



## Underway spectrophotometry along the Atlantic Meridional Transect reveals high performance in satellite chlorophyll retrievals



Robert J.W. Brewin<sup>a,b,\*</sup>, Giorgio Dall'Olmo<sup>a,b</sup>, Silvia Pardo<sup>a</sup>, Virginie van Dongen-Vogels<sup>c</sup>, Emmanuel S. Boss<sup>d</sup>

<sup>a</sup> Plymouth Marine Laboratory (PML), Prospect Place, The Hoe, Plymouth PL1 3DH, UK

<sup>b</sup> National Centre for Earth Observation, PML, Plymouth PL1 3DH, UK

<sup>c</sup> University of Technology Sydney, Plant Functional Biology & Climate Change Cluster, Sydney, Australia

<sup>d</sup> School of Marine Sciences, University of Maine, Orono, Maine 04469-5741, US

### ARTICLE INFO

#### Article history:

Received 20 December 2015

Received in revised form 4 May 2016

Accepted 14 May 2016

Available online 1 June 2016

#### Keywords:

Phytoplankton  
Ocean colour  
Remote sensing  
Chlorophyll  
Validation  
Atlantic Ocean

### ABSTRACT

To evaluate the performance of ocean-colour retrievals of total chlorophyll-a concentration requires direct comparison with concomitant and co-located *in situ* data. For global comparisons, these *in situ* match-ups should be ideally representative of the distribution of total chlorophyll-a concentration in the global ocean. The oligotrophic gyres constitute the majority of oceanic water, yet are under-sampled due to their inaccessibility and under-represented in global *in situ* databases. The Atlantic Meridional Transect (AMT) is one of only a few programmes that consistently sample oligotrophic waters. In this paper, we used a spectrophotometer on two AMT cruises (AMT19 and AMT22) to continuously measure absorption by particles in the water of the ship's flow-through system. From these optical data continuous total chlorophyll-a concentrations were estimated with high precision and accuracy along each cruise and used to evaluate the performance of ocean-colour algorithms. We conducted the evaluation using level 3 binned ocean-colour products, and used the high spatial and temporal resolution of the underway system to maximise the number of match-ups on each cruise. Statistical comparisons show a significant improvement in the performance of satellite chlorophyll algorithms over previous studies, with root mean square errors on average less than half ( $\sim 0.16$  in  $\log_{10}$  space) that reported previously using global datasets ( $\sim 0.34$  in  $\log_{10}$  space). This improved performance is likely due to the use of continuous absorption-based chlorophyll estimates, that are highly accurate, sample spatial scales more comparable with satellite pixels, and minimise human errors. Previous comparisons might have reported higher errors due to regional biases in datasets and methodological inconsistencies between investigators. Furthermore, our comparison showed an underestimate in satellite chlorophyll at low concentrations in 2012 (AMT22), likely due to a small bias in satellite remote-sensing reflectance data. Our results highlight the benefits of using underway spectrophotometric systems for evaluating satellite ocean-colour data and underline the importance of maintaining *in situ* observatories that sample the oligotrophic gyres.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### 1. Introduction

Phytoplankton are an essential component of the ocean, modifying its biological, chemical and physical environment. The majority of light absorbed by phytoplankton is transferred to heat, which can modify the temperature and physical structure of the water column (Sathyendranath, Gouveia, Shetye, Ravindran, & Platt, 1991; Zhai, Tang, Platt, & Sathyendranath, 2011), with a smaller component used in photosynthesis, the conversion of inorganic carbon (carbon dioxide) to organic carbon. Photosynthesis by phytoplankton is responsible for roughly half of net primary production on Earth (Longhurst,

Sathyendranath, Platt, & Caverhill, 1995), helping to modulate the total CO<sub>2</sub> concentration in water and its pH, influencing CO<sub>2</sub> air-sea gas exchange, carbon cycling and consequently Earth's climate. Organic carbon produced by phytoplankton is made available to most marine species as an energy source, and ultimately, influences global fish catch (Chassot et al., 2010). In addition to carbon, phytoplankton contribute to the biogeochemical cycling of a variety of climatically-important elements, such as silica, nitrate and phosphate. It is for these reasons phytoplankton are recognised as an Essential Climate Variable in the implementation plan of the Global Climate Observing System (GCOS, 2011).

Total chlorophyll-a concentration (hereafter denoted chlorophyll), refers to the sum of key photosynthetic pigment concentrations including: monovinyl chlorophyll-a; divinyl chlorophyll-a; and chlorophyllide-a (Brotas et al., 2013; Claustre et al., 2004). Partly due to difficulties in

\* Corresponding author at: Plymouth Marine Laboratory (PML), Prospect Place, The Hoe, Plymouth PL1 3DH, UK.

E-mail address: [robr@pml.ac.uk](mailto:robr@pml.ac.uk) (R.J.W. Brewin).

measuring phytoplankton carbon biomass, chlorophyll is used as a simple proxy of phytoplankton biomass (acknowledging biomass-independent changes in chlorophyll can occur from variations in light and nutrients), as it can be routinely estimated *in situ* (e.g. fluorometrically, using High Performance Liquid Chromatography (HPLC) or absorption line height) or through satellite remote-sensing of ocean colour (O'Reilly et al., 1998). Collectively chlorophyll in phytoplankton cells has a huge impact on ocean colour, visible from outer space.

Since the launch of the NASA Coastal Zone Color Scanner (CZCS) in 1978, and subsequent ocean-colour satellite missions (e.g. the Ocean Color and Temperature Sensor (OCTS), the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) of NASA; the Medium Resolution Imaging Spectrometer (MERIS) of ESA; two Moderate Resolution Imaging Spectro-radiometers (MODIS-Aqua and MODIS-Terra) of NASA; and the NASA-NOAA Visible Infrared Imager Radiometer Suite (VIIRS)), blue-to-green ratios of water reflectance and model-based algorithms have been developed to derive chlorophyll from satellite ocean-colour data. The synoptic coverage, quality and continuity of satellite chlorophyll data has led to many scientific advances (see McClain, 2009, for a review on this topic), and it is now widely regarded as the main source of data for assessing recent and future change in pelagic ecosystems (Siegel & Franz, 2010).

Despite these advances, our understanding of the accuracy and precision of ocean-colour chlorophyll data has been impeded by the limited number of, and geographic coverage of, *in situ* measurements co-incident with the satellite data (O'Reilly et al., 1998). Relative to other satellite-derived oceanic variables, such as sea-surface temperature (e.g. see Table 3 of Merchant et al., 2014), the number of *in situ* chlorophyll measurements co-incident with the satellite data is low. In addition to being limited in number, the distribution of *in situ* data available is often biased toward coastal and eutrophic waters, with limited data available in the remote oligotrophic waters, despite representing the majority of the surface ocean (Werdell & Bailey, 2005).

The Atlantic Meridional Transect (AMT) is a multidisciplinary programme designed to undertake biological, chemical, and physical measurements across a transect of >12,000 km through the centre of the Atlantic Ocean, ranging from the eutrophic shelf seas and upwelling systems, to the mid-ocean oligotrophic gyres (Aiken et al., 2000; Robinson et al., 2006). Established in 1995, the programme is now in its 20th year (Rees et al., 2015), having completed (to date) 25 cruises. It was originally conceived to test and ground truth satellite algorithms of ocean colour (in particular the SeaWiFS sensor; Hooker & McClain, 2000) as the transect crosses a wide range of ocean provinces and conditions, and importantly, crosses the sparsely sampled oligotrophic gyres. Measurements collected on AMT have contributed to global bio-optical datasets used for satellite algorithm development and validation (O'Reilly et al., 1998; Werdell & Bailey, 2005), and have been used to assess the performance of satellite ocean-colour chlorophyll retrievals (Aiken et al., 2009; Brewin et al., 2010; Brewin, Sathyendranath, Jackson, et al., 2015). AMT is widely recognised as an ideal platform to evaluate satellite ocean-colour data (Aiken & Hooker, 1997; Hooker & McClain, 2000; Rees et al., 2015).

Traditionally, satellite chlorophyll data has been validated using co-incident discrete point measurements of chlorophyll, such as that acquired through HPLC or through fluorometric chlorophyll extraction. When comparing co-incident *in situ* point measurements with the satellite data, errors can occur due to vast differences in the observational scales of the two types of measurements. Discrete *in situ* measurements of chlorophyll typically represent volumes of sea water of the order of 5 l or less, whereas satellite ocean-colour pixels are typically between 1 km and 4 km in size, which if one assumes an optical depth of 10 m (often much deeper in clear waters), equates to a volume of water in the region of  $1 \times 10^{10}$  and  $16 \times 10^{10}$  l. Furthermore, the spatial variability within a satellite pixel is very difficult to sample using discrete point measurements.

One option to address this mismatch in spatial scales between satellite and discrete point measurements is by using continuous *in situ* data

collected by a moving ship. Since the 1930's plankton measurements have been collected using the Continuous Plankton Recorder towed on research vessels and voluntary ships, and have been used for comparison with satellite ocean-colour data (Batten, Walne, Edwards, & Groom, 2003; Brewin et al., 2011; Raitos, Reid, Lavender, Edwards, & Richardson, 2005). Continuous chlorophyll data derived from a lidar fluorosensor on a moving research vessel have also been compared with ocean-colour chlorophyll data over wide spatial scales (Barbini, Colao, Fantoni, Fiorani, & Palucci, 2003; Barbini et al., 2004). A common approach to collecting continuous *in situ* chlorophyll measurements is through calibrated *in vivo* fluorescence data collected by sampling surface water from the flow-through system of a moving ship (Lorenzen, 1966). This approach has been used to validate ocean-colour data in coastal (Folkestad, Pettersson, & Durand, 2007; Harding, Magnuson, & Mallonee, 2005; Petersen, Wehde, Krasemann, Colijn, & Schroeder, 2008) and shelf regions (Hu et al., 2003, 2005; Zhang et al., 2006), and has demonstrated the importance of validating a satellite pixel using multiple samples in heterogeneous conditions (Hu, Nababan, Biggs, & Muller-Karger, 2004). Yet, this approach has its caveats. For instance, the fluorescence yield can vary between species of phytoplankton (Kiefer, 1973b; Strickland, 1968) and within a single species subjected to different environmental conditions (Kiefer, 1973a; Slovacek & Bannister, 1973). *In vivo* fluorescence in surface waters can also be affected by non-photochemical quenching during daytime (e.g. Cullen & Lewis, 1995). These problems impact the accuracy and precision of the *in situ* chlorophyll data, particularly in the oligotrophic gyres (Strass, 1990).

Recently, techniques have been developed to continuously measure the inherent optical properties of particles in surface sea water sampled by the flow-through system of a moving ship (Boss et al., 2013; Dall'Olmo, Boss, Behrenfeld, & Westberry, 2012; Dall'Olmo, Westberry, Behrenfeld, Boss, & Slade, 2009; Dall'Olmo et al., 2011; Koponen et al., 2007; Slade et al., 2010; Westberry, Dall'Olmo, Behrenfeld, & Moutin, 2010). Using these techniques, chlorophyll concentration can be estimated with remarkable accuracy and continuously, using particulate absorption measurements and the line-height of the absorption peak at red wavelengths (Boss, Collier, Larson, Fennel, & Pegau, 2007; Dall'Olmo et al., 2012; Davis, Moore, Zaneveld, & Napp, 1997; Westberry et al., 2010). These techniques have been used to evaluate ocean-colour data (Brewin, Raitos, et al., 2015; Werdell, Proctor, Boss, Leeuw, & Ouhssain, 2013) and have the potential to overcome some of the caveats of estimating chlorophyll using *in vivo* fluorescence.

In this paper, we use an optical set-up on two AMT cruises (AMT19 and AMT22) to continuously measure chlorophyll concentration from the ship's flow-through system. The chlorophyll datasets are then used to evaluate the performance of satellite chlorophyll algorithms on different ocean-colour sensors.

## 2. Methodology

### 2.1. Statistical tests

To test the performance of satellite chlorophyll algorithms we used a series of univariate statistical tests commonly used in comparisons between modelled and *in situ* data (e.g. Brewin, Sathyendranath, Müller, et al., 2015; Doney et al., 2009; Friedrichs et al., 2009), including: the Pearson correlation coefficient ( $r$ ); the root mean square error ( $\Psi$ ); the average bias between model and measurement ( $\delta$ ); the unbiased root mean square error ( $\Delta$ ); the slope ( $S$ ) and intercept ( $I$ ) of a Type-2 regression; and the percentage of possible retrievals ( $\eta$ ). For  $S$  and  $I$ , we used Type-2 regression (Glover, Jenkins, & Doney, 2011, MATLAB function `lsqfitm.m`). The equations used for each of these statistical tests are provided in Table 1. All statistical tests were performed in  $\log_{10}$  space, considering chlorophyll is approximately log-normally distributed in the ocean (Campbell, 1995, see also Fig. 3b).

**Table 1**  
Summary of statistical tests used in the study.

Symbol	Description	Equation <sup>a</sup>
$r$	Pearson correlation coefficient	$\frac{1}{N-1} \sum_{i=1}^N \left[ \frac{C_i^M - (\frac{1}{N} \sum_{j=1}^N C_j^M)}{(\frac{1}{N-1} \sum_{k=1}^N [C_k^M - (\frac{1}{N} \sum_{l=1}^N C_l^M)]^2)^{1/2}} \right] \left[ \frac{C_i^E - (\frac{1}{N} \sum_{m=1}^N C_m^E)}{(\frac{1}{N-1} \sum_{n=1}^N [C_n^E - (\frac{1}{N} \sum_{o=1}^N C_o^E)]^2)^{1/2}} \right]$
$\Psi$	Root mean square error	$\left[ \frac{1}{N} \sum_{i=1}^N (C_i^E - C_i^M)^2 \right]^{1/2}$
$\delta$	Bias	$\frac{1}{N} \sum_{i=1}^N (C_i^E - C_i^M)$
$\Delta$	Unbiased root mean square error	$\left( \frac{1}{N} \sum_{i=1}^N \left\{ \left[ C_i^E - \left( \frac{1}{N} \sum_{j=1}^N C_j^E \right) \right] - \left[ C_i^M - \left( \frac{1}{N} \sum_{k=1}^N C_k^M \right) \right] \right\}^2 \right)^{1/2}$
$S$	Slope of Type-2 regression <sup>b</sup>	$\frac{C^E - I}{C^M}$
$I$	Intercept of Type-2 regression <sup>b</sup>	$C^E - S C^M$
$\eta$	Percentage of possible retrievals	$\frac{N^E}{N^M} 100$

<sup>a</sup>  $C$  denotes the variable (chlorophyll concentration) and  $N$  is the number of samples with both estimated and measured data. The superscript  $E$  denotes the estimated variable (e.g. using satellite data) and the superscript  $M$  denotes the measured variable (e.g. measured *in situ*).

<sup>b</sup> Type-2 regression between  $C^M$  and  $C^E$  (where  $C^E = C^M S + I$ ) was used to derive  $S$  and  $I$ .

## 2.2. Underway optical sampling

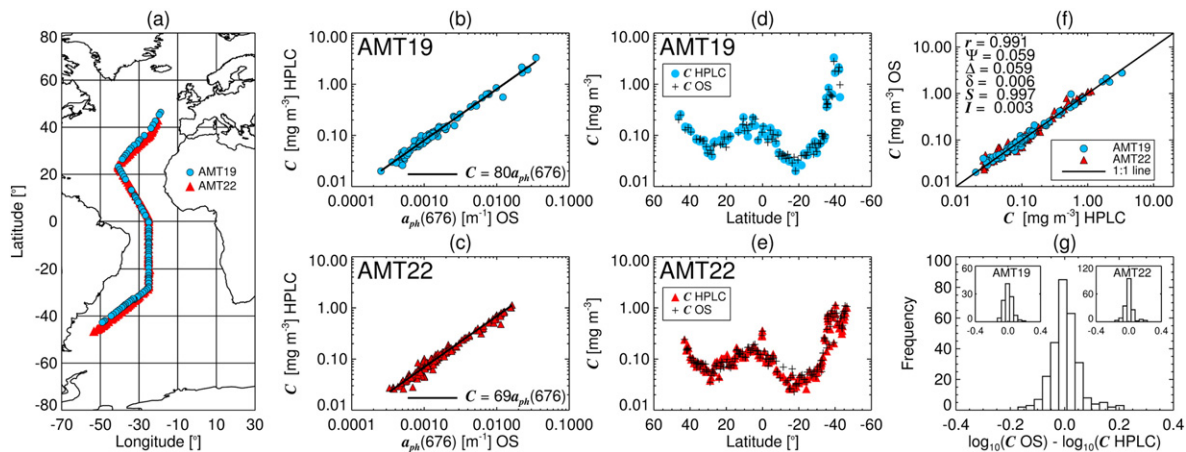
AMT19 and AMT22 underway data were collected on board the *RRS James Cook* from the 14th of October to the 28th of November 2009, and the 15th October to the 20th of November 2012, respectively. Both cruises followed a very similar cruise track, spanning 50°N to 50°S (Fig. 1).

On both cruises, optical instruments were attached to the ship's clean flow-through system, continuously pumping seawater from a nominal depth of about 5 m. The methods of Dall'Olmo et al. (2009), Slade et al. (2010) and Dall'Olmo et al. (2012) were followed, which involved first passing water through a Vortex debubbler, then either passing water directly through the optical instruments (50 min for every hour) or diverting seawater through a Cole Parmer 0.2  $\mu\text{m}$ -cartridge filter (for 10 min every hour), the later used to provide a baseline for particulate absorption measurements. Either a WET Labs AC-S hyperspectral spectrophotometer (hyperspectral between ~400 and ~750 nm, with a spectral resolution of 5 nm and a band pass of 15 nm), or a WET Labs AC-9 (nine wavelengths between 412 and 715 nm, with a band pass of 10 nm) were used to measure spectral absorption. Spectral particulate absorption ( $a_p(\lambda)$ ) were calculated by subtracting the 0.2  $\mu\text{m}$  filtered measurements from the unfiltered

measurements, providing calibration-independent estimates of  $a_p(\lambda)$  accounting for instrumental drifts and residual calibration errors. Following Dall'Olmo et al. (2009), data were converted into 1-min median bins. The 1-min binned data were medians of higher frequency data (~240 measurements): when the ship moves at ~18 km h<sup>-1</sup> (typically), each 1-min binned average is representative of approximately 0.3 km. For AMT19, an AC-S was used at the beginning of the cruise but after 13 days the lamp of the attenuation channel (i.e., C-channel) failed, and an AC9 meter was added to the flow-through system (Dall'Olmo et al., 2012). As the "band pass" of the AC9 is narrower (10 nm) than that of the AC-S (15 nm), concurrent AC-9 and AC-S  $a_p(\lambda)$  data were calibrated to ensure no systematic bias in  $a_p(\lambda)$  between instruments (see Section 2.1.3 of Dall'Olmo et al., 2012, for further details). For AMT22, a WET Labs AC-S was used during the entire cruise.

For both AMT19 and AMT22, discrete water samples (2 to 4 l) were collected along the transects from the underway flow-through system. The water samples were filtered onto Whatman GF/F filters (nominal pore size of 0.7  $\mu\text{m}$ ) and stored in liquid nitrogen. Chlorophyll (C) was determined after the cruise in the laboratory using HPLC analysis.

To estimate chlorophyll from the underway optical system, data for  $a_p(\lambda)$  were extracted at 650, 676 and 715 nm. The phytoplankton absorption coefficient at 676 nm ( $a_{ph}(676)$ ) was then estimated using



**Fig. 1.** (a) Cruise tracks for AMT19 (Oct–Nov 2009) and AMT22 (Oct–Nov 2012) for underway HPLC data; (b) and (c) show the relationship between co-located (20 min time window)  $a_{ph}(676)$ , derived from the underway optical system (OS) and underway HPLC chlorophyll (C), for AMT19 (106 samples) and AMT22 (176 samples) respectively; (d) and (e) show chlorophyll (C) from HPLC and the underway optical system (OS) with latitude for the two cruises; (f) shows the relationship between HPLC and OS-inferred chlorophyll; and (g) absolute  $\log_{10}$ -transformed difference between HPLC and OS-inferred chlorophyll.

the line height method of Davis et al. (1997) as modified by Boss et al. (2007), such that

$$a_{ph}(676) = a_p(676) - [39/65a_p(650) + 26/65a_p(715)]. \quad (1)$$

To convert  $a_{ph}(676)$  into chlorophyll concentrations ( $C$ ), we extracted concurrent data on  $a_{ph}(676)$  from the optical system to that of the discrete HPLC chlorophyll data. The  $a_{ph}(676)$  data were averaged in  $\log_{10}$  space over a 20-min period centered on the time the discrete HPLC water samples were collected ( $\pm 10$  min), then back-transformed. We then fitted a non-linear relationship (Bricaud, Babin, Morel, & Claustre, 1995; Bricaud, Morel, Babin, Allali, & Claustre, 1998) between  $a_{ph}(676)$  and  $C$ , such that

$$C = Aa_{ph}(676)^B. \quad (2)$$

The parameters  $A$  and  $B$  were determined for each cruise separately, by fitting Eq. (2) using HPLC chlorophyll and corresponding data on  $a_{ph}(676)$  from the optical system. For AMT19, estimates of  $A$  and  $B$  were 88 and 1.02 ( $N = 106$ ), respectively. For AMT22,  $A$  and  $B$  were 62 and 0.99 ( $N = 176$ ), respectively. In both cases, the slope of the power-law function ( $B$ ) was not significantly different from 1.0, suggesting a linear relationship between  $a_{ph}(676)$  and  $C$  along the two AMT cruise tracks, in contrast to previous studies using global datasets (e.g. Brewin, Devred, Sathyendranath, Hardman-Mountford, & Lavender, 2011; Bricaud et al., 1995, 1998; Werdell et al., 2013). To abide by the law of parsimony, Eq. (2) was replaced with a linear relationship between  $a_{ph}(676)$  and  $C$ , such that

$$C = Aa_{ph}(676). \quad (3)$$

The parameter  $A$  was estimated as  $80 \pm 2.0$  and  $69 \pm 1.3$  for AMT19 and AMT22 respectively (Fig. 1b and c), where the uncertainties are the 95% confidence intervals of the means. Note that the parameter  $A$  is influenced not only by the chl-specific absorption coefficient at 676 nm, but also by differences in the optical set-up on the two cruises (e.g. different instruments with different spectral responses). Fig. 1d–g show a comparison of chlorophyll estimated from the optical set-up (Eq. (3)) with the corresponding HPLC chlorophyll data. The optical set-up is shown to estimate chlorophyll with very good accuracy along the two AMT transects. Eqs. (1) and (3) were used to reconstruct chlorophyll for all  $a_p$  data collected on AMT19 and AMT22, resulting in 45,171 1-min binned chlorophyll samples for AMT19 and 34,934 for AMT22.

### 2.3. In situ hyperspectral radiometry

To aid interpretation of the satellite chlorophyll validation results, we used *in situ* above-water hyperspectral radiometry data collected on AMT19 and AMT22. An above-water Hyperspectral Surface Acquisition Remote Sensing System (SATLANTIC HYPERSAS) was installed on a fixed pole on the bow of the ship on both AMT19 and AMT22. Hyperspectral downwelling irradiance ( $E_s$ ), sky radiance ( $L_i$ ) and water-leaving radiance ( $L_t$ ) were recorded at a number of stations along the AMT track. These stations all occurred around local noon, where the ship stopped for CTD profiles. On both cruises, the  $E_s$  sensor was pointed toward zenith, with the  $L_i$  and  $L_t$  sensors deployed at fixed angles facing the sky and water respectively ( $\sim 40^\circ$  and  $\sim 130^\circ$  respectively from nadir, assuming a horizontal ship). Sensor windows were regularly cleaned during both cruises with lens paper.

$E_s$ ,  $L_i$  and  $L_t$  were extracted from the HYPERSAS for a 1-h period over the duration of each station, using SATLANTIC SatView and SatCon software. The HYPERSAS data were processed as follows:

- On each instrument, a shutter closes periodically to record dark values. The  $E_s$ ,  $L_i$  and  $L_t$  data were first dark corrected, by interpolating the dark value data in time to match the light measurements for each sensor, then subtracting the dark values from the light measurements at each wavelength.

- The  $E_s$ ,  $L_i$  and  $L_t$  were then interpolated to the same set of wavelengths (every 3.5 nm from 350–800 nm), which coincides roughly with the wavelengths of the  $E_s$  sensor, the instrument with the smallest number of channels.
- As the three sensors have different integration times and thus collect data at slightly different time stamps, the  $E_s$ ,  $L_i$  and  $L_t$  data were interpolated to the same set of time stamps, which was selected based on the sensor with the slowest integration time (typically the  $L_t$  sensor). This resulted in  $E_s$ ,  $L_i$  and  $L_t$  data at the same time and same sets of wavelengths.
- For each station, only spectra with a sun zenith angle of  $<60^\circ$ , an azimuth angle between either  $100$  and  $170^\circ$  (centered at  $135^\circ \pm 35^\circ$ ) were used (Mobley, 1999). Any spectra with negative values at 443 nm (which can occur when cleaning the sensor) were removed.
- To minimise sun glint contamination we exploited the near-infrared portion of the  $L_t$  reflectance spectrum which, in open ocean waters, should be close to zero. The statistical distribution of  $L_t$ (NIR) data, where NIR represents the average of  $L_t$  in the region 750–800 nm, at each station was analysed and spectra were only retained in the lower 5th percentile of  $L_t$ (NIR) (Hooker, Lazin, Zibordi, & McLean, 2002).
- Remote-sensing reflectance ( $R_{rs}(\lambda)$ ) was then computed according to  $R_{rs}(\lambda) = [L_t(\lambda) - \rho L_i(\lambda)]/E_s(\lambda)$ , where  $\rho$  was computed for each station following (Mobley, 2015), using the median wind speed, azimuth angle, and sun zenith angle over the duration of each station, and assuming a viewing angle of  $40^\circ$ .
- $R_{rs}(\lambda)$  data in the near-infrared were computed (averaged in the region 750–800 nm) and subtracted from each spectra, to remove any additional contamination by sky and sun glint.
- For the remaining spectra at each station, remote-sensing reflectance ratios ( $R_{rs}(443)/R_{rs}(547)$  and  $R_{rs}(488)/R_{rs}(547)$ ) were computed. Median  $R_{rs}(443)/R_{rs}(547)$  and  $R_{rs}(488)/R_{rs}(547)$  values were extracted for each station, after degrading the hyperspectral  $R_{rs}$  data to 11 nm averages centered on each wavelength, to be consistent with NOMAD data used to parameterise the NASA OC-series of chlorophyll algorithms (Werdell & Bailey, 2005). To remove noisy station data and maximise the consistency of the dataset, only station data were used where the maximum coefficient of variation of  $R_{rs}(443)/R_{rs}(547)$  and  $R_{rs}(488)/R_{rs}(547)$  was less than 0.15.
- Finally, chlorophyll data from the optical system were extracted for the same time period at each station (1 h) and median concentrations were computed. This resulted in 8 stations on AMT19 and 22 stations on AMT22 with concurrent *in situ*  $R_{rs}(443)/R_{rs}(547)$ ,  $R_{rs}(488)/R_{rs}(547)$  and chlorophyll.

### 2.4. Satellite ocean-colour datasets

We conducted our ocean-colour evaluation using daily, level 3, 4 km binned satellite ocean-colour products. The choice to use level 3 (4 km) ocean-colour products for the evaluation, as opposed to level 2 (1 km) typically used in satellite validation protocols (Bailey & Werdell, 2006), stems from: (i) the continuous underway sampling method used allows for many samples to be collected within a 4 km pixel, to account for the effects of sub-pixel variability over a larger pixel area; (ii) the merged ocean-colour products evaluated here are merged at level 3 rather than level 2; and (iii) in a recent user requirements survey (Sathyendranath, 2011), it was found that ecosystem modellers and earth observation scientists using ocean-colour data have a preference for level 3 products over level 2. Nonetheless, we investigate the impact of using level 3 data for validation by comparing results with a validation using level 2 data.

For the AMT19 period (14th of October to the 28th of November 2009), MODIS-Aqua (R2014.0 and R2013.1) and MERIS (processed with SeaDAS, R2012.1) daily, global, level 3 spectral remote-sensing

reflectance  $R_{rs}(\lambda)$  data were downloaded from the NASA website (<http://oceancolor.gsfc.nasa.gov/>). Level 3, global, 4 km MERIS ocean-colour  $R_{rs}(\lambda)$  products were also produced using the POLYMER atmospheric-correction algorithm (version 3.0; Steinmetz, Deschamps, & Ramon, 2011) over the AMT19 period. SeaWiFS (R2010.0) level 3, global, 4 km data were produced at Plymouth Marine Laboratory (PML), by re-binning Level 2 SeaWiFS data (acquired from NASA) to a 4 km grid, as the NASA level 3 SeaWiFS products are provided at 9 km. In addition to single-sensor datasets, we also downloaded merged ocean-colour products from the European Space Agency (ESA) Ocean Colour Climate Change Initiative (OC-CCI) over the AMT19 period. OC-CCI data constitute merged (level 3, 4 km binned) MERIS (POLYMER), MODIS-Aqua and SeaWiFS products, and are available at <http://www.oceancolour.org/> (Sathyendranath et al., 2012). Both versions 1.0 and 2.0 of the OC-CCI products were downloaded and used in the study. Level-2, 1 km MODIS-Aqua (R2014.0) data were also acquired from NASA for AMT19, for cross-comparison of Level 2 and Level 3 MODIS-Aqua data. Level 2 data were processed using standard NASA flags.

For the AMT22 period (15th October to the 20th of November 2012), MODIS-Aqua (R2014.0 and R2013.1) and VIIRS (R2014.1) daily, global, level 3 spectral remote-sensing reflectance  $R_{rs}(\lambda)$  data were downloaded from the NASA website (<http://oceancolor.gsfc.nasa.gov/>). OC-CCI version 1.0 and 2.0 data were also acquired for AMT22, with version 2.0 downloaded from the OC-CCI website (<http://www.oceancolour.org/>) and version 1.0 produced at Plymouth Marine Laboratory for the study, noting that the publicly available OC-CCI version 1.0 dataset ends in July 2012. Table 2 highlights the satellite datasets used in this study.

### 2.5. Satellite and *in situ* match-up procedure

The following procedure was implemented to ensure high quality match-ups between Level 3 satellite  $R_{rs}(\lambda)$  and *in situ* chlorophyll data.

- To minimise effects of sub-pixel variability on the validation, satellite  $R_{rs}(\lambda)$  data were matched in time (same day of year) and space (latitude and longitude, closest 4 km pixel) with the *in situ* chlorophyll data for AMT19 and AMT22. When one or more *in situ* samples were matched to the same satellite pixel, the *in situ* chlorophyll concentrations were averaged (using  $\log_{10}$  transformation) and considered as a single match-up. Only samples were used where there were >5 *in situ* data points within a satellite pixel (Fig. 2a), and where the standard deviation of the *in situ*  $\log_{10}(C)$  measurements was less than 0.1 (~95 percentile of data, see Fig. 2b).
- To test for homogeneity of the region surrounding the satellite match-up, eight other satellite pixels surrounding the centre pixel (total of 9, 3×3) were also extracted from the satellite data. The coefficient of variation (median coefficient of variation for  $R_{rs}$  bands between 412 and 555 nm) for each box of nine pixels was then computed. Match-ups were excluded if the coefficient of variation was >0.15 (Fig. 2c, similar to Bailey & Werdell, 2006, acknowledging that they used 5×5 km level 2 data) and when <50% of the pixels were available in the surrounding region.
- Finally, the average solar zenith angle for the *in situ* chlorophyll data within each satellite match-up was computed, using the time and location of *in situ* data collection. To minimise the time difference between satellite and *in situ* data collection, only match-ups with a solar zenith angle <90° were used (Fig. 2d), meaning that only underway data collected during daylight hours were used in the study. Regarding this latter point, the time difference between satellite and *in situ* data will vary depending on satellite overpass time, which is different for each satellite. For merged products, this becomes more complicated to compute, as a merged product contains a combination of information from different satellite sensors, that collect data at

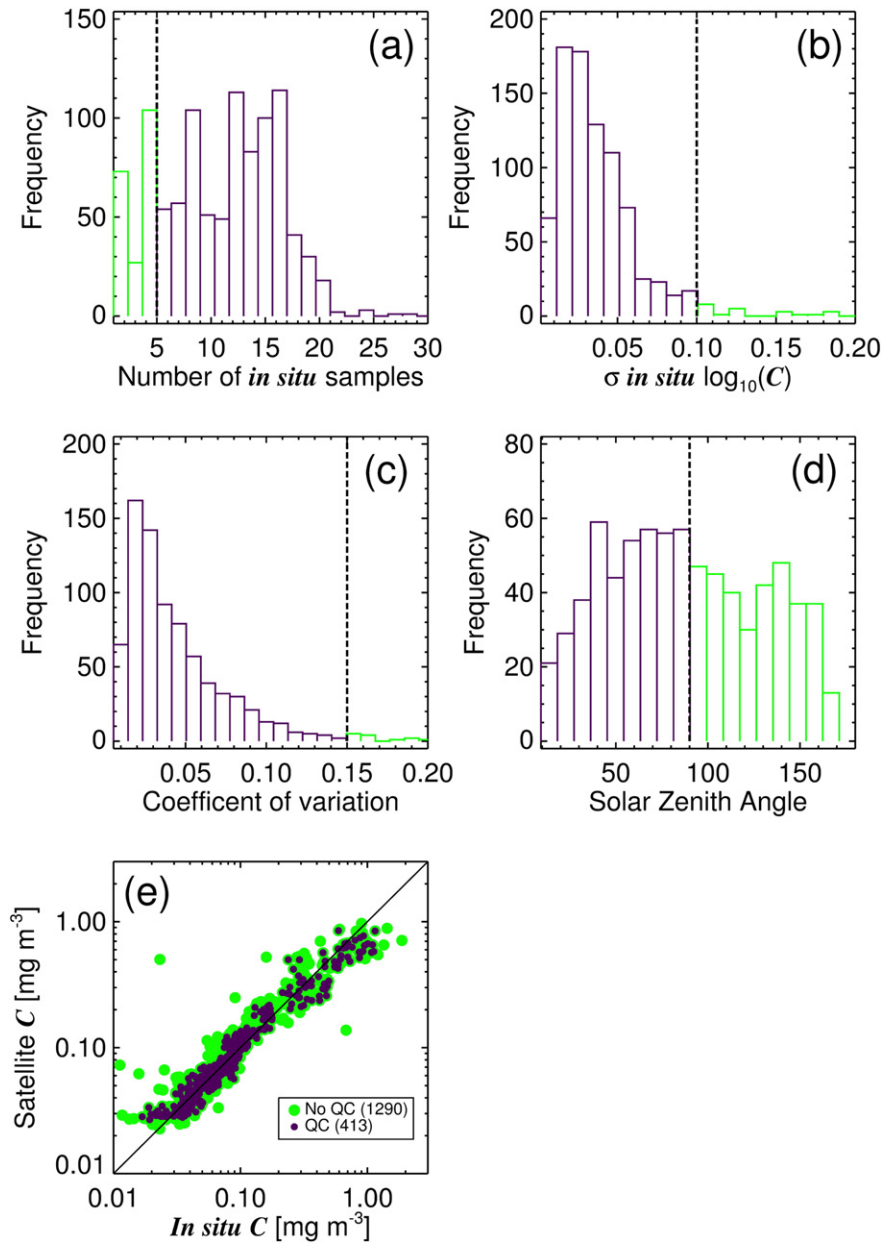
**Table 2**  
Information on the satellite datasets used in this study.

AMT	Sensor	Product	Standard chlorophyll algorithm used	Version/Reprocessing	Temporal <sup>a</sup> resolution	Spatial resolution	Processing level	Projection	Atmospheric correction algorithm used	N <sup>b</sup>
19	MODIS-Aqua	$R_{rs}(\lambda)$	OC3M-547	R2013.1	Daily	~4 km	Level 3	Sinusoidal	SeaDAS	139
19	MODIS-Aqua	$R_{rs}(\lambda)$	OC3M-547	R2014.0	Daily	~4 km	Level 3	Sinusoidal	SeaDAS	142
19	MODIS-Aqua	$R_{rs}(\lambda)$	OC3M-547	R2014.0	Swath time	~1 km	Level 2	Native <sup>c</sup>	SeaDAS	231
19	SeaWiFS	$R_{rs}(\lambda)$	OC4	R2010.0	Daily	~4 km	Level 3	Sinusoidal	SeaDAS	56
19	MERIS	$R_{rs}(\lambda)$	OC4E	R2012.1	Daily	~4 km	Level 3	Sinusoidal	SeaDAS	176
19	MERIS	$R_{rs}(\lambda)$	OC4E	Version 3.0	Daily	~4 km	Level 3	Sinusoidal	POLYMER	289
19	OC-CCI v1.0	$R_{rs}(\lambda)$	OC4	Version 1.0	Daily	~4 km	Level 3	Sinusoidal	SeaDAS/POLYMER	401
19	OC-CCI v2.0	$R_{rs}(\lambda)$	OC4	Version 2.0	Daily	~4 km	Level 3	Sinusoidal	SeaDAS/POLYMER	413
22	MODIS-Aqua	$R_{rs}(\lambda)$	OC3M-547	R2013.1	Daily	~4 km	Level 3	Sinusoidal	SeaDAS	156
22	MODIS-Aqua	$R_{rs}(\lambda)$	OC3M-547	R2014.0	Daily	~4 km	Level 3	Sinusoidal	SeaDAS	155
22	VIIRS	$R_{rs}(\lambda)$	OC3V	R2014.0	Daily	~4 km	Level 3	Sinusoidal	SeaDAS	167
22	OC-CCI v1.0	$R_{rs}(\lambda)$	OC4	Version 1.0	Daily	~4 km	Level 3	Sinusoidal	SeaDAS/POLYMER	155
22	OC-CCI v2.0	$R_{rs}(\lambda)$	OC4	Version 2.0	Daily	~4 km	Level 3	Sinusoidal	SeaDAS/POLYMER	156

<sup>a</sup> Data acquired from 14th of October to 28th of November 2009 for AMT19, and 15th October to 20th of November 2012 for AMT22.

<sup>b</sup> Number of satellite and *in situ* match-ups used in the evaluation.

<sup>c</sup> Satellite data collected at the native resolution of the sensor.



**Fig. 2.** The satellite and *in situ* match-up procedure using OC-CCI v2.0 data for AMT19 and the OCI chlorophyll algorithm. (a) Step 1: histogram of the number of *in situ* data points within a 4 km satellite match-up pixel, showing the threshold of 5 samples with data included in purple and excluded in green. (b) Step 2: histogram of the standard deviation of the *in situ*  $\log_{10}(C)$  measurements within each satellite pixel, showing the threshold of 0.1 (~95 percentile of data) with data included in purple and excluded in green. Step 3: (c) histogram of the median coefficient of variation for  $R_{rs}$  bands between 412 and 555 nm, for a box of nine pixels (with  $\geq 50\%$  coverage) surrounding the satellite match-up pixel. Match-ups were excluded (green) if there was a coefficient of variation  $> 0.15$  (Bailey & Werdell, 2006). Step 4: (d) histogram of solar zenith angles for remaining data, match-ups were excluded (green) with a solar zenith angle greater than  $90^\circ$ , and hence remaining match-ups were collected during daylight hours. (e) Scatter plot of satellite and *in situ* match-ups showing samples before (green) and after (purple) applying quality control (Steps 1–4).

slightly different periods of the day. By eliminating data collected during the night, and considering the satellite overpass times of SeaWiFS, MODIS-Aqua, VIIRS and MERIS vary between around  $\pm 2.5$  h local noon, the maximum time difference between *in situ* and satellite data used for a match-up is not likely exceed  $\sim 8$  h.

Fig. 2a–e illustrate the satellite match-up procedure used on OC-CCI v2.0 data for AMT19 using the OCI chlorophyll algorithm (see following section for description of OCI algorithm). The percentage of match-up data retained following application of the quality control step was  $\sim 32\%$  for this example (Fig. 2). For the level 2 match-ups, the same exact procedure was used, with the only difference being that match-ups were kept where there were  $> 3$  *in situ* data points within a satellite

pixel (rather than  $> 5$  for the level 3 data), to account for the fact that the level 2 pixels are smaller in size. We also computed the time difference between satellite overpass and *in situ* data collection for the level 2 data.

## 2.6. Ocean-colour chlorophyll algorithms

The satellite chlorophyll ( $C$ ) algorithms incorporated into the comparison are described in this section. Each algorithm uses  $R_{rs}(\lambda)$  as input, and was applied to satellite  $R_{rs}(\lambda)$  to compute chlorophyll.

### 2.6.1. OC-series

The NASA OC-series of algorithms refer to a series of polynomial, band-ratio chlorophyll algorithms (O'Reilly et al., 2000) that relate the

log-transformed ratio of blue to green remote-sensing reflectances ( $X$ ) to the chlorophyll concentration ( $C$ ). For the NASA OC4 algorithm,  $X$  is computed as

$$X = \log_{10}\{\max[R_{rs}(443), R_{rs}(490), R_{rs}(510)]/R_{rs}(555)\}. \tag{4}$$

Depending on the band set of the satellite, and on whether a maximum band-ratio (maximum of 2 or 3 pairs of wavebands as in Eq. (4)) or a single band-ratio algorithm is used,  $X$  can be computed in different ways. Table 3 shows the band-ratio algorithms used in this study, their identifier, their associated satellite datasets, the wavebands used to compute  $X$ , and whether the algorithm was considered as the standard algorithm for the associated satellite dataset (see also Table 2). Once  $X$  is known, chlorophyll ( $C$ ) can be estimated according to:

$$C = 10^{(q_0+q_1X+q_2X^2+q_3X^3+q_4X^4)}, \tag{5}$$

where  $q_0, q_1, q_2, q_3$  and  $q_4$  are empirical coefficients that vary according to the particular OC band-ratio algorithm used (see Table 3).

### 2.6.2. OCI

The band-difference algorithm of Hu, Lee, and Franz (2012) was also tested in this study. This algorithm has been found to perform well in oligotrophic environments ( $<0.25 \text{ mg m}^{-3}$  Hu et al., 2012; Brewin, Raitso, Pradhan, & Hoteit, 2013; Brewin, Raitso, et al., 2015). The approach uses a Colour Index (denoted here as  $\xi$ ), based on a band-difference between remote-sensing reflectance in the green part of the visible spectrum and a base-line formed linearly between the blue and red wavebands, such that:

$$\xi = R_{rs}(555) - 0.5[R_{rs}(443) + R_{rs}(670)]. \tag{6}$$

Chlorophyll is then related to  $\xi$  using the following equation:

$$C = 10^{A+B\xi}, \tag{7}$$

where  $A = -0.4909$  and  $B = 191.659$ . Since Eq. (7) was designed specifically for waters with low chlorophyll ( $\leq 0.25 \text{ mg m}^{-3}$ ), at higher chlorophyll concentrations ( $>0.3 \text{ mg m}^{-3}$ ) a standard band-ratio algorithm (e.g. OC4 for SeaWiFS) is used (Eqs. (4) and (5)), whereas for chlorophyll concentrations between 0.25 and  $0.3 \text{ mg m}^{-3}$ , a combination of Eq. (7) and a standard band-ratio algorithm is used to allow a smooth transition between algorithms. For OC-CCI data, the OCI algorithm is expressed as

$$C = \begin{cases} 10^{A+B\xi} & \text{if } [10^{A+B\xi}] \leq 0.25 \text{ mg m}^{-3} \\ \alpha [10^{(q_0+q_1X+q_2X^2+q_3X^3+q_4X^4)}] + (1-\alpha) [10^{A+B\xi}] & \text{if } 0.25 < [10^{A+B\xi}] \leq 0.3 \text{ mg m}^{-3} \\ 10^{(q_0+q_1X+q_2X^2+q_3X^3+q_4X^4)} & \text{if } [10^{A+B\xi}] > 0.3 \text{ mg m}^{-3}, \end{cases} \tag{8}$$

where  $\alpha$  serves to provide a linear transition from Eq. (7) to Eq. (5) as chlorophyll increases from 0.25 to  $0.3 \text{ mg m}^{-3}$ , with  $q_0 = 0.3272, q_1 = -2.9940, q_2 = 2.7218, q_3 = -1.2259$  and  $q_4 = -0.5683$ . The  $\alpha$  parameter is computed as  $\alpha = (10^{A+B\xi} - 0.25)/(0.3 - 0.25)$ . Whereas we used the OC series of algorithms (OC4, OC4E, OC3M-547 and OC3V) as the standard algorithm for the NASA, ESA and OC-CCI datasets, we acknowledge that NASA are now processing ocean-colour datasets using the OCI algorithm as the standard algorithm of choice, in addition to the standard OC-series of algorithms.

### 2.6.3. GSM

The semi-analytical Garver-Siegel-Maritorena (GSM) model, initially developed by Garver and Siegel (1997) and later updated by

**Table 3**  
The NASA OC-series of empirical, band-ratio chlorophyll algorithms used in the study, together with their associated coefficients (NASA, 2010).

Satellite data	Identifier	Equation used to compute X (see Eq. (4))	$q_0$	$q_1$	$q_2$	$q_3$	$q_4$	Standard algorithm used <sup>a</sup>
SeaWiFS & OC-CCI	OC4	$X = \log_{10}\{\max[R_{rs}(443), R_{rs}(490), R_{rs}(510)]/R_{rs}(555)\}$	0.3272	-2.9940	2.7218	-1.2259	-0.5683	Y
SeaWiFS & OC-CCI	OC3S	$X = \log_{10}\{\max[R_{rs}(443), R_{rs}(490)]/R_{rs}(555)\}$	0.2515	-2.3798	1.5823	-0.6372	-0.5692	N
SeaWiFS & OC-CCI	OC2S	$X = \log_{10}\{R_{rs}(490)/R_{rs}(555)\}$	0.2511	-2.0853	1.5035	-3.1747	0.3383	N
MODIS-Aqua	OC3M-547	$X = \log_{10}\{\max[R_{rs}(443), R_{rs}(488)]/R_{rs}(547)\}$	0.2424	-2.7423	1.8017	0.0015	-1.2280	Y
MODIS-Aqua	OC2M-547	$X = \log_{10}\{R_{rs}(488)/R_{rs}(547)\}$	0.2500	-2.4752	1.4061	-2.8233	0.5405	N
MERIS	OC4E	$X = \log_{10}\{\max[R_{rs}(443), R_{rs}(490), R_{rs}(510)]/R_{rs}(560)\}$	0.3255	-2.7677	2.4409	-1.1288	-0.4990	Y
VIIRS	OC3V	$X = \log_{10}\{\max[R_{rs}(443), R_{rs}(486)]/R_{rs}(551)\}$	0.2228	-2.4683	1.5867	-0.4275	-0.7768	Y

<sup>a</sup> Refers to whether the algorithm is used as the standard algorithm for the associated satellite dataset (Y = Yes and N = No). Whereas we used the OC series of algorithms (OC4, OC4E, OC3M-547 and OC3V) as the standard algorithms for the NASA, ESA and OC-CCI datasets, we acknowledge that NASA are now processing ocean-colour datasets using the OCI algorithm as the standard algorithm of choice, in addition to the standard OC-series of algorithms.

Maritorena, Siegel, and Peterson (2002), was also used in this study. GSM is based on an underlying bio-optical model, where

$$R_{rs}(\lambda) = \left\{ \sum_{i=1}^2 g_i \left[ \frac{b_{bw} + b_{bp}(\lambda_0) \left(\frac{\lambda}{\lambda_0}\right)^{-\gamma}}{b_{bw} + b_{bp}(\lambda_0) \left(\frac{\lambda}{\lambda_0}\right)^{-\gamma} + a_w(\lambda) + C a_{ph}^*(\lambda) + a_{dg}(\lambda_0) \exp[-S_{dg}(\lambda - \lambda_0)]} \right]^i \right\} 0.5238, \quad (9)$$

$\lambda_0 = 443$ , and  $g_i$ ,  $\gamma$ ,  $S_{dg}$  and  $a_{ph}^*(\lambda)$  are predefined input parameters (Maritorena et al., 2002). The value 0.5238 represent a conversion from below-water ( $R_{rs}(\lambda, -0)$ ) to above-water remote-sensing reflectance ( $R_{rs}(\lambda)$ ). Using non-linear optimisation, the GSM model retrieves simultaneous estimates of chlorophyll ( $C$ ), absorption by combined detrital and dissolved matter at 443 nm ( $a_{dg}(\lambda_0)$ ) and particle backscattering at 443 nm ( $b_{bp}(\lambda_0)$ ) from  $R_{rs}(\lambda)$ . This method was designed to estimate chlorophyll independent of influence from  $a_{dg}(443)$  and  $b_{bp}(443)$ , and output chlorophyll is constrained to lie within the range that was used to parameterise the model ( $0.01 < C < 64 \text{ mg m}^{-3}$ ).

### 2.7. OC-CCI algorithm intercomparison

In addition to testing standard satellite chlorophyll products (Table 2), we also conducted a chlorophyll algorithm comparison. For this comparison, we chose to use the OC-CCI dataset given an increase in data coverage, and consequently satellite match-ups, typically observed when using merged ocean-colour products (Brewin, Raitso, et al., 2015; Maritorena, Fanton d'Andon, Mangin, & Siegel, 2010; Racault et al., 2015).

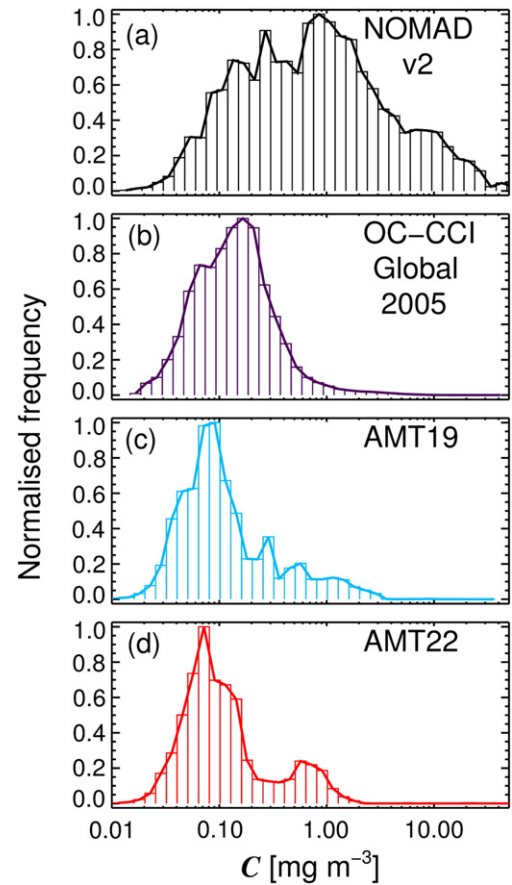
To rank algorithm performance we used the classification method of Brewin, Sathyendranath, Müller, et al. (2015), which scores algorithm performance by comparing each statistical test ( $r$ ,  $\psi$ ,  $\delta$ ,  $\Delta$ ,  $S$ ,  $I$  and  $\eta$ ) of an algorithm with the average values of all algorithms, to determine whether the statistic in question is significantly worse (0 points), similar (1 point) or better (2 points) than the average of all algorithms. All points for each statistic are then summed to give a total score, which is normalised to the average score of all algorithms. A score of one indicates the performance of an algorithm is average with respect to all algorithms tested, a score greater than one indicates algorithm performance is better than average, and a score less than one indicates algorithm performance is worse than average. Using the method of bootstrapping (Efron, 1979; Efron & Tibshirani, 1993), involving random re-sampling with replacement to create ~1000 new datasets of the same size as the original dataset but not identical to it and re-running the classification (Monte-Carlo approach), the stability of the scoring system and the sensitivity of the scores were tested using confidence intervals on the classification output. For further details on this approach, the reader is referred to Section 4 of Brewin, Sathyendranath, Müller, et al. (2015).

## 3. Results and Discussion

### 3.1. AMT chlorophyll distribution

The most well-known and accepted bio-optical datasets, designed for evaluating satellite ocean-colour data, is the NASA bio-Optical Marine Algorithm Data set (NOMAD), developed and updated by NASA (Werdell & Bailey, 2005). A tradition of outstanding support has been established at NASA to deal with queries and comments from NOMAD users, and to integrate bio-optical data from a variety of campaigns into this unique dataset, including some earlier AMT cruises.

Notwithstanding the remarkable efforts by NASA to produce this dataset, when comparing the normalised frequency distribution of *in situ* chlorophyll samples in NOMAD (Fig. 3a, Version 2.0 ALPHA) with that from the global ocean (Fig. 3b, estimated from an annual 2005 OC-CCI composite of chlorophyll), it is clear that oligotrophic waters (with low chlorophyll concentrations) are under-represented. This is



**Fig. 3.** Normalised frequency distribution of *in situ* chlorophyll samples (normalised by maximum value) from (a) the NASA NOMAD version 2.0 dataset (both HPLC and fluorometry chlorophyll), (b) the global ocean (estimated from an annual 2005 OC-CCI composite of chlorophyll), (c) the optical system on AMT19 (45,171) and (d) the optical system on AMT22 (34,934).

likely a reflection of oligotrophic waters being generally less accessible, and hence under-sampled, when compared with coastal and eutrophic waters. In contrast, the distribution of *in situ* chlorophyll samples in AMT19 and AMT22 (Fig. 3c and d) are more in-line with that from the global ocean (Fig. 3b), with a slight bias towards the oligotrophic regions, emphasising the value of data collected on AMT for assessing satellite chlorophyll algorithms designed for application in the global ocean.

### 3.2. Relationship between *in situ* reflectance ratios and chlorophyll

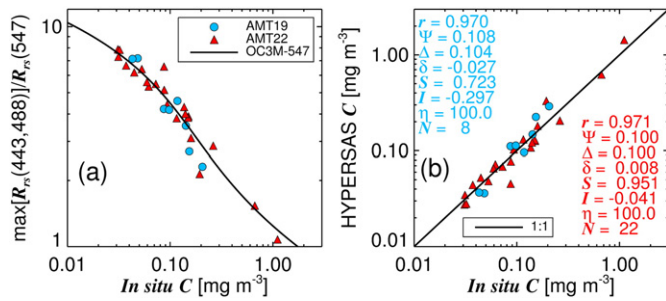
Relationships between *in situ* reflectance ratios derived from the HYPERSAS and *in situ* chlorophyll from the optical system are plotted in Fig. 4a and b. Data from both AMT19 and AMT22 stations show close resemblance with standard relationships between maximum blue-green band reflectance ratios (OC3M-547) and chlorophyll (Fig. 4a). Systematic biases ( $\delta$ ) between *in situ* chlorophyll and chlorophyll estimated from the HYPERSAS data, using the OC3M-547 algorithm, are negligible (Fig. 4b), suggesting no biases should be observed between satellite estimates of chlorophyll based on reflectance ratios, using standard algorithms like the OC3M-547, and *in situ* chlorophyll.

### 3.3. AMT19

#### 3.3.1. Standard algorithms AMT19

The number of match-ups collected on AMT19, for MERIS, MODIS-Aqua and OC-CCI V1 and V2 data vary between 139 and 413 (Fig. 5). It is worth noting that for SeaWiFS, in October 2009 (during AMT19)





**Fig. 4.** Relationships between *in situ* reflectance ratios derived from the HYPERSAS and *in situ* chlorophyll from the optical system, at stations along the AMT19 and AMT22 transects. (a) Shows the maximum band-ratio of  $\max[R_{rs}(443), R_{rs}(488)]/R_{rs}(547)$  plotted as a function of chlorophyll, with the OC3M-547 algorithm overlain. (b) Shows a scatter plot of chlorophyll estimated from the HYPERSAS using the OC3M-547 algorithm against *in situ* chlorophyll from the underway optical system. Blue statistics refer to AMT19 data and red AMT22.

there were various spacecraft and communication issues which resulted in many days of missing observations, and hence there are less match-ups (56, see Fig. 5). The high number of match-ups obtained on a single AMT cruise (~45 days long) illustrates the benefits in using continuous along-track flow-through spectrophotometric systems to maximise the number of point-to-point comparisons between *in situ* and satellite data. Furthermore, the large number of match-ups obtained using merged ocean-colour products (401 and 413 for OC-CCI V1 and V2), during the period in which MERIS, MODIS and SeaWiFS were operating, illustrates the improvement in spatial coverage obtained when merging ocean-colour data from different platforms, as compared with single sensor data.

With the exception of SeaWiFS, which as mentioned was suffering from spacecraft and communication issues, the standard algorithms perform remarkably well in statistical tests for all satellite datasets on AMT19, with correlation coefficients ( $r$ ) ranging from 0.867 to 0.979, and root mean square errors ( $\psi$ ) ranging from 0.105 to 0.186 (Fig. 5). In particular, standard algorithms on MERIS data processed with POLYMER and MODIS-Aqua (both R2013.1 and R2014.0), perform remarkably well ( $\psi$  ranging from 0.105 to 0.121, Fig. 5). Standard algorithms on OC-CCI V1 and V2 also perform well ( $\psi$  ranging from 0.140 to 0.152 and  $r$  from 0.955 to 0.960, Fig. 5), though with a slightly higher  $\psi$  compared with MERIS processed with POLYMER and MODIS-Aqua data, likely due to the inclusion of SeaWiFS in the merged dataset, which performed less accurately in the AMT19 comparison (Fig. 5).

### 3.3.2. Algorithm comparison AMT19

Results from the AMT19 chlorophyll algorithm comparison, using OC-CCI V2 data, are shown in Fig. 6. In general, all algorithms are found to perform well, with correlation coefficients ( $r$ ) greater than 0.95 and  $\psi$  values ranging from 0.091 to 0.160. The band ratio algorithms (OC2S, OC3S and OC4) show slightly larger variability at low chlorophyll concentrations ( $<0.05 \text{ mg m}^{-3}$ ) and tend to deviate from the 1:1 line, when compared with OCI and GSM algorithms. The GSM algorithm has a lower unbiased root mean square error ( $\Delta$ ) than the band ratio algorithms (OC2S, OC3S and OC4), however, it tends to underestimate chlorophyll at higher concentrations ( $>0.2 \text{ mg m}^{-3}$ ), as illustrated by a negative bias ( $\delta = -0.100$ ) and a slope lower than one ( $S = 0.835$ ).

According to the Brewin, Sathyendranath, Müller, et al. (2015) points classification (see bar chart in Fig. 6), and from a visual inspection on the scatter plots in Fig. 6, the OCI algorithm is found to out-perform other algorithms on AMT19 OC-CCI V2 data. The algorithm has the highest correlation coefficient ( $r$ ), lowest  $\Delta$  and  $\Psi$ , and a slope ( $S$ ) close to one. In fact, the OCI root mean square error is remarkably low ( $\Psi = 0.091$ ), suggesting a ~9% average error on log-transformed OC-CCI chlorophyll for AMT19.

## 3.4. AMT22

### 3.4.1. Standard algorithms AMT22

The number of match-ups collected on AMT22, for MODIS-Aqua, VIIRS and OC-CCI V1 and V2 data vary between 155 and 167 (Fig. 7). For the OC-CCI data (V1 and V2), fewer match-ups are available when compared with AMT19, as only MODIS-Aqua data is included in the merged product. This is because MERIS ceased to operate in April 2012 and SeaWiFS in December 2010. Both OC-CCI datasets (V1 and V2) also use the earlier version of MODIS-Aqua reprocessed data (R2013.1). Furthermore, both OC-CCI datasets band-shift and bias-correct the R2013.1 MODIS-Aqua data to be representative of, and consistent with, SeaWiFS data, so are not exactly the same as MODIS-Aqua R2013.1 data. The VIIRS data is found to have the highest number of match-ups (167) when compared with the other satellite datasets for AMT22.

Consistent with AMT19 data (Fig. 5), standard algorithms perform well in statistical tests for the AMT22 data, with  $r$  ranging from 0.972 to 0.978, and  $\Psi$  ranging from 0.155 to 0.214 (Fig. 7). However, in contrast with AMT19 (Fig. 5), there appears to be a negative bias in all satellite estimates of chlorophyll at low concentrations ( $<0.2 \text{ mg m}^{-3}$ ). This underestimate (negative  $\delta$ ) also occurs at higher concentrations in the R2013.1 MODIS-Aqua data and the OC-CCI V1 and V2 data, but is less evident in MODIS-Aqua R2014.0. Considering both OC-CCI versions use the MODIS-Aqua R2013.1 reprocessing, it is not surprising that results are more consistent with MODIS-Aqua R2013.1 rather than R2014.0 (Fig. 7). The systematic bias is also evident in the VIIRS data at low chlorophyll ( $<0.1 \text{ mg m}^{-3}$ ), but not at higher concentrations. Considering this chlorophyll validation constitutes one of the first independent validations of VIIRS chlorophyll in open ocean oligotrophic waters, the performance of VIIRS in the statistical results is very encouraging ( $r = 0.978$  and  $\Psi = 0.182$ ), in agreement with validation results from VIIRS in other regions (e.g. Kahru, Kudela, Anderson, Manzano-Sarabia, & Mitchell, 2014).

### 3.4.2. Algorithm comparison AMT22

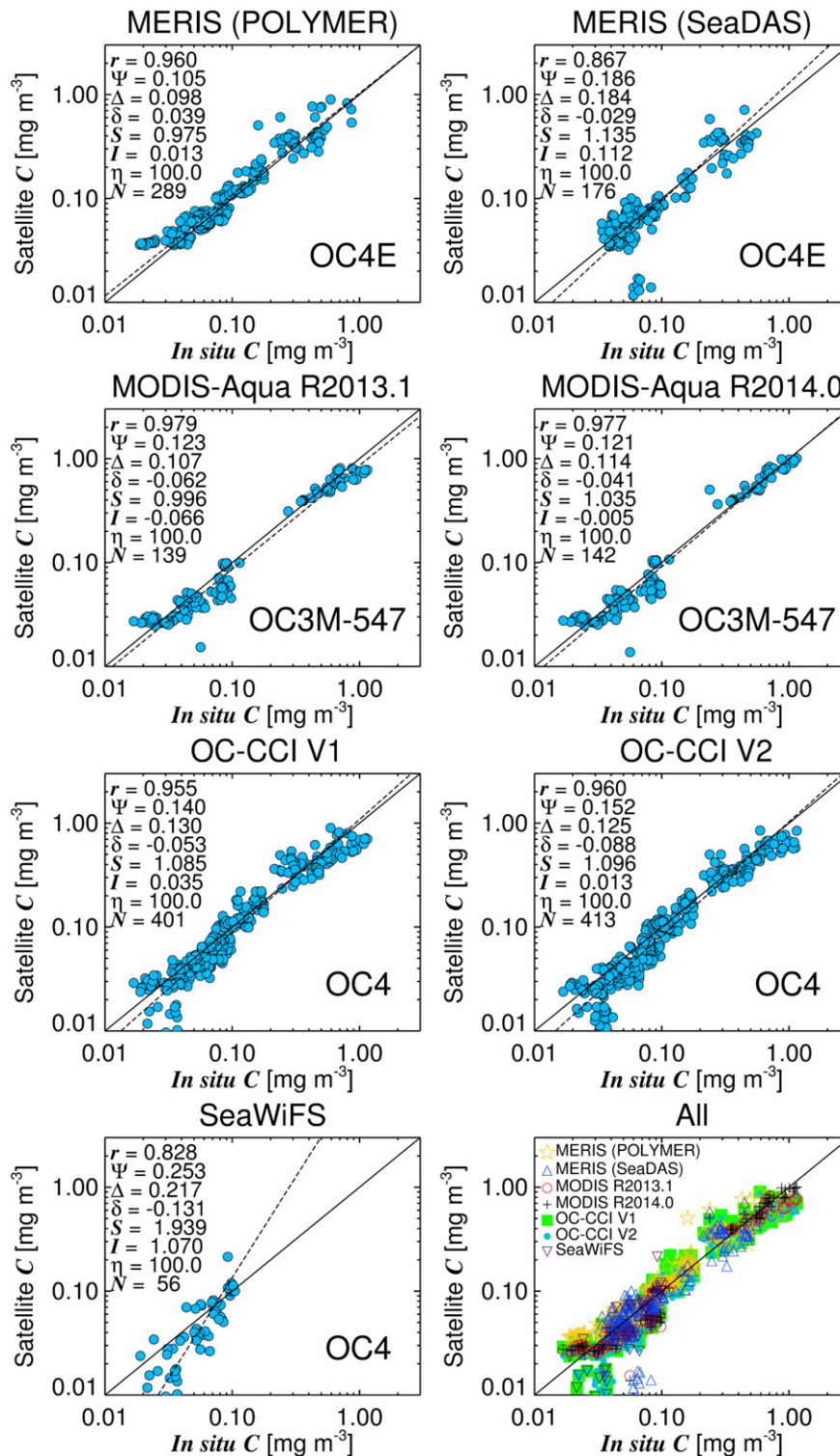
The AMT22 chlorophyll algorithm comparison, using OC-CCI V2 data, is shown in Fig. 8. Consistent with AMT19, all algorithms have high correlation coefficients ( $r > 0.95$ ) and low unbiased root mean square errors ( $\Delta$  ranging from 0.129 to 0.153). However, the systematic bias observed using standard algorithms on OC-CCI V2 (Fig. 7) is evident in all satellite algorithms at low concentrations ( $<0.2 \text{ mg m}^{-3}$ ), with all algorithms seen to underestimate chlorophyll. With the exception of OC2S, all algorithms also underestimate chlorophyll at higher concentrations.

In contrast to results from AMT19 (Fig. 6), the Brewin, Sathyendranath, Müller, et al. (2015) points classification (see bar chart in Fig. 8) suggests the band-ratio algorithms (OC4, OC3S and OC2S) have the highest performance, followed by the OCI and GSM algorithms. The systematic bias observed (underestimate in chlorophyll) is most striking in the GSM algorithm (Fig. 8), and almost disappears in the OC2S, which is seen to perform best in the algorithm comparison (Fig. 8). In the following section we consider the reasons for the observed bias between satellite and *in situ* AMT22 chlorophyll data.

### 3.4.3. Bias in satellite and *in situ* chlorophyll on AMT22

There are three possible causes of the observed bias ( $\delta$ ) in satellite and *in situ* chlorophyll on AMT22: (i) a positive bias in the *in situ* chlorophyll measurements; (ii) a change in the relationship between  $R_{rs}$  and chlorophyll in the Atlantic between AMT19 (2009) and AMT22 (2012); and (iii) a bias in the satellite  $R_{rs}$  data causing a negative bias in satellite chlorophyll estimates.

The *in situ* chlorophyll measurements from the optical system on AMT22 were carefully calibrated with 176 concurrent HPLC measurements and found to be in excellent agreement (Fig. 1). The relationship between HPLC chlorophyll ( $C$ ) and  $a_{ph}(676)$  from the optical system was found to be slightly different on AMT22 when compared with

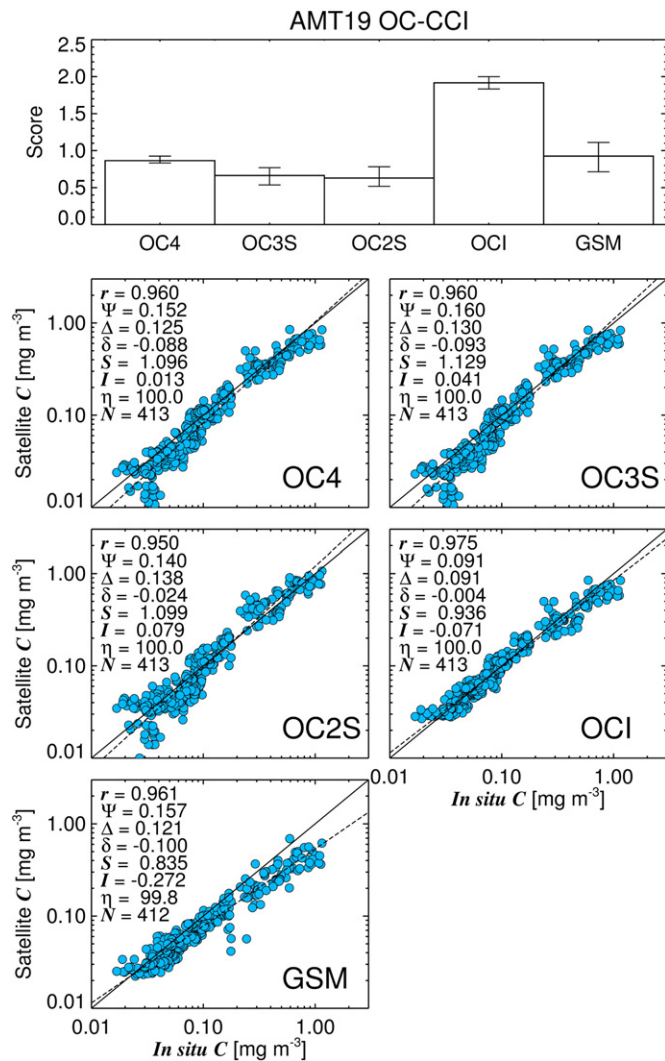


**Fig. 5.** Scatter plots of satellite and *in situ* chlorophyll match-ups collected on AMT19, for a variety of satellite datasets and standard (blue-green band-ratio) NASA algorithms. Solid line refers to 1:1 line, and dashed line a Type-2 regression.

AMT19 (Fig. 1). If we use the AMT19 relationship between  $C$  and  $a_{ph}(676)$  (Fig. 1b) on AMT22 data, *in situ* chlorophyll increases further, as opposed to decreasing to be more in line with the satellite data. The normalised frequency distribution of *in situ* chlorophyll samples on AMT22 is also in good agreement with AMT19 (Fig. 3c and d), with the peak of this distribution actually slightly lower on AMT22 ( $\sim 0.07 \text{ mg m}^{-3}$ ) than AMT19 ( $\sim 0.09 \text{ mg m}^{-3}$ ), which again is inconsistent with there being a positive bias in the *in situ* AMT22 chlorophyll

data, assuming AMT19 chlorophyll is correct and considering the similar cruise tracks and time of year (Fig. 1a). Therefore, it is unlikely that the observed bias ( $\delta$ ) in satellite and *in situ* chlorophyll on AMT22 is related to a positive bias in the *in situ* chlorophyll.

It is possible a change in the relationship between  $R_{rs}$  ratios and chlorophyll along the Atlantic Meridional Transect occurred between AMT19 (2009) and AMT22 (2012). In fact, there is evidence of changes in the relationship between phytoplankton size structure, inferred from



**Fig. 6.** Comparison of different chlorophyll algorithms applied to OC-CCI V2 data with *in situ* chlorophyll from AMT19. NASA band-ratio algorithms include OC4, OC3S and OC2S; OCI refers to the band-difference algorithm of Hu et al. (2012); and GSM the semi-analytical algorithm of Maritorena et al. (2002). Solid line refers to 1:1 line, and dashed line a Type-2 regression. The bar chart at the top of the figure shows ranking of algorithms based on the objective classification of Brewin, Sathyendranath, Müller, et al. (2015) for AMT19 OC-CCI V2 match-up data.

phytoplankton pigment analysis, and chlorophyll concentration along the AMT transect between 2003 and 2010 (Agirbas et al., 2015). Chlorophyll algorithms, that empirically relate chlorophyll to changes in reflectance, implicitly assume a fixed relationship between phytoplankton community structure and chlorophyll (Dierssen, 2010; IOCCG, 2014). Modifications in this relationship will likely impact the performance of these empirical algorithms when applied to satellite  $R_{rs}$  data. However, *in situ*  $R_{rs}$  and chlorophyll from both AMT19 and AMT22 stations show close resemblance with standard empirical relationships between maximum blue-green band reflectance ratios and chlorophyll (Fig. 4). Therefore, it is unlikely that any change in the relationship between  $R_{rs}$  and chlorophyll between cruises is responsible for the observed bias ( $\delta$ ) in satellite and *in situ* chlorophyll on AMT22.

The most likely cause of the observed bias ( $\delta$ ) in satellite and *in situ* chlorophyll on AMT22 is a bias in the satellite  $R_{rs}$  data, causing a negative bias in satellite chlorophyll estimates. It is well known that the MODIS-Aqua sensor, now in its 14th year of operation, has been degrading and requires ongoing calibration and reprocessing (Meister & Franz, 2014). In particular, the blue channels (412 nm and 443 nm) require the largest corrections, are the most difficult to calibrate due

to their degradation pattern, and require the largest vicarious calibration (NASA, 2015). The OC3M-547 results on AMT22 show significant improvements between the R2013.1 and R2014.0 reprocessing (Fig. 7), with a decrease in the root-mean square error ( $\Psi$ ) relative to the R2013.1 reprocessing, emphasising the great work NASA are doing to correct for MODIS-Aqua  $R_{rs}$  degradation.

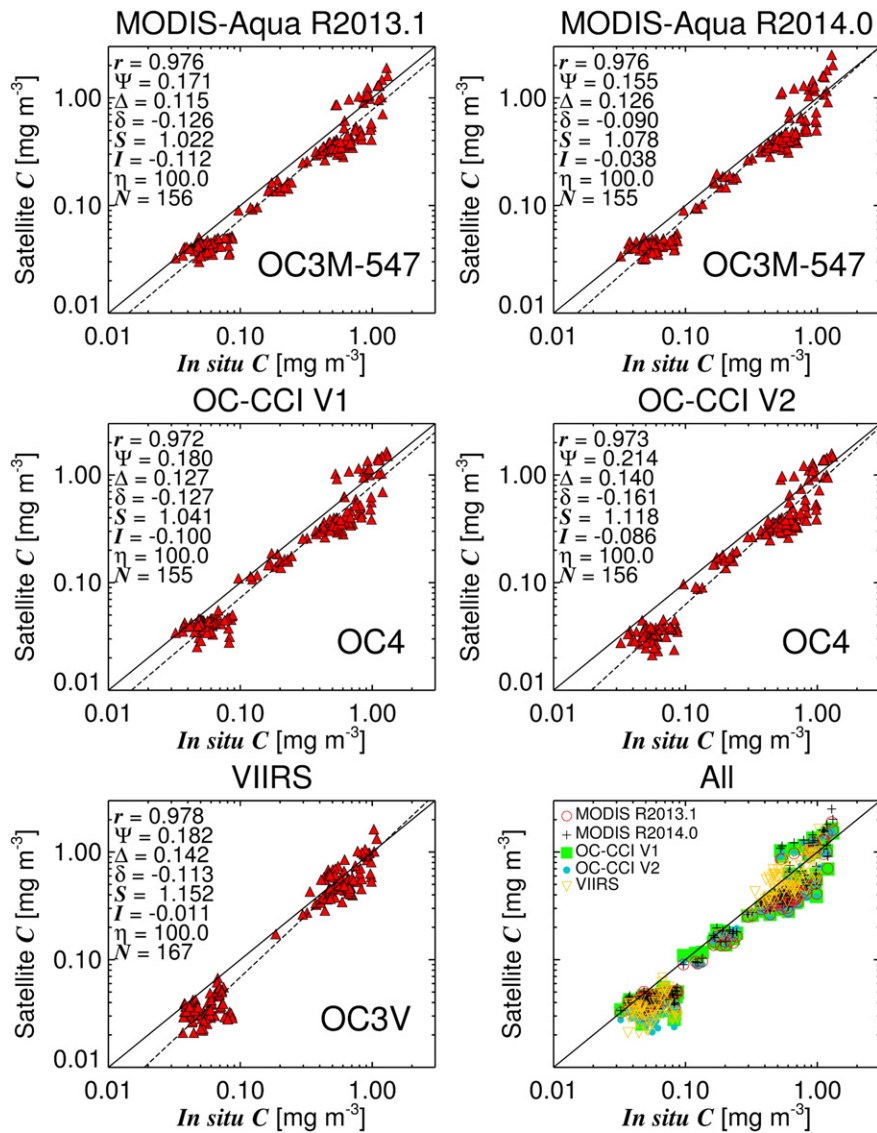
Issues with MODIS-Aqua degradation patterns at blue channels (412 nm and 443 nm) impacting chlorophyll retrievals are further emphasised when assessing the OC-CCI AMT22 algorithm comparison results (Fig. 8). The OC2S algorithm is found to perform best according to the Brewin, Sathyendranath, Müller, et al. (2015) points classification on AMT22. It has the highest  $r$  and lowest  $\Psi$  and  $\Delta$ , and the smallest bias ( $\delta$ ). Of all algorithms tested in the comparison, this is the only chlorophyll algorithm that does not use the 412 nm and 443 nm bands (see Table 3), utilising only the 490 nm and 555 nm bands. When applying the OC2M-547 algorithm to AMT22 MODIS-Aqua R2014.0, which uses the 488 nm and 547 nm bands,  $\Psi$  decreases (Fig. 9) relative to the OC3M-547 algorithm. The OC3M-547 algorithm uses in addition to the 488 nm and 547 nm band the 443 nm band, especially in oligotrophic waters as the 443 nm band generally has a higher signal. Issues with MODIS-Aqua degradation patterns at blue channels (412 nm and 443 nm) also explain the large bias in GSM for OC-CCI on AMT22 relative to AMT19 (Figs. 6 and 8). The GSM is the only algorithm in the comparison that uses the entire spectrum to estimate chlorophyll, including both the 412 nm and 443 nm channel. Furthermore, as the non-linear minimisation used in the version of the GSM algorithm tested is based on a minimisation to the absolute values of the reflectance spectrum, the 412 nm waveband has the highest weighting as its signal is generally the highest of all wavelengths in the oligotrophic waters sampled by AMT. The impact of MODIS-Aqua degradation at blue channels on GSM chlorophyll has been reported elsewhere (Maritorena et al., 2010), and illustrates the importance of weighing the minimisation of model-based algorithms such as GSM according to the uncertainty in  $R_{rs}$  data (Maritorena et al., 2010).

Given this is one of the first validations of VIIRS chlorophyll in oligotrophic waters, it is difficult to ascertain the cause of the low chlorophyll bias seen in Fig. 7. Future validation exercises are required to continue monitoring the performance of VIIRS in oligotrophic waters and ascertain the causes of any such bias.

### 3.5. Comparison of level 2 and level 3 match-ups

To investigate the impact of using level 3 rather than level 2 data for validation along the AMT cruise track, level 3 results are compared with a validation using level 2 data for MODIS-Aqua R2014.0 on AMT19 data in Fig. 10. In general, there is very good agreement between the level 2 and 3 data as indexed by similar results in statistical tests (Fig. 10), supporting the level 3 match-up analysis conducted in this study. Even when only including level 2 data within  $\pm 3$  h of the satellite data, statistical tests do not improve relative to those using the level 3 data. Considering level 3 (4 km) data are aggregates of level 2 (1 km) data, and that more *in situ* measurements are likely to be included in a level 3 match-up when compared with a level 2 match-up, one may have expected the random component of the error ( $\Delta$ ) be lower for the level 3 data which was not observed, except when comparing with the  $\pm 3$  h level 2 match-ups (see Fig. 10). This may be related to the inclusion of procedures such as the homogeneity test, designed to remove noisy level 2 match-ups. Furthermore, although fewer *in situ* measurements are included in a level 2 match-up compared with a level 3 match-up, we set a minimum criteria of three 1-min bins for a level 2 match-up, meaning a substantial number of *in situ* measurements (minimum of  $3 \times 240 = 720$ ) were still used, covering a significant area in a 1 km pixel ( $3 \times 0.3 \text{ km} = 0.9 \text{ km}$ , assuming the ship was moving at  $\sim 18 \text{ km h}^{-1}$ ).

In order to ensure good quality satellite and *in situ* level 3 match-ups, we have applied a series of procedures building on earlier studies based



**Fig. 7.** Scatter plots of satellite and *in situ* chlorophyll match-ups collected on AMT22, for a variety of satellite datasets and standard (blue-green band-ratio) NASA algorithms. Solid line refers to 1:1 line, and dashed line a Type-2 regression.

on level 2 data and discrete point match-ups (Bailey & Werdell, 2006), that we have adapted for use with *in situ* underway and satellite level 3 data. However, there may be cases where some of these procedures require further adaptation, or may no longer be needed. For instance take the homogeneity test, when using a  $3 \times 3$  box of pixels on 4 km level 3 data (area of 144 km<sup>2</sup>), in certain environments (e.g. productive or shelf regions), heterogeneity may be due to real oceanographic features such as fronts that could be inadvertently excluded using this test. Furthermore, when averaging many continuous underway measurements within a 4 km satellite pixel, sub-pixel variability will be accounted for, so the homogeneity test may no longer be needed. Of course this assumes there is enough underway measurements within a pixel to capture the sub-pixel variability, and that the ship track adequately covered the pixel. Future validation efforts using flow-through measurements should focus on these types of considerations, and build on the procedures suggested here.

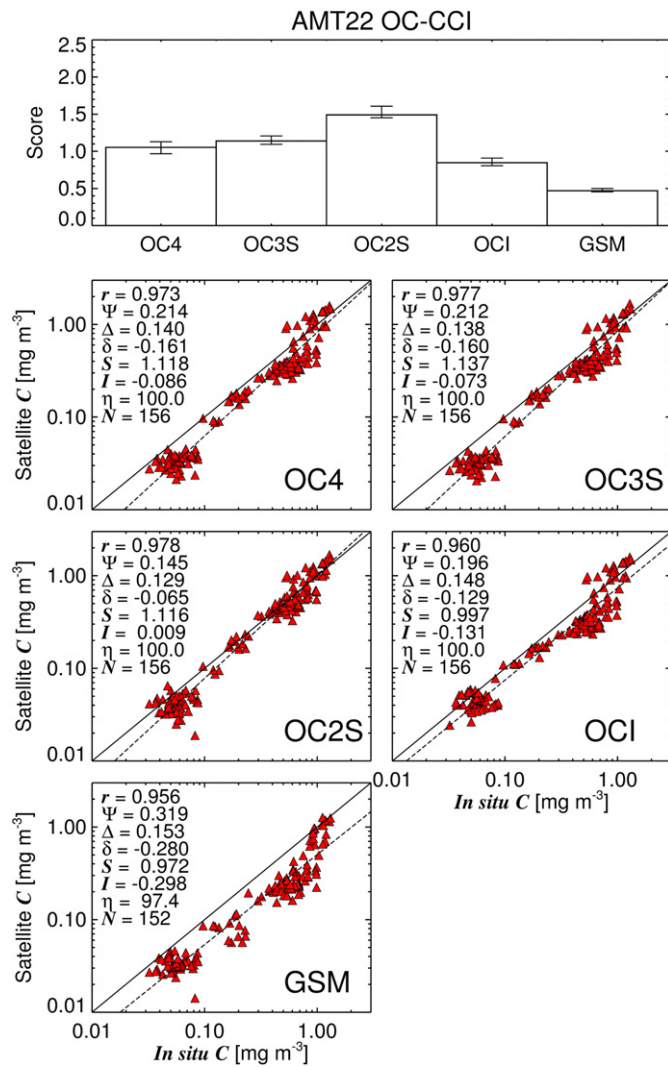
### 3.6. Performance of satellite chlorophyll algorithms

Table 4 compares statistical results ( $r$  and  $\Psi$ ) of standard satellite chlorophyll algorithms (see Figs. 5 and 7) derived in this study with those from previous studies, that used global datasets of discrete point

measurements (either using HPLC or fluorometry derived *in situ* chlorophyll). In this comparison, we excluded the results of SeaWiFS from AMT19 (Fig. 5), considering the low number of match-ups ( $N=56$ ) relative to other sensors and the fact that SeaWiFS was suffering from various spacecraft and communication issues during this period.

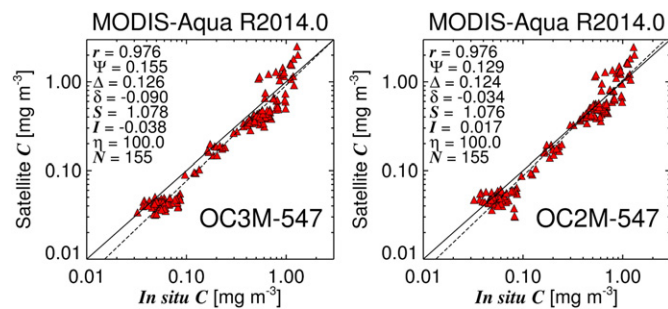
Statistical results from our study show an improvement in the performance of satellite chlorophyll algorithms over previous studies. The average correlation coefficient ( $r=0.880 \pm 0.029$ , Table 4) reported in previous studies is significantly lower than the average values reported here ( $r=0.961 \pm 0.033$ , Table 4), and the average root mean square error reported in previous studies ( $\Psi=0.337 \pm 0.056$ , Table 4) is greater than twice that reported here ( $\Psi=0.157 \pm 0.033$ , Table 4), and significantly higher.

The better performance of satellite chlorophyll data from AMT19 and AMT22 could be due to the AMT cruise occurring at a specific time of year, whereas global datasets of discrete point measurements include data from a variety of locations at different times of year, and hence are likely to include a wider range of variability in optical properties for a given trophic environment. For instance, depending on location, variability in the ratio of CDOM to chlorophyll and backscattering to chlorophyll are likely to change with season, which will impact satellite chlorophyll retrievals. It could also be that the AMT datasets used

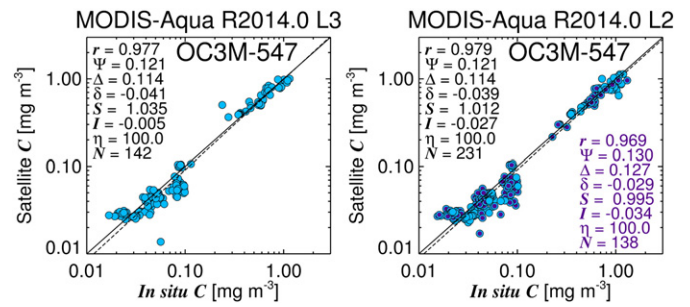


**Fig. 8.** Comparison of different chlorophyll algorithms applied to OC-CCI V2 data with *in situ* chlorophyll from AMT22. NASA band-ratio algorithms include OC4, OC3S and OC2S; OCI refers to the band-difference algorithm of Hu et al. (2012); and GSM the semi-analytical algorithm of Maritorena et al. (2002). Solid line refers to 1:1 line, and dashed line a Type-2 regression. The bar chart at the top of the figure shows ranking of algorithms based on the objective classification of Brewin, Sathyendranath, Müller, et al. (2015) for AMT22 OC-CCI V2 match-up data.

here are more inclusive of the oligotrophic waters relative to other global datasets (Fig. 3), where algorithm performance may improve. In addition, it may be that these algorithms perform better in the Atlantic waters, relative to other oceans. Better performance in our study may also be related to minimising the methodological variability which is



**Fig. 9.** Scatter plots of satellite and *in situ* chlorophyll match-ups collected on AMT22, for MODIS-Aqua R2014.0 using the OC3M-547 and OC2M-547 algorithms. Solid line refers to 1:1 line, and dashed line a Type-2 regression.



**Fig. 10.** Scatter plots of satellite and *in situ* chlorophyll match-ups collected on AMT19, for MODIS-Aqua R2014.0 and using level 3 (L3) and level 2 (L2) datasets and the OC3M-547 algorithm. Solid line refers to 1:1 line, and dashed line a Type-2 regression. Level 2 data within  $\pm 3$  h of the satellite data and related statistical tests are shown in purple.

inherent in global HPLC datasets, when combining observations from many different investigators (Claustre et al., 2004).

Notwithstanding the aforementioned reasons, it is worth considering the advantages of the continuous spectrophotometric sampling used in this study, relative to traditional comparisons using discrete point measurements. The underway optical system is automated and thus produces highly-consistent datasets and introduces very little human error, unlike HPLC or fluorometry, where uncertainties can occur from the moment water enters a Niskin bottle on the CTD to the final pigment extraction and quantification. These automatic measurements can be easily collected for long time periods and over vast areas of the ocean, resulting in large datasets for satellite validation at a fraction of the cost of HPLC methods. Comparing continuous ACS 0.2  $\mu\text{m}$  filtered measurements with unfiltered measurements provides estimates of  $a_p(\lambda)$  accounting for instrumental drifts, residual calibration errors and biofouling. The use of absorption line height on  $a_p(\lambda)$  data to estimate chlorophyll is remarkably accurate when carefully compared with discrete HPLC measurements not just along the AMT cruise track (Fig. 1b and c), but also in other oceans (Boss et al., 2013; Dall'Olmo et al., 2009; Werdell et al., 2013; Westberry et al., 2010), and using phytoplankton cultures (Roesler & Barnard, 2013).

One of the most important characteristics of underway optical data is the feasibility to integrate many observations collected over a satellite pixel, to quantify and account for sub-pixel variability. This is very difficult to do using discrete measurements of HPLC, fluorometry or filter-pad  $a_{ph}$  measurements. These underway optical systems are simply better suited to evaluate satellite observations, when compared with discrete point measurements.

Despite the benefits of using underway optical systems for satellite validation, there are some caveats. It is vital that these systems are deployed on clean and well-maintained flow-through systems, to ensure the optical sensors are sampling uncontaminated seawater unaffected from biota growing in the ship's plumbing system. The tubing and instruments used should be cleaned and regularly checked for fouling. As highlighted in Fig. 1, the relationship between ACS derived  $a_{ph}(676)$  and chlorophyll was different between cruises, suggesting it is important to collect discrete HPLC samples in the same range of conditions sampled by the optical underway system. Differences may be caused by either: i) changes in the chlorophyll specific absorption coefficient, for instance, from modifications in community structure; and ii) differences in the optical set-up used, for instance, from different instruments with different spectral responses.

Ships underway water intake is typically at a nominal depth of 5 m. In very clear waters the satellite signal can be representative of a water layer as deep as 40 m. Ocean-colour algorithms can therefore be affected by vertical variations in chlorophyll (Stramska & Stramski, 2003). Considering the agreement between satellite and *in situ* chlorophyll demonstrated here this is unlikely to have had significant impact on our results. Finally, there may be artificial effects on optical properties from pumping water. To quantify such effects, comparisons should be made

**Table 4**  
Comparison of statistical results of standard chlorophyll algorithms in this study, with some previous studies.

Study	Satellite dataset	Algorithm	r	ψ	N <sup>a</sup>	Locations
Bregg and Casey (2004)	SeaWiFS	OC4	0.872	0.310	4168	Global
Bailey and Werdell (2006)	SeaWiFS	OC4	0.913	0.406	271	Global
Maritorena et al. (2010)	GloboColour	GSM	0.823	0.399	778	Global
Siegel et al. (2013)	SeaWiFS	OC4	0.892	0.310	1543	Global
Siegel et al. (2013)	SeaWiFS	GSM	0.880	0.363	1380	Global
Brewin, Sathyendranath, Jackson, et al. (2015)	OC-CCI V1	OC4	0.880	0.250	598	Global
NASA (2015)	MODIS-Aqua R2014.0	OCI	0.903	0.324	657	Global
		Average	0.880	0.337		
		Standard deviation	0.029	0.056		
This study	MERIS (POLYMER)	OC4E	0.960	0.105	289	AMT19
This study	MERIS (SeaDAS)	OC4E	0.867	0.186	176	AMT19
This study	MODIS-Aqua R2013.1	OC3M-547	0.979	0.123	139	AMT19
This study	MODIS-Aqua R2014.0	OC3M-547	0.977	0.121	142	AMT19
This study	OC-CCI V1	OC4	0.955	0.140	401	AMT19
This study	OC-CCI V2	OC4	0.960	0.152	413	AMT19
This study	MODIS-Aqua R2013.1	OC3M-547	0.976	0.171	156	AMT22
This study	MODIS-Aqua R2014.0	OC3M-547	0.976	0.155	155	AMT22
This study	VIIRS	OC3V	0.978	0.182	167	AMT22
This study	OC-CCI V1	OC4	0.972	0.180	155	AMT22
This study	OC-CCI V2	OC4	0.973	0.214	156	AMT22
		Average	0.961	0.157		
		Standard deviation	0.032	0.033		

<sup>a</sup> Number of satellite and *in situ* match-ups.

with discrete *in situ* optical measurements (Dall'Olmo et al., 2012; Westberry et al., 2010). To harness the benefits of underway optical sampling, it is important that the community establishes rigorous protocols to ensure data consistency, compatibility and accuracy.

#### 4. Summary

We used an optical set-up to continuously measure absorption by particles on two AMT cruises (AMT19 and AMT22). Continuous estimates of *in situ* chlorophyll concentration on the two AMT cruises were computed from the optical set-up using a calibration between coincident measurements of HPLC chlorophyll and absorption by particles in the red portion of the visible spectrum. The chlorophyll distribution of the two resulting *in situ* datasets (AMT19 and AMT22) were found to be similar to that observed in the global ocean, with a slight bias towards the under-sampled oligotrophic gyres. The two *in situ* datasets were used to evaluate the performance of satellite chlorophyll algorithms applied to different level 3 binned ocean-colour datasets.

Statistical comparisons between *in situ* measurements co-incident with the satellite data indicate the performance of satellite chlorophyll algorithms is better than that described in previous studies. We find that the root mean square error between satellite and *in situ* chlorophyll data to be on average, less than half that reported previously using global datasets. We hypothesise that this improvement is due to the underway spectrophotometric sampling method being better suited to evaluate satellite observations, when compared with discrete point measurements and *in vivo* fluorescence (Werdell et al., 2013). We observed a bias (underestimate) in satellite chlorophyll at low concentrations on the AMT22 cruise for some satellite algorithms. This was likely due to a small bias in satellite remote-sensing reflectance data, considering the relationship between chlorophyll and *in situ* remote-sensing reflectance on AMT22 was found to follow a standard relationship, and no biases were observed in the *in situ* chlorophyll data. Our results support the use of underway optical systems for evaluating satellite ocean-colour data, and emphasise the benefits of maintaining *in situ* observatories in oligotrophic regions, such as the Atlantic Meridional Transect. These have implications for the validation of recently-launched and future ocean-colour missions (e.g. the ESA Ocean and Land Colour Instrument (OLCI) on-board Sentinel-3, NASA's Pre-Aerosol Clouds and ocean Ecosystem (PACE) mission, and the Japan Aerospace Exploration Agency (JAXA) Second generation Global Imager (SGLI) on-board the Global Change Observation Mission - Climate (GCOM-C)).

#### Acknowledgments

The authors would like to thank the Ocean Biology Processing Group of NASA for the processing and distribution of the SeaWiFS, MODIS-Aqua and MERIS data, without which this work would not have been feasible. We thank ESA for the distribution of the MERIS data and for OC-CCI data. We thank Chris Gallienne for deploying the HYPERSAS on AMT19 and AMT22 and Tim Smyth for processing AMT19 data using SatCon software. We thank Mike Grant for processing OC-CCI V1 data for AMT19 and for the processing of 4 km binned SeaWiFS data. We also thank Francois Steinmetz for providing the POLYMER code used to process MERIS. We thank the editor and two reviewers for providing useful comments that helped improve the manuscript. We acknowledge the NERC Earth Observation Data Acquisition and Analysis Service (NEODAAS) for near-real time EO data support during AMT cruises and assistance with post-cruise EO processing. We also acknowledge support from the Copernicus Marine Environment Monitoring Service (CMEMS) Ocean Colour Thematic Assembly Centre. SP was supported by NEODAAS and CMEMS. This work is supported by the UK National Centre for Earth Observation, and is a contribution to the ESA Ocean Colour Climate Change Initiative, ESA AMT4SentinelFRM and the international IMBER project. AMT data were supported by the UK Natural Environment Research Council National Capability funding to Plymouth

Marine Laboratory and the National Oceanography Centre, Southampton. This is contribution number 282 of the AMT programme.

## References

- Agirbas, E., Martinez-Vicente, V., Brewin, R. J. W., Racault, M. -F., Aïrs, R. L., & Llewellyn, C. A. (2015). Temporal changes in total and size-fractionated chlorophyll-a in surface waters of three provinces in the Atlantic Ocean (September to November) between 2003 and 2010. *Journal of Marine Systems*, 150, 56–652.
- Aiken, J., & Hooker, S. B. (1997). The Atlantic Meridional Transect: Spatially extensive calibration and validation of optical properties and remotely-sensed measurements of ocean color. *Backscatter*, 8, 8–11.
- Aiken, J., Pradhan, Y., Barlow, R., Lavender, S., Poulton, A., Holligan, P., & Hardman-Mountford, N. J. (2009). Phytoplankton pigments and functional types in the Atlantic Ocean: A decadal assessment, 1995–2005. *Deep Sea Research, Part I*, 56(15), 899–917.
- Aiken, J., Rees, N., Hooker, S., Holligan, P., Bale, A., Robins, D., ... Pilgrim, D. (2000). The Atlantic Meridional Transect: overview and synthesis of data. *Progress in Oceanography*, 45, 257–312.
- Bailey, S. W., & Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, 102, 12–23.
- Barbini, R., Colao, F., De Dominicis, L., Fantoni, R., Fiorani, L., Palucci, A., & Artamonov, E. S. (2004). Analysis of simultaneous chlorophyll measurements by lidar fluorosensor, MODIS and SeaWiFS. *International Journal of Remote Sensing*, 25(11), 2095–2110.
- Barbini, R., Colao, F., Fantoni, R., Fiorani, L., & Palucci, A. (2003). Lidar fluorosensor calibration of the SeaWiFS chlorophyll algorithm in the Ross Sea. *International Journal of Remote Sensing*, 24, 3205–3218.
- Batten, S. D., Walne, A. W., Edwards, M., & Groom, S. B. (2003). Phytoplankton biomass from continuous plankton recorder data: an assessment of the phytoplankton colour index. *Journal of Plankton Research*, 25, 697–702.
- Boss, E. S., Collier, R., Larson, G., Fennel, K., & Pegau, W. S. (2007). Measurements of spectral optical properties and their relation to biogeochemical variables and processes in Crater Lake, Crater Lake National Park, OR. *Hydrobiologia*, 574, 149–159.
- Boss, E., Picherai, M. P., Leeuw, T., Chase, A., Karsenti, E., Gorsky, G., ... Claustre, H. (2013). The characteristics of particulate absorption, scattering and attenuation coefficients in the surface ocean; Contribution of the Tara Oceans expedition. *Methods in Oceanography*, 7, 52–62.
- Brewin, R. J. W., Devred, E., Sathyendranath, S., Hardman-Mountford, N. J., & Lavender, S. J. (2011a). Model of phytoplankton absorption based on three size classes. *Applied Optics*, 50(22), 4535–45492.
- Brewin, R. J. W., Hardman-Mountford, N. J., Lavender, S., Raitsos, D. E., Hirata, T., Uitz, J., ... Gentili, B. (2011b). An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sensing of Environment*, 115(2), 325–339.
- Brewin, R. J. W., Sathyendranath, S., Jackson, T., Barlow, R., Brotas, V., Aïrs, R., & Lamont, T. (2015b). Influence of light in the mixed layer on the parameters of a three-component model of phytoplankton size structure. *Remote Sensing of Environment*, 168, 437–450.
- Brewin, R. J. W., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P. -Y., Devred, E., ... White, G. N., III (2015c). The Ocean Colour Climate Change Initiative: III. A round-robin comparison on in-water bio-optical algorithms. *Remote Sensing of Environment*, 162, 271–294.
- Brewin, R. J. W., Raitsos, D. E., Pradhan, Y., & Hoteit, I. (2013). Comparison of chlorophyll in the Red Sea derived from MODIS-Aqua and *in vivo* fluorescence. *Remote Sensing of Environment*, 136, 218–224.
- Brewin, R. J. W., Raitsos, D. E., Dall'Olmo, G., Zarokanellos, N., Jackson, T., Racault, M. -F., ... Hoteit, I. (2015a). Regional ocean-colour chlorophyll algorithms for the Red Sea. *Remote Sensing of Environment*, 165, 64–85.
- Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R., & 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., Morel, A., Babin, M., Allali, K., & Claustre, H. (1998). Variations of light absorption by suspended particles with the chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models. *Journal of Geophysical Research*, 103, 31,033–31,044.
- Brotas, V., Brewin, R. J. W., Sá, C., Brito, A. C., Silva, A., Mendes, C. R., ... Sathyendranath, S. (2013). Deriving phytoplankton size classes from satellite data: Validation along a trophic gradient in the eastern Atlantic Ocean. *Remote Sensing of Environment*, 134, 66–77.
- Campbell, J. W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research*, 100(C7), 13237–13254.
- Chassot, E., Bonhommeau, S., Dulvy, N. K., Mélin, F., Watson, R., Gascuel, D., & Le Pape, O. (2010). Global marine primary production constrains fisheries catches. *Ecology Letters*, 13, 495–505.
- Claustre, H., Hooker, S. B., Van Heukelem, L., Berthon, J. -F., Barlow, R., Ras, J., ... Marty, J. -C. (2004). An intercomparison of HPLC phytoplankton pigment methods using *in situ* samples: Application to remote sensing and database activities. *Marine Chemistry*, 85, 41–61.
- Cullen, J. J., & Lewis, M. R. (1995). Biological processes and optical measurements near the sea-surface: some issues relevant to remote sensing. *Journal of Geophysical Research*, 100, 13,255–13,266.
- Dall'Olmo, G., Boss, E., Behrenfeld, M., & Westberry, T. K. (2012). Particulate optical scattering coefficients along an Atlantic Meridional Transect. *Optics Express*, 20, 21532–21551.
- Dall'Olmo, G., Boss, E., Behrenfeld, M., Westberry, T. K., Courties, C., Prieur, L., ... Moutin, T. (2011). Inferring phytoplankton carbon and eco-physiological rates from diel cycles of spectral particulate beam-attenuation coefficient. *Biogeosciences*, 8, 3423–3439.
- Dall'Olmo, G., Westberry, T. K., Behrenfeld, M. J., Boss, E., & Slade, W. H. (2009). Significant contribution of large particles to optical backscattering in the open ocean. *Biogeosciences*, 6(6), 947–967.
- Davis, R. F., Moore, C. C., Zaneveld, J. R. V., & Napp, J. M. (1997). Reducing the effects of fouling on chlorophyll estimates derived from long-term deployments of optical instruments. *Journal of Geophysical Research*, 102, 5851–5855.
- Dietersen, H. M. (2010). Perspectives on empirical approaches for ocean color remote sensing of chlorophyll in a changing climate. *Proceedings of the National Academy of Sciences of the United States of America*, 107(40), 17073–17078.
- Doney, S. C., Lima, I. D., Moore, J. K., Lindsay, K., Behrenfeld, M. J., Westberry, T. K., ... Takahashi, T. (2009). Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data. *Journal of Marine Systems*, 76, 95–112.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7, 1–26.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman and Hall.
- Folkestad, A., Pettersson, L. H., & Durand, D. D. (2007). Inter-comparison of ocean colour data products during algal blooms in the Skagerrak. *International Journal of Remote Sensing*, 28, 569–592.
- Friedrichs, M. A. M., Carr, M. -E., Barber, R. T., Scardi, M., Antoine, D., Armstrong, R. A., ... Winguth, A. (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. (1997). Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: 1. Time series from the Sargasso Sea. *Journal of Geophysical Research*, 102, 18,607–18,625.
- GCOS (2011). Systematic observation requirements from satellite-based data products for climate. *Technical report*. 7 bis, avenue de la Paix, CH-1211 Geneva 2, Switzerland: World Meteorological Organisation (WMO).
- Glover, D. M., Jenkins, W. J., & Doney, S. C. (2011). *Modeling methods for marine science*. Cambridge University Press.
- Gregg, W. W., & Casey, N. W. (2004). Global and regional evaluation of the SeaWiFS chlorophyll data set. *Remote Sensing of Environment*, 93, 463–479.
- Harding, L. W., Jr., Magnuson, A., & Mallonee, M. E. (2005). SeaWiFS retrievals of chlorophyll in Chesapeake Bay and the mid-Atlantic bight. *Estuarine, Coastal and Shelf Science*, 62, 75–94.
- Hooker, S. B., & McClain, C. R. (2000). The calibration and validation of SeaWiFS data. *Progress in Oceanography*, 45, 427–465.
- Hooker, S. B., Lazin, G., Zibordi, G., & McLean, S. (2002). An evaluation of above- and in-water methods for determining water-leaving radiances. *Journal of Atmospheric and Oceanic Technology*, 19, 486–515.
- Hu, C., Lee, Z., & Franz, B. (2012). Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *Journal of Geophysical Research*, 117, C01011.
- Hu, C., Muller-Karger, F. E., Biggs, D. C., Carder, K. L., Nababan, B., Nadeau, D., & Vanderbloemen, J. (2003). Comparison of ship and satellite bio-optical measurements on the continental margin of the NE Gulf of Mexico. *International Journal of Remote Sensing*, 24, 2597–2612.
- Hu, C., Muller-Karger, F. E., Taylor, C., Carder, K. L., Kelble, C., Johns, E., & Heil, C. A. (2005). Red tide detection and tracing using MODIS fluorescence data: A regional example in SW Florida coastal waters. *Remote Sensing of Environment*, 97, 311–321.
- Hu, C., Nababan, B., Biggs, D. C., & Muller-Karger, F. E. (2004). Variability of bio-optical properties at sampling stations and implications for remote sensing: a case study in the north-east Gulf of Mexico. *International Journal of Remote Sensing*, 25(11), 2111–2120.
- IOCCG (2014). Phytoplankton Functional Types from Space. Tech. rep. In S. Sathyendranath (Ed.), *Reports of the International Ocean-Colour Coordinating Group*, No. 15. Dartmouth, Canada: IOCCG.
- Kahru, M., Kudela, R. M., Anderson, C. R., Manzano-Sarabia, M., & Mitchell, B. G. (2014). Evaluation of satellite retrievals of ocean chlorophyll-a in the California Current. *Remote Sensing*, 6(9), 8525–8540.
- Kiefer, D. A. (1973a). Chlorophyll a fluorescence in marine centric diatoms: Responses of chloroplasts to light and nutrient stress. *Marine Biology*, 23, 39–46.
- Kiefer, D. A. (1973b). Fluorescence properties of natural phytoplankton populations. *Marine Biology*, 22, 263–269.
- Koponen, S., Attila, J., Pulliainen, J., Kallio, K., Pyhälähti, T., Lindfors, A., ... Hallikainen, M. (2007). A case study of airborne and satellite remote sensing of a spring bloom event in the Gulf of Finland. *Continental Shelf Research*, 27, 228–244.
- Longhurst, A., Sathyendranath, S., Platt, T., & Caverhill, C. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *Journal of Plankton Research*, 17, 1245–1271.
- Lorenzen, C. J. (1966). A method for the continuous measurement of *in vivo* chlorophyll concentration. *Deep-Sea Research*, 13, 223–227.
- Maritorena, S., Fanton d'Andon, O. H., Mangin, A., & Siegel, D. A. (2010). Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues. *Remote Sensing of Environment*, 114, 1791–1804.
- Maritorena, S., Siegel, D. A., & Peterson, A. R. (2002). Optimization of a semi-analytical ocean color model for global-scale applications. *Applied Optics*, 41(15), 2705–2714.
- McClain, C. R. (2009). A decade of satellite ocean color observations. *Annual Review of Marine Science*, 1, 19–42.
- Meister, G., & Franz, B. A. (2014). Corrections to the MODIS aqua calibration derived from MODIS aqua ocean color products. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 6534–6541.

- Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgina, C. E., Corlett, G. K., ... Donlon, C. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change initiative (SST CCI). *Geoscience Data Journal*, 1, 179–191.
- Mobley, C. D. (1999). Estimation of the remote-sensing reflectance from above-surface measurements. *Applied Optics*, 38, 7442–7455.
- Mobley, C. D. (2015). Polarized reflectance and transmittance properties of windblown sea surfaces. *Applied Optics*, 54, 4828–4849.
- NASA (2010, March). Ocean Color Chlorophyll (OC) v6. URL <http://oceancolor.gsfc.nasa.gov/REPROCESSING/R2009/ocv6/>
- NASA (2015, July). MODIS-Aqua Ocean Color Reprocessing 2014.0. URL <http://oceancolor.gsfc.nasa.gov/cms/reprocessing/OCReproc20140MA>
- O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., ... McClain, C. (1998). Ocean chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, 103(C11), 24,937–24,953.
- O'Reilly, J. E., Maritorena, S., Siegel, D., O'Brien, M. C., Toole, B. G., Mitchell, D., ... Culver, M. (2000). Ocean color chlorophyll algorithms for SeaWiFS, OC2, and OC4. Technical report. In S. B. Hooker, & E. R. Firestone (Eds.), *SeaWiFS postlaunch technical report series. Vol. 11. SeaWiFS postlaunch calibration and validation analyses, part 3* (pp. 9–23). Greenbelt, Maryland: NASA, Goddard Space Flight Center.
- Petersen, W., Wehde, H., Krasemann, H., Colijn, F., & Schroeder, F. (2008). *FerryBox* and MERIS – Assessment of coastal and shelf sea ecosystems by combining in situ and remotely sensed data. *Estuarine, Coastal and Shelf Science*, 77, 296–307.
- Racault, M. -F., Raitos, D. E., Berumen, M. L., Brewin, R. J. W., Platt, T., Sathyendranath, S., & Hoteit, I. (2015). Phytoplankton phenology indices in coral reef ecosystems: Application to ocean-colour observations in the Red Sea. *Remote Sensing of Environment*, 160, 222–234.
- Raitos, D. E., Reid, P. C., Lavender, S. J., Edwards, M., & Richardson, A. J. (2005). Extending the SeaWiFS chlorophyll data set back 50 years in the northeast Atlantic. *Geophysical Research Letters*, 32, L06603.
- Rees, A., Robinson, C., Smyth, T., Aiken, J., Nightingale, P., & Zubkov, M. (2015). 20 years of the Atlantic Meridional Transect – AMT. *Limnology and Oceanography Bulletin*, 24, 101–107.
- Robinson, C., Poulton, A. J., Holligan, P. M., Baker, A. R., Forster, G., Gist, N., ... Zubkov, M. V. (2006). The Atlantic Meridional Transect (AMT) Programme: a contextual view 1995–2005. *Deep Sea Research, Part II*, 53, 1485–1515.
- Roesler, C., & Barnard, A. H. (2013). Optical proxy for phytoplankton biomass in the absence of photophysiology: Rethinking the absorption line height. *Methods in Oceanography*, 7, 79–94.
- Sathyendranath, S. (2011). Ocean colour climate change initiative (OC-CCI) – Phase one user requirements document. URL [http://www.esa-oceancolour-cci.org/?q=webfm\\_send/60/1](http://www.esa-oceancolour-cci.org/?q=webfm_send/60/1)
- Sathyendranath, S., Brewin, R. J. W., Müller, D., & Brockmann, C. (2012). Ocean Colour Climate Change Initiative: Approach and Initial Results. *IEEE International Geoscience and Remote Sensing Symposium*, 2024–2027. <http://dx.doi.org/10.1109/IGARSS.2012.6350979>.
- Sathyendranath, S., Gouveia, A. D., Shetye, S. R., Ravindran, P., & Platt, T. (1991). Biological control of surface temperature in the Arabian Sea. *Nature*, 349, 54–56.
- Siegel, D. A., & Franz, B. A. (2010). Century of phytoplankton change. *Nature*, 466, 569–571.
- Siegel, D. A., Behrenfeld, M. J., Maritorena, S., McClain, C. R., Antoine, D., Bailey, S. W., ... Yoder, J. A. (2013). Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission. *Remote Sensing of Environment*, 135, 77–91.
- Slade, W. H., Boss, E., Dall'Olmo, G., Langner, M. R., Loftin, J., Behrenfeld, M. J., ... Westberry, T. K. (2010). Underway and moored methods for improving accuracy in measurement of spectral particulate absorption and attenuation. *Journal of Atmospheric and Oceanic Technology*, 27, 1733–1746.
- Slovacek, R., & Bannister, T. (1973). NH<sub>4</sub>Cl activation of the fluorescence yield in CO<sub>2</sub> starved *Chlorella pyrenoidosa*. *Biochimica et Biophysica Acta*, 325, 114–119.
- Steinmetz, F., Deschamps, P. -Y., & Ramon, D. (2011). Atmospheric correction in presence of sun glint: application to MERIS. *Optics Express*, 19, 9783–9800.
- Stramska, M., & Stramski, D. (2003). Effects of a nonuniform vertical profile of chlorophyll concentration on remote-sensing reflectance of the ocean. *Applied Optics*, 44, 1735–1747.
- Strass, V. (1990). On the calibration of large-scale fluorometric chlorophyll measurements from towed undulating vehicles. *Deep Sea Research, Part I*, 37, 525–540.
- Strickland, J. (1968). Continuous measurement of in vivo chlorophyll: a precautionary note. *Deep Sea Research*, 15, 225–227.
- Werdell, P. J., & Bailey, S. W. (2005). An improved in-situ bio-optical data set for ocean colour algorithm development and satellite data production validation. *Remote Sensing of Environment*, 98, 122–140.
- Werdell, P. J., Proctor, C. W., Boss, E., Leeuw, T., & Ouhssain, M. (2013). Underway sampling of marine inherent optical properties on the Tara Oceans expedition as a novel resource for ocean color satellite data product validation. *Methods in Oceanography*, 7, 40–51.
- Westberry, T. K., Dall'Olmo, G., Behrenfeld, M., & Moutin, T. (2010). Coherence of particulate beam attenuation and backscattering coefficients in diverse open ocean environments. *Optics Express*, 18, 15,419–15,425.
- Zhai, L., Tang, C., Platt, T., & Sathyendranath, S. (2011). Ocean response to attenuation of visible light by phytoplankton in the Gulf of St. Lawrence. *Journal of Marine Systems*, 88, 285–297.
- Zhang, C., Hu, C., Shang, S., Müller-Karger, F. E., Li, Y., Dai, M., ... Hong, H. (2006). Bridging between SeaWiFS and MODIS for continuity of chlorophyll-a concentration assessments off Southeastern China. *Remote Sensing of Environment*, 102, 250–263.