaryotes. Two models were then developed relating either phytoplankton absorption at 443 nm ($a_{ph}(443)$) or [chl] to the spectral slope of absorption in the range 443–510 nm. The models were then validated against *in situ* data and contemporary SeaWiFS 8-day composite data, which indicated good agreement. The distributions were found to be consistent with previous basin-scale observations (Aiken et al., 2008).

Abundance-based algorithms assume that a given biomass, either the surface [chl] or $a_{ph}(\lambda)$, covaries with the dominance of, or the fraction of, a particular PFT or PSC. While such approaches can be robust at extremes (high and low), they may be limited in intermediate regions where variations in the proportions of different groups could confound the signal. In such regions, abundance-based algorithms may fail to distinguish between blooms of different PFTs or PSCs that have the same biomass. Furthermore, variability of optical properties within-species or within functional types according to growth conditions could introduce additional classification errors. Despite this, abundance-based approaches have relevance to primary production models that produce estimates based on phytoplankton biomass.

2.3. Ecological-based approaches

Ecological approaches blend spatio-temporal and physical data, in addition to bio-optical information, to help detect different phyto-

plankton groups. Raitos et al. (2008) developed an approach based on knowledge of the physical and biological regime to infer PFTs in the North Atlantic. The dominance of different groups was determined from 3732 Continuous Plankton Recorder (CPR) match-ups based on cell counts, and then compared to spatio-temporal information, sea surface temperature (SST), [chl], photosynthetically-active radiation (PAR), wind stress and $L_{wm}(\lambda)$.

The approach used an Artificial Neural Network (ANN) to discriminate among diatoms, dinoflagellates, coccolithophorids and silicoflagellates. Results showed that 70% of PFT dominance derived from the CPR was explained by the input data, and that specific PFTs dominate based on a different blend of physical, ecological and biological factors. The chemical regime was not assessed in this study as no satellite offers such data. However, it may have been indirectly tested through wind stress which is partly responsible for vertical mixing of the water column and may be an indication of nutrient availability (Raitos et al., 2008). Overall, the model indicated that spatio-temporal information (latitude, longitude, and month) and SST were the most important factors determining PFTs.

An advantage of the ecological approach is that it utilises additional information to bio-optics to detect different phytoplankton groups. This could potentially lead to better results. Furthermore, this approach may provide an insight into how different phytoplankton groups react to changing environmental conditions and which environmental parameters have the largest influence on specific phytoplankton groups. However, advanced statistical approaches are intricate, involving hidden layers and complex interactions. Therefore, from an analytical perspective they can be difficult to interpret and are heavily dependent on the quality and quantity of the input data.

3. Data

3.1. *In situ* data

Phytoplankton groups can be identified from various types of *in situ* measurements with each method exhibiting limitations and advantages. Common approaches include microscopic analysis, flow cytometry, HPLC analysis of marker pigments, and DNA sequencing. HPLC analysis has the advantage to be comprehensive in terms of phytoplankton size range, it is the only method for which a sufficient amount of globally representative data is available, and despite having limitations (see Section 5), in recent years it has been extensively used as a proxy for size class (e.g. Aiken et al., 2008, 2009; Bricaud et al., 2004, 2007; Claustre et al., 2005; Devred et al., 2006; Hirata et al., 2008; Ras et al., 2008; Uitz et al., 2009, 2006, 2008; Vidussi et al., 2001). Nair et al. (2008) highlighted that the use of any *in situ* method in isolation could imply identifications of phytoplankton groups that may not be entirely dependable, hence incorporating different *in situ* methodologies would lead to a more accurate diagnosis. Therefore, in addition to four HPLC datasets, two *in situ* cell count datasets were used in this study.

- Atlantic Meridional Transect (AMT) HPLC pigment data from 1997 to 2004 (AMT 5–15) was obtained and quality assured by statistical methods according to Aiken et al. (2009).
- NASA SeaBASS HPLC pigment data was obtained from the NASA website from 1997 to 2007 (<http://seabass.gsfc.nasa.gov/>; Werdell et al., 2003). This data was accessed on the 5th September 2008 after the removal of the CHORS HPLC pigment data.
- Hawaiian Ocean Time Series (HOTS) HPLC pigment data acquired between 1997 and 2006 (<http://hahana.soest.hawaii.edu/hot/>; Karl & Lukas, 1996).
- Bermuda Atlantic Time Series (BATS) HPLC pigment data acquired from 2002 to 2004 (<http://bats.bios.edu/data/>; Michaels & Knap, 1996).

- Phytoplankton cell count data from the Continuous Plankton Recorder (CPR) was obtained for the North Atlantic from 1997 to 2003 (Richardson et al., 2006).
- Phytoplankton cell count data was obtained from the Western Channel Observatory for the L4 site from 1997 to 2007 (<http://www.westernchannelobservatory.org.uk/>; Southward et al., 2005).

3.1.1. HPLC data

All HPLC data were classified using Diagnostic Pigment Analysis (DPA) according to two different methods: the method of Vidussi et al. (2001) extended by Uitz et al. (2006); and the method of Hirata et al. (2008). The dominant size class was established based on whether a size class (pico-, nano- or microplankton) had a diagnostic pigment to [chl] ratio of greater than 0.45, i.e. representing >45% of the total population in terms of pigment concentration (Hirata et al., 2008). For the Hirata et al. (2008) DPA method, the pigment chlorophyll-b is included in the nanoplankton size class and all samples with [chl] < 0.25 mg m⁻³ are defined as picoplankton.

Originally 2176 AMT HPLC measurements, 2131 HPLC SeaBASS measurements, 305 HPLC HOTS measurements and 34 HPLC BATS measurements were utilised. This was reduced to 1093 AMT HPLC measurements, 761 HPLC SeaBASS measurements, 96 HPLC HOTS measurements, and 34 HPLC BATS measurements by only including data taken in the top 10 m of the water column. Where there were two or more measurements within the 10 m surface layer, either the dominant size class closest to the surface was selected, or if there were more than two measurements within the surface layer, the most frequent dominant size class was selected. Any HPLC data used to develop and train any of the satellite algorithms used in this intercomparison (see Section 4) were eliminated from the database.

3.1.2. CPR data

Measurements of phytoplankton abundance (cell counts) from the CPR in the North Atlantic were collected by a high-speed plankton recorder towed behind 'ships-of-opportunity' in the surface layer of the ocean ~6–10 m deep. Each measurement represents ~18 km of tow. During the SeaWiFS period (post 1997) ships typically moved at 14.8 knots meaning each measurement is representative of ~32 min of tow, and each sample equates to ~3 m³ of filtered seawater (Richardson et al., 2006). The CPR device filters plankton on a constantly moving band of mesh silk (mesh size 270 μm). Theoretically, CPR samples are more effective at filtering microplankton and less effective at filtering nano- or picoplankton, which can slip between the silk mesh. However, species smaller than 10 μm have been identified repeatedly in the CPR samples (Hays et al., 1995) which is thought to be a result of plankton clogging up the filter and the capture on the finer threads of silk that constitute the mesh-weave (Raitso et al., 2006). The proportion of cells captured by the silk has been shown to reflect the major changes in abundance, distribution and community composition of large-celled (>10 μm) phytoplankton (Robinson, 1970), and has been shown to be consistent and comparable over time (Batten et al., 2003). Furthermore, molecular analysis of the CPR samples has indicated its use for genotyping smaller phytoplankton sizes (Ripley et al., 2008) at a resolution comparable to traditional sampling techniques (D. Schroeder, pers. comm.).

Despite the CPR approach not sampling small cells, we propose a method to retrieve the fraction of small cells from that of large cells. This method has uncertainty (see Section 5), but we think that another source of data than HPLC would add value to the comparison. For the purposes of this study, the CPR dataset was used to infer the dominant microplankton group, and samples were categorised as dominated by microplankton or not. The total number of species per sample for four PFTs (diatoms, dinoflagellates, silicoflagellates, and coccolithophores) were determined, and a final category of no-dominance was allocated to samples with no cell counts (see Raitso

et al. (2008) for details). The dominant groups were then determined using the Z-factor standardised method (Raitso et al., 2008),

$$Z_i = \frac{n_i - \bar{x}_i}{s_i}, \quad (1)$$

where, n_i is the cell count for phytoplankton type i in a sample, \bar{x}_i is the overall mean of all cell counts for each type i , and s_i is the standard deviation of all samples for type i . The largest Z_i for each sample was used as the dominant phytoplankton type. The dominant phytoplankton type can be derived from this standardised method because the number of cells between each of the four groups was substantially different (Raitso et al., 2008). Where either diatoms or dinoflagellates were dominant, samples were allocated as dominated by microplankton, and the rest of the samples (including no-dominance samples) were allocated as not dominated by microplankton. We therefore assume that the pixels not dominated by microplankton are dominated by either pico- or nanoplankton that constitute the remaining autotrophic pool. Originally 17,061 measurements were used in this study spanning 1997–2003.

3.1.3. L4 data

The L4 time series (10 nautical miles offshore of Plymouth, UK in the English Channel, 50°15'N, 4°13'W) was established by Plymouth Marine Laboratory (PML) in 1988 and, since 1992, phytoplankton species, abundance and biomass have been collected on a weekly basis. Paired samples are collected from a depth of 10 m and preserved with 2% Lugol's iodine solution (Thronsdon, 1978) and 4% buffered formaldehyde. The taxonomic groups identified in the quantitative sample analysis include diatoms, dinoflagellates, coccolithophores, flagellates and picoplankton. Diatoms and dinoflagellates were classified as microplankton and coccolithophores and flagellates as nanoplankton. Between 10 and 100 ml of the sample was settled for about 48 h, and cells identified by inverted microscopy to the species level (Southward et al., 2005). For picoplankton, however, samples were settled for >6 days and picoplankton enumerated using high magnification (e.g. ×900 mag) in appropriate numbers of fields-of-view, such as 20 or 50. Data from 1997 to 2007 were downloaded from the Western Channel Observatory website (<http://www.westernchannelobservatory.org.uk/>). The dominant phytoplankton type for each sample was estimated using the Z-factor standardised method as for the CPR data. Table 1 shows the number of dominant PSC samples in each *in situ* dataset.

3.2. Satellite data

The HPLC and CPR data were matched to Level 3 SeaWiFS daily products acquired from the NASA OceanColor website (<http://oceancolor.gsfc.nasa.gov/>), at 9 km resolution, and plus or minus one pixel (3 × 3 window). This criterion, although less restricting than NASA's 3-hour window for data and algorithm validation, was adopted to maximise the number of match-ups. Mean data for $L_{wn}(\lambda)$ (mW cm⁻² μm⁻¹ sr⁻¹), [chl] (mg m⁻³), PAR (E m⁻² d⁻¹), and optical aerosol thickness (T865 dimensionless), as well as the associated standard deviations, were calculated across the 9 pixels. The IOP models of Lee et al. (2002) (QAA v5) and Smyth et al. (2006) were used to calculate the $a_{ph}(\lambda)$ at these data points. For the Hirata et al. (2008) model, we used two IOP models in this study to minimise any potential biases that could influence the retrieval of size class when adopting a single IOP model. The match-ups resulted in 250 HPLC AMT, 305 HPLC SeaBASS, 39 HPLC HOTS, 14 HPLC BATS and 5664 CPR measurements spanning from 1997 to 2007, and shown in Fig. 1.

Advanced Very High Resolution Radiometer (AVHRR) Pathfinder 5 daily mean SST data at 4 km resolution were obtained from the NASA PO.DAAC website (<http://poet.jpl.nasa.gov/>) and matched to all the HPLC data points. Night time SST products were used to avoid the

Table 1
Intercomparison results showing the percentage accuracy of each methodological approach when compared with *in situ* datasets using method 1.

PFT technique	HPLC data Vidussi et al. (2001) DPA			HPLC data Hirata et al. (2008) DPA			L4 data			CPR data		
	Pico %	Nano %	Nano/pico %	Pico %	Nano %	Nano/pico %	Pico %	Nano %	Nano/pico %	Micro %	Nano/pico %	Micro %
Model A	78.6 ± 9.0	55.5 ± 10.9	98.2 ± 2.0	73.5 ± 7.0	71.4 ± 27.1	98.2 ± 2.0	33.3 ± 54.2	–	–	–	94.2 ± 1.3	12.8 ± 2.9
Model B	52.0 ± 9.8	87.7 ± 6.4	98.8 ± 1.5	38.0 ± 7.0	87.5 ± 15.4	98.5 ± 1.5	52.2 ± 15.0	1.4 ± 2.8	77.4 ± 15.6	61.3 ± 18.2	85.5 ± 1.7	51.7 ± 3.0
Model C	70.3 ± 5.7	61.0 ± 6.5	95.0 ± 1.9	69.7 ± 4.5	88.2 ± 7.8	95.0 ± 1.9	91.9 ± 4.4	1.3 ± 1.8	54.7 ± 11.5	74.3 ± 8.4	80.8 ± 1.2	58.5 ± 2.3
Model D	93.1 ± 3.2	22.1 ± 5.4	96.0 ± 1.8	90.7 ± 2.8	37.3 ± 13.1	96.3 ± 1.7	37.6 ± 7.9	73.4 ± 10.0	39.9 ± 11.3	54.8 ± 9.6	95.7 ± 0.6	23.3 ± 2.0
Model E	95.0 ± 2.7	65.5 ± 6.4	99.2 ± 0.8	95.9 ± 1.9	73.6 ± 11.3	99.2 ± 0.8	75.8 ± 7.0	73.4 ± 10.0	68.9 ± 10.8	53.8 ± 9.6	99.5 ± 0.2	87.4 ± 1.5
Model F	87.4 ± 4.1	37.1 ± 6.4	95.1 ± 1.9	88.7 ± 3.1	84.5 ± 9.3	95.2 ± 1.9	91.9 ± 4.4	4.5 ± 4.6	57.4 ± 11.5	66.7 ± 9.1	85.5 ± 1.1	49.8 ± 2.3
Model G	–	–	98.2 ± 1.1	–	–	98.0 ± 1.2	93.3 ± 4.1	–	–	80.5 ± 7.6	72.6 ± 1.4	60.8 ± 2.3
Model H	–	–	94.0 ± 2.1	–	–	93.8 ± 2.1	92.6 ± 4.2	–	–	80.5 ± 7.6	77.8 ± 1.3	60.4 ± 2.3
Model I	64.7 ± 7.1	45.5 ± 9.2	88.4 ± 3.7	61.8 ± 5.9	66.7 ± 18.3	88.5 ± 3.6	78.9 ± 8.6	–	–	77.4 ± 15.6	74.1 ± 1.8	67.0 ± 2.8
Model B2	77.2 ± 8.3	76.5 ± 8.3	98.3 ± 1.7	68.4 ± 6.7	81.3 ± 19.2	98.0 ± 1.8	84.4 ± 11.0	4.2 ± 6.2	37.1 ± 17.7	105	3863	1801
Number of samples	239	213	452	404	55	459	149	77	74	151	1233 ^A	489 ^A
	84 ^A	82 ^A	166 ^A	155 ^A	14 ^A	169 ^A	6 ^A	36 ^B	31 ^B	67 ^B	2223 ^B	1061 ^B
	101 ^B	102 ^B	203 ^B	187 ^B	16 ^B	203 ^B	45 ^B	–	–	–	–	–
	173 ^I	112 ^I	285 ^I	263 ^I	24 ^I	287 ^I	90 ^I	–	–	–	–	–

^ADenotes the number of samples used to test model A.

^BDenotes the number of samples used to test models B and B2.

^IDenotes the number of samples used to test model I.

solar radiation bias from daily surface heating. European Remote-Sensing Satellites (ERS-2) and NASA-QuikSCAT (QS) weekly composites of mean wind stress data (ERS-2), and daily mean wind stress data (QS) (0.5° by 0.5° spatial resolution) were obtained from CERSAT, IFREMER (<http://www.ifremer.fr/cersat/en/index.htm>). All SST and wind stress data were interpolated to 9 km resolution and matched to all HPLC data points.

At the L4 station, SeaWiFS daily 1 km mapped data (L_{wn} , [chl], and T865) were acquired from the NERC Earth Observation Data Acquisition and Analysis Service (NEODAAS) in order to reduce potential interference from the adjacent land. This data was processed over the L4 station. The $a_{ph}(\lambda)$ coefficients were calculated as previously described, and the satellite data were matched to the phytoplankton cell counts which resulted in 256 matching data points between 1997 and 2004 (Fig. 1). Samples between 2002 and 2004 were matched to daily satellite data. Because samples collected before 2002 were logged with only the start day of the week as opposed to the sample day, satellite data were extracted on the Monday of every sampling week (usual sampling day), whether the *in situ* were actually sampled on that day or not (see Section 6.1 regarding the validity of this approach).

4. Methods

4.1. PFT and PSC techniques

In this intercomparison we incorporate published PFT satellite approaches designed for global application using the SeaWiFS sensor (Table 2). In this section we describe how the following PFT and PSC algorithms were reformulated or directly implemented to detect dominant phytoplankton size class (micro-, nano-, and picoplankton) and applied to the satellite *in situ* match-up data:

- Model A: Alvain et al. (2008) spectral-response approach (PHYSAT).
- Model B: Ciotti and Bricaud (2006) spectral-response approach.
- Model C: Uitz et al. (2006) abundance-based approach.
- Model D: Hirata et al. (2008) abundance-based approach ($a_{ph}(443)$; Lee et al., 2002).
- Model E: Hirata et al. (2008) abundance-based approach ($a_{ph}(443)$; Smyth et al., 2006).
- Model F: Hirata et al. (2008) abundance-based approach using [chl].
- Model G: Devred et al. (2006) abundance-based approach using regional parameters.
- Model H: Devred et al. (2006) abundance-based approach using global parameters.
- Model I: Raitos et al. (2008) ecological-based approach.

4.1.1. Model A (Alvain et al., 2008)

The PHYSAT method determines six dominant PFTs: diatoms, nanoeukaryotes (separately *Phaeocystis* and coccolithophores), *Prochlorococcus* and *Synechococcus*. In this study, we assumed that these major PFTs can be divided into the three phytoplankton size classes: microplankton = diatoms; nanoplankton = nanoeukaryotes, *Phaeocystis* and coccolithophores; and picoplankton = *Prochlorococcus* and *Synechococcus*. It is acknowledged that the approach does not identify all the phytoplankton groups within each size class. The PHYSAT algorithm (Alvain et al., 2008) was applied to SeaWiFS data with [chl] ranges between 0.04 and 4 mg m⁻³ and aerosol thickness lower than 0.15. The SeaWiFS $L_{wn}(\lambda, [chl])$ LUT was then implemented (see Table 1 of Alvain et al. (2005) for details) to determine the dominant group.

4.1.2. Model B (Ciotti & Bricaud, 2006)

The Ciotti and Bricaud (2006) approach involved initially running the updated Loisel and Stramski (2000) IOP model (Loisel & Poteau, 2006, Loisel et al. in prep) to derive total absorption $a(\lambda)$ from the remotely sensed reflectance ($R_{rs}(\lambda)$). Once $a(\lambda)$ had been retrieved, a

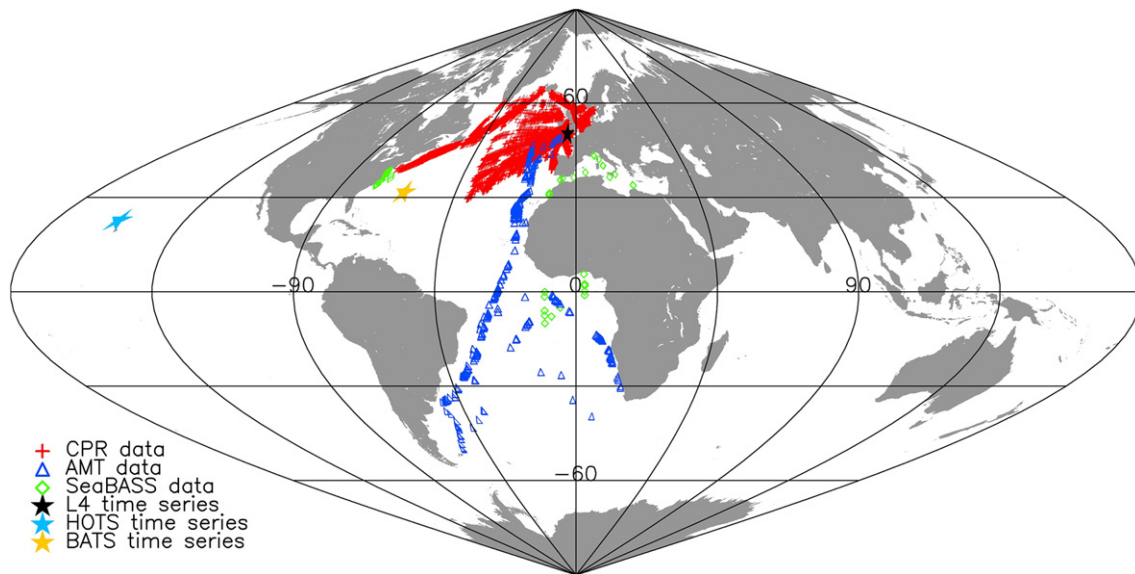


Fig. 1. Geographic distribution of *in situ* data used in the study.

nonlinear optimisation technique was used to split $a(\lambda)$ into the contributions from coloured detrital matter ($a_{CDM}(\lambda)$) and phytoplankton ($a_{ph}(\lambda)$) (see Eqs. (3) and (8) of Ciotti & Bricaud, 2006). Using the model of Ciotti et al. (2002), and following the procedure described in Ciotti and Bricaud (2006) (see ‘Method 2: nonlinear optimisation technique’ of Ciotti & Bricaud, 2006), a generalised reduced gradient nonlinear optimisation code was set up to retrieve $a_{CDM}(443)$, the spectral slope of CDM absorption (S_{CDM}) and a size parameter of phytoplankton (S_f).

When running this technique, samples with $[chl] \leq 0$ or remote sensing reflectance (R_{rs}) ≤ 0 at 412, 443, 490 or 510 nm were removed. When running the reflectance inversion any samples which gave out-of-range values for the diffuse attenuation coefficient (K_d) (and hence invalid $a(\lambda)$ values) or where the optimisation method did not converge were removed. Samples for which the retrieved S_f was dependent on the initial value (several minima in the function to minimise) were also removed. This was determined by

Table 2
Description of models used in the intercomparison.

Model	Type	Satellite input variables							Reference		
		[chl]	L_{wm} (λ)	R_{rs} (λ)	a_{ph}^1 (λ)	a_{ph}^2 (λ)	T865	PAR		WS	SST
A	SR	×	×				×			Alvain et al. (2005, 2008)	
B	SR	×		×						Ciotti and Bricaud (2006)	
C	AB	×								Uitz et al. (2006)	
D	AB				×					Hirata et al. (2008)	
E	AB					×				Hirata et al. (2008)	
F	AB	×								Hirata et al. (2008)	
G	AB	×								Devred et al. (2006)	
H	AB	×								Devred et al. (2006)	
I	EB	×	×					×	×	×	Raitsos et al. (2008)

SR = a spectral-response model, AB = an abundance-based model, EB = an ecological-based model, WS = wind stress.

$a_{ph}^1(\lambda)$ refers to $a_{ph}(\lambda)$ calculated according to Lee et al. (2002) and $a_{ph}^2(\lambda)$ refers to $a_{ph}(\lambda)$ calculated according to Smyth et al. (2006).

comparing three sets of results, with initial values of S_f equal to 0, 0.5 and 1, and discarding the samples for which the standard deviation divided by the mean exceeded 10%.

The size parameter of phytoplankton (S_f) is fixed to vary continuously between two extremes of 0 and 1, that represents the extremes in size (i.e. the largest microplankton and smallest picoplankton). For our comparison, it became necessary to establish interval values for S_f that could represent the micro-, nano- and picoplankton. To do that, we first assigned a size value for the two extremes of S_f , representative of the smallest picoplankton (referred to as a picovector) and the largest microplankton (referred to as a microvector). We then chose an equation to interpolate between these two extremes in order to estimate S_f values at 2 and 20 μm which could then be used to determine a dominant micro-, nano- or picoplankton pixel. Note that these assumptions ignore that part of the S_f variability was due to pigment packaging independent of cell size (i.e. variations in intracellular pigment concentrations resulting from photoacclimation).

The selection of a general shape for the curve of the expected decay of S_f with cell size is not a simple choice, as a large amount of noise around any curve will be expected, due to the distinct degrees of packaging possible in a given cell size (Bricaud et al., 2004) that in the field may vary in time and space. Based on the theory of Van de Hulst (1957), we can expect an exponential decay in pigment packaging as the product of the diameter and the internal concentration of the pigments increases. Therefore, this would suggest either a log-linear or exponential relationship between packaging and cell size.

Table 3 shows a range of possible values of S_f at 2 and 20 μm using both log-linear and exponential interpolations. In producing Table 3, a variety of diameter values for the fixed pico- and microvectors were used, ranging from 0.2 to 1 μm for the picovector and 30 to 120 μm for the microvector. Each picovector was set to an S_f value of 0.999 and each microvector set to an S_f value of 0.001 when interpolating. Depending on the representative sizes for the vectors and mathematical interpolation, the 2 μm S_f value varies between 0.540 and 0.943 and the 20 μm S_f value varies between 0.010 and 0.376. This large variability highlights the importance of assigning extreme cell sizes, and interpolation methods, that are appropriate for a particular region or alternatively, for large scale application, developed using globally representative data.

For this study, we used a picovector of 0.6 μm and a microvector of 60 μm (each assigned to S_f values of 0.999 and 0.001 respectively). The picovector was chosen based on *Prochlorococcus* data (see Ciotti and Bricaud (2006)) and the microvector was chosen based on samples

Table 3

S_f boundaries at 2 and 20 μm using logarithmic (log) and exponential (exp) interpolations and for a variety of assumed microvectors and picovectors set to the extremes of S_f (0.999 and 0.001 respectively).

Picovector (μm)	Size (μm)	S_f (microvector 30 μm)		S_f (microvector 60 μm)		S_f (microvector 90 μm)		S_f (microvector 120 μm)	
		log	exp	log	exp	log	exp	log	exp
		0.2	2	0.540	0.658	0.596	0.812	0.623	0.870
0.4	2	0.627	0.688	0.679	0.830	0.703	0.883	0.717	0.910
0.6	2	0.692	0.719	0.738	0.849	0.759	0.897	0.772	0.921
0.8	2	0.747	0.751	0.787	0.868	0.806	0.897	0.817	0.932
1.0	2	0.796	0.788	0.830	0.889	0.845	0.924	0.855	0.943
0.2	20	0.082	0.010	0.193	0.102	0.248	0.217	0.281	0.317
0.4	20	0.095	0.010	0.220	0.103	0.279	0.221	0.314	0.320
0.6	20	0.105	0.010	0.238	0.105	0.301	0.228	0.340	0.324
0.8	20	0.114	0.011	0.255	0.106	0.320	0.224	0.358	0.328
1.0	20	0.121	0.011	0.268	0.108	0.334	0.227	0.376	0.332

taken during a bloom of *Gonyaulax digitale* (Ciotti et al., 2002). We used a logarithmic interpolation to calculate S_f -values at 2 and 20 μm (0.74 and 0.24 respectively) which were then used to determine a pixel dominated by pico- (<2 μm), nano- (2–20 μm) and microplankton (>20 μm) after retrieving S_f from the inversion. Note that these S_f boundaries must be interpreted with caution and their limitations must be considered when discussing the results from model B.

4.1.3. Model C (Uitz et al., 2006)

The Uitz et al. (2006) approach involved dividing global oceanic waters into stratified and mixed environments based on the ratio of the euphotic depth (Z_p) (Morel & Maritorena, 2001) to the mixed-layer depth (Z_m) (de Boyer Montégut et al., 2004). As the satellite signal only penetrates the surface layer of the ocean, for the stratified environment, surface PSC percentages (see Fig. 6c of Uitz et al., 2006) were used. For the mixed environment, mixed water PSC percentages (see Table 6 of Uitz et al., 2006) representing the euphotic layer were used, as according to Uitz et al. (2006) these percentages were found to be uniform with depth. The Uitz et al. (2006) approach involves partitioning stratified and mixed environments into a small number of trophic classes according to intervals of chlorophyll-*a* values, and associating each interval with a size structure (PSC %). Here we interpolated between the mean chlorophyll-*a* values of each trophic class to avoid discontinuities. These percentages were then applied to the satellite [chl] values in order to calculate the percentage contribution of the three size classes for each pixel. The percentage of the size class at a given pixel was then converted into a dominant size class if the relative chlorophyll contribution of a particular size class was greater than 45%, as in the DPA analysis.

4.1.4. Models D, E and F (Hirata et al., 2008)

The two approaches of Hirata et al. (2008) use ranges of variation in [chl] (pico < 0.25; nano 0.25 to 1.3; and micro > 1.3 mg m^{-3} , as implemented in Aiken et al. (2008)) and in $a_{ph}(443)$ (pico < 0.024; nano 0.024 to 0.060; and micro > 0.060 m^{-1} , as implemented in Aiken et al., 2008) to distinguish dominant size class. Both methods were applied to satellite-derived [chl] and $a_{ph}(443)$ in order to determine the dominant size class. Of all the models used in this intercomparison, the Hirata et al. (2008) model was the only approach that did not need to be reformulated to detect the dominant PSC. Model D uses $a_{ph}(443)$ calculated according to Lee et al. (2002), model E uses $a_{ph}(443)$ calculated according to Smyth et al. (2006) and model F uses the satellite-derived [chl].

4.1.5. Models G and H (Devred et al., 2006)

The absorption model of Devred et al. (2006), based on the work of Sathyendranath et al. (2001), yields information about the two main optically significant phytoplankton populations in a dataset. The

information derived from fitting the model to the data consists of the specific absorption of both populations ($a_{ph1}(\lambda)$ and $a_{ph2}(\lambda)$), the rate of increase of chlorophyll in the small-celled population as a function of total chlorophyll (S_1) and the maximum chlorophyll concentration of the small-cell population (C_1^m). If the dataset used to fit the model covers a wide range of chlorophyll, the two populations of phytoplankton are assumed to be (1) a combination of pico- and nanoplankton and (2) microplankton (Devred et al., 2006).

For a given sample (pixel in our case), the proportion of each population to the total biomass can be derived either by linear combination of the derived specific absorption coefficients of the two populations, which will yield the concentration of both populations (spectral-response-based approach), or by applying Eq. (2) in Devred et al. (2006), using the chlorophyll concentration, the derived rate of increase (S_1) and the maximum concentration of the small-cell population (C_1^m) (abundance-based approach). Ideally, the fitted parameters should be computed for any given dataset (*in situ* or remotely sensed). In this study, we have used the second method (abundance-based approach) with the regional (model G) and global (model H) parameters from Devred et al. (2006) which were both derived from data that included pico-, nano- and microplankton, such that the proportion of the large-cell population can be seen as exclusively microplankton. The Devred et al. (2006) method focused on two size classes, the microplankton size class and the combined nano-picoplankton size class. Following the Uitz et al. (2006) technique, when the percentage of microplankton was greater than 45%, the pixel was allocated as dominated by microplankton, and the rest of the pixels allocated as dominated by combined nano-picoplankton.

4.1.6. Model I (Raitso et al., 2008)

The Neural Network approach of Raitso et al. (2008) was run using the spatial (latitude and longitude), temporal (month), bio-optical ([chl], L_{wn} 555), and physical (PAR, SST, and wind stress) match-up data. The approach is designed to determine the probability of four PFTs (diatoms, dinoflagellates, silicoflagellates, and coccolithophores) occurring in a satellite pixel (probability ranged from 0 to 1, 0 being not present 1 being present) in addition to the probability of none of these PFTs occurring (referred to as 'non-dominance', see Raitso et al. (2008)).

The Raitso et al. (2008) approach was developed using CPR data. Here we make an arbitrary assumption that the 'non-dominance' category can be referred to as the probability of picoplankton occurring in a satellite pixel. It can be assumed that the dominant phytoplankton size was smaller than 10 μm when a 'non-dominance' response occurred in the CPR filter (Raitso et al., 2008; Richardson et al., 2006). While it is acknowledged that nanoplankton can range from 2 to 10 μm , we base our assumption on the basis that silicoflagellate and coccolithophore cells can range from 2 to 10 μm , and that these groups may also be identified using this satellite approach and classified as nanoplankton. Nonetheless, we acknowledge that this assumption has to be used cautiously when analysing the results.

A dominant microplankton satellite pixel was allocated where either the diatom or dinoflagellate group had the highest probability of occurring, a dominant nanoplankton pixel was allocated where either the silicoflagellate or coccolithophore group had the highest probability of occurring, and a dominant picoplankton pixel was allocated where the 'non-dominance' category had the highest probability of occurring.

4.2. Comparison with *in situ* data

All methods were applied to the satellite data and compared to the *in situ* HPLC data. As a sub-set of the CPR data was used to develop model I, this model was not tested on the CPR data to avoid potential biases in the intercomparison. As model A and model I were

developed using 9 km as opposed to 1 km resolution SeaWiFS data, they were not applied to the L4 dataset. Only pixels that met the selection criteria for model A, model B and model I were used when testing these approaches in the intercomparison. Furthermore, we only tested model A using pixels where the model detected a dominant phytoplankton group as opposed to including pixels where an unidentified group was determined. This reduced the number of HPLC comparison data points from 608 to 377 for model I, 248 for model B and 175 for model A. For the L4 dataset the number of comparison data points was reduced from 256 to 98 for model B, and for the CPR dataset from 5561 to 1722 for model A and 3284 for model B.

Two methods were used to compare the satellite approaches with the *in situ* data. The first method, referred to as method 1, was designed to provide a robust calculation of the probability of detection of each size for each model. This method was designed to account for potential uncertainty in both the satellite and *in situ* measurements. The second method, referred to as method 2, was used to test inter-class errors and misclassifications in the satellite approaches.

4.2.1. Method 1: probability of detection

A method similar to Hirata et al.'s (2008) was used to analyse the performance of the satellite-derived PSCs when compared with *in situ* match-ups. This method is based on a scoring technique, with a correct classification indicating 2 points, a near-correct classification 1 point, and an incorrect classification indicating 0 points. For each match-up, the satellite approaches were run on the mean, the mean plus the standard deviation and the mean minus the standard deviation of the 9 pixels using their respective satellite input, yielding three results for each model. If any of the three results matched the dominant *in situ* size class a correct classification (2 points) was assigned. In the few cases where the three results span more than one size class, providing one of these results matched the dominant *in situ* size class, a correct classification (2 points) was still assigned, as deemed appropriate given uncertainty due to the contrasting observational scales (temporal and spatial) between *in situ* and satellite data and considering the variability around the *in situ* sample, as indexed by the satellite measurements.

For the near-correct classification criteria, if no correct classification was recorded the *in situ* data were re-analysed to assess a more mixed environment where there could be co-dominance of two size classes. For the HPLC DPA data, where the dominant size had a DPA ratio greater than 0.45 the data was also assessed to find if another size class had a DPA ratio of greater than 0.4 (based on an uncertainty estimate of 9.3% for the ratio of accessory pigment to total chlorophyll; Claustre et al., 2004; i.e. ± 0.05) at the same point; if so, a second dominant size class was recorded. For CPR and L4 cell counts, if a second size class had a Z-factor within 0.025 of the dominant size class (based on calculated 95% confidence levels when pooling CPR and L4 data), it was recorded as a second dominant size class. If any of the three satellite results matched the second dominant size class, a near-correct classification (1 point) was recorded. Otherwise, where there were no matches in any of the three satellite results, an incorrect classification (0 points) was recorded. The results were then converted into a percentage for each size class by dividing the number of points calculated for each technique by the maximum possible number of points and multiplying by one hundred. This methodology was applied to all the datasets. A flow chart of the validation procedure is shown in Fig. 2.

For each model and each size class, 95% confidence intervals were derived from the standard error of the mean percentage and the *t*-distribution of the sample size. Confidence levels provide a very powerful way of showing differences and similarities between many groups (Dythan, 2003). If the 95% confidence intervals of two or more models overlapped then we assumed that the models performed statistically similar in the comparison. If the 95% confidence intervals

of two or more models did not overlap, then we assumed that the models performed statistically different in the comparison.

4.2.2. Method 2: misclassification matrices

In order to test inter-class errors in the satellite approaches, misclassification matrices were adopted (Guptil, 1989; Nathanail & Rosenbaum, 1995). Figure 3 shows an example of a misclassification matrix. In the matrix (Fig. 3a), points on the leading diagonal have been correctly classified (Nathanail & Rosenbaum, 1995). An error of omissions occurs when a satellite prediction fails to recognize a size class that should have been identified according to the *in situ* sample. This is calculated according to the sum of the column less the leading diagonal cell value, divided by the sum of the column and multiplied by one hundred (see Fig. 3a). An error of commission occurs when a satellite prediction incorrectly identifies a pixel as a different size class. This is calculated according to the sum of the row less the leading diagonal cell value, divided by the sum of the row and multiplied by one hundred (see Fig. 3a).

A scatter plot of the errors of omission and commission in the dataset (Fig. 3b) allows the size classes that have been poorly defined by the satellite approach to be readily identified (Nathanail & Rosenbaum, 1995). Each size class in the matrix is represented by a single point on the plot with the error of omission as the ordinate and the error of commission as the abscissa. Points lying above the 45° line represent classes whose definition is too narrow leading to false exclusion of members of that size class, whereas points lying below the 45° line represent classes whose definition is too broad leading to false inclusion of members of other size classes. Points lying far from the origin reflect higher error, and points lying closer to the origin reflect lower error.

For method 2, only the *in situ* data for which a single dominant size class occurred were used (samples from the near-correct criteria in method 1 were eliminated from the datasets); this reduced the number of samples to 547 HPLC samples using the Vidussi et al. (2001) DPA (571 HPLC using the DPA of Hirata et al. (2008)), 5575 CPR samples and 246 L4 samples. For each match-up, a single dominant size class was determined for each model by calculating the most frequent size class from the three match-up results. The satellite approaches were then compared with the *in situ* data using the misclassification matrix.

5. Methodological uncertainties

There are four main areas of methodological uncertainty within the analysis. Firstly, there are measurement errors. In this study, the *in situ* data is essentially deemed to be the truth, whereas, in reality *in situ* measurements also have associated errors. Measurement outliers were minimised for the HPLC analysis through robust quality control procedures (see Aiken et al. (2009)), however, an intercomparison of HPLC pigment methods indicates instrument error of 7% for [chl] and on average 21.5% for other pigments (ranging from 11.5% for fucoxanthin to 32.5% for peridinin; Claustre et al., 2004). When comparing satellite data with *in situ* data, errors can occur due to the observational scales of the two types of measurements. The L4 cell count data was typically analysed using 10 to 100 ml samples, which are then compared with 3 km × 3 km satellite data, assuming the satellite penetrates to 10 m depth; this equates to a volume of water of 0.09 km³. The HPLC data was typically taken in volumes of sea water in the order of 5 l, whereas satellite measurements used for this study were typically representative of 27 km × 27 km, equating to an approximate volume of water of 7.29 km³. This is quite a contrast in volume when compared with the *in situ* measurements. Furthermore, there are additional errors with the satellite approaches associated with atmospheric correction and the performance of the satellite sensor itself.

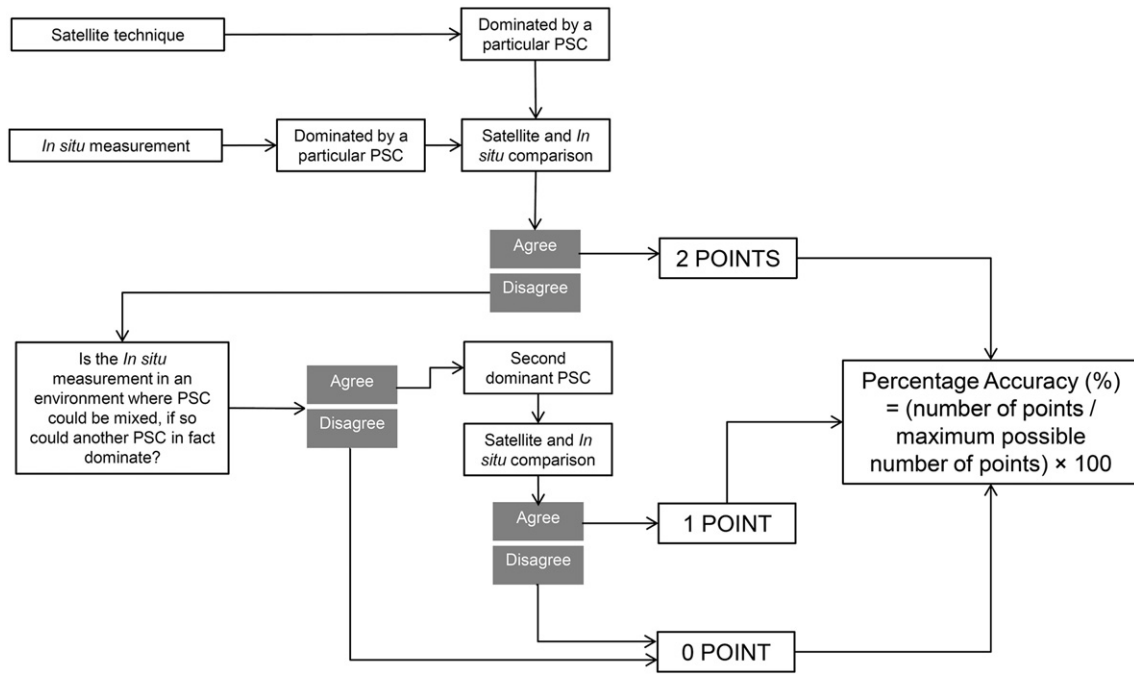


Fig. 2. Flow chart describing the validation procedure used in method 1 to test the probability of detection by the different satellite approaches.

Secondly, there are errors associated with the use of pigment concentration to determine size class. The HPLC DPA, as highlighted by Vidussi et al. (2001) and Uitz et al. (2006), does not strictly reflect the true size of phytoplankton. Diatoms, for example, have been observed in the nano-size range, whereas, in this procedure, they are only identified as microplankton. Some taxonomic pigments might be shared by various phytoplankton groups, such as fucoxanthin (the main indicator of diatoms) which may also be found in some prymnesiophytes. PFT or PSC techniques that have been developed using a specific *in situ* method, such as HPLC, are expected to perform better when compared to *in situ* data measured in the same way. This is the case of all models except B, G, H and I and will need to be taken into account when analysing the results from the HPLC data. Due to the size of the CPR mesh silk, and the fact that the CPR samples cannot quantitatively calculate the contribution of each size class, the CPR dataset is essentially a semi-quantitative estimate of dominant microplankton samples, which will also have to be considered when analysing the results. Furthermore, with regard to the L4 data, there is

expected to be higher uncertainty in the picoplankton cell counts in comparison with nano- and microplankton cell counts, due to the difficulty in counting smaller size classes when using an inverted microscope.

Thirdly, with regard to the satellite algorithms, each method is very different in its approach and it is thus very difficult to make a quantitative comparison with the *in situ* data. This study has focused primarily on size class, whereas models A and I look at specific taxonomic groups. They do not attempt to account for all the taxonomic groups within a size class that this study is assuming, although model A is based on specific diagnostic pigments as in the other HPLC based approaches.

Finally, this study assesses dominance, and some of the approaches have been adapted to fit this criterion in order to make the satellite techniques inter-comparable. Therefore, methods 1 and 2 are specifically designed to test dominance based approaches. Approaches that derive fractional contributions (e.g. models C, G and H) may fare differently in an intercomparison based on fractional

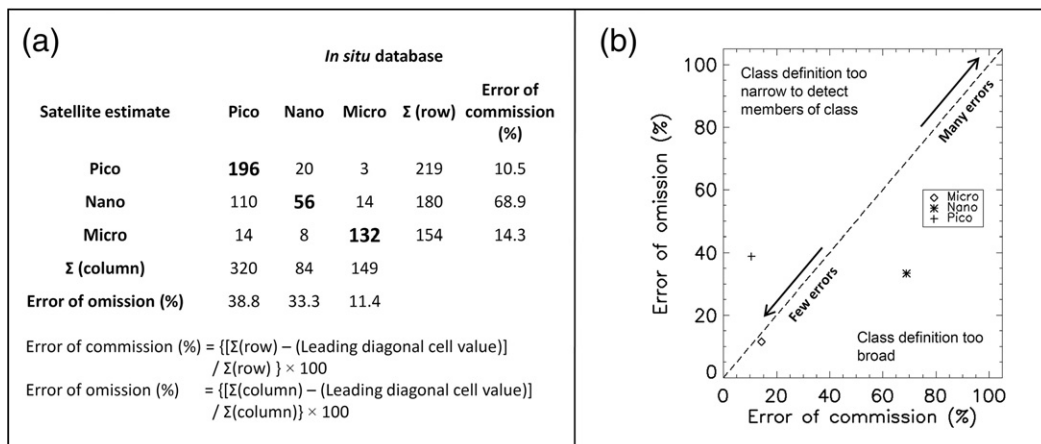


Fig. 3. Example of a misclassification matrix used in method 2 (a) and a scatter plot of omission against commission for each class in the misclassification matrix (b).

contributions. It is important to bear in mind these methodological uncertainties when discussing the performance of the algorithms.

6. Results

6.1. Method 1 results

Table 1 and Fig. 4 show the results from method 1. The error bars in Fig. 4 represent the 95% confidence levels. In the case of the HPLC results, using the Vidussi et al. (2001) DPA procedure and concerning microplankton alone, models C, F, G and H were found to perform with higher accuracy (90.1–92.9%) than models A, B, D and E (27.8–75.6%). Model I however, was not significantly different from models C, E, F, G and H. Furthermore, models E and I performed with higher accuracy (75.6–79.3%) than models A, B and D (36.5–52.2%), and models A, B and D were not statistically different. Results using the Hirata et al. (2008) DPA procedure and concerning microplankton alone, indicated that models C, E, F, G, H and I were found to perform with higher accuracy (75.8–93.3%) than models A, B and D (33.3–52.2%) although they were not significantly different from model A. Concerning combined nano-picoplankton, in both DPA procedures, all models were found to perform with similar accuracy (>88.4%).

Concerning nanoplankton, in the Vidussi et al. (2001) DPA procedure, model B performed with the highest accuracy (87.7%), followed by models C and E (61.0–65.5%). However, model A was not significantly different from models C and E, and model I was not

significantly different from models A and C. Models A, F and I performed with higher accuracy than model D. Concerning nanoplankton, using the Hirata et al. (2008) DPA procedure, models B, C, E and F performed with higher accuracy (73.6–88.2%) than model D. However, models A, B, C, E, F and I were not statistically different, and models A, D and I were not statistically different. Concerning picoplankton, in both the Vidussi et al. (2001) and Hirata et al. (2008) DPA procedures, models D, E and F performed with higher accuracy (87.4–95.9%) than models A, B, C and I. However, in Vidussi et al. (2001) DPA procedure, model F was not significantly different from model A and models C and I were not statistically different. In the Hirata et al. (2008) DPA procedure, model I performed with higher accuracy than model B. The mean percentage of all the models combined, using the HPLC dataset and the Vidussi et al. (2001) DPA procedure, was 70.7%, 95.9%, 53.8% and 77.3% for microplankton, combined nano-picoplankton, nanoplankton and picoplankton respectively. The mean percentage of all the models combined, using the HPLC dataset and the Hirata et al. (2008) DPA procedure, was 71.9%, 95.9%, 72.7% and 74.0% for microplankton, combined nano-picoplankton, nanoplankton and picoplankton respectively.

Regarding the L4 comparison, and concerning microplankton, models C, G and H performed with slightly higher accuracy than models D and E (74.3–80.5%). Concerning combined nano-picoplankton, model E was found to perform with higher accuracy than all other models ($98.7 \pm 1.8\%$), model D also was found to perform with higher accuracy than models C, F, G and H, but was not statistically different from model B. Concerning nanoplankton all models were similar (39.9–77.4%

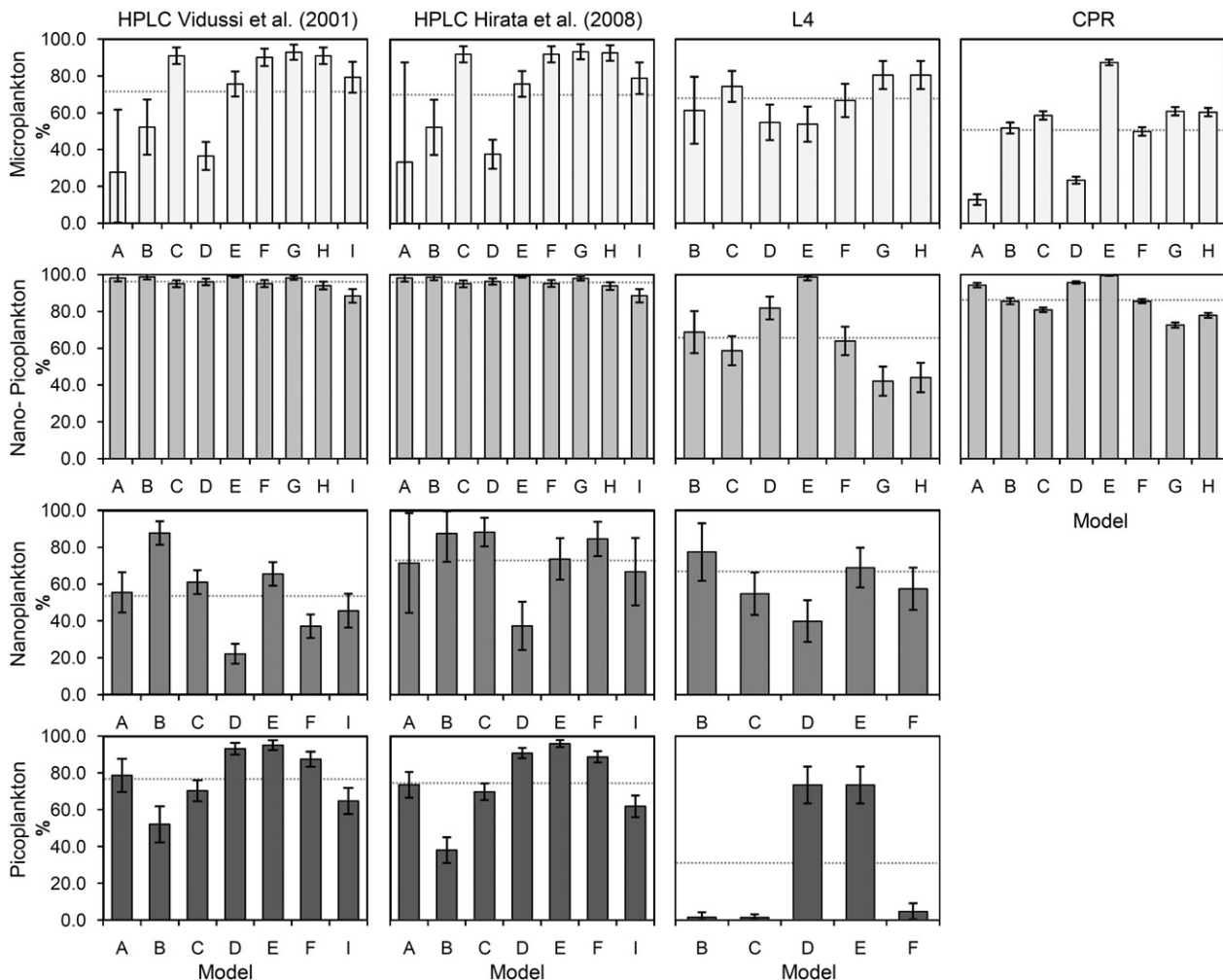


Fig. 4. Histograms showing the score (%) of satellite-derived versus *in situ* dominant PSCs for several algorithms using method 1. The error bars represent the 95% confidence intervals and the dotted line represents the mean of all models.

accuracy) with models B and E performing significantly higher than model D. Regarding picoplankton, models D and E were found to perform with the highest accuracy ($73.4 \pm 10.0\%$). The mean percentage of all the models combined, using the L4 dataset, was 67.4%, 65.4%, 59.7% and 30.8% for microplankton, combined nano-picoplankton, nano-plankton and picoplankton respectively. PSC percentages retrieved pre-2002 were compared with post-2002 percentages and a significant statistical correlation was found (p -value < 0.05) supporting the pre-2002 L4 match-up procedure described in Section 3.2.

Regarding the CPR comparison, and concerning microplankton, model E performed with the highest accuracy ($87.4 \pm 1.5\%$), followed by models G, H, C, B and F which performed with higher accuracy than models A and D. Model D performed with higher accuracy than model A. Concerning combined nano-picoplankton, model E performed with the highest accuracy ($99.5 \pm 0.2\%$) followed by models D and A, then models B and F. Models C and H performed with higher accuracy than model G. The mean percentage of all the models combined, using the CPR dataset, was 50.6% and 86.4% for microplankton and combined nano-picoplankton respectively.

6.2. Method 2 results

Figure 5 shows the scatter plots of omission against commission for each size class for (a) HPLC derived using Vidussi et al. (2001), (b) HPLC derived using Hirata et al. (2008), (c) L4 data and (d) CPR data. Consistent with method 1 the models are generally found to detect combined nano-picoplankton with the highest accuracy as indicated by their representative points lying closer to the origin when compared with other size classes.

When comparing the two DPA procedures (Fig. 5a and b), while the results for the micro- and combined nano-picoplankton are similar, the results from the pico- and nanoplankton were different. Figure 5a indicates that when using the Vidussi et al. (2001) DPA

technique, all the picoplankton data points (with the exception of models B and I) lie below the 45° line, implying that the satellite models' detection of this size class is too broad and that they are incorrectly trapping members of other size classes. Alternatively, the nanoplankton data points appear to lie above the 45° line, implying that the satellite models' are poorly identifying nanoplankton and that their detection of this size class is too narrow. When using the Hirata et al. (2008) DPA (Fig. 5b) all the nanoplankton data points appear to lie below the 45° line implying that all the models' detection of this size class is too broad. Both DPA procedures indicate that the models appear to detect picoplankton with higher accuracy than nanoplankton, as the picoplankton points lie closer to the origin, consistent with the mean percentages of all models shown in method 1 (HPLC data using Vidussi et al. (2001) DPA).

Regarding the L4 dataset and taking all the results from the models into account, there appears to be no obvious bias with all points lying evenly around the 45° line (Fig. 5c). However, individually, models D and E appear to be detecting combined nano-picoplankton too broadly and microplankton too narrowly, conversely, models B, C, F, G and H appear to be detecting microplankton too broadly and combined nano-picoplankton too narrowly. Model F appears to also classify picoplankton too narrowly. Regarding the CPR dataset (Fig. 5d), models A, D, and E appear to classify combined nano-picoplankton too broadly and microplankton too narrowly, and models B, C, F, G and H lie evenly around the 45° line implying no obvious bias.

With regard to model B, by placing limits on the S_f values we found that the model appears to be detecting pico- and microplankton too narrowly and nano- and combined nano-picoplankton too broadly, when compared with the HPLC data (see Fig. 5a and b). Note that this result is reflected in the HPLC data in method 1 (Fig. 4) as model B performs accurately at detecting nanoplankton (87.5–87.7%) and combined nano-picoplankton (~98.5%), and less accurately at detecting picoplankton (38.0–52.0%) and microplankton (~52.2%). By detecting nanoplankton too broadly, model B appears to misclassify pico- and microplankton pixels as nanoplankton. S_f values derived from model B were proposed as a continuum for co-varying pigment packaging and cell size, not for detecting dominant PSC, and placing limits was not intended.

Results from Fig. 5a–b indicate that the S_f value of 0.74 was too large to accurately split the pico- and nanoplankton population (at 2 μm) and the S_f value of 0.24 was too small to accurately split the combined nano-picoplankton and microplankton population (at 20 μm). Therefore, using the Vidussi et al. (2001) DPA procedure on the HPLC data, we consecutively reduced the S_f value from 0.74 (at 2 μm) by 0.01 in each iteration until the nano- and picoplankton data points converged on the 45° line in Fig. 5a and we consecutively increased the S_f value from 0.24 (at 20 μm) by 0.01 in each iteration until the combined nano-picoplankton and microplankton population data points converged on the 45° line in Fig. 5a. This indicated that S_f values of 0.64 at 2 μm and 0.32 at 20 μm were more adequate and prevented model B detecting pico- and microplankton too narrowly and nano- and combined nano-picoplankton too broadly. Furthermore, these values compliment comparisons of S_f to the proportion of micro- and picoplankton > 0.45 (using the DPA procedure of Vidussi et al. (2001) and Uitz et al. (2006)) using a variety of data gathered by the Laboratoire d'Océanographie de Villefranche (see Bricaud et al. (2006)).

We re-ran model B in method 1 (referred to as model B2) using the new S_f values at 2 and 20 μm . Results for model B2 are shown in Table 1. Regarding the HPLC data, results for picoplankton are shown to improve significantly, increasing from $52.0 \pm 9.8\%$ to $77.2 \pm 8.3\%$ for the Vidussi et al. (2001) DPA procedure and from $38.0 \pm 7.0\%$ to $68.4 \pm 6.7\%$ for the Hirata et al. (2008) DPA procedure. For both DPA procedures, results for microplankton are also shown to improve significantly from $52.2 \pm 15.0\%$ to $84.4 \pm 11.0\%$ and results for the combined nano-picoplankton and nanoplankton did not change significantly. Such

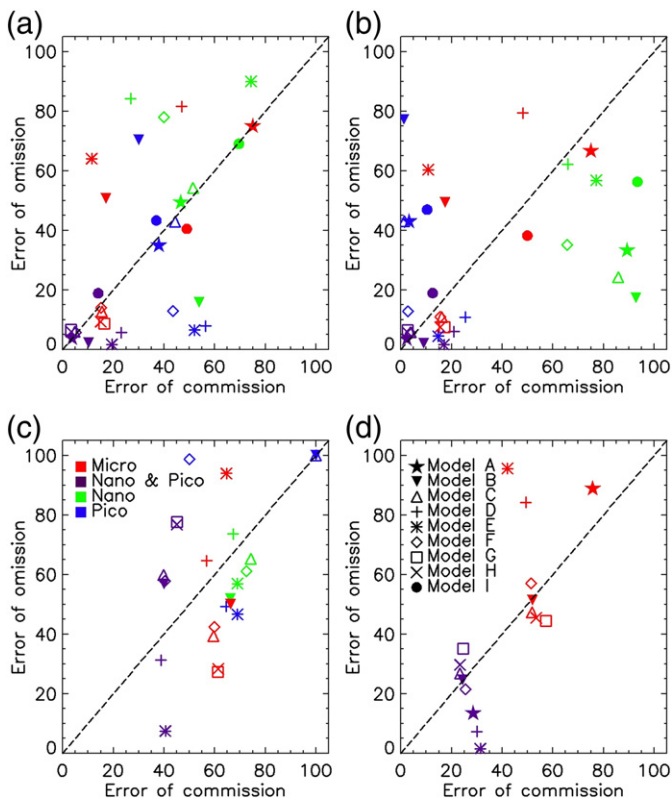


Fig. 5. Scatter plots of omission against commission for each size class in method 2 for (a) HPLC derived *in situ* using Vidussi et al. (2001), (b) HPLC derived *in situ* using Hirata et al. (2008), (c) L4 *in situ* data and (d) CPR *in situ* data.

results clearly emphasise how misclassification matrices may be used to improve model designs.

When comparing the results from methods 1 and 2 certain discrepancies arise. In method 1 (Fig. 4), model E generally performs above the average of all models and consistently performs better than model D at detecting nano- and microplankton in the HPLC and CPR datasets. However, according to method 2, model E performs similar to model D across the three size classes and appears to perform less accurately than one would expect after examining results from method 1 (points do not lie very close to the origin in Fig. 5). Closer analysis revealed that the variability in $a_{ph}(443)$ around each sample site (9 satellite pixels), when using the Smyth et al. (2006) model, was consistently higher than when using the Lee et al. (2002) model, particularly at higher $a_{ph}(443)$ values. According to method 1, a correct classification (2 points) was assigned when any of the mean, the mean plus the standard deviation and the mean minus the standard deviation of the 9 pixels matched the *in situ* sample. In method 1, model E may have benefited from higher variability in $a_{ph}(443)$ around each sample site, which would explain the discrepancies between the performance of model E in method 1 when compared with the performance of model E in method 2, particularly considering that method 2 did not account for model input variability around each sample site.

7. Discussion

The HPLC data used in this study was taken from a wide geographical area, incorporating the North and South Atlantic Oceans, the North Pacific gyre and the Mediterranean Sea, and thereby covering a number of trophic regimes. When analysing the HPLC results in methods 1 and 2 from an integrative perspective, it becomes apparent that dominant pico-, combined nano-pico-plankton and microplankton pixels are more easily detected than nanoplankton pixels from satellite, as indexed by the mean percentages of all models using the Vidussi et al. (2001) DPA (95.9% for combined nano-pico-plankton, 70.7% for microplankton, 77.3% for picoplankton and 53.8% for nanoplankton) and considering that nanoplankton data points lie further from the origin in Fig. 5a and b than other size classes. Irigoien et al. (2004) investigated global biodiversity patterns in marine phytoplankton and found their diversity to be a unimodal function of phytoplankton biomass, with maximum diversity at intermediate levels of biomass and minimum diversity at high and low levels of biomass. Assuming nanoplankton prevail at intermediate biomass (Aiken et al., 2008; Irigoien et al., 2004) one may expect higher diversity in nanoplankton than pico- or microplankton, which may result in greater variability in their optical characteristics making them harder to detect, as a community, from satellite.

Conversely, however, at the L4 site micro- and nanoplankton had a higher mean percentage when combining all models (67.4% and 59.7%) than picoplankton (30.8%). This result, to a certain extent, can be attributed to the location of the L4 site which is essentially in a case 2 region (depending on seasonal physical forcing). The [chl]-based models (C and F) performed particularly poorly at detecting picoplankton, lowering the integrative mean percentage. Furthermore, model B, which uses a non-linear relationship between [chl] and $a_{ph}(505)$ to normalise the derived absorption spectrum (see Eqs. (10) of Ciotti & Bricaud, 2006), also performed poorly at detecting picoplankton. This may be linked to the fact that the satellite estimate of [chl], acquired using the OC4v4 SeaWiFS algorithm, is known to be frequently overestimated in case 2 regions. Considering that both the [chl]-based models rely on low [chl] values to detect picoplankton, and that the derived absorption spectrum used in model B is normalised using an $a_{ph}(505)$ to [chl] relationship, this would be expected to influence the models' results significantly. Improvements in the two IOP-based models (D and E) at detecting picoplankton could be due to the advantage of using an IOP model in optically complex case 2 regions where it is important to

partition and quantify the influence various organic and inorganic water constituents have on the reflectance spectrum. All the models used are designed for open ocean waters (case 1), so their application to case 2 waters has to be handled cautiously.

When comparing spectral-response, ecological and abundance-based models, it firstly becomes apparent that, in general, models A, B and I (two spectral-response-based models and an ecological-based model) perform with similar accuracy to the abundance-based models (models C to H). Model I was shown to perform moderately well when compared with the HPLC data. This result is encouraging for ecological methods in general, as the Raitso et al. (2008) technique is not a size-class-based approach and it is different from the other models in that it relies on additional information to that of bio-optics. Considering that model I is specifically trained for the North Atlantic, it was shown to perform with fair accuracy in areas outside the North Atlantic (Note: the Artificial Neural Network failed on latitudes south of the equator). This is particularly interesting considering that Raitso et al. (2008) found spatial and temporal information to be two of the three most important variables for PFT discrimination. When constraining the *in situ* pigment data to the North Atlantic, percentage accuracy did not substantially increase (results not shown). The accuracy of the Neural Network algorithm may improve if it was trained on globally representative HPLC and cell count data. Results support the conclusions in Raitso et al. (2008) that by introducing additional information besides bio-optical information, improved satellite PSC detection may be achieved. However, further analysis needs to be conducted to verify this assumption. Use of advanced statistics such as Neural Network, self-organising maps and multilayered perceptrons, to identify additional biological information to that of [chl] from optical measurements, is a developing area of research (see Chazottes et al. (2006, 2007) and Bricaud et al. (2007)).

Model A performed with moderate accuracy at detecting pico- and nanoplankton, however, with lower accuracy at detecting dominant microplankton pixels. It should be noted that the algorithm is designed to detect specific PFTs rather than all the PFTs within a specific size class (e.g. model A does not detect dinoflagellates in the microplankton size class), the diagnostic pigment fucoxanthin is the primary pigment that was used to identify diatoms in the development of model A, as with the other HPLC abundance-based satellite approaches for detecting microplankton. Model B was also seen to perform moderately well throughout the intercomparison, particularly at detecting nanoplankton, and improved at detecting pico- and microplankton when adjusting the S_f boundaries (model B2). This result is encouraging for model B considering that the model was validated on relatively limited data (see Ciotti and Bricaud (2006)) and that the S_f value was proposed as a continuum for co-varying pigment packaging and cell size, not to detect dominant size class.

When comparing the number of samples used to test the algorithms in Table 1 it becomes apparent that the abundance-based models (C to H) were tested against approximately twice the amount of samples tested on models I and B and approximately thrice the amount of samples tested using model A. This was a consequence of the two spectral-response algorithms (models A and B) functioning on a specific criterion and the ecological-based approach (model I) limited to its geographical region (Northern hemisphere). The results from method 1 indicated that the abundance-based models performed with high accuracy across the three size classes; models C, E, F, G and H all performed with high accuracy at detecting microplankton in the HPLC dataset (75.6–93.3%) and in the L4 and CPR datasets (49.8–87.4%). Models C and E also performed with high accuracy at detecting nanoplankton in both the HPLC and L4 datasets (54.7–88.2%) and models C, D, E and F performed with high accuracy at detecting picoplankton in the HPLC dataset (69.7–95.9%). Therefore, it can be concluded from this study that abundance-based algorithms generally provide good accuracy in detecting PSC from satellite remote sensing, suitable for wide scale application.

These abundance-based approaches were also found to perform well across the entire trophic range, indicating that information on the size structure of the phytoplankton may in fact be latent in the satellite-derived biomass fields. This result supports a variety of *in situ* evidence also suggesting that the trophic status of a pelagic ecosystem may be indicative of the phytoplankton community (Bouman et al., 2005; Chisholm, 1992; Platt et al., 2005; Yentsch & Phinney, 1989) and complements Margalef's (1967, 1978) Mandala theory. Aiken et al. (2008) postulated that the bioenergetic status (BE), i.e. the transformation of light energy (photosynthesis) through intermediate stages to the synthesis of cells, is the definitive phytoplankton functional process that determines phytoplankton taxa, size classes and ecosystem trophic status, and that this bioenergetic status is quantitatively linked to phytoplankton bio-optical traits (BOTs) that are specific properties of phytoplankton size and taxa (BE-BOT hypothesis). Aiken et al. (2008) go on to suggest that [chl] and the slope of the phytoplankton absorption spectrum between 443 and 510 nm (directly linked to [chl]) are the two principle phytoplankton bio-optical traits, are indicative of phytoplankton size, and from a remote sensing perspective, in the open ocean, are the dominant organic influence on the visible light spectrum. The results from the abundance-based algorithms support such a theory.

Surrounding the general relationship between size and [chl] there is biological noise (Bricaud et al., 2004; Uitz et al., 2006), which very likely contains information on the phytoplankton. If abundance-based models were to improve, research needs to focus on understanding the source of this biological noise and accounting for it in PFT predictions. This may include results from photoacclimation of populations to various light conditions (various incident irradiances), or alternatively a decrease of irradiance with depth (see Brewin et al. (2010)).

While it is difficult to compare model I with other models, as it is based on a complex interaction of ecological as well as bio-optical input, commonalities in the other approaches can be investigated. The abundance-based models C, E, F, G and H generally performed well at detecting dominant microplankton. All these models assume that high levels of [chl] or $a_{ph}(443)$ (which are well correlated; Bricaud et al., 2004) are indicative of microplankton. Model B2 also detects microplankton well. This model relies on the assumption that microplankton have a flatter spectral shape to nano- or picoplankton. It is well known that with increasing [chl] the spectral shape of $a_{ph}(\lambda)$ becomes flatter due to changes in pigment packaging (Morel & Bricaud, 1981) and from a decrease in the relative concentration of accessory pigments (Bricaud et al., 1995). Therefore, similarities in these two approaches were expected.

Abundance-based approaches attribute lower magnitudes of [chl] to the determination of nano- and picoplankton whereas model B attributes a steepening in the spectral shape of $a_{ph}(\lambda)$ in this regard. Again these two features can be linked, as with decreasing [chl] the spectral shape of $a_{ph}(\lambda)$ is expected to become steeper due to an increase in the relative concentration of accessory pigments. This is also consistent with a number of studies suggesting that the absorption efficiency of pigments increases in the blue wavelengths with decreasing [chl] (Duysens, 1956; Hoepffner & Sathyendranath, 1991; Kirk, 1975; Morel & Bricaud, 1981; Sathyendranath et al., 1987), which would ultimately steepen the spectral shape of $a_{ph}(\lambda)$. Regarding model A, a study by Brown et al. (2008) suggests that model A is indirectly attributing lower backscattering anomalies to the determination of smaller sizes (i.e. nanoeukaryotes) in comparison to higher backscattering anomalies for larger size classes (diatoms). Similar arguments may apply to models B to H considering that nano-picoplankton generally prevail in stratified, stable environments which exhibit low heterogeneity, where one may expect lower backscattering anomalies, lower levels of [chl] and a steeper spectral shape of $a_{ph}(\lambda)$. Alternatively, higher backscattering anomalies may be expected in more dynamic heterogeneous regions, characteristic of eutrophic environments.

The robustness of abundance-based approaches appears to be a result of their utilisation of the first order variability in the satellite ocean-colour signal, whereas spectral-response-based algorithms use the second order variability which is weaker, in some respects harder to infer (Garver et al., 1994) and more vulnerable to errors in atmospheric processing despite arguably containing additional information on phytoplankton accessory pigments (Alvain et al., 2005). However, with improved sensors and atmospheric correction algorithms, spectral-response-based algorithms may become more robust.

Comparing and validating different PFT and PSC satellite algorithms in both case 1 and case 2 regions using *in situ* data is a critical issue with regard to improving synoptic estimates of these groups. Such information has the potential to be used to verify ecosystem models, improve remotely sensed primary production estimates by accounting for community composition and advance our understanding of the biogeochemical interactions between phytoplankton and their environment. Here we have attempted to validate different PFT and PSC satellite algorithms on a global scale, a procedure which is of particular importance for a novel field of study such as the detection of PFTs from satellite data. This is especially relevant when other researchers seek to use such synoptic information in different areas of biogeochemical research, including validating other models.

Throughout this comparison we have made the assumption that the *in situ* data is the truth, when in fact (as stated in section 5) there is uncertainty in the *in situ* observations, which are not direct measurements of phytoplankton size class. In order to produce a more robust intercomparison, future efforts need to focus on gathering more *in situ* data over larger spatial scales and clarify uncertainty in using *in situ* proxies to infer phytoplankton size. With extensive field measurements that combine both biological and optical measurements, more quantitative development, testing and validation of the satellite PFT algorithms can be achieved. Future comparisons may benefit from focusing on specific functional groups, in addition to size class, and by comparing the satellite results with the output from a variety of biogeochemical models, which may give an indication into how well PFT biogeochemical models are performing in comparison with the satellite approaches. Furthermore, in this rapidly expanding field, novel methods are constantly being developed to identify PFTs from remote sensing (e.g. Bracher et al., 2009; Kostadinov et al., 2009) which should also be incorporated into future comparisons. This will ultimately result in improved approaches and a better understanding of the role of different phytoplankton groups in the global carbon cycle.

8. Summary and conclusions

Nine models, designed to detect multi-phytoplankton communities from satellite data, participated in this, the first satellite-derived phytoplankton community comparison exercise. The models were reformulated to detect dominant phytoplankton size class in order to make the approaches comparable and a co-located satellite and *in situ* match-up dataset was developed. Model performance was tested by: firstly, developing a robust procedure to assess the models probability of detecting a dominant phytoplankton size class pixel and, secondly, using misclassification matrices in order to test inter-class errors in the satellite models.

Results indicate that spectral-response, ecological and abundance-based approaches can all perform with similar accuracy, though individual model performance varied according to the size of the phytoplankton, the input satellite data sources and *in situ* validation data types. Detection of microplankton and picoplankton were generally better than detection of nanoplankton and abundance-based approaches were shown to provide better spatial retrieval of PSCs. The potential use of misclassification matrices as a design aid to modify models for detecting dominant PSC was highlighted.

Uncertainties in the comparison procedure and data sources were considered and indicate that improved availability of *in situ* observations is required to advance research in this field.

Acknowledgments

The authors acknowledge all contributors to the SeaBASS dataset for *in situ* pigment data. We thank present and past staff who are involved with the maintenance, collection, analysis and processing of the BATS and HOTS pigment databases. We thank the officers and crew of RRS James Clark Ross and RRS Discovery and all scientists who helped in the acquisition and analyses of AMT data. We thank present and past staffs of SAHFOS who have contributed to the maintenance of the CPR time series and thank the present and past PML staff that have contributed to the maintenance, collection, analysis and processing of the L4 data. We would also like to thank Séverine Alvain for providing the PHYSAT look-up table and useful comments on previous versions of the manuscript, Declan Schroeder of the Marine Biological Association and Abigail McQuatters-Gollop based at SAHFOS for comments on the CPR data, Claire Widdicombe based at PML for help on the L4 data, and NEODAAS, in particular Peter Miller based at PML, for help processing satellite data over the L4 site. This work is funded by the National Environmental Research Council, UK, through a PhD studentship, the National Centre for Earth Observation and NERC Oceans 2025 programme (Themes 6 and 10). SeaWiFS data used in this publication were produced by the SeaWiFS project at the Goddard Space Flight Centre. The data were obtained from the Goddard Earth Sciences Distributed Active Archive Center under the auspices of the National Aeronautics and Space Administration (NASA). Use of this data is in accord with the SeaWiFS Research Data Use Terms and Agreements. Finally, we would like to thank the three anonymous reviewers for very constructive comments which we feel significantly improved earlier versions of the manuscript. This publication comprises AMT publication 194.

References

- Ackleson, S. G., Balch, W. M., & Holligan, P. M. (1994). Response of water-leaving radiance to particulate calcite and chlorophyll-*a* concentrations: A model for Gulf of Maine coccolithophore blooms. *Journal of Geophysical Research*, *99*(C4), 7483–7499.
- Aiken, J., Fishwick, J. R., Lavender, S., Barlow, R., Moore, G. F., Sessions, H., et al. (2007). Validation of MERIS reflectance and chlorophyll during the BENCAL cruise October 2002: Preliminary validation of new demonstration products for phytoplankton functional types and photosynthetic parameters. *International Journal of Remote Sensing*, *28*(3–4), 497–516.
- Aiken, J., Hardman-Mountford, N. J., Barlow, R., Fishwick, J. R., Hirata, T., & Smyth, T. J. (2008). Functional links between bioenergetics and bio-optical traits of phytoplankton taxonomic groups: An overarching hypothesis with applications for ocean colour remote sensing. *Journal of Plankton Research*, *30*(2), 165–181.
- Aiken, J., Pradhan, Y., Barlow, R., Lavender, S., Poulton, A., Holligan, P., et al. (2009). Phytoplankton pigments and functional types in the Atlantic Ocean: A decadal assessment, 1995–2005. AMT Special Issue. *Deep-Sea Research II*, *56*, 899–917. doi:10.1016/j.dsr2.2008.09.017.
- Alvain, S., Moulin, C., Dandonneau, Y., & Breon, F. M. (2005). Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep-Sea Research I*, *52*, 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. *Global Biogeochemical Cycles*, *22*, GB3001. doi:10.1029/2007GB003154.
- Anderson, T. R. (2005). Plankton functional type modelling: Running before we can walk? *Journal of Plankton Research*, *27*(11), 1073–1081.
- Batten, S. D., Clark, R., Flinkman, J., Hays, G. C., John, E., John, A. W. G., et al. (2003). CPR sampling: The technical background, materials and methods, consistency and comparability. *Progress in Oceanography*, *58*(2–4), 193–215.
- Bouman, H., Platt, T., Sathyendranath, S., & Stuart, V. (2005). Dependence of light-saturated photosynthesis on temperature and community structure. *Deep Sea Research I*, *52*, 1284–1299.
- Boyd, P. W., & Newton, P. (1999). Does planktonic community structure determine downward particulate organic carbon flux in different oceanic provinces? *Deep Sea Research I*, *46*, 63–91.
- 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.
- Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S., Barciela, R. M., & Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecological Modelling*, *221*, 1472–1483.
- Bricaud, A., Babin, M., Morel, A., & Claustre, H. (1995). Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization. *Journal of Geophysical Research*, *100*, 13,321–13,332.
- Bricaud, A., Ciotti, A. M., Silió-Calzada, A., & Gentili, B. (2006). Retrievals of a size parameter for phytoplankton and light absorption by colored detrital matter from ocean color measurements: Validation and application to SeaWiFS images. *Ocean Optics XVIII, Montréal, Québec, Oct. 9–13, 2006, CD-ROM*.
- Bricaud, A., Claustre, H., Ras, J., & Oubelkheir, K. (2004). Natural variability of phytoplankton absorption in oceanic waters: Influence of the size structure of algal populations. *Journal of Geophysical Research*, *109*, C11010. doi:10.1029/2004JC002419.
- Bricaud, A., Meija, C., Biondeau-Patissier, D., Claustre, H., Crepon, M., & Thiria, S. (2007). Retrieval of pigment concentrations and size structure of algal populations from their absorption spectra using multilayered perceptrons. *Applied Optics*, *46*(8), 1251–1260.
- Brown, C. A., Huot, Y., Werdell, P., Gentili, B., & Claustre, H. (2008). The origin and global distribution of second order variability in satellite ocean color and its potential applications to algorithm development. *Remote Sensing of Environment*, *112*, 4186–4203.
- Brown, C. W., & Podesta, G. P. (1997). Remote sensing of coccolithophore blooms in the Western South Atlantic Ocean. *Remote Sensing Environment*, *60*, 83–91.
- Brown, C. W., & Yoder, J. A. (1994). Coccolithophorid blooms in the global ocean. *Journal of Geophysical Research*, *99*, 7467–7482.
- Campbell, J., Antoine, D., Armstrong, R., Arrigo, K., Balch, W., Barber, R., et al. (2002). Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance. *Global Biogeochemical Cycles*, *16*(3), doi:10.1029/2001GB001444.
- Carr, M. E., Friedrichs, M. A. M., Schmeltz, M., Aita, M. N., Antoine, D., Arrigo, K. R., et al. (2006). A comparison of global estimates of marine primary production from ocean color. *Deep-Sea Research I*, *53*(5–7), 741–770.
- Chazottes, A., Bricaud, A., Crepon, M., & Thiria, S. (2006). Statistical analysis of a database of absorption spectra of phytoplankton and pigment concentrations using self-organizing maps. *Applied Optics*, *45*(31), 8102–8115.
- Chazottes, A., Crepon, M., Bricaud, A., Ras, J., & Thiria, S. (2007). Statistical analysis of absorption spectra of phytoplankton and of pigment concentrations observed during three POMME cruises using a neural network clustering method. *Applied Optics*, *46*(18), 3790–3799.
- Chisholm, S. W. (1992). Phytoplankton size. In P. G. Falkowski, & A. D. Woodhead (Eds.), *Primary productivity and biogeochemical cycles in the sea* (pp. 213–237). New York: Springer.
- Ciotti, A. M., & Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by coloured detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf off Brazil. *Limnology and Oceanography: Methods*, *4*, 237–253.
- Ciotti, A. M., Lewis, M. R., & Cullen, J. J. (2002). Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnology and Oceanography*, *47*(2), 404–417.
- Claustre, H., Babin, M., Merien, D., Ras, J., Prieur, L., Dallot, S., et al. (2005). Towards a taxon-specific parameterization of bio-optical models of primary production: A case study in the North Atlantic. *Journal of Geophysical Research*, *110*, C07S12. doi:10.1029/2004JC002634.
- Claustre, H., Hooker, S. B., Van Heukelem, L., Berthon, J.-F., Barlow, R., Ras, J., et al. (2004). An intercomparison of HPLC phytoplankton pigment methods using *in situ* samples: Application to remote sensing and database activities. *Marine Chemistry*, *85*, 41–61.
- de Boyer Montégut, C., Madec, G., Fisher, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research*, *109*, C12003. doi:10.1029/2004JC002378.
- Devred, E., Sathyendranath, S., Stuart, V., Maas, H., Ulloa, O., & Platt, T. (2006). A two-component model of phytoplankton absorption in the open ocean: Theory and applications. *Journal of Geophysical Research*, *111*, C03011. doi:10.1029/2005JC002880.
- Duysens, L. N. M. (1956). The flattening of the absorption spectrum of suspensions as compared to that of solutions. *Biochimica et Biophysica Acta*, *19*, 1–12.
- Dythan, C. (2003). Choosing and using statistics: A biologist's guide, Second edition. Oxford, UK: Blackwell Science Ltd.
- Falcatore, A., d'Alcalá, M. R., Croot, P., & Bowler, C. (2000). Perception of environmental signals by a marine diatom. *Science*, *288*(5475), 2363–2366.
- Friedrichs, M. A. M., Carr, M. E., Barber, R. T., Scardi, M., Antoine, D., Armstrong, R. A., et al. (2009). Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean. *Journal of Marine Systems*, *76*(1–2), 113–133.
- Garver, S. A., Siegel, D. A., & Mitchell, B. G. (1994). Variability in near-surface particulate absorption spectra: What can a satellite ocean color imager see? *Limnology and Oceanography*, *39*, 1349–1367.
- Guptil, S. C. (1989). Inclusion of accuracy data in a feature based object-oriented data model. In M. Goodchild, & S. Gopal (Eds.), *The accuracy of spatial databases* (pp. 91–97). London: Taylor & Francis.
- Hays, G. C., Warner, A. J., John, A. W. G., Harbour, D. S., & Holligan, P. M. (1995). Coccolithophores and the continuous plankton recorder survey. *Journal of the Marine Biological Association of the UK*, *75*, 503–506.
- Hirata, T., Aiken, J., Hardman-Mountford, N. J., Smyth, T. J., & Barlow, R. G. (2008). An absorption model to derive phytoplankton size classes from satellite ocean colour. *Remote Sensing of Environment*, *112*(6), 3153–3159.
- Hirata, T., Hardman-Mountford, N. J., Barlow, R., Lamont, T., Brewin, R. J. W., Smyth, T., et al. (2009). An inherent optical property approach to the estimation of size-specific photosynthetic rates in eastern boundary upwelling zones from satellite ocean colour: An initial assessment. *Progress in Oceanography*, *83*, 393–397.
- Hoepffner, N., & Sathyendranath, S. (1991). Effect of pigment composition on absorption properties of phytoplankton. *Marine Ecological Progress Series*, *73*(1), 11–23.

- Holligan, P. M., Viollier, M., Harbour, D. S., Camus, P., & Champagne-Philippe, M. (1983). Satellite and ship studies of coccolithophore production along the continental shelf edge. *Nature*, *304*, 339–342.
- Irigoin, X., Huisman, J., & Harris, R. P. (2004). Global biodiversity patterns of marine phytoplankton and zooplankton. *Nature*, *429*(6994), 863–867.
- Karl, D. M., & Lukas, R. (1996). The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation. *Deep Sea Research II*, *43*(2–3), 129–156.
- Kirk, J. T. O. (1975). A theoretical analysis of the contribution of algal cells to the attenuation of light within waters, II. Spherical cells. *New Phytologist*, *75*, 21–36.
- Kostadinov, T. S., Siegel, D. A., & Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *Journal of Geophysical Research*, *114*, C09015, doi:10.1029/2009JC005303.
- Laws, E. A., Falkowski, P. G., Smith, W. O., Jr., Ducklow, H., & McCarth, J. J. (2000). Temperature effects on export production in the open ocean. *Global Biogeochemical Cycles*, *14*, 1231–1246.
- Le Quéré, C., Harrison, S. P., Prentice, C. I., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, *11*(11), 2016–2040.
- Lee, Z. P., Carder, K. L., & Arnone, R. A. (2002). Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, *41*, 5755–5772.
- Loisel, H., & Poteau, A. (2006). Inversion of IOP based on R_{rs} and remotely retrieved K_d . In Z. P. Lee (Ed.), *Remote sensing of inherent optical properties: Fundamentals, tests of algorithms, and applications*. Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada.
- Loisel, H., & Stramski, D. (2000). Estimation of the inherent optical properties of natural waters from irradiance attenuation coefficient and reflectance in the presence of Raman scattering. *Applied Optics*, *41*, 2705–2714.
- Margalef, R. (1967). Some concepts relative to the organisation of plankton. *Oceanography and Marine Biology Annual Review*, *5*, 257–289.
- Margalef, R. (1978). Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanologica Acta*, *1*, 493–509.
- Michaels, A. F., & Knap, A. H. (1996). Overview of the U.S. JGOFS BATS and Hydrostation S program. *Deep Sea Research II*, *43*(2–3), 157–198.
- Michaels, A. F., & Silver, M. W. (1988). Primary production, sinking fluxes and the microbial food web. *Deep-Sea Research I*, *35*, 473–490.
- Morel, A., & Bricaud, A. (1981). Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. *Deep-Sea Research*, *28*, 1375–1393.
- Morel, A., & Maritorena, S. (2001). Bio-optical properties of oceanic waters: A reappraisal. *Journal of Geophysical Research*, *106*(C4), 7163–7180.
- Morel, A., & Prieur, L. (1977). Analysis of variations in ocean color. *Limnology and Oceanography*, *22*, 709–722.
- Mouw, C. B., & Yoder, J. A. (2005). Primary production calculations in the Mid-Atlantic Bight, including effects of phytoplankton community size structure. *Limnology and Oceanography*, *50*(4), 1232–1243.
- Nair, A., Sathyendranath, S., Platt, T., Morales, J., Stuart, V., Forget, M.-H., et al. (2008). Remote sensing of phytoplankton functional types. *Remote Sensing of Environment*, *112*(8), 3366–3375.
- Nathanail, C. P., & Rosenbaum, M. S. (1995). The misclassification matrix: A tool for validating geotechnical data. *Quarterly Journal of Engineering Geology*, *28*, 381–384.
- Nelson, D. M., Treguer, P., Brzezinski, M. A., Leynaert, A., & Queguiner, B. (1995). Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationships to biogenic sedimentation. *Global Biogeochemical Cycles*, *9*, 359–372.
- O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., et al. (1998). Ocean chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, *103*(C11), 24,937–24,953.
- Platt, T., Bouman, H., Devred, E., Fuentes-Yaco, C., & Sathyendranath, S. (2005). Physical forcing and phytoplankton distributions. *Scientia Marina*, *69*(1), 55–73.
- Platt, T., Sathyendranath, S., Forget, M. H., White, G. N., Caverhill, C., Bouman, H., et al. (2008). Operational estimation of primary production at large geographical scales. *Remote Sensing of Environment*, *112*(8), 3437–3448.
- Platt, T., Sathyendranath, S., & Stuart, V. (2006). Why study biological oceanography? *Aquabiology*, *28*, 542–557.
- Prieur, L., & Sathyendranath, S. (1981). An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter and other particulate materials. *Limnology and Oceanography*, *26*, 617–689.
- Probyn, T. A. (1985). Nitrogen uptake by size-fractionated phytoplankton populations in the southern Benguela upwelling system. *Marine Ecological Progress Series*, *22*, 249–258.
- Raitsos, D. E., Lavender, S. J., Maravelias, C. D., Haralambous, J., Richardson, A. J., & Reid, P. C. (2008). Identifying four phytoplankton functional types from space: An ecological approach. *Limnology and Oceanography*, *53*(2), 605–613.
- Raitsos, D. E., Lavender, S. J., Pradhan, Y., Tyrrell, T., Reid, P. C., & Edwards, M. (2006). Coccolithophore bloom size variation in response to the regional environment of the subarctic North Atlantic. *Limnology and Oceanography*, *51*(5), 2122–2130.
- Ras, J., Claustre, H., & Uitz, J. (2008). Spatial variability of phytoplankton pigment distributions in the Subtropical South Pacific Ocean: A comparison between *in situ* and predicted data. *Biogeosciences*, *5*, 353–369.
- Richardson, A. J., Walne, A. W., John, A. W., Jonas, T. D., Lindley, J. A., Sims, D. W., et al. (2006). Using continuous plankton recorder data. *Progress in Oceanography*, *68*(1), 27–74.
- Ripley, S. J., Baker, A. C., Miller, P. I., Walne, A. W., & Schroeder, D. C. (2008). Development and validation of a molecular technique for the analysis of archived formalin-preserved phytoplankton samples permits retrospective assessment of *Emiliania huxleyi* communities. *Journal of Microbiological Methods*, *73*(2), 118–124.
- Robinson, G. A. (1970). Continuous plankton records: Variation in the seasonal cycle of phytoplankton in the North Atlantic. *Bulletin of Marine Ecology*, *6*, 333–345.
- Saba, V. S., Friedrichs, M. A. M., Carr, M. E., Antoine, D., Armstrong, R. A., Asanuma, I., Aumont, O., Bates, N. R., Behrenfeld, M. J., Bennington, V., Bopp, L., Bruggeman, J., Buitenhuis, E. T., Church, M. J., Ciotti, A. M., Doney, S. C., Dowell, M., Dunne, J., Dutkiewicz, S., Gregg, W., Hoepfner, N., Hyde, K. J. W., Ishizaka, J., Kameda, T., Karl, D. M., Lima, I., Lomas, M. W., Marra, J., McKinley, G. A., Mélin, F., Moore, J. K., Morel, A., O'Reilly, J., Salihoglu, B., Scardi, M., Smyth, T. J., Tang, S., Tjiputra, J., Uitz, J., Vichi, M., Waters, K., Westberry, T. K., Yool, A., (2010). The challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT. *Global Biogeochemical Cycles*, *24*, doi:10.1029/2009GB003655.
- Sathyendranath, S., Lazzara, S. L., & Prieur, L. (1987). Variations in the spectral values of specific absorption of phytoplankton. *Limnology and Oceanography*, *32*, 403–415.
- Sathyendranath, S., Stuart, V., Cota, G., Maas, H., & Platt, T. (2001). Remote sensing of phytoplankton pigments: A comparison of empirical and theoretical approaches. *International Journal of Remote Sensing*, *22*, 249–273.
- Sathyendranath, S., Watts, L., Devred, E., Platt, T., Caverhill, C., & Maass, H. (2004). Discrimination of diatoms from other phytoplankton using ocean-colour data. *Marine Ecological Progress Series*, *272*, 59–68.
- Sieburth, J. M., Smetacek, V., & Lenz, J. (1978). Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnology and Oceanography*, *23*, 1256–1263.
- Silió-Calzada, A., Bricaud, A., Uitz, J., & Gentili, B. (2008). Estimation of new primary production in the Benguela upwelling area, using ENVISAT satellite data and a model dependent on the phytoplankton community size structure. *Journal of Geophysical Research*, *113*, C11023, doi:10.1029/2007JC004588.
- Smyth, T. J., Moore, G. F., Hirata, T., & Aiken, J. (2006). Semianalytical model for the derivation of ocean color inherent optical properties: Description, implementation, and performance assessment. *Applied Optics*, *45*(31), 1–16.
- Southward, A. J., Langmead, O., Hardman-Mountford, N. J., Aiken, J., Boalch, G. T., Dando, P. R., et al. (2005). Long-term oceanographic and ecological research in the western English Channel. *Advances In Marine Biology*, *47*, 1–105.
- Subramaniam, A., Carpenter, E. J., Karentz, D., & Falkowski, P. G. (1999). Bio-optical properties of the marine diazotrophic cyanobacteria *Trichodesmium* spp. I. Absorption and photosynthetic action spectra. *Limnology and Oceanography*, *44*(3), 608–617.
- Sunda, W. G., & Huntsman, S. A. (1997). Interrelated influence of iron, light and cell size on marine phytoplankton growth. *Nature*, *390*, 389–392.
- Thronsen, J. (1978). Preservation and storage. In A. Sournia (Ed.), *Phytoplankton manual* (pp. 69–74). Paris: UNESCO.
- Uitz, J., Claustre, H., Brian Griffiths, F., Ras, J., Garcia, N., & Sandroni, V. (2009). A phytoplankton class-specific primary production model applied to the Kerguelen Islands region (Southern Ocean). *Deep-Sea Research I*, *56*, 541–560.
- Uitz, J., Claustre, H., Morel, A., & Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *Journal of Geophysical Research*, *111*, C08005, doi:10.1029/2005JC003207.
- Uitz, J., Huot, Y., Bruyant, F., Babin, M., & Claustre, H. (2008). Relating phytoplankton photophysiological properties to community structure on large scales. *Limnology and Oceanography*, *52*(2), 614–630.
- 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 Biogeochemical Cycles*, *24*, doi:10.1029/2009GB003680.
- Van de Hulst, H. C. (1957). Light scattering by small particles. New York: Wiley.
- Vidussi, F., Claustre, H., Manca, B. B., Luchetta, A., & Marty, J. C. (2001). Phytoplankton pigment distribution in relation to upper thermocline and ecological circulation in the eastern Mediterranean Sea during winter. *Journal of Geophysical Research*, *106*(C9), 19,939–19,956.
- Werdell, J. (2009). Global bio-optical algorithms for ocean color satellite applications. *EOS Transactions of the American Geophysical Union*, *90*(1), doi:10.1029/2009E0010005.
- Werdell, P., Bailey, S. W., Fargion, G. S., Pietras, C., Knobelspiesse, K. D., Feldman, G. C., et al. (2003). Unique data repository facilitates ocean color satellite validation. *EOS Transactions of the American Geophysical Union*, *84*(38), 337.
- Yentsch, C. S., & Phinney, D. A. (1989). A bridge between ocean optics and microbial ecology. *Limnology and Oceanography*, *34*, 1694–1704.