

The University of Maine

DigitalCommons@UMaine

Marine Sciences Faculty Scholarship

School of Marine Sciences

8-1-2015

Regional ocean-colour chlorophyll algorithms for the Red Sea

Robert J.W. Brewin

Plymouth Marine Laboratory

Dionysios E. Raitsos

Plymouth Marine Laboratory

Giorgio Dall'Olmo

Plymouth Marine Laboratory

Nikolaos Zarokanellos

King Abdullah University of Science and Technology

Thomas Jackson

Plymouth Marine Laboratory

See next page for additional authors

Follow this and additional works at: https://digitalcommons.library.umaine.edu/sms_facpub



Part of the [Oceanography and Atmospheric Sciences and Meteorology Commons](#)

Repository Citation

Brewin, Robert J.W.; Raitsos, Dionysios E.; Dall'Olmo, Giorgio; Zarokanellos, Nikolaos; Jackson, Thomas; Racault, Marie Fanny; Boss, Emmanuel S.; Sathyendranath, Shubha; Jones, Burt H.; and Hoteit, Ibrahim, "Regional ocean-colour chlorophyll algorithms for the Red Sea" (2015). *Marine Sciences Faculty Scholarship*. 233.

https://digitalcommons.library.umaine.edu/sms_facpub/233

This Article is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Marine Sciences Faculty Scholarship by an authorized administrator of DigitalCommons@UMaine. For more information, please contact um.library.technical.services@maine.edu.

Authors

Robert J.W. Brewin, Dionysios E. Raitsos, Giorgio Dall'Olmo, Nikolaos Zarokanellos, Thomas Jackson, Marie Fanny Racault, Emmanuel S. Boss, Shubha Sathyendranath, Burt H. Jones, and Ibrahim Hoteit



Regional ocean-colour chlorophyll algorithms for the Red Sea



Robert J.W. Brewin^{a,b,*}, Dionysios E. Raitsos^a, Giorgio Dall'Olmo^{a,b}, Nikolaos Zarokanellos^c, Thomas Jackson^a, Marie-Fanny Racault^a, Emmanuel S. Boss^d, Shubha Sathyendranath^{a,b}, Burt H. Jones^c, Ibrahim Hoteit^e

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

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

^c King Abdullah University for Science and Technology (KAUST), Red Sea Research Center, Thuwal 23955-6900, Saudi Arabia

^d School of Marine Sciences, University of Maine, Orono, ME 04469-5741, USA

^e King Abdullah University for Science and Technology (KAUST), Earth Sciences and Engineering Division, Thuwal 23955-6900, Saudi Arabia

ARTICLE INFO

Article history:

Received 3 October 2014

Received in revised form 30 March 2015

Accepted 22 April 2015

Available online xxxx

Keywords:

Phytoplankton

Ocean colour

Remote sensing

Chlorophyll

Red Sea

Validation

Coloured dissolved organic matter

ABSTRACT

The Red Sea is a semi-enclosed tropical marine ecosystem that stretches from the Gulf of Suez and Gulf of Aqaba in the north, to the Gulf of Aden in the south. Despite its ecological and economic importance, its biological environment is relatively unexplored. Satellite ocean-colour estimates of chlorophyll concentration (an index of phytoplankton biomass) offer an observational platform to monitor the health of the Red Sea. However, little is known about the optical properties of the region. In this paper, we investigate the optical properties of the Red Sea in the context of satellite ocean-colour estimates of chlorophyll concentration. Making use of a new merged ocean-colour product, from the European Space Agency (ESA) Climate Change Initiative, and *in situ* data in the region, we test the performance of a series of ocean-colour chlorophyll algorithms. We find that standard algorithms systematically overestimate chlorophyll when compared with the *in situ* data. To investigate this bias we develop an ocean-colour model for the Red Sea, parameterised to data collected during the Tara Oceans expedition, that estimates remote-sensing reflectance as a function of chlorophyll concentration. We used the Red Sea model to tune the standard chlorophyll algorithms and the overestimation in chlorophyll originally observed was corrected. Results suggest that the overestimation was likely due to an excess of CDOM absorption per unit chlorophyll in the Red Sea when compared with average global conditions. However, we recognise that additional information is required to test the influence of other potential sources of the overestimation, such as aeolian dust, and we discuss uncertainties in the datasets used. We present a series of regional chlorophyll algorithms for the Red Sea, designed for a suite of ocean-colour sensors, that may be used for further testing.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The Red Sea is a narrow, semi-enclosed oceanic basin situated between the continents of Africa and Asia. At its southern end, it is connected to the Gulf of Aden and Arabian Sea, through the strait of Bab-el-Mandeb, and at its northern end to the Mediterranean Sea through the Suez Canal. Situated between 12°N and 28°N, it provides the shortest commercial shipping route between the Atlantic and Indian Ocean and thus is a major economic asset to the region (Johns & Sofianos, 2012). The Red Sea is also the world's northernmost tropical sea and among the warmest and most saline seas on the planet (Belkin, 2009; Longhurst, 2007; Raitsos et al., 2011). These unique environmental conditions (high temperature and salinity) reflect those predicted in other marine regions several decades from now (Christensen et al., 2007).

The Red Sea is categorised as a large marine ecosystem (Belkin, 2009) and sustains coral reefs that provide habitat for a diverse range of marine organisms (Baars, Schalk, & Veldhuis, 1998), including sponges, bi-valves, pelagic larvae, fish, crustaceans, mollusks and echinoderms. At the base of the marine food-web, phytoplankton act as an integral component of these coral reef ecosystems transferring energy to higher levels of the marine food-web, sustaining fisheries and providing sustenance to many inhabitants of the region. Despite the economic and ecological importance of the Red Sea, despite extensive knowledge on its physical characteristics (e.g. Sofianos & Johns, 2003; Yao, Hoteit, Pratt, Bower, Köhl, et al., 2014; Yao, Hoteit, Pratt, Bower, Zhai, et al., 2014) given its strategic position as a commercial shipping route, and despite extensive studies analysing the bio-optical properties of the Gulf of Eilat (Iluz, Yacobi, & Gitelson, 2003; Labiosa, Arrigo, Genin, Monismith, & Van Dijken, 2003; Sokoletsky, Dubinsky, Shoshany, & Stambler, 2003; Sokoletsky, Dubinsky, Shoshany, & Stambler, 2004; Stambler, 2005, 2006) located at the northern tip of the Red Sea, knowledge on large-scale biological dynamics in the region is limited to knowledge on the phytoplankton seasonal cycle, rates of uptake of

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

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

carbon and nitrogen by phytoplankton, and the influence of coral reef ecosystems on Red Sea productivity (Acker, Leptoukh, Shen, Zhu, & Kempler, 2008; Qurban, Balala, Kumar, Bhavya, & Wafar, 2014; Racault et al., 2015; Raitos, Pradhan, Hoteit, Brewin, & Stenchikov, 2013).

The main source of data used to investigate large-scale biological dynamics in the Red Sea has been synoptic estimates of chlorophyll concentration (denoted here as *C* and referred to in this paper as the chlorophyll concentration, being the sum of monovinyl chlorophyll-a, divinyl chlorophyll-a, chlorophyllide-a, and chlorophyll-a epimers and allomers (Werdell & Bailey, 2005), a measure of phytoplankton biomass) derived using satellite ocean-colour data (Acker et al., 2008; Brewin, Raitos, Pradhan, & Hoteit, 2013; Raitos et al., 2013). The temporal and spatial coverages of ocean-colour data surpass that of any *in situ* biological datasets currently available for the Red Sea. Since the advent of the first ocean-colour sensor, NASA's Coastal Zone Color Scanner, satellite ocean-colour data have been used to understand the optical properties of the Red Sea (e.g. Kirby, Parmeter, Arnone, & Oriol, 1993) and other biogeochemically-relevant variables such as the chlorophyll concentration. More recently, Acker et al. (2008) used ocean-colour data from the SeaWiFS and MODIS-Aqua sensors to investigate variations in chlorophyll concentration in the northern Red Sea, and Raitos et al. (2013) used data from MODIS-Aqua to describe the seasonal succession of chlorophyll and its relationship to the physical forcing.

A difficulty with biological interpretation of ocean-colour data from the Red Sea has been the lack of *in situ* data required for validation and uncertainty characterisation. There have been some validation efforts: for instance, Barbini et al. (2004) showed reasonable agreement between satellite-derived chlorophyll and *in situ* lidar fluorescence-derived chlorophyll in the Red Sea. Using *in vivo* fluorometric chlorophyll measurements collected over large spatial scales, Brewin et al. (2013) demonstrated that the performance of standard MODIS-Aqua chlorophyll products in the Red Sea was comparable to that in other regions of the global ocean. Nonetheless, conclusions drawn on biological variability using ocean-colour data in the Red Sea still remain uncertain, due primarily to lack of suitable *in situ* datasets and limited understanding of the optical properties. A further complication in using ocean-colour data for the Red Sea is difficulties in atmospheric correction, for instance, due to the presence of high concentrations of atmospheric dust aerosols from the surrounding arid continents which can render many satellite ocean-colour pixels unusable for analysis of chlorophyll concentration (Acker et al., 2008).

Between September 2009 and March 2012, the Tara Oceans expedition conducted an ~91,000 km voyage to capture the global distribution of marine planktonic organisms (Boss et al., 2013; Werdell, Proctor, Boss, Leeuw, & Ouhssain, 2013). A hyperspectral absorption and

attenuation meter (WETLabs, Inc. AC-S) together with a flow-through system (Boss et al., 2013; Slade et al., 2010) was used for continuous measurements of absorption and attenuation by marine particles along the entire Tara cruise track (Boss et al., 2013). During January 2010, the Tara cruise conducted a meridional transect of the Red Sea (Fig. 1) providing the first comprehensive dataset of absorption and attenuation by particles at large spatial scales in the Red Sea.

Recently, efforts have also been made to improve coverage of ocean-colour data through the merging of data from different ocean-colour platforms (e.g. the GlobColour project; Maritorena, Fanton d'Andon, Mangin, & Siegel, 2010). In 2010 the European Space Agency launched the Ocean Colour Climate Change Initiative (OC-CCI; Brewin et al., 2015) with the goal of creating a long-term, consistent, error-characterised time series of merged ocean-colour products (MODIS-Aqua, SeaWiFS and MERIS), for use in climate-change studies. One of the advantages of the OC-CCI dataset (version 1, see <http://www.oceancolour.org/>) is that, by using an atmospheric correction algorithm called POLYMER (Steinmetz, Deschamps, & Ramon, 2011) on MERIS, which performs retrievals in the presence of sun glint, significant increases in ocean-colour coverage were attained (Steinmetz et al., 2011).

In this paper, we make use of *in situ* bio-optical data collected in the Red Sea as part of the Tara Oceans expedition (Boss et al., 2013; Werdell et al., 2013), together with previous *in situ* datasets (Barbini et al., 2004; Brewin et al., 2013), to investigate the optical properties of the Red Sea in the context of satellite ocean-colour estimates of chlorophyll concentration. Using satellite (OC-CCI data) and *in situ* match-ups, we evaluate the performance of a suite of chlorophyll algorithms in the Red Sea. Discrepancies between satellite and *in situ* chlorophyll are investigated by developing an ocean-colour model for the Red Sea, parameterised to data from the Tara Oceans expedition. The model describes relationships between inherent optical properties of water constituents (absorption and backscattering) and chlorophyll concentration. The ocean-colour model is used to: i) improve our understanding of the optical properties of the region and ii) tune empirical satellite chlorophyll algorithms for use in the Red Sea. Uncertainties in our approach are discussed and conclusions are cautiously stated given the under-sampled nature of this region and its unique atmospheric properties.

2. Methodology

2.1. Study site

The Red Sea is an elongated basin with a mean depth of 524 m and a surface area of $\sim 4.5 \times 10^{11} \text{ m}^2$ (Patzert, 1974). A deep trench runs along the centre of the Red Sea from north to south reaching a

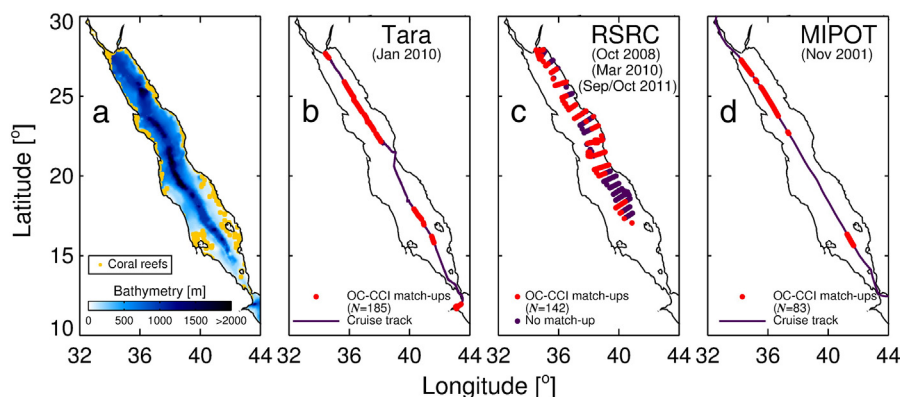


Fig. 1. (a) Bathymetry of the Red Sea (using ETOPO5 sea-floor elevation data obtained from the National Oceanographic and Atmospheric Administration (NOAA) at <http://www.ngdc.noaa.gov/mgg/global/etopo5.html>, regridded to 4 km spatial resolution) with the locations of coral reefs overlain (from the Global Distribution of Coral Reefs dataset (version 2010) compiled by UNEP-WCMC available at <http://data.unep-wcmc.org/datasets/13> in ArcGIS vector format, which was then converted to NetCDF and gridded to 4 km spatial resolution). Figures (b) to (d) show the distribution of the *in situ* and satellite match-up data used in the study.

maximum depth of about ~2300 m, with shallower waters generally found at the southern end and deeper waters toward the north (Fig. 1a). The Red Sea splits into two gulfs at the northern end, the Gulf of Aqaba and the Gulf of Suez. The transport through these gulfs is extremely small, meaning the only significant connection between the Red Sea and the open ocean is through the Strait of Bab el Mandeb at the south (Sofianos & Johns, 2003), where the Red Sea interacts with the Gulf of Aden (Fig. 1a) and seasonal water exchange occurs (Yao, Hoteit, Pratt, Bower, Zhai, et al., 2014). Horizontal circulation in the Red Sea is dominated by eddies (Yao, Hoteit, Pratt, Bower, Köhl, et al., 2014; Zhan, Subramanian, Yao, & Hoteit, 2014) and overturning is influenced by cyclonic recirculation and by overturning circulation in the northern Red Sea, with sinking occurring along the eastern boundary and upwelling along the western boundary (Yao, Hoteit, Pratt, Bower, Köhl, et al., 2014). The Red Sea is surrounded by arid land masses with low precipitation, little riverine input (Patzert, 1974) and high evaporation rates (Sofianos & Johns, 2003). The atmospheric properties over the Red Sea (e.g. large dust storms: Prakash, Stenchikov, Kalenderski, Osipov, & Bangalath, 2015) challenge remote-sensing of ocean colour (Acker et al., 2008).

The seasonal cycles of phytoplankton, as estimated from remote-sensing of ocean colour, indicate higher concentrations during the winter months, attributed to vertical mixing in the north and horizontal advection of nutrient-rich water in the south, and minimum concentrations during the summer, associated with strong seasonal stratification (Raitos et al., 2013). Seasonal changes in chlorophyll may also reflect changes in the carbon to chlorophyll ratio in response to seasonal variations in light and nutrients, with high light (lower chlorophyll) in summer and lower light (higher chlorophyll) in winter. However, during summer months higher concentrations of phytoplankton can occur in some regions. The Red Sea is surrounded by productive coral reefs (Fig. 1a) which contrast in phytoplankton seasonality to open ocean waters (peaking during summer months; Racault et al., 2015) and are thought to contribute to the horizontal transfer of nutrients and phytoplankton by eddies to open waters (Acker et al., 2008; Raitos et al., 2013). For further details on phytoplankton seasonal cycles, spatial structures, trophic regimes and inter-annual variations, inferred using ocean-colour data, the reader is referred to the recent works of Raitos et al. (2013, 2015), Triantafyllou et al. (2014) and Racault et al. (2015).

2.2. In situ data

Three sources of *in situ* data were used in this study (Fig. 1):

- *In situ* hyperspectral absorption and attenuation data collected on a flow-through system (Boss et al., 2013; Werdell et al., 2013) during the Tara Oceans expedition in the Red Sea in January 2010 (hereafter denoted Tara).
- *In vivo* fluorometric data on chlorophyll concentration collected from three research cruises during 2008, 2010 and 2011 (Brewin et al., 2013), as part of the Research Cruises expedition programme of the Red Sea Research Center (RSRC) of King Abdullah University of Science and Technology (KAUST).
- *In vivo* Lidar fluorescence data on chlorophyll concentration in the Red Sea (Barbini et al., 2004) collected as part of the Mediterranean Sea, Indian and Pacific Oceans Transect (MIPOT) oceanographic campaign between Italy and New Zealand in November 2001.

All three sources of *in situ* data have been used previously for comparison with satellite ocean-colour observations (Barbini et al., 2004; Brewin et al., 2013; Werdell et al., 2013), though their inter-consistency is subject to caution, as discussed in Section 5.1.1. These datasets are explained in more detail in the following sections.

2.2.1. Tara data

During January 2010, hyperspectral particulate absorption ($a_p(\lambda)$, where λ is the wavelength) and particulate attenuation ($c_p(\lambda)$) data were collected in the Red Sea on the R/V Tara Oceans expedition using a WET Labs AC-S hyper-spectral spectrophotometer and Sea-Bird Electronics SBE45 MicroTSG unit (Boss et al., 2013; Picheral et al., 2014; Werdell et al., 2013). Water from the ship's flow-through system (~2 m depth) was passed through a Vortex debubbler and then through a WET Labs AC-S. The flow-through system sent the water either directly to the AC-S instrument or through a 0.2 μ m cartridge filter preceding the AC-S instrument. Spectral $a_p(\lambda)$ and $c_p(\lambda)$ in the wavelength interval 400–740 nm were then calculated by subtracting the filtered measurements from the unfiltered measurements, providing calibration-independent estimates of $a_p(\lambda)$ and $c_p(\lambda)$ accounting for instrumental drifts and residual calibration errors. These data were downloaded from the NASA SeaBASS website (<http://seabass.gsfc.nasa.gov/>; Werdell et al., 2003) and processed following the methods described in Slade et al. (2010). The system has been evaluated previously in a wide variety of oceanic waters (Boss et al., 2013; Dall'Olmo, Boss, Behrenfeld, & Westberry, 2012; Dall'Olmo, Westberry, Behrenfeld, Boss, & Slade, 2009; Dall'Olmo et al., 2011; Slade et al., 2010; Westberry, Dall'Olmo, Behrenfeld, & Moutin, 2010) and used to evaluate satellite ocean-colour products (Brewin, Dall'Olmo, Sathyendranath, & Hardman-Mountford, 2012; Werdell et al., 2013). Boss et al. (2013) and Werdell et al. (2013) provide a more detailed description of the processing methods used during Tara.

In total, 9169 spectra (one-minute binned averages) of $a_p(\lambda)$ and $c_p(\lambda)$ were extracted from the Tara dataset spanning the Red Sea (Fig. 1). The hyperspectral data were available at ~4 nm intervals between 404 and 736 nm. Data on $a_p(\lambda)$ and $c_p(\lambda)$ were extracted at the following wavelengths representative of recent ocean-colour sensors (e.g. SeaWiFS, MODIS, MERIS and VIIRS): 410, 412, 443, 486, 488, 490, 510, 530, 547, 551, 555, 560, 620, 665 and 670 nm. Where data at the exact wavelengths were not available, spectral interpolation was used. Data for $a_p(\lambda)$ were also extracted at 650, 676 and 715 nm, for estimation of the chlorophyll concentration. To derive chlorophyll concentration (C) we used the method of Boss et al. (2013) and Werdell et al. (2013), whereby the phytoplankton absorption coefficient at 676 nm ($a_{ph}(676)$) is first estimated using the line height method of Davis, Moore, Zaneveld, and Napp (1997) as modified by Boss, Collier, Larson, Fennel, and Pegau (2007), such that

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

The chlorophyll concentrations (C) were then estimated from $a_{ph}(676)$ following Werdell et al. (2013), according to

$$C = \left[\frac{a_{ph}(676)}{0.0152} \right]^{0.9055}. \quad (2)$$

The coefficients for Eq. (2) were computed from 52 corresponding High Performance Liquid Chromatography (HPLC) samples of chlorophyll and $a_{ph}(676)$ (computed using Eq. (1), samples taken within 1 h) collected along the entire Tara Oceans expedition (see Fig. 1 of Werdell et al., 2013). The method of using the line-height of the absorption peak at red wavelengths to estimate chlorophyll concentration has been found to perform with high accuracy when compared with discrete *in situ* HPLC chlorophyll data in a wide range of natural optical environments and diverse phytoplankton cultures (Dall'Olmo et al., 2009, 2012; Roesler & Barnard, 2013; Westberry et al., 2010). Furthermore, the line height determination is also effective in removing the contributions of absorption by coloured dissolved organic matter (CDOM) and non-algal particles, relatively independent of the effects of biofouling, and unlike the *in vivo* fluorometric method for estimating chlorophyll, the absorption line height is insensitive to incident irradiance and non-photochemical quenching (Roesler & Barnard, 2013).

Following Werdell et al. (2013), the particulate backscattering coefficient (b_{bp}) was estimated from the a_p and c_p data by using the scattering-to-backscattering ratio as a function of chlorophyll, according to the method of Twardowski et al. (2001), such that

$$b_{bp}(\lambda) = (0.0096C^{-0.253}) [c_p(\lambda) - a_p(\lambda)]. \quad (3)$$

Whereas b_{bp} modelled from a_p and c_p is less ideal than direct measurement of b_{bp} , the scattering-to-backscattering ratio has been shown to vary consistently according to trophic conditions for a variety of open-ocean waters (Dall'Olmo et al., 2012; Twardowski et al., 2001; Whitmire, Boss, Cowles, & Pegau, 2007).

During the Tara cruise a HyperPro (Satlantic Free-Falling Optical Profile with two HyperOCR optical sensors) was deployed at various locations during sunny or mostly sunny periods. The instrument was deployed in surface mode, where a float was attached to the profiler at the surface allowing for measurements of upwelling radiance beneath the surface and downwelling irradiance just above the water, which are then used to compute remote-sensing reflectance (see Chase, 2013 for data processing methods). Data on *in situ* remote-sensing reflectance were available at two stations in the Red Sea.

2.2.2. RSRC data

During 2008, 2010 and 2011, oceanographic data were collected from three research cruises as part of the Research Cruises expedition programme of the RSRC of KAUST (Fig. 1). Continuous *in vivo* fluorescence vertical profiles were collected at each station using a WET Labs ECO-FLNTUs (FLNTURTD-964) fluorometer attached to a CTD. A total of 401 *in situ* samples from the three RSRC cruises were used in the analysis, consisting of average chlorophyll concentrations within the first optical depth. Brewin et al. (2013) provide a detailed description of the processing methods used during the three RSRC cruises (see Section 2.1 of Brewin et al., 2013) and additional details are also provided as Supplementary material. Of the 401 *in situ* samples, 33 were from 2008, 107 from 2010 and 261 from the 2011 cruise.

2.2.3. MIPOT data

As part of the MIPOT oceanographic campaign, conducted between Italy and New Zealand, measurements of surface chlorophyll concentration were derived using a lidar fluorosensor aboard the RV *Italica* (Barbini, Colao, Fantoni, Fiorani, & Palucci, 2001; Barbini, Colao, Fantoni, Fiorani, & Palucci, 2003; Barbini, Colao, Fantoni, Palucci, & Ribezzo, 1999; Barbini, Colao, Fantoni, Palucci, & Ribezzo, 2001; Barbini et al., 2004). Between 18th and 22nd November the RV *Italica* passed through the centre of the Red Sea (Fig. 1), resulting in 505 measurements of *in situ* chlorophyll concentration. This data were downloaded from the NASA SeaBASS website <http://seabass.gsfc.nasa.gov/> (Werdell et al., 2003).

2.3. Satellite data

2.3.1. OC-CCI data

The OC-CCI project is one of 14 ESA funded CCI projects and focuses specifically on creating a consistent, error-characterised time-series of ocean-colour products, for use in climate-change studies (Brewin et al., 2015). From 2014, Version 1 of the OC-CCI dataset has been available for community use (<http://www.oceancolour.org/>). The dataset consists of a time-series of merged and bias-corrected MERIS, MODIS-Aqua and SeaWiFS data at 4 km-by-4 km resolution. Briefly, top-of-atmosphere reflectance data for SeaWiFS, MODIS-Aqua and MERIS, at visible wavebands, were processed using atmospheric-correction models, SeaDAS (Fu, Baith, & McClain, 1998) for SeaWiFS and MODIS-Aqua (SeaDAS version 7.0 using standard SeaDAS flags) and POLYMER (Steinmetz et al., 2011) for MERIS. Remote-sensing reflectance (R_{rs}) data from MODIS-Aqua and MERIS are then band-shifted to match the

wavelengths of SeaWiFS using an in-water bio-optical model (e.g. see Mélin & Zibordi, 2007). MODIS-Aqua and MERIS are then bias-corrected to SeaWiFS at each wavelength, using a pixel-by-pixel bias-correction model developed using daily match-up data between the respective sensors during over-lapping time periods, and finally the data are merged into a single product. For further information on OC-CCI processing, extensive documentation can be found at the following website <http://www.esa-oceancolour-cci.org/>.

We chose to use OC-CCI products primarily due to improved coverage in the Red Sea region when compared with individual sensors (see Fig. 2 of Racault et al., 2015) and other merged products (e.g. GlobColour, see Section 8 of Sathyendranath & Krasemann, 2014). Daily level 3 sinusoidal projected R_{rs} data at 412, 443, 490, 510, 555 and 670 nm, were downloaded from the OC-CCI website covering the period September 1997 to July 2012. For regional verification of the OC-CCI products, daily level 3 binned R_{rs} data for SeaWiFS processed with SeaDAS, MODIS processed with SeaDAS, MERIS processed with SeaDAS and MERIS processed with POLYMER were also used for the January 2010 period, which corresponds to the time of the Tara Expedition passing through the Red Sea.

2.3.2. *In situ* satellite match-up data

Daily $R_{rs}(\lambda)$ data from OC-CCI were matched in time (daily temporal matchup) and space (latitude and longitude, closest 4 km pixel) with *in situ* chlorophyll data. For all *in situ* datasets, but particularly the underway cruises (Tara and MIPOT), 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. No data were used for more than one matchup. A total of 410 OC-CCI match-ups were obtained, 185 for Tara, 142 for RSRC and 83 for MIPOT. The geographical distribution of these match-ups is shown in Fig. 1. For Tara, an additional 84 SeaWiFS match-ups were available, 152 for MODIS, 96 for MERIS processed with SeaDAS and 154 for MERIS processed with POLYMER.

2.4. Satellite chlorophyll algorithms

2.4.1. OC4

Satellite chlorophyll (C) algorithms incorporated into the comparison are described in the following section. The standard NASA OC4 (v6) band-ratio chlorophyll algorithm (O'Reilly et al., 2000) is a polynomial algorithm which relates the ratio of remote-sensing reflectances at three pairs of wavebands with chlorophyll concentration (C), using the following algorithm:

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

The chlorophyll (C) is estimated according to:

$$C = 10^{(q_0 + q_1 X + q_2 X^2 + q_3 X^3 + q_4 X^4)}, \quad (5)$$

where $q_0 = 0.3272$, $q_1 = -2.9940$, $q_2 = 2.7218$, $q_3 = -1.2259$ and $q_4 = -0.5683$ (NASA, 2010).

2.4.2. OCI

The band-difference algorithm of Hu, Lee, and Franz (2012) was also explored. This algorithm has been found to perform well at low chlorophyll concentrations ($<0.25 \text{ mg m}^{-3}$ Hu et al., 2012), which are typical of the Red Sea (Brewin et al., 2013). The approach uses a Colour Index (denoted here as ξ), 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:

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

Chlorophyll is then related to ξ using the following equation:

$$C = 10^{A+B\xi}, \quad (7)$$

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 ($> 0.3 \text{ mg m}^{-3}$) OC4 is used (Eq. 5), whereas for chlorophyll concentrations between 0.25 and 0.3 mg m^{-3} , a combination of Eqs. (7) and (5) are used to facilitate a smooth transition between algorithms. The OCI algorithm is expressed as

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

where α serves to provide a linear transition from Eq. (7) to Eq. (5) as chlorophyll increases from 0.25 to 0.3 mg m^{-3} . This parameter is computed as $\alpha = (10^{A+B\xi} - 0.25)/(0.3 - 0.25)$.

2.4.3. GSM

The semi-analytical Garver–Siegel–Maritorena (GSM) model, initially developed by Garver and Siegel (1997) and later updated by Maritorena, Siegel, and Peterson (2002), was also used in this study. The GSM model retrieves simultaneous estimates of chlorophyll, absorption by combined detrital and dissolved matter at 443 nm ($a_{dg}(443)$) and particle backscattering at 443 nm ($b_{bp}(443)$) from $R_{rs}(\lambda)$, assuming an underlying bio-optical model and using non-linear optimisation. 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 ($0.01 < C < 64 \text{ mg m}^{-3}$).

2.4.4. OC4-MG

The OC4 model is designed for applications in Case-1 waters, in which absorption and scattering by substances other than phytoplankton are assumed to covary with chlorophyll concentration. Morel and Gentili (2009b) developed a correction that can be applied to OC4-type of algorithms, to account for deviations in absorption by gelbstoff (coloured dissolved organic matter, CDOM, denoted a_g) from the Case-1 assumption. The OC4 algorithm corrected according to Morel and Gentili (2009b) is referred to here as OC4-MG.

This approach is based on the assumption that the ratio of reflectance at 412 nm to that at 443 nm ($R(412)/R(443)$) is mainly sensitive to CDOM, albeit influenced to some extent by chlorophyll, and that the ratio of reflectance at 490 nm to that at 555 nm ($R(490)/R(555)$) is essentially dependent on chlorophyll, although also influenced to some extent by CDOM. The approach uses the bio-optical model of Morel and Maritorena (2001), that implicitly includes a prescribed relationship between CDOM absorption and chlorophyll, and thus produces a unique set of curves relating $R(412)/R(443)$ to $R(490)/R(555)$. Deviations from the prescribed relationship are introduced using a factor ϕ , with $\phi > 1$ indicating an excess and $\phi < 1$ a deficit of a_g per unit chlorophyll. Morel and Gentili (2009b) produced a 2D lookup table relating $R(412)/R(443)$ to $R(490)/R(555)$ for specific discrete values of ϕ . Relative anomalies in CDOM (ϕ) with respect to its standard (chlorophyll-related) values can then be computed efficiently using reflectance ratios derived from ocean colour, by first converting $R_{rs}(\lambda)$ to $R(\lambda)$ (see Appendix B of Morel & Gentili, 2009b). Once ϕ is obtained, chlorophyll from standard band-ratio algorithms (OC4) can be corrected for the effect of excess or deficit of CDOM (OC4-MG), by using another 2D lookup table (Morel & Gentili, 2009b) and inputting the initial chlorophyll estimate (OC4) and the retrieved ϕ value.

2.5. Statistical tests and algorithm ranking

To compare the performance of the satellite chlorophyll algorithms with the *in situ* data in the Red Sea, a suite of univariate statistical tests were used following Brewin et al. (2015). These included: the Pearson correlation coefficient (r); the root mean square error (Ψ); the average bias between model and measurement (δ); the centrepattern (or unbiased) root mean square error (Δ), noting that $\Delta^2 = \Psi^2 - \delta^2$; the slope (S) and intercept (I) of a Type-2 regression; and the percentage of possible retrievals (η). The equations used for each of these statistical tests are provided in Section 4.1 of Brewin et al. (2015). All statistical tests were performed in \log_{10} space, considering chlorophyll is approximately log-normally distributed (Campbell, 1995), at least in the range of 0.01 to 10 mg m^{-3} typically observed in the open-ocean.

To rank the performance of the algorithms used in this study we adopted the classification presented in Brewin et al. (2015). Briefly, for each statistical test (r , Ψ , δ , Δ , S , I and η) the statistic of a model is compared with the average of all models, and a score is assigned based (predominately) on whether the statistic in question is significantly worse (0 points), similar (1 point) or better (2 points) than the average of all models. All points are then summed over each statistical test and the total score for each model is normalised by the average score of all models being tested. A score of one indicates the performance of a model is average with respect to all models, a score greater than one indicates model performance is better than the average, and a score less than one indicates model performance is worse than average.

Following Brewin et al. (2015), the stability of the scoring system and the sensitivity of the scores were tested using the method of bootstrapping (Efron, 1979; Efron & Tibshirani, 1993). This involved random re-sampling with replacement of the *in situ* data (1000 times) to create 1000 new datasets of the same size as the original dataset, but not identical to it (Monte-Carlo approach). The points classification was then re-run for each new dataset and from the resulting distribution of scores, a mean score for each model was computed. Additionally, 15.9% and 84.1% confidence intervals (equivalent to one standard deviation for a normal distribution) on the bootstrap distribution were taken and assumed to be the error-bars or confidence limits on the mean score for each model. For further details on the points classification method the reader is referred to Section 4 of Brewin et al. (2015).

3. Ocean-colour model for the Red Sea

In this section we present an ocean-colour model that is tuned to data in the Red Sea. It is based on the classic Case-1 bio-optical assumption (Morel & Prieur, 1977) that ocean-colour reflectance can be related to total chlorophyll concentration (C) and co-varying material. We consider this assumption a reasonable starting point given that: the focus of the study is on ocean-colour chlorophyll algorithms; the Red Sea is oligo- to meso-trophic in nature (Raitos et al., 2013) with negligible terrestrial or riverine input; and little is known about the optical properties in the region.

We begin by assessing the relationship between absorption by particles (including both phytoplankton and detritus) at 443 nm, denoted $a_p(443)$, and chlorophyll concentration (C) estimated from a_p in the red during Tara, shown in Fig. 2a. Whereas chlorophyll is estimated from a_p between 650 and 715 nm following Eqs. (1) and (2), the relationship between $a_{ph}(676)$ and $a_{ph}(443)$ varies due to changes in accessory pigmentation among phytoplankton that influences the blue region of the absorption spectrum with minor influence on the red region (Bricaud, Claustre, Ras, & Oubelkheir, 2004), and also because the flattening effect on the absorption spectra, related to the size of the particles and variations in pigment concentration for a cell of a given size, is not the same in the red and in the blue-green parts of the spectrum. As measured during Tara in January 2010, chlorophyll varies between 0.04 and 1.0 mg m^{-3} (Fig. 2a), confirming the Red Sea is oligo- to meso-

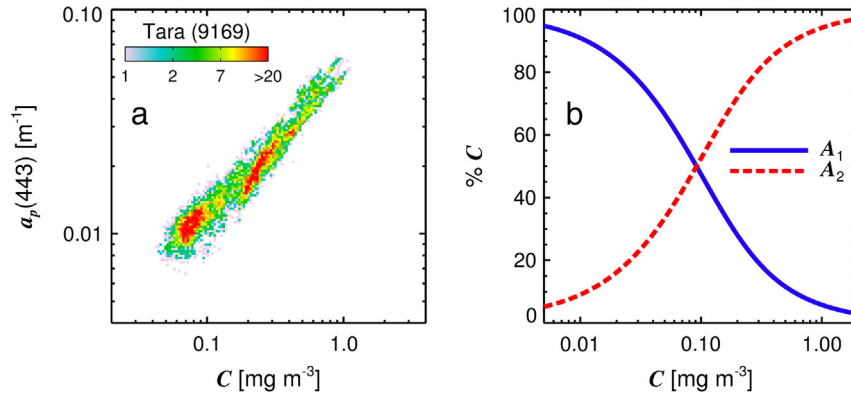


Fig. 2. (a) Relationship between chlorophyll and $a_p(443)$ measured during Tara and (b) the percentage contribution of assemblages 1 (A_1) and 2 (A_2) to total chlorophyll as a function of total chlorophyll (C) following the parameterisation of Brotas et al. (2013) for pico-phytoplankton.

trophic. Fig. 2a shows two dense groups (or assemblages) of particles: one group (hereafter denoted assemblage 1) is found between 0.06 and 0.1 mg m^{-3} of chlorophyll ($a_{ph}(443)$ between 0.009 and 0.013 m^{-1}) and the other group (hereafter denoted assemblage 2) is found between 0.2 and 0.6 mg m^{-3} of chlorophyll ($a_{ph}(443)$ between 0.013 and 0.030 m^{-1}). These assemblages are also separated geographically, with assemblage 1 found mainly in the northern oligotrophic waters of the Red Sea and assemblage 2 in the more mesotrophic southern waters. The optical properties of the two assemblages (with respect to $a_{ph}(443)$) are significantly different, with the particulate specific absorption coefficient at 443 nm ($a_p^*(443) = a_p(443)/C$) of assemblage 1 considerably higher ($\sim 0.22 \text{ m}^2 (\text{mg C})^{-1}$) than that of assemblage 2 ($\sim 0.05 \text{ m}^2 (\text{mg C})^{-1}$).

3.1. Two-component model of chlorophyll

Based on the finding that the Red Sea is primarily dominated by two assemblages of particles with different optical properties, we develop a two-component ocean-colour model for the Red Sea. For this model, assemblage 1 is representative of more oligotrophic conditions and assemblage 2 of mesotrophic conditions. Note that other studies (e.g. Brewin, Devred, Sathyendranath, Hardman-Mountford, & Lavender, 2011; Devred, Sathyendranath, Stuart, & Platt, 2011) have highlighted the benefits of including a third assemblage representative of more eutrophic conditions. However, this is not a requirement for the Red Sea considering chlorophyll rarely exceeds 1 mg m^{-3} . We start by relating the chlorophyll concentration of assemblage 1 (C_1) to the total chlorophyll concentration following an equation first presented by Sathyendranath, Stuart, Cota, Maass, and Platt (2001)

$$C_1 = C_1^m [1 - \exp(-S_1 C)], \quad (9)$$

where C_1^m represents the asymptotic maximum chlorophyll concentration for assemblage 1, and S_1 determines the initial increase in C_1 with C . The chlorophyll concentration of assemblage 2 (C_2) can then be computed according to

$$C_2 = C - C_1^m [1 - \exp(-S_1 C)]. \quad (10)$$

The parameters of Eq. (9) (C_1^m and S_1) can be computed with knowledge on C_1 and C , acquired from either High Performance Liquid Chromatography (HPLC, Brewin, Dall'Olmo, et al., 2012; Brewin, Hirata, et al., 2012b; Brewin et al., 2010, 2011; Brotas et al., 2013), fluorometric analysis using size-fractionated filtration (Brewin, Sathyendranath, Tilstone, Lange, & Platt, 2014), or from a_{ph} and chlorophyll data, by making assumptions on the relationship between C_1^m and S_1 (Devred et al., 2006, 2011). As there is currently little information in the Red Sea on HPLC or fluorometric analysis using size-fractionated

filtration, we chose to use the parameters of Brotas et al. (2013) for pico-phytoplankton within the 1st optical depth, such that $C_1^m \sim 0.06$ and $S_1 \sim 17.06$ (see Table 1). We chose to use these parameters for the

Table 1
Parameters of the Red Sea ocean-colour model.

Parameters for the 2-component assemblage model (Eqs. 9 and 10)			
Reference	C_1^m [mg m ⁻³]	S_1 [dimensionless]	
Brotas et al. (2013)	0.058 (0.054–0.063)	17.056 (14.693–18.646)	
Parameters for particle absorption model (Eq. 12)			
Wavelength [nm]	$a_{p,1}^*$ [m ² (mg C) ⁻¹]	$a_{p,2}^*$ [m ² (mg C) ⁻¹]	
410	0.1823 (0.1807–0.1840)	0.0464 (0.0459–0.0468)	
412	0.1858 (0.1841–0.1874)	0.0472 (0.0467–0.0477)	
443	0.2132 (0.2120–0.2142)	0.0461 (0.0457–0.0465)	
486	0.1341 (0.1335–0.1347)	0.0347 (0.0345–0.0350)	
488	0.1313 (0.1307–0.1319)	0.0340 (0.0337–0.0342)	
490	0.1284 (0.1278–0.1290)	0.0332 (0.0330–0.0334)	
510	0.0745 (0.0740–0.0749)	0.0242 (0.0241–0.0244)	
530	0.0344 (0.0340–0.0347)	0.0196 (0.0195–0.0197)	
547	0.0195 (0.0192–0.0198)	0.0170 (0.0169–0.0172)	
551	0.0183 (0.0181–0.0187)	0.0154 (0.0153–0.0156)	
555	0.0168 (0.0166–0.0171)	0.0138 (0.0137–0.0140)	
560	0.0156 (0.0153–0.0159)	0.0117 (0.0116–0.0119)	
620	0.0110 (0.0108–0.0112)	0.0056 (0.0055–0.0057)	
665	0.0251 (0.0250–0.0253)	0.0173 (0.0173–0.0174)	
670	0.0291 (0.0290–0.0293)	0.0201 (0.0200–0.0201)	
Phytoplankton assemblage parameters for particle backscattering model (Eq. 14)			
Parameter	$i = 1$	$i = 2$	
$b_{pp,i}^*(\lambda_0)$ [m ² (mg C) ⁻¹]	0.0016 (0.0003–0.0028)	0.0023 (0.0022–0.0024)	
γ_i [dimensionless]	4.48 (2.59–16.18)	0.43 (0.38–0.50)	
Background component parameters for particle backscattering model (Eq. 14)			
$b_{bp}^k(\lambda_0)$ [m ⁻¹]	γ_k [dimensionless]		
0.00096 (0.00091–0.00102)	0.54 (0.45–0.68)		
Parameters for CDOM absorption model (Eq. 18)			
Parameter	$i = 1$	$i = 2$	
$a_{g,i}(\lambda_0)$ [m ² (mg C) ⁻¹]	0.0907 (0.0655–0.1425)	0.2022 (0.1791–0.2754)	
$S_{g,i}$ [nm ⁻¹]	0.031 (0.028–0.038)	0.014 (0.013–0.015)	
Parameters of Eq. (21) of Lee et al. (2009) ^a			
$G_w^0(\Omega)$	$G_w^1(\Omega)$	$G_p^0(\Omega)$	$G_p^1(\Omega)$
0.0604	0.0406	0.0402	0.1310

95% confidence intervals are in brackets and were estimated using a Monte-Carlo approach and $\lambda_0 = 443 \text{ nm}$.

^a Parameters are for solar zenith angle in air equal to zero, sensor nadir-view angle in air equal to zero, and a sensor azimuth angle in relation to the solar plane equal to zero, where Ω collectively represents the three sun-sensor angular geometries.

following reasons: (1) when fitting Eq. (12) (shown below) to data in the Red Sea, these parameters resulted in the lowest bias and the highest correlation coefficient between modelled $a_p(443)$ and measured $a_p(443)$ when compared with results obtained using other published parameters (e.g. Brewin, Dall'Olmo, et al., 2012; Brewin, Hirata, et al., 2012; Brewin et al., 2010, 2011, 2014); (2) with these model parameters, the dominant assemblage transitions from 1 to 2 at around 0.1 mg m^{-3} of chlorophyll (Fig. 2b), consistent with results in Fig. 2a; and (3) these parameters were computed using data within the penetration depth of the satellite signal (Brotas et al., 2013). The percentage contribution of each assemblage to total chlorophyll as a function of total chlorophyll is illustrated in Fig. 2b.

3.2. Particulate absorption model

The particulate absorption ($a_p(\lambda)$) model we adopt follows a similar model presented by Sathyendranath et al. (2001) for phytoplankton absorption. Here, $a_p(\lambda)$ is modelled as a function of chlorophyll (C) according to

$$a_p(\lambda) = a_{p,1}^*(\lambda)C_1 + a_{p,2}^*(\lambda)C_2, \quad (11)$$

where $a_{p,1}^*(\lambda)$ and $a_{p,2}^*(\lambda)$ represent the chlorophyll-specific particulate absorption coefficients of assemblages 1 and 2, respectively. Expanding

Eq. (11) by inserting Eqs. (9) and (10) results in the following expression:

$$a_p(\lambda) = C_1^m \left[a_{p,1}^*(\lambda) - a_{p,2}^*(\lambda) \right] \left[1 - \exp(-S_1 C) \right] + a_{p,2}^*(\lambda) C. \quad (12)$$

By setting C_1^m and S_1 according to Brotas et al. (2013) (Table 1, Fig. 2), Eq. (12) was fitted to $a_p(\lambda)$ and C from Tara to derive $a_{p,1}^*(\lambda)$ and $a_{p,2}^*(\lambda)$. The fitting procedure used a standard, non-linear least squares method (Levenberg–Marquardt, IDL Routine MPFITFUN Moré, 1978; Markwardt, 2008), and was fitted in \log_{10} space, assuming a log-normal distribution for $a_p(\lambda)$ (Bricaud, Morel, Babin, Allali, & Claustre, 1998). The retrieved parameters, provided in Table 1 and plotted in Fig. 3a, were obtained by taking the median and 95% confidence interval on the 1000 bootstraps. This involved: randomly re-sampling the data with replacement to create 1000 new datasets of the same sample size as the original dataset but not identical to it or each other; fitting Eq. (12) to each dataset to obtain a distribution of parameters; and then taking the median and 95% confidence intervals on the estimated parameter distribution, the latter to provide an indication of uncertainty on the retrieved parameter. Assemblage 1 has a higher chlorophyll-specific particulate absorption in the blue region of the spectrum when compared with assemblage 2 (Fig. 3a). Eq. (12) captures the variability in a_p as a function of C with low error (Table 2), and the proposed relationship between $a_p(443)$ and C is similar to that presented

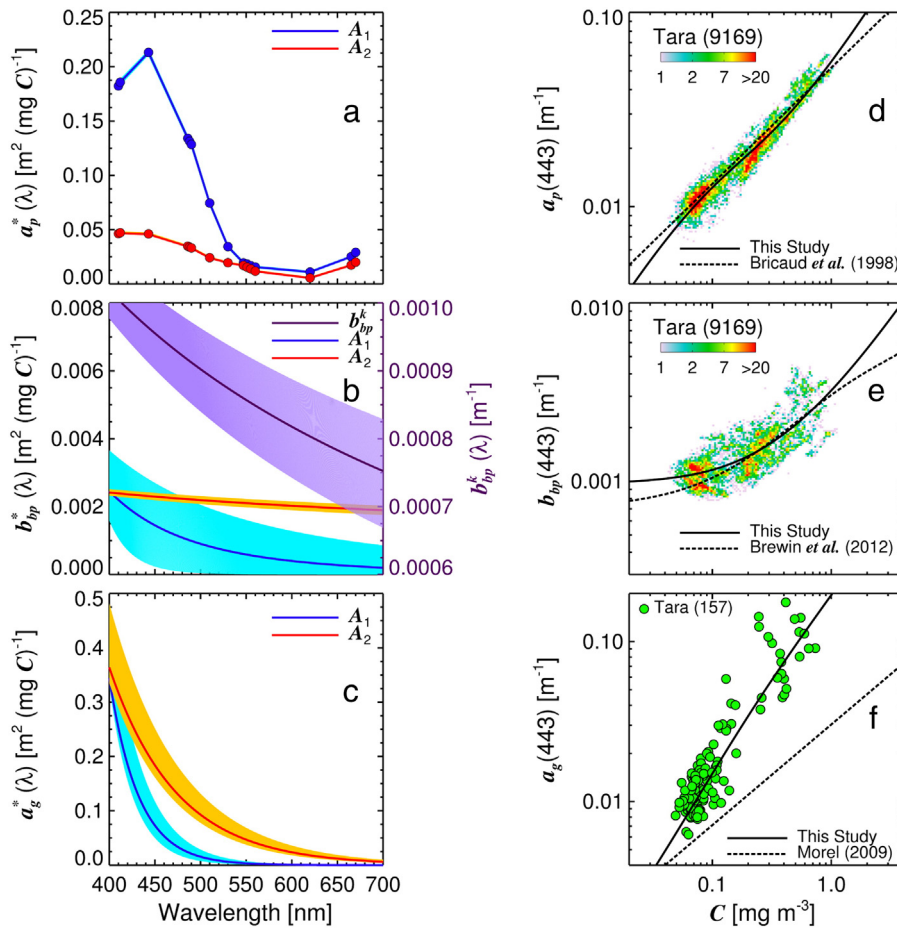


Fig. 3. Parameters of the Red Sea two-component bio-optical model: (a) chlorophyll-specific particulate absorption coefficients (a_p) as a function of wavelength, (b) chlorophyll-specific particulate backscattering coefficients (b_{bp}) as a function of wavelength and the background backscattering component (b_{bp}^k), and (c) chlorophyll-specific CDOM absorption coefficients (a_g) as a function of wavelength for the two assemblages (A_1 and A_2). Figures d–f show the relationship between $a_p(443)$ and C , $b_{bp}(443)$ and C , and $a_g(443)$ and C respectively, for the Tara data and the fitted model.

Table 2
Performance of Eqs. (12), (14) and (18).

Wavelength [nm]	a_p		a_g		b_{bp}	
	r	Ψ	r	Ψ	r	Ψ
410	0.937	0.079	0.919	0.122	0.778	0.091
412	0.937	0.078	0.919	0.123	0.778	0.091
443	0.969	0.050	0.920	0.142	0.776	0.091
486	0.979	0.045	0.920	0.162	0.787	0.090
488	0.979	0.045	0.920	0.162	0.788	0.090
490	0.978	0.045	0.920	0.163	0.790	0.090
510	0.977	0.051	0.920	0.168	0.795	0.090
530	0.970	0.070	0.920	0.172	0.796	0.091
547	0.959	0.092	0.919	0.177	0.799	0.090
550	0.957	0.094	0.919	0.178	0.800	0.090
555	0.953	0.098	0.919	0.180	0.800	0.090
560	0.948	0.101	0.919	0.182	0.799	0.090
620	0.904	0.129	0.919	0.226	0.784	0.094
665	0.994	0.033	0.919	0.278	0.763	0.102
670	0.996	0.026	0.919	0.284	0.763	0.102

r refers to the Pearson correlation coefficient and Ψ the root mean square error. All statistical tests were performed in \log_{10} space.

by Bricaud et al. (1998) using a global dataset, but with no measurements in the Red Sea (Fig. 3d).

3.3. Particulate backscattering model

Various models have been presented that relate particulate backscattering (b_{bp}) to chlorophyll (C) in open-ocean (Case-1) waters (e.g. Huot, Morel, Twardowski, Stramski, & Reynolds, 2008; Morel & Maritorena, 2001; Sathyendranath et al., 2001; Twardowski et al., 2001). To maintain consistency with the two-component model of particulate absorption (Eq. 12), we used the approach of Brewin, Dall'Olmo, et al. (2012), such that

$$b_{bp}(\lambda) = C_1^m \left[b_{bp,1}^*(\lambda) - b_{bp,2}^*(\lambda) \right] [1 - \exp(-S_1 C)] + b_{bp,2}^*(\lambda) C + b_{bp}^k(\lambda), \quad (13)$$

where $b_{bp,1}^*(\lambda)$ and $b_{bp,2}^*(\lambda)$ represent the chlorophyll-specific particulate backscattering coefficients of assemblage 1 and 2 respectively, and $b_{bp}^k(\lambda)$ represents a constant background. It is worth noting that the two assemblages used in this study (assemblage 1 and 2) are different from those used in Brewin, Dall'Olmo, et al. (2012). In the Brewin, Dall'Olmo, et al. (2012) study, where the focus was on developing a model from oligo- to eu-trophic waters, small cells were grouped as pico-nanoplankton (cells <20 μm) and larger cells as microplankton (cells >20 μm). Alternatively in this study, where the focus is on oligo- to meso-trophic waters, assemblage 1 is likely more representative of the picoplankton size class and assemblage 2 of the nanoplankton size class. This difference is reflected in the choice of parameters for C_1^m and S_1 in Eq. (13).

Eq. (13) was fitted using $b_{bp}(\lambda)$ and C from Tara to derive $b_{bp,1}^*(\lambda)$, $b_{bp,2}^*(\lambda)$ and $b_{bp}^k(\lambda)$ using the same fitting procedure as that used in Eq. (12) and setting C_1^m and S_1 according to Brotas et al. (2013) (Table 1). Following Brewin, Dall'Olmo, et al. (2012), the spectral dependency of $b_{bp,1}^*(\lambda)$, $b_{bp,2}^*(\lambda)$ and $b_{bp}^k(\lambda)$ was assumed to follow a power function, such that

$$b_{bp}(\lambda) = b_{bp,1}^*(\lambda_0)(\lambda/\lambda_0)^{-\gamma_1} \{ C_1^m [1 - \exp(-S_1 C)] \} + b_{bp,2}^*(\lambda_0)(\lambda/\lambda_0)^{-\gamma_2} \{ C - C_1^m [1 - \exp(-S_1 C)] \} + b_{bp}^k(\lambda_0)(\lambda/\lambda_0)^{-\gamma_k}, \quad (14)$$

where γ_1 , γ_2 and γ_k represent the exponents of the power function describing the spectral shape of each component of the model, and $\lambda_0 = 443$ nm. Having obtained a distribution of parameters for $b_{bp,1}^*(\lambda)$, $b_{bp,2}^*(\lambda)$ and $b_{bp}^k(\lambda)$ from fitting Eq. (13) to the $b_{bp}(\lambda)$ and C data, the

remaining parameters (γ_1 , γ_2 and γ_k) were obtained by fitting a wavelength-dependent power function to $b_{bp,1}^*(\lambda)$, $b_{bp,2}^*(\lambda)$ and $b_{bp}^k(\lambda)$. Retrieved parameters for Eq. (14) are provided in Table 1 together with their 95% confidence intervals and plotted in Fig. 3b.

In general, assemblage 2 has a higher (and better constrained) chlorophyll-specific particulate backscattering than assemblage 1 at most wavelengths (Fig. 3b), with assemblage 1 having a steeper power exponent than assemblage 2, consistent with the expectation that smaller particles have a steeper power exponent than larger particles (Kostadinov, Siegel, & Maritorena, 2009; Loisel, Nicolas, Sciandra, Stramski, & A.P., 2006). However, it is worth noting that the background component γ_k is flatter than that observed by Brewin, Dall'Olmo, et al. (2012) and that γ_1 is also very high (close to that of molecular scattering by pure-water, see Table 1). It is likely that more data at very low chlorophyll concentrations would help to separate better the spectral shape of backscattering by the background component and that by assemblage 1 when fitting Eq. (14). Nonetheless, Eq. (14) is seen to capture most of the variability in b_{bp} as a function of chlorophyll (C) with reasonable error statistics (Table 2). The model captures the general trend in $b_{bp}(443)$ with increasing chlorophyll, and is in good agreement with that proposed by Brewin, Dall'Olmo, et al. (2012) using a global dataset that did not contain any measurements from the Red Sea (Fig. 3e).

3.4. CDOM absorption model

Considering that no direct measurements of CDOM absorption were taken during Tara, we used satellite match-ups between *in situ* chlorophyll and satellite OC-CCI R_{rs} to develop a dataset of CDOM absorption (a_g) and chlorophyll, from which we could parameterise a model. Using the 185 OC-CCI match-ups for Tara, we first converted $R_{rs}(\lambda)$ to $R(\lambda)$ (see Appendix B of Morel & Gentili, 2009b) and implemented the Morel and Gentili (2009b) 2D lookup table to obtain ϕ . Of the 185 match-ups, there were 157 realistic estimates of ϕ with 28 ϕ estimates falling outside the realistic range (Morel & Gentili, 2009b). The values of ϕ were found to vary between 1.29 and 8.76, with a median value of 2.11 and a standard deviation of 1.25, and indicate that absorption by CDOM per unit chlorophyll during Tara was on average twice that of standard Case-1 relationships. Absorption by CDOM at 400 nm ($a_g(400)$) was computed for each *in situ* sample using *in situ* chlorophyll data and the estimate of ϕ (Morel, 2009; Morel & Gentili, 2009b), such that

$$a_g(400) = \phi 0.065 C^{0.63}. \quad (15)$$

To estimate $a_g(\lambda)$ from $a_g(400)$, Morel and Gentili (2009b) used a fixed value of 0.018 nm^{-1} for the exponential slope of CDOM with wavelength (S_g). Whereas the 2D lookup table used to obtain ϕ implicitly includes this hypothesis, a sensitivity study in Morel and Gentili (2009b) indicates the choice of the S_g is not crucial when deriving ϕ . Recent evidence from *in situ* and satellite inversion data indicates S_g is negatively correlated with a_g at a reference wavelength (Bricaud, Ciotti, & Gentili, 2012; Swan, Nelson, Siegel, & Fields, 2013), likely related to the selective bleaching of some CDOM components by sunlight (Whitehead et al., 2000), and possibly important in regions exposed to high doses of solar radiation such as the Red Sea. Using data from the NASA NOMAD dataset (Werdell & Bailey, 2005), from only Case-1 waters, following Lee and Hu (2006), and data from the Bermuda Bio-Optics Project (Siegel et al., 2001), S_g was found to be significantly correlated with $a_g(400)$ following a non-linear relationship ($S_g = 0.0086 a_g(400)^{-0.2024}$, $r = 0.61$, $p < 0.0001$, see Supplementary Fig. S1). This relationship was used to estimate spectral a_g from the 157 estimates of $a_g(400)$, such that

$$a_g(\lambda) = a_g(400) \exp \left[-0.0086 a_g(400)^{-0.2024} (\lambda - 400) \right]. \quad (16)$$

For the 157 measurements, S_g was found to vary between 0.011 and 0.020 nm^{-1} , with a median value of 0.018 nm^{-1} and a standard deviation of 0.002 nm^{-1} , and was consistent with the value of 0.018 nm^{-1} originally used by Morel and Gentili (2009b).

Having developed this dataset of inferred $a_g(\lambda)$ and C characteristic of conditions in the Red Sea during the Tara Expedition, this dataset was used to develop a Red Sea model that related $a_g(\lambda)$ directly to C , given that no direct measurements of CDOM absorption were available. For consistency with the two-component model of particulate absorption (Eq. 12), we used a similar approach to describe the relationship between chlorophyll (C) and $a_g(\lambda)$, where

$$a_g(\lambda) = C_1^m [a_{g,1}^*(\lambda) - a_{g,2}^*(\lambda)] [1 - \exp(-S_1 C)] + a_{g,2}^*(\lambda) C, \quad (17)$$

and $a_{g,1}^*(\lambda)$ and $a_{g,2}^*(\lambda)$ represent the chlorophyll-specific absorption coefficient by CDOM of assemblage 1 and 2 respectively. Eq. (17) was fitted using the 157 corresponding $a_g(\lambda)$ and C measurements from the Tara Expedition to derive $a_{g,1}^*(\lambda)$ and $a_{g,2}^*(\lambda)$ using the same fitting procedure as that used in Eq. (12) and setting C_1^m and S_1 according to Brotas et al. (2013) (Table 1). Statistically, the performance of Eq. (17) was similar to that of a traditional power function fitted to the same dataset (Z-test at 443 nm, $p = 0.95$; Cohen & Cohen, 1983). Since C_1^m and S_1 were prescribed (Brotas et al., 2013), there are two unknown parameters in Eq. (17), which is the same as the number of parameters required to fit a power function. Given the exponential function (Eq. 16) that $a_g(\lambda)$ follows, Eq. (17) can be expressed according to

$$a_g(\lambda) = a_{g,1}^*(\lambda_0) \exp[-S_{g,1}(\lambda - \lambda_0)] \{C_1^m [1 - \exp(-S_1 C)]\} + a_{g,2}^*(\lambda_0) \exp[-S_{g,2}(\lambda - \lambda_0)] \{C - C_1^m [1 - \exp(-S_1 C)]\}, \quad (18)$$

where $S_{g,1}$ and $S_{g,2}$ represent the exponential slopes of CDOM for each assemblage, and $\lambda_0 = 443 \text{ nm}$. Having obtained the values of $a_{g,1}^*(\lambda)$ and $a_{g,2}^*(\lambda)$ at every wavelength from fitting Eq. (17), the parameters $S_{g,1}$ and $S_{g,2}$ were then obtained by fitting the exponential function to $a_{g,1}^*(\lambda)$ and $a_{g,2}^*(\lambda)$. The retrieved parameters for Eq. (18) are provided in Table 1 together with their 95% confidence intervals and are plotted in Fig. 3c.

From the resulting parameterisation of Eq. (18), waters where assemblage 2 reside are seen to have higher CDOM absorption for a given chlorophyll concentration, at most wavelengths, than waters where assemblage 1 reside, and a lower exponential slope (Fig. 2c). The model is seen to capture the trend in $a_g(443)$ with increasing chlorophyll (Fig. 3f, see also Table 2 for statistical tests). However, the Red Sea model estimates a significantly higher amount of $a_g(443)$ for a given chlorophyll concentration (Fig. 3f) when compared with standard global relationships (Morel, 2009; Morel & Gentili, 2009b).

3.5. Remote sensing reflectance (R_{rs}) model

Once the IOPs of a_p , a_g and b_{bp} are known, they can be combined with pure seawater IOPs to estimate total absorption (a) and backscattering (b_b), such that

$$a(\lambda) = a_p(\lambda) + a_g(\lambda) + a_w(\lambda), \quad (19)$$

and

$$b_b(\lambda) = b_{bp}(\lambda) + b_{bw}(\lambda), \quad (20)$$

where $a_w(\lambda)$ and $b_{bw}(\lambda)$ are pure water IOPs for absorption and backscattering respectively. For the Red Sea model, pure water a_w from Pope and Fry (1997) and b_{bw} from Zhang and Hu (2009) and Zhang, Hu, and He (2009) were used. The b_{bw} values were computed assuming a salinity of 40 psu and a water temperature of 27°C , typical of the northern oligotrophic waters of the Red Sea (Raitos et al., 2013;

Triantafyllou et al., 2014; Yao, Hoteit, Pratt, Bower, Zhai, et al., 2014) where b_{bw} has a high contribution to b_b .

Various methods have been proposed to relate $a(\lambda)$ and $b_b(\lambda)$ to $R_{rs}(\lambda)$ (e.g. Gordon et al., 1988; Lee, Carder, & Arnone, 2002). Here we used the approach of Lee, Lubac, Werdell, and Arnone (2009), Lee, Arnone, Hu, Werdell, and Lubac (2010), Lee, Shang, et al. (2010), Lee et al. (2013) that explicitly separates the phase-function effects of molecular and particle scattering, such that

$$R_{rs}(\lambda, \Omega) = \left(G_0^w(\Omega) + G_1^w(\Omega) \frac{b_{bw}(\lambda)}{a(\lambda) + b_b(\lambda)} \right) \frac{b_{bw}(\lambda)}{a(\lambda) + b_b(\lambda)} + \left(G_0^p(\Omega) + G_1^p(\Omega) \frac{b_{bp}(\lambda)}{a(\lambda) + b_b(\lambda)} \right) \frac{b_{bp}(\lambda)}{a(\lambda) + b_b(\lambda)}, \quad (21)$$

where parameters $G_0^w(\Omega)$, $G_1^w(\Omega)$, $G_0^p(\Omega)$, and $G_1^p(\Omega)$ were derived from Hydrolight simulations (Lee et al., 2009) for various Sun angles and viewing geometries, with Ω collectively representing these geometries. Parameters are provided in Table 1 for a solar zenith angle in air equal to zero, sensor nadir-view angle in air equal to zero, and a sensor azimuth angle in relation to the solar plane equal to zero (i.e. fully-normalised). Eqs. (12), (14), (18), (19), (20) and (21) can be used together with model parameters in Table 1 and pure water IOPS (Pope & Fry, 1997; Zhang & Hu, 2009; Zhang et al., 2009) to reconstruct $R_{rs}(\lambda)$ as a function of chlorophyll (C).

4. Results

4.1. Performance of the Red Sea ocean-colour model

The Red Sea ocean-colour model was first evaluated using the 185 OC-CCI satellite R_{rs} and chlorophyll match-ups during Tara. Using *in situ* chlorophyll, R_{rs} from 412 to 555 nm was reconstructed using the model (Eqs. 12, 14, 18, 19, 20 and 21). Reconstructed R_{rs} are compared with observed R_{rs} from satellite data (OC-CCI) in Fig. 4. The top row of plots show the reconstructed and observed absolute values of R_{rs} . Between 412 and 510 nm there is good agreement between model and observations, with correlation coefficients (r) exceeding 0.8 and low errors ($\Psi < 0.001 \text{ sr}^{-1}$, $\Delta < 0.0009 \text{ sr}^{-1}$). There is a slight tendency for the model to underestimate $R_{rs}(412)$ at larger values (Fig. 4), and for $R_{rs}(555)$ the model is relatively invariant and struggles to reproduce the variability shown in the OC-CCI observations.

Whereas the magnitude of R_{rs} is important for deriving IOPs, it is the shape of the R_{rs} (the spectral values of R_{rs} normalised to the corresponding value at a single wavelength) that is particularly sensitive to changes in chlorophyll (O'Reilly et al., 2000). The bottom row of plots in Fig. 4 compares the R_{rs} from the model, normalised at 555 nm, with the corresponding R_{rs} from OC-CCI. In general, there is good agreement between model and observations at all wavelengths, with high correlation coefficients ($r > 0.94$), low errors ($\Psi < 0.49$, $\Delta < 0.45$) and biases (δ) close to zero. Similar conclusions can be drawn when comparing the Colour Index (ξ , see Eq. 6) reconstructed from the model with the observed ξ from OC-CCI (Fig. 4).

We also conducted a similar comparison on *in situ* spectra collected using the HyperPro at two stations in the Red Sea (see Supplementary Fig. S2). Results resonate with the comparison in Fig. 4, with the model underestimating at high values of R_{rs} in the blue region of the spectrum (found at both stations), and in better agreement with the shape of the *in situ* R_{rs} (normalised at 555 nm).

4.2. Performance of satellite chlorophyll algorithms

In general, all four algorithms performed reasonably at estimating chlorophyll when compared with *in situ* values ($r > 0.60$ and $\Psi < 0.29$, Fig. 5). According to the objective classification, two empirical algorithms (OC4 and OC1) have the best overall performance (highest points score in the bar chart in Fig. 5). This is mainly due to a combination of a

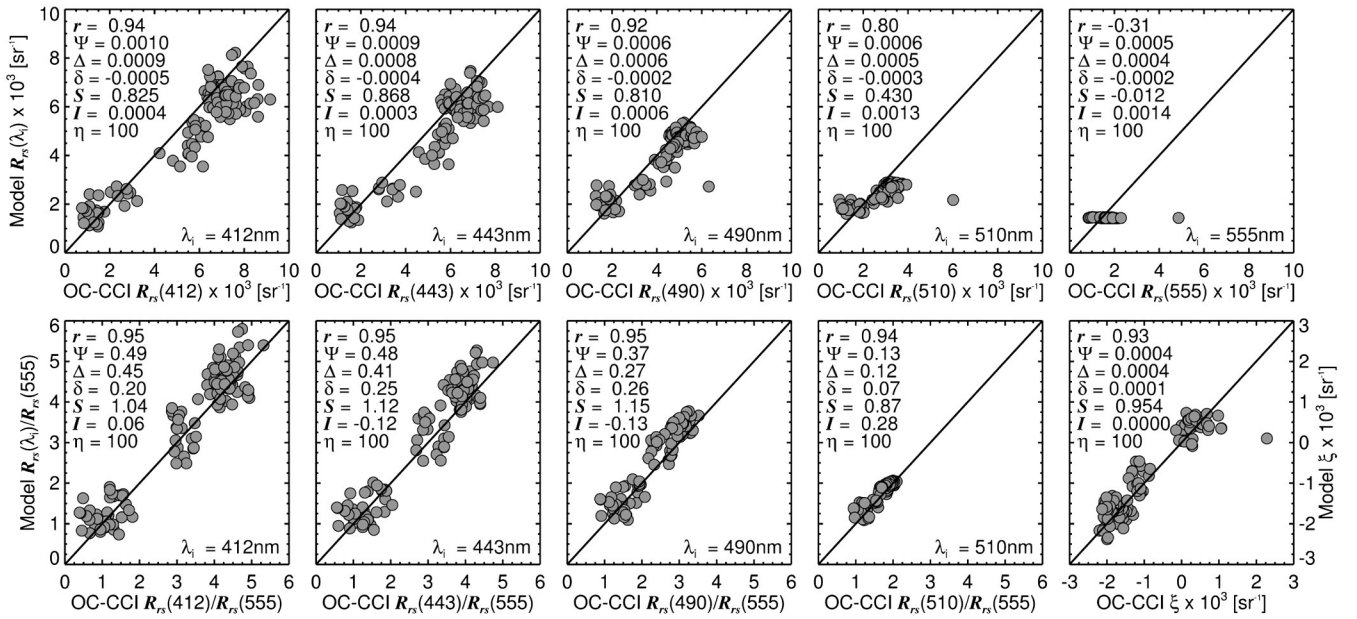


Fig. 4. Comparison of reconstructed remote sensing reflectance (R_{rs}) using *in situ* chlorophyll (C) and OC-CCI observed R_{rs} . The top row of figures show the absolute values of R_{rs} and the bottom row the shape of R_{rs} , normalised at 555 nm, and the Colour Index (ξ). Black line refers to 1:1 line.

higher correlation coefficient (r) for these two algorithms when compared with GSM and OC4-MG, a lower centre-pattern (or unbiased) root mean square error (Δ), a Type-2 regression slope (S) close to one, the maximum percentage of possible retrievals (η), and high confidence in the retrieved bias (δ) and intercept (I). However, despite the better performance of OC4 and OCI according to the classification, both algorithms systematically overestimate chlorophyll, as indexed by a positive bias significantly different from zero (δ , Fig. 5), which was not observed for the GSM and OC4-MG algorithms.

4.3. Tuning of empirical satellite chlorophyll algorithms

Consistent with previous ocean-colour models (e.g. Gordon et al., 1988; Morel, 1988; Morel & Maritorena, 2001; Sathyendranath et al., 2001) as the chlorophyll concentration decreases the Red Sea ocean-colour model predicts an increase in R_{rs} at blue wavelengths and a slight decrease in green and red wavelengths (Fig. 6a). Fig. 6b shows a plot of chlorophyll (C) as a function of the maximum band-ratio ($R_{rs}(443 > 490 > 510)/R_{rs}(555)$) using output from the Red Sea ocean-colour model with the OC4 algorithm overlain. For a given maximum band-ratio of R_{rs} , the Red Sea ocean-colour model predicts significantly lower chlorophyll than the OC4 algorithm. Similarly, when we plot chlorophyll (C) as a function of the Colour Index (ξ) using output from the Red Sea ocean-colour model (Fig. 6c) with the relationship used in the OCI algorithm overlain, we observe that the Red Sea ocean-colour model predicts a higher ξ than the OCI algorithm at chlorophyll concentrations less than 0.3 mg m^{-3} .

To investigate if the bias observed in the OC4 and OCI algorithms (Fig. 5) can be explained using the Red Sea ocean-colour model, we reparameterised the OC4 and OCI algorithms using output from the Red Sea ocean-colour model (hereafter denoted OC4-RG and OCI-RG respectively, where RG refers to regionally-tuned). The Red Sea ocean-colour model was run for 2560 logarithmically-space bins between 0.01 and 10.0 mg m^{-3} chlorophyll to derive the corresponding $R_{rs}(\lambda)$ values.

4.3.1. OC4-RG

For the OC4-RG algorithm, Eq. (5) was fitted to output from the Red Sea model and the tuned parameters are provided in Table 3. As the Red Sea ocean-colour model outputs at a variety of wavelengths

representative of recent and future ocean-colour sensors (NASA, 2010), parameters for standard band-ratio algorithms that function at slightly different wavelengths (for application to MODIS, MERIS, OLCI and VIIRS) are also provided in Table 4. Note that for these algorithms Eq. (4) is modified depending on the R_{rs} ratio provided in Table 3, and Eq. (5) is then used along with the corresponding values for q_0, q_1, q_2, q_3 and q_4 provided in Table 3. The OC4-RG algorithm is overlain in Fig. 6b and matches the relationship between chlorophyll (C) and the maximum band-ratio with which it was parameterised (Red Sea forward model).

4.3.2. OCI-RG

Fig. 6c shows that for ξ values between -0.005 and -0.001 , the relationship between \log_{10} chlorophyll (C) and ξ in the Red Sea forward model is linear, however, above -0.001 ξ this relationship breaks down. Using radiative transfer modelling, Hu et al. (2012) showed that above a ξ value of -0.0005 sr^{-1} , assuming a linear relationship between \log_{10} chlorophyll (C) and ξ results in a significant underestimation in chlorophyll. For the Red Sea, the forward model predicts that the ξ value above which the linear relationship breaks down is slightly lower at around -0.001 sr^{-1} (Fig. 6c). Therefore, Eq. (7) was tuned using output from the Red Sea forward model for ξ values lower than -0.001 sr^{-1} , and the resulting parameters are provided in Table 4. In comparison to the original parameters of Hu et al. (2012), where $A = -0.491$ and $B = 191.66$, there was a significant decrease in A (-0.802) and a slight increase in B (197.74). A reduction in the A parameter is consistent with the results of Brewin et al. (2013), using a dataset entirely independent of any measurements used to parameterise the Red Sea ocean-colour model.

As the ξ value at which the linear relationship between \log_{10} chlorophyll (C) and ξ breaks down is slightly lower for the Red Sea (Fig. 6c), the boundaries at which the algorithm switches from a band-difference (ξ) to a band-ratio (OC4-RG) were modified accordingly, such that the OCI-RG algorithm is expressed as

$$C = \begin{cases} 10^{A+B\xi} & \text{if } [10^{A+B\xi}] \leq 0.10 \text{ mg m}^{-3} \\ \alpha [10^{q_0+q_1X+q_2X^2+q_3X^3+q_4X^4}] + (1-\alpha) [10^{A+B\xi}] & \text{if } 0.10 < [10^{A+B\xi}] \leq 0.15 \text{ mg m}^{-3} \\ 10^{(q_0+q_1X+q_2X^2+q_3X^3+q_4X^4)} & \text{if } [10^{A+B\xi}] > 0.15 \text{ mg m}^{-3}, \end{cases} \quad (22)$$

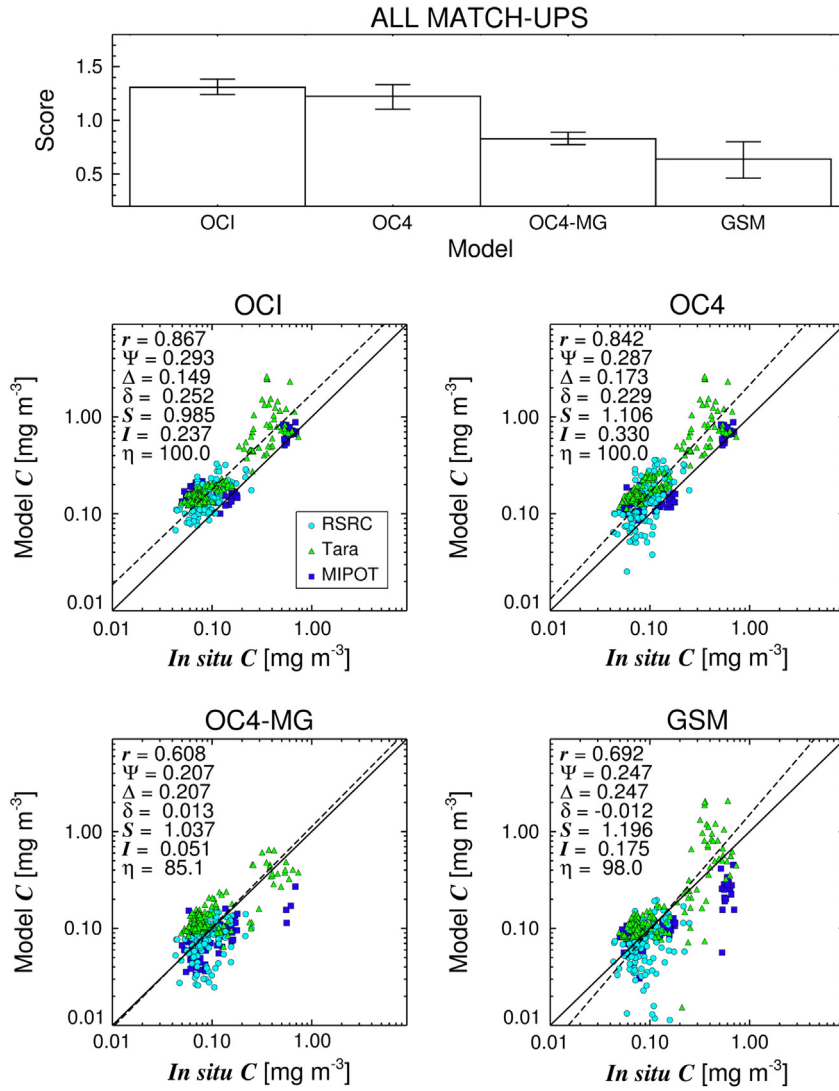


Fig. 5. Results from the chlorophyll (C) model comparison. The bar chart at the top shows the results from the point classification and scatter plots of modelled and *in situ* chlorophyll are shown below for the OCI, OC4, OC4-MG and GSM. Black line refers to 1:1 line and dashed line Type-2 regression.

where $\alpha = (10^A + B\xi - 0.10)/(0.15 - 0.10)$. The parameters for A and B are provided in Table 4 and the corresponding values for q_0 , q_1 , q_2 , q_3 and q_4 are provided in Table 3. The model parameters for wavelengths relevant to MODIS, MERIS, OLCI and VIIRS are also provided in Table 4. Note that for implementation of these algorithms, Eqs. (6) and (22) are used with the R_{rs} band-difference (Table 4), values for A and B (Table 4), and the band-ratio algorithm (Table 3) that are appropriate

for the sensor used. The linear relationship between \log_{10} chlorophyll and ξ tuned to the Red Sea (denoted ξ_{RG}), is laid over the forward model and the ξ model of Hu et al. (2012). We see that the forward model and the Red-Sea-tuned ξ model match each other very well for ξ values less than -0.001 sr^{-1} , for which it was parameterised (Fig. 6c).

To verify further that the OC4-RG and OCI-RG algorithms reproduce the relationships between R_{rs} and chlorophyll in the Red Sea model, we

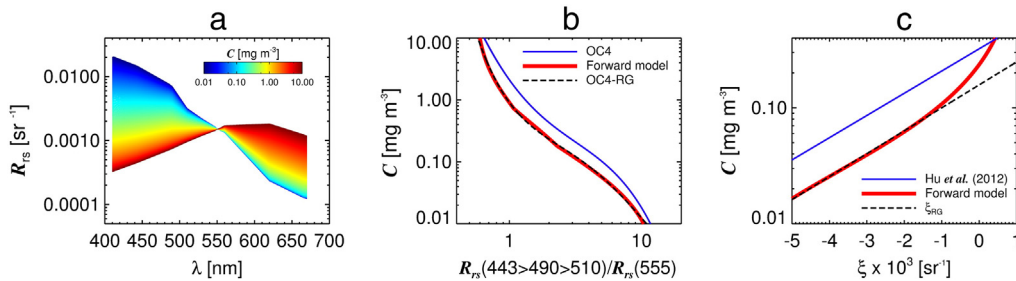


Fig. 6. (a) $R_{rs}(\lambda)$ plotted as a function of wavelength for a variety of chlorophyll concentrations (C) using the Red Sea ocean-colour model. (b) Shows C plotted as a function of the maximum band-ratio for the OC4 model, output from the Red Sea ocean-colour model (forward model), and a tuned OC4 model, denoted OC4-RG, tuned using output from the Red Sea ocean-colour model. (c) Shows C plotted as a function of the Colour Index (ξ) using the parameterisation of Hu et al. (2012), output from the Red Sea ocean-colour model (forward model), and a tuned ξ relationship, denoted ξ_{RG} , tuned using output from the Red Sea ocean-colour model below a ξ value of -0.001 sr^{-1} .

Table 3

Values of the coefficients for standard, empirical band-ratio chlorophyll algorithms tuned to the Red Sea ocean-colour model for a suite of ocean-colour sensor wavelengths.

Sensor/dataset	Identifier ^a	R_{rs} ratio	q_0	q_1	q_2	q_3	q_4
SeaWiFS/OC-CCI	OC4-RG	$\max(443, 490, 510)/555$	−0.0381	−2.9297	4.6447	−5.5384	1.9556
MERIS/OLCI	OC4ME-RG	$\max(443, 490, 510)/560$	−0.0472	−2.5860	3.4994	−3.9545	1.2466
MODIS	OC3MO-RG	$\max(443, 488)/547$	−0.1333	−2.4079	2.7585	−3.4081	1.1122
VIIRS	OC3VI-RG	$\max(443, 486)/551$	−0.1307	−2.1605	2.1482	−2.6768	0.8301

^a ME = MERIS, MO = MODIS and VI = VIIRS. The reader is also referred to NASA (2010) for more information on the algorithms. RG refers to a tuning of the algorithm to the Red Sea.

conducted a sensitivity analysis (Supplementary Fig. S3). For chlorophyll concentrations ranging from 0.01 to 10.0 mg m^{−3} we used the Red Sea ocean-colour model to estimate R_{rs} and then applied the OC4-RG and OCI-RG algorithms to the R_{rs} to compute chlorophyll concentration. If both algorithms were parameterised correctly, initial chlorophyll should agree with chlorophyll derived from the empirical models, as confirmed by the results in Supplementary Fig. S3 for both OC4-RG and OCI-RG. We also introduced random noise into the R_{rs} values prior to applying OC4-RG and OCI-RG algorithms, whereby for a given R_{rs} value, wavelength independent random noise between $\pm 10\%$ (or $\pm 20\%$) was applied to the R_{rs} estimated from the Red-Sea ocean-colour model, then this spectrum was used to estimate chlorophyll using OC4-RG and OCI-RG and compared with initial chlorophyll values. Results are shown in Supplementary Fig. S3 and indicate both OC4-RG and OCI-RG cope well with the addition of random noise in R_{rs} up to 20%.

4.4. Performance of tuned empirical satellite chlorophyll algorithms

The two tuned algorithms perform significantly better, as indexed by higher scores than for the other algorithms (Fig. 7). The systematic overestimation in chlorophyll observed in the untuned OC4 and OCI algorithms is no longer apparent in the tuned algorithms (Fig. 7). The Red Sea ocean-colour model used to parameterise OC4-RG and OCI-RG relates changes in a_p , b_{bp} and a_g to changes in chlorophyll concentration. The relationships between a_p and chlorophyll and the relationships between b_{bp} and chlorophyll from the Tara dataset (Fig. 3) were found to be similar to those observed using globally representative datasets (Brewin, Dall'Olmo, et al. 2012; Bricaud et al., 1998). However, the relationship between a_g and chlorophyll (Fig. 3f) was significantly different, with the Red Sea model requiring a higher amount of a_g for a given chlorophyll concentration when compared with standard global relationships (Morel, 2009; Morel & Gentili, 2009b), possibly suggesting an excess of CDOM absorption per unit chlorophyll in the Red Sea as the possible cause of the systematic overestimation in chlorophyll observed in the original OC4 and OCI algorithms.

To test this hypothesis, we replaced Eq. (18) with the a_g model of Morel (2009) in our Red Sea model (see Fig. 3f) and re-parameterised the OC4 algorithm (Eq. 5, denoted OC4-RG-M09) where $q_0 = 0.4010$, $q_1 = -2.9973$, $q_2 = 3.6843$, $q_3 = -4.6653$ and $q_4 = 1.6263$. Fig. 8 shows a comparison of scatter plots of modelled and *in situ* chlorophyll for OC4, OC4-RG-M09 and OC4-RG. The systematic overestimation in chlorophyll observed in OC4 (positive bias (δ)) is also observed with OC4-RG-M09, emphasising that a higher-than-average CDOM absorption per unit chlorophyll in the Red Sea ocean-colour model (Eq. 18)

appears to explain why the bias (δ) is closer to zero for OC4-RG and why it performs better in the algorithm evaluation than OC4 (Fig. 7).

4.5. Application to OC-CCI satellite images

The OC4, OCI, OC4-RG and OCI-RG algorithms were run on daily OC-CCI data for the year 2010 and used to produce monthly composites (after log₁₀ transformation), which were then used to produce annual chlorophyll composites (Fig. 9). The spatial patterns in the 2010 annual chlorophyll composites are consistent with the description of Raitso et al. (2013) using MODIS-Aqua data. The northern Red Sea is the most oligotrophic with higher chlorophyll values in the southern regions, and around coral reef-bound coastal waters (Fig. 1a). The log₁₀ differences between annual chlorophyll composites of OC4 and OC4-RG, and OCI and OCI-RG, are also plotted in Fig. 9. Differences between empirical algorithms are roughly systematic (positive bias) over the entire chlorophyll range, and similar for both the band-ratio (OC4 and OC4-RG) and band-difference (OCI and OCI-RG) algorithms. Differences are slightly larger at higher chlorophyll concentrations, around shallow regions (<200 m, see Fig. 1a) in the southern Red Sea and near coral reefs. In very shallow waters, estimates of chlorophyll using empirical algorithms have high uncertainty, due to the potential influence of bottom reflectance on the shape of the R_{rs} spectrum. In more oligotrophic waters, such as the northern Red Sea, differences are slightly larger between the OCI and OCI-RG algorithms, when compared with differences between OC4 and OC4-RG.

Fig. 10 shows two OC-CCI daily images of chlorophyll concentration for 11th January and 31st May 2010 produced using the OCI-RG algorithm. The percentage contribution of assemblage 1 and 2 (denoted A_1 and A_2 respectively) are also plotted for each daily image, by using the OCI-RG algorithm (Eq. 22) and Eqs. (9) and (10). Note that the use of Eqs. (9) and (10) with OCI-RG derived chlorophyll is consistent with the framework of the Red Sea model, considering OCI-RG was parameterised using the Red Sea model in which Eqs. (9) and (10) are embed explicitly. The 11th January 2010 shows a clear OC-CCI image during the period of Tara. The geographical distribution of chlorophyll on the 11th January 2010 is similar to the annual composites shown in Fig. 9. The percentage contribution of assemblage 1 and assemblage 2 to total chlorophyll indicates assemblage 1 dominates the northern waters of the Red Sea and assemblage 2 the southern waters (Fig. 10). However, comparison with the 31st May 2010 illustrates just how dynamic the patterns in chlorophyll concentration can be in the Red Sea. During this day, blooms of phytoplankton, dominated by assemblage 2 (P_2) and likely influenced by enhanced meso-scale eddy activity during this period (Acker et al., 2008; Raitso et al., 2013; Zhan et al., 2014), or

Table 4

Values of the coefficients for the band-difference algorithm of Hu et al. (2012) tuned to the Red Sea ocean-colour model for a suite of ocean-colour sensor wavelengths.

Sensor/dataset	Identifier ^a	R_{rs} band-difference (ξ)	A	B
SeaWiFS/OC-CCI	OCI-RG	$R_{rs}(555) - 0.50(R_{rs}(443) + R_{rs}(670))$	−0.8021	197.7366
MERIS/OLCI	OCIME-RG	$R_{rs}(560) - 0.53(R_{rs}(443) + R_{rs}(665))$	−0.7625	188.2083
MODIS	OCIMO-RG	$R_{rs}(547) - 0.46(R_{rs}(443) + R_{rs}(670))$	−0.8843	212.5575
VIIRS	OCIVI-RG	$R_{rs}(551) - 0.48(R_{rs}(443) + R_{rs}(670))$	−0.8417	204.7011

Note that for implementation of Eq. (22), $\alpha = (10^A + B\xi - 0.10)/(0.15 - 0.10)$.^a ME = MERIS, MO = MODIS and VI = VIIRS. RG refers to a tuning of the algorithm to the Red Sea.

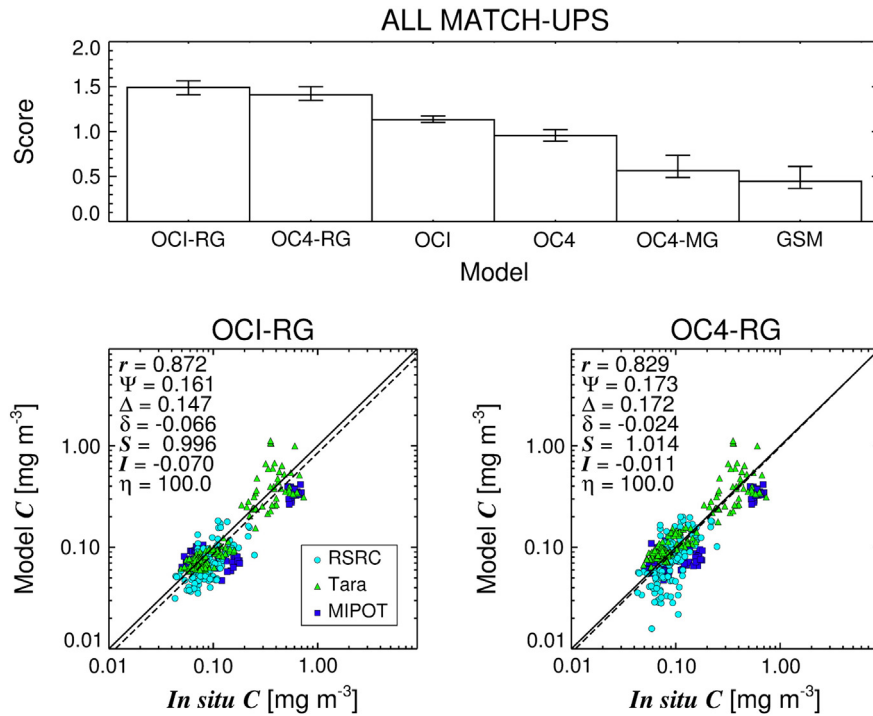


Fig. 7. Results from the chlorophyll (C) model comparison when including the two regionally tuned empirical algorithms (OC4-RG and OCI-RG). The bar chart at the top shows the results from the point classification and scatter plots of modelled and *in situ* chlorophyll are shown below for the OC4-RG and OCI-RG. Black line refers to 1:1 line and dashed line Type-2 regression.

wind-driven upwelling and downwelling (Labiosa et al., 2003), swirl around the northern half of the Red Sea (Fig. 10).

Seasonal climatologies of chlorophyll concentration and the percentage contribution of assemblage 1 (A_1) and assemblage 2 (A_2) to total chlorophyll, processed using the OCI-RG algorithm over the entire OC-CCI time-series (1997–2012), are shown in Fig. 11 to illustrate the typical seasonal succession. The temporal variations in chlorophyll in open ocean regions of the Red Sea are consistent with those described by Raitos et al. (2013), with higher concentrations during the winter and lower concentrations during the summer. However, coral reef-bound coastal waters (Fig. 1a) display equal or higher chlorophyll during the summer period relative to that in winter (Racault et al., 2015). Over the majority of the northern Red Sea, assemblage 1 (A_1) contributes highly to the chlorophyll concentration in summer and autumn with a lower contribution during spring and winter. The southern Red Sea and coral reef-bound coastal waters are dominated by assemblage 2 (A_2) all year around, with coral reef-bound coastal waters having a lower contribution of A_2 in spring, and open-ocean waters having a lower contribution of A_2 in autumn.

5. Discussion

5.1. Uncertainties in the analysis

5.1.1. *In situ* chlorophyll data

Chlorophyll data used in the study came from three different sources, *in vivo* fluorescence, Lidar fluorescence, and particulate absorption line height. Each of these methods are subject to uncertainties. *In vivo* fluorescence in surface waters can be affected by daytime-fluorescence quenching (Cullen & Lewis, 1995). The fluorescence yield can also vary among species (Kiefer, 1973b; Strickland, 1968) and within a single species, subjected to different environmental conditions (Kiefer, 1973a; Slovacek & Bannister, 1973). Using the same *in vivo* fluorescence data, Brewin et al. (2013) observed a positive bias in the OCI algorithm using MODIS-Aqua, in both high- and low-light samples, suggesting that the positive bias in OC4 and OCI observed in the *in vivo* fluorescence data in this study (Fig. 5) was not related to daytime-fluorescence quenching. Nonetheless, despite the fluorometer being laboratory calibrated prior to each cruise, it was not field calibrated as

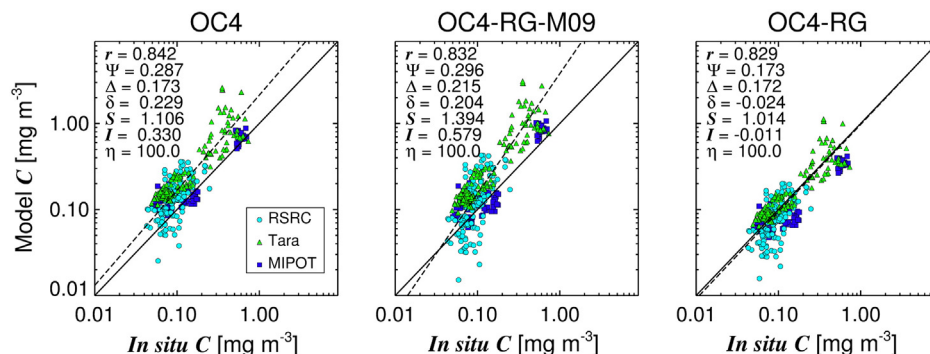


Fig. 8. Scatter plots of modelled and *in situ* chlorophyll for OC4, OC4-RG-M09 and OC4-RG. Black line refers to 1:1 line and dashed line Type-2 regression.

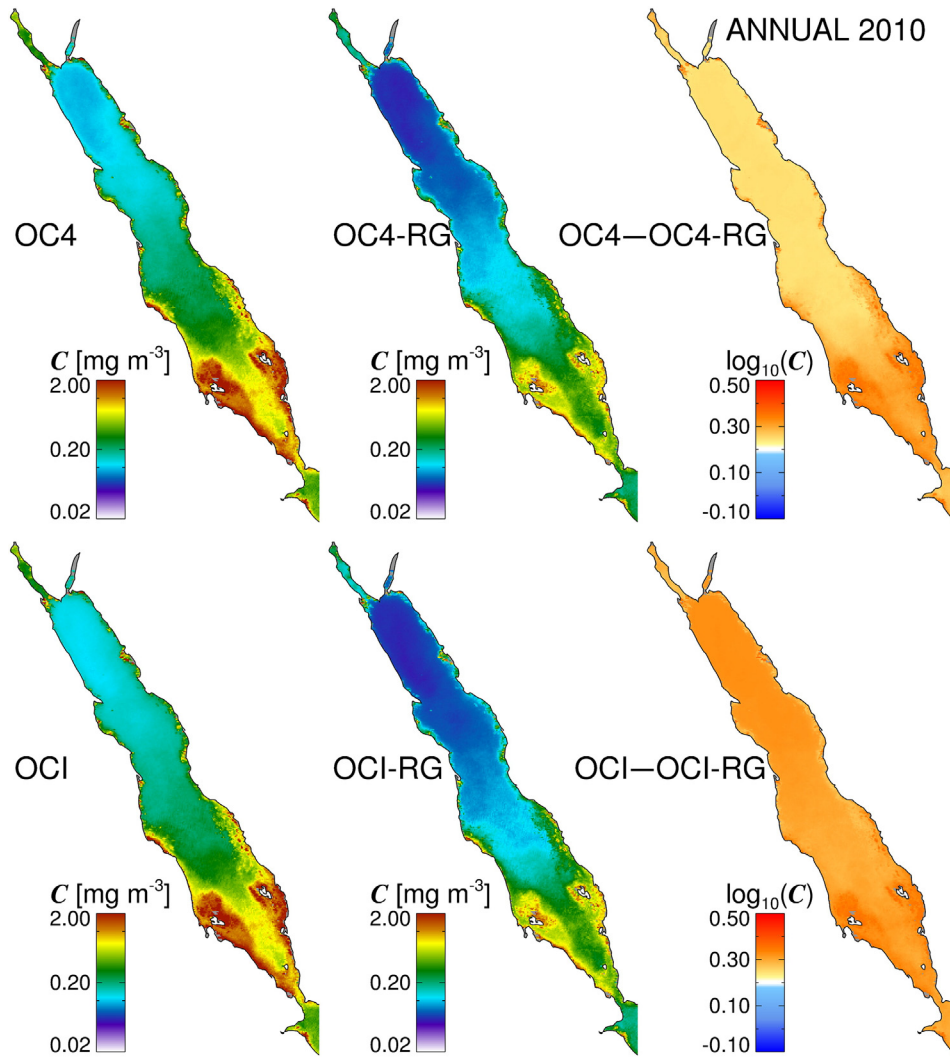


Fig. 9. Annual OC-CCI Red Sea chlorophyll (C) composites for 2010 estimated from the OC4, OCI, OC4-RG and OCI-RG algorithms. The \log_{10} differences between annual chlorophyll composites of OC4 and OC4-RG, and OCI and OCI-RG, are also plotted.

no independent measurements of chlorophyll (e.g. from HPLC or from *in vitro* fluorometry) were taken during the three RSRC cruises.

Lidar fluorescence is also subject to similar issues as *in vivo* fluorescence. The method assumes a linear regime for the laser excitation and a low chromophore density for all the species present, accurate calibration of fluorescence signal against concurrent Raman signal from water, and accurate calibration of Raman units to chlorophyll concentration (Barbini, Colao, Fantoni, Fiorani et al., 2001). During the MIPOT oceanographic campaign, conventional analyses were performed according to the spectrofluorometric technique and the fluorescence-to-Raman ratio and absolute concentrations were found to be well correlated (Barbini et al., 2004), lending some confidence to the MIPOT data.

Of the three chlorophyll methods, the use of absorption line height to estimate chlorophyll concentration is likely to be the most reliable. It is relatively insensitive to instrument drift, incident irradiance and non-photochemical quenching (Roesler & Barnard, 2013). The method has also been found to perform well when compared with discrete *in situ* HPLC chlorophyll data and diverse phytoplankton cultures (Boss et al., 2013; Chase et al., 2013; Dall'Olmo et al., 2009, 2012; Roesler & Barnard, 2013; Westberry et al., 2010). Nonetheless the method is still subject to uncertainties: for instance, Eq. (2) was calibrated using concurrent HPLC data and $a_p(\lambda)$ taken during the entire Tara Oceanographic campaign in a wide range of bio-optical environments over the global ocean. Measurements of HPLC were taken at 4 m depth (similar to

underway sampling (~2 m depth)) at four stations in the Red Sea. When comparing chlorophyll from HPLC with chlorophyll estimated from the AC-S (Eqs. 1 and 2, averaged over a 30 minute window when the HPLC data were collected), one measurement was in reasonable agreement, but the other three observations showed higher chlorophyll from HPLC ($\delta \sim 0.3$), similar to the average bias between OCI and *in situ* chlorophyll (Fig. 5). These three HPLC estimates were not only higher than those estimated from the AC-S, but also higher than: (i) chlorophyll estimated using a fluorometer (pre-calibrated with HPLC data) mounted to a CTD at the three stations and (ii) chlorophyll estimated from a suite of inversion algorithms applied to the *in situ* R_{rs} data at two of the three stations (data not shown). Furthermore, when dividing HPLC chlorophyll by $a_{ph}(676)$ derived from the AC-S, to compute the chlorophyll-specific phytoplankton absorption coefficient at 676 nm at the three stations, the average value was $0.0086 \text{ m}^2 [\text{mg C}]^{-1}$, unusually low for an oligotrophic to mesotrophic environment (Bricaud, Babin, Morel, & Claustre, 1995). Conclusions cannot be drawn from only four HPLC observations, which themselves are subject to uncertainty (Claustre et al., 2004), but the result emphasises the importance of taking concurrent observations of chlorophyll from difference sources, to reduce ambiguity and quantify uncertainty.

When considering the uncertainties in all three sources of *in situ* chlorophyll data, one may question whether standard algorithms do in fact systematically overestimate chlorophyll in the Red Sea, and

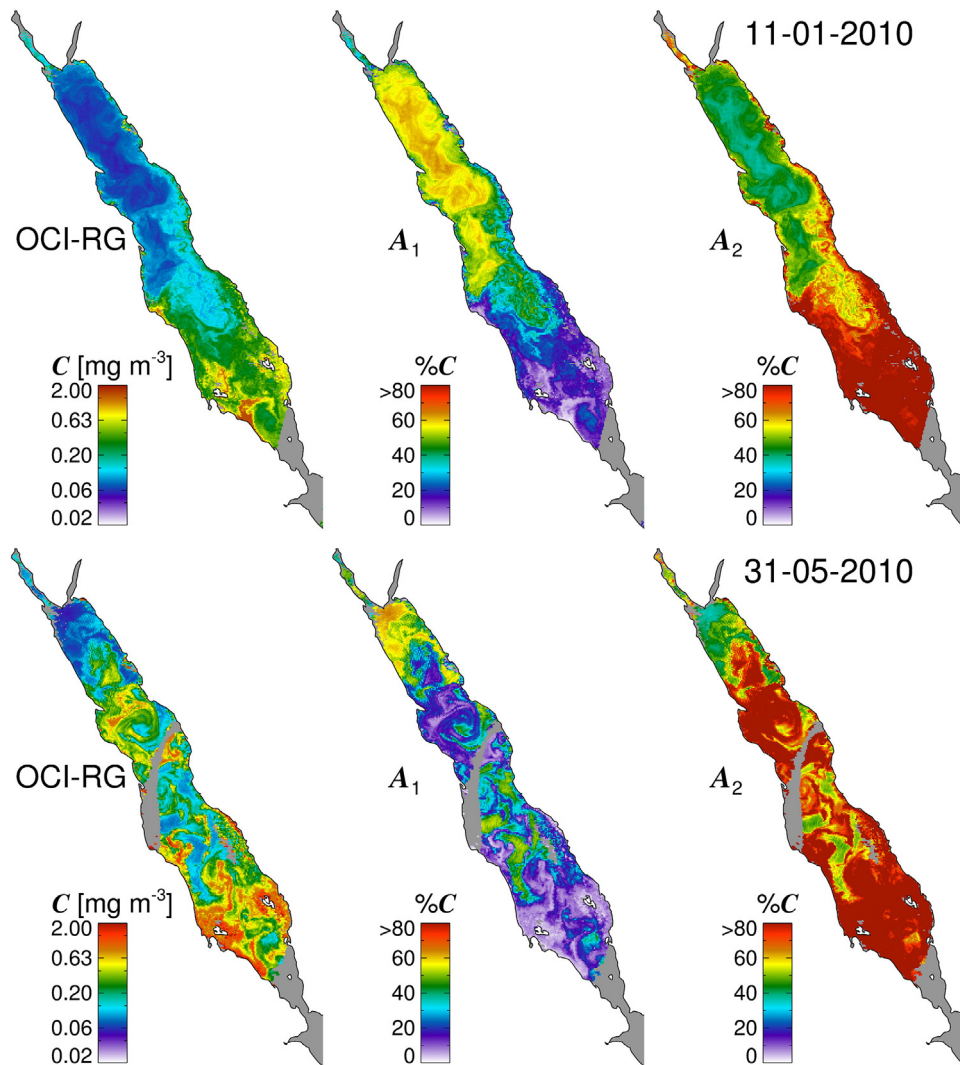


Fig. 10. Two OC-CCI daily images of chlorophyll concentration for 11th January and 31st May 2010 produced using the OCI-RG algorithm. The percentage contribution of assemblage 1 (A_1) and assemblage 2 (A_2) are also plotted for each daily image, using the OCI-RG algorithm and Eqs. (9) and (10).

whether this overestimation is instead an artefact of uncertainty in the *in situ* data itself. At this stage it is worth emphasising that the systematic overestimate of OC4 and OCI was observed individually for all three sources of *in situ* chlorophyll data (Fig. 5), thus lending support to their reliability and to the tuning of the algorithms (Fig. 7). To further illustrate this point, and to emphasise the consistency and inter-compatibility of the three *in situ* datasets, we created a satellite OC4 composite of chlorophyll (OC-CCI data) over the duration of each cruise, and compared satellite chlorophyll estimates with *in situ* data spatially (see Supplementary Fig. S4). Results show that: i) there is good agreement in spatial variability between *in situ* data and satellite chlorophyll for all three datasets; and ii) there is a consistent overestimate (positive bias) in satellite OC4 chlorophyll across all three *in situ* datasets.

Labiosa et al. (2003) compared SeaWiFS-derived chlorophyll concentrations (OC4v4) with *in situ* chlorophyll, derived fluorometrically (*in vitro*), at Eilat located at the northern tip of the Gulf of Aqaba, at a pier close to the shore (Genin, Lazar, & Brenner, 1995). Consistent with our findings, they observed that standard SeaWiFS algorithms overestimated chlorophyll at high concentrations (bloom periods). However, at lower concentrations they observed a slight underestimation of SeaWiFS derived chlorophyll. This underestimation was likely related to differences in the location of the satellite data (just offshore of Eilat) with respect to the *in situ* data (taken from a Pier), considering Labiosa et al. (2003) observed a small positive bias between *in situ*

chlorophyll data collected from the Pier with that 1 km offshore, despite very good correlation (see Fig. 2d of Labiosa et al., 2003). Additional datasets on *in situ* chlorophyll concentration in the Red Sea are clearly required to scrutinise our findings further.

5.1.2. *In situ* measurements of IOPs used for bio-optical modelling

Measurements of particulate absorption (a_p) and attenuation (c_p), taken using the flow-through set-up of Slade et al. (2010), have been well validated over a variety of oceanographic environments (Dall'Olmo et al., 2009, 2011, 2012; Slade et al., 2010; Westberry et al., 2010). Boss et al. (2013) analysed a_p and c_p data over the entire Tara Oceanographic campaign and found the data to be consistent with other published data. Following the method of Werdell et al. (2013), b_{bp} was estimated from a_p and c_p using the method of Twardowski et al. (2001), which assumes the scattering-to-backscattering ratio varies as function of chlorophyll (Eq. 3). Whereas the scattering-to-backscattering ratio has been shown to vary consistently according to trophic conditions (Dall'Olmo et al., 2012; Twardowski et al., 2001; Whitmire et al., 2007), the reported variability around the relationship is high, and it remains to be established whether the Twardowski et al. (2001) relationship holds in the Red Sea. Validation of this approach would require concurrent measurements of a_p , c_p and b_{bp} along with chlorophyll, currently not available for the Red Sea. Direct measurement of b_{bp} and chlorophyll at large scales are clearly

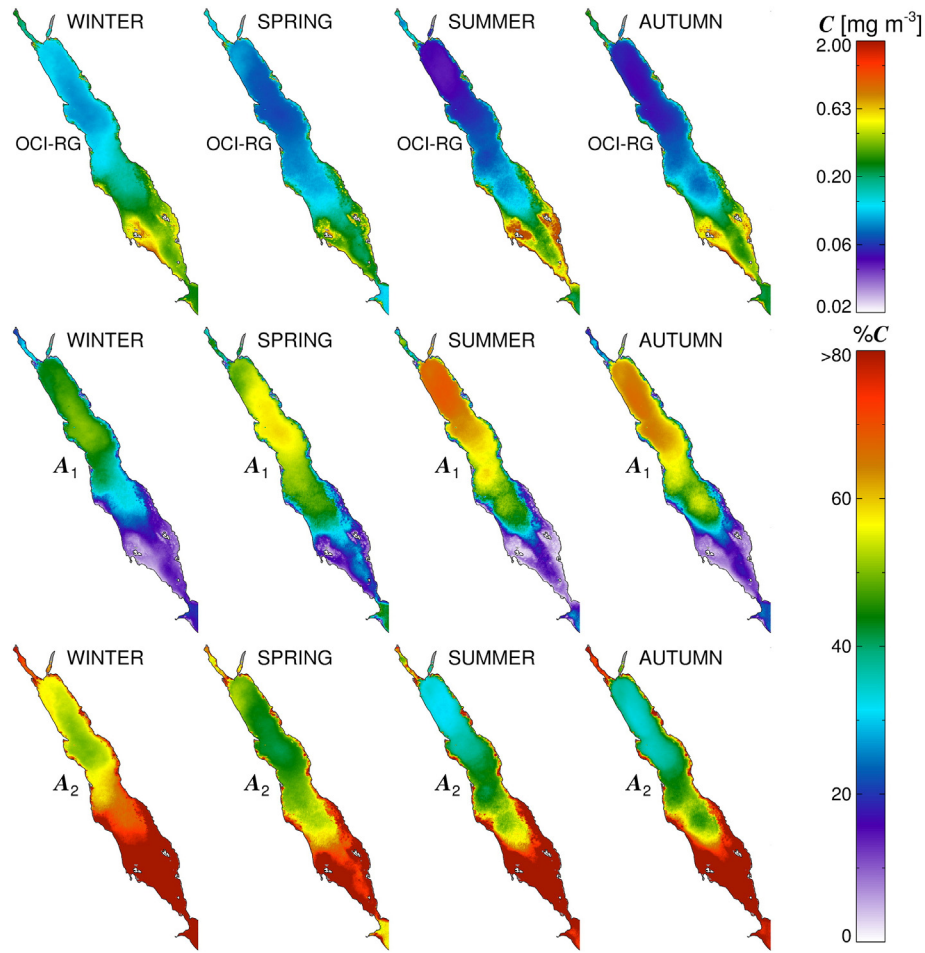


Fig. 11. Seasonal climatologies of chlorophyll concentration (C) and the percentage contribution of assemblage 1 (A_1) and assemblage 2 (A_2), over the entire OC-CCI time-series (1997–2012) processed using the OCI-RG algorithm, for Winter (December, January and February), Spring, (March, April and May), Summer (June, July and August) and Autumn (September, October and November).

preferable, though some comfort can be taken from the good agreement between b_{bp} estimated using Eq. (3) and independent satellite observations of b_{bp} (Werdell et al., 2013). When comparing Tara Red Sea match-ups of \log_{10} -transformed $b_{bp}(443)$ estimated using GSM with \log_{10} -transformed $b_{bp}(443)$ estimated using Eq. (3), we found a low bias ($\delta \sim -0.08$) and root mean square error ($\Psi \sim 0.25$), but the GSM model did struggle to reproduce the variability in $b_{bp}(443)$ primarily as a consequence of a few outliers and a low range of variability in the match-ups.

Considering that no direct measurements of $a_g(\lambda)$ were used in the study, and that $a_g(\lambda)$ was estimated using *in situ* chlorophyll (Eq. 15), the corresponding estimates of ϕ from satellite data (Morel & Gentili, 2009b) and assuming a relationship between $a_g(400)$ and S_g derived from measurements outside Red Sea waters (Eq. 16), uncertainty in the derivation of $a_g(\lambda)$ requires particular attention. The inverse relationship between $a_g(400)$ and S_g used in Eq. (16) is consistent with satellite observations of S_{dg} (where the subscript “dg” refers to combined CDOM and detrital absorption) derived by Bricaud et al. (2012), and values of S_{dg} observed from satellite (between 0.01 and 0.02 nm^{-1} in the Red Sea, see Fig. 1 of Bricaud et al., 2012) are consistent with values of S_g used in this study. Estimates of ϕ , used in Eq. (15), are the primary source of data describing region-specific CDOM absorption content in the Red Sea, and thus are given particular attention in the next section.

5.1.3. Satellite measurements of ϕ

Values of ϕ were estimated from satellite observations (OC-CCI) using the method of Morel and Gentili (2009b). As discussed by Morel

and Gentili (2009b), see also Morel and Gentili (2009a), the quality of retrieved $R_{rs}(412)$ is crucial for accurate estimation of ϕ , and yet estimation of $R_{rs}(412)$ is difficult, and match-ups based on satellite and *in situ* R_{rs} data often find this wavelength to have the highest errors (e.g. Mélin & Zibordi, 2007; Mélin et al., 2011). Retrievals of $R_{rs}(412)$ rely to a certain degree on accurate vicarious calibration, and errors can increase when the sun zenith angle and viewing angle increase (IOCCG, 2010, though sun zenith angle is perhaps not such a problem in the Red Sea, due to its close proximity to the tropics). *In situ* measurements of $R_{rs}(\lambda)$ over large areas are currently not available for the Red Sea, but are required to ascertain potential biases in ϕ , and validate any region-specific bio-optical relationship based on $R_{rs}(\lambda)$. Values of ϕ were computed on the *in situ* spectra collected using the HyperPro at two stations in the Red Sea (see Supplementary Fig. S2). Both stations showed values of ϕ greater than one (1.4 and 1.5), consistent with observations from satellite data. At one station, a corresponding satellite daily match-up of $R_{rs}(\lambda)$ from OC-CCI was available, which was found to be in good agreement with the *in situ* spectrum collected using the HyperPro (see Supplementary Fig. S2c and S2d). When estimating ϕ using both satellite and *in situ* $R_{rs}(\lambda)$ spectra, values of ϕ were greater than one (1.5 and 2.3, respectively) in both cases, suggesting a higher amount of $a_g(443)$ for a given chlorophyll concentration when compared with average global relationships. However, satellite estimates of ϕ were higher (2.3) than *in situ* estimates (1.5).

Additional uncertainties in satellite measurements of ϕ might arise from the unique atmospheric properties of the region. The Red Sea is surrounded by deserts and frequently influenced by dust storms

(Chen et al., 2007; Edwards, 1987), common during spring and summer. Whereas cloud cover is generally low in the region, dust aerosols make atmospheric correction difficult, cause a reduction in ocean-colour retrievals (Acker et al., 2008; Steinmetz et al., 2011) and may systematically influence $R_{rs}(412)$. Desert dust has also been shown to influence the retrieved $R_{rs}(\lambda)$ spectrum, leading to biases in the retrieval of the chlorophyll concentration (Claustre et al., 2002; Moulin, Gordon, Chomko, Banzon, & Evans, 2001). When absorbing aerosols are present in the atmosphere, atmospheric correction in the blue bands becomes less accurate (Moulin et al., 2001). Claustre et al. (2002) demonstrated that Saharan dust deposition can lead to enhanced absorption in the blue and backscattering in the green parts of the visible spectrum, directly resulting in an over-estimation of chlorophyll concentration in oligotrophic waters, and likely, a high ϕ index. The influence of dust on the observed systematic overestimation of chlorophyll by OC4 and OCI (Fig. 5), and on the estimates of ϕ used to tune the Red Sea bio-optical model (Eq. 18), cannot be ruled out.

Whereas (to our knowledge) no direct measurements of $a_g(\lambda)$ are available at large spatial scales in the Red Sea, two recent cruises, conducted by RSRC of KAUST university in March and November 2013, made measurements of coincident fluorometric-CDOM concentration (FCDOM, in ppb) and chlorophyll concentration (*in vivo* fluorescence), using a WET Labs, Inc. Eco Sensor mounted onto a CTD profiler, in the central-northern Red Sea (see Fig. 12a). Conversion from measured FCDOM concentration in ppb to $a_g(\lambda)$ is difficult (Xing, Morel, Claustre, D'Ortenzio, & Poteau, 2012). Differences in sensor calibrations used to convert FCDOM output values (counts) into ppb units vary, as do conversions from ppb to $a_g(\lambda)$, even in similar regions (Xing et al., 2012). However, values of FCDOM in ppb may give some indication as to whether CDOM is relatively high or not. A total of 107 profiles were made during the two cruises (70 in March and 37 in November). A detailed description of the processing of the March and November 2013 cruises is provided in the Supplementary material.

For both the March and November cruises FCDOM was positively correlated with chlorophyll ($p < 0.001$, Fig. 12b). However, for the March cruise the amount of FCDOM per unit chlorophyll was higher than the November cruise, likely as result of the March cruise sampling close to the coastline of Jeddah, near coral reefs and anthropogenic influence (Fig. 12a). When comparing the ratio $R(412)/R(443)$ to $R(490)/R(555)$ derived from a MODIS-Aqua composite computed over the period of each cruise and related ϕ values, the March match-ups show a significantly lower $R(412)/R(443)$ for a given $R(490)/R(555)$ than the November match-ups, and thus higher ϕ values (see Fig. 12c). Whereas the absolute ϕ values derived from MODIS cannot be verified, results from this analysis suggest that the ϕ index captured the relative changes in FCDOM among the March and November cruises, lending support to the use of the Morel and Gentili (2009b) ϕ index in the Red Sea.

Another way to verify relative changes in the ϕ index proposed by Morel and Gentili (2009b) is to check whether seasonal variations in ϕ (and hence CDOM) are consistent with knowledge of interactions

between CDOM and marine physical processes. The November 2013 cruise was located primarily in offshore waters in the central Red Sea (see Fig. 12a and Supplementary Fig. S5a), away from the coastline of Jeddah (unlike the March 2013 cruise). Supplementary Fig. S5b shows the average vertical profile of FCDOM for the November cruise and the average mixed layer depth (MLD), computed as the depth at which the temperature changed by 0.5 °C relative to the surface temperature (Monterey & Levitus, 1997; Raitos et al., 2013). Consistent with other regions in the global ocean (Nelson & Siegel, 2013), FCDOM concentration increases below the mixed layer. Assuming no effect of horizontal advection or changes in light, an increase in MLD is likely to result in an increase in the average FCDOM concentration in the mixed-layer, as deep CDOM-rich waters will become entrained to the surface mixed layer. Consistent with this hypothesis, we observed a significant ($p < 0.05$) positive correlation between MLD and average FCDOM concentration in the mixed-layer (see Supplementary Fig. S5c) during the November cruise.

To test whether seasonal variations in the ϕ index are coupled to seasonal variations in MLD, we focused on the Northern Red Sea (above 24°N) as this region is controlled primarily by seasonal variations in vertical convection (Sofianos & Johns, 2003; Triantafyllou et al., 2014; Yao, Hoteit, Pratt, Bower, Köhl, et al., 2014). When comparing average monthly climatological values of the ϕ index above 24°N, derived from the SeaWiFS sensor (downloaded from the NASA ocean-colour website <http://oceancolor.gsfc.nasa.gov/>), with modelled seasonal variations in MLD from Yao, Hoteit, Pratt, Bower, Köhl, et al. (2014), we find a significant ($p < 0.05$) positive correlation ($r = 0.85$, $p = 0.003$, Fig. 13a). Furthermore, average monthly climatological values of the ϕ index above 24°N are significantly negatively correlated ($r = -0.88$, $p = 0.002$, Fig. 13b) with climatological values of sea surface temperature (SST) derived from MODIS-Aqua, consistent with the assumption that increases in MLD bring deep, cooler CDOM-rich water into the surface mixed layer. The CDOM index is also significantly negatively correlated ($r = -0.76$, $p = 0.002$, Fig. 13c) with monthly climatological values of photosynthetically available radiation (PAR) above 24°N derived from SeaWiFS, consistent with knowledge of bleaching of some CDOM components by enhanced sunlight (Nelson & Siegel, 2013; Vodacek, Blough, DeGrandpre, Peltzer, & Nelson, 1997; Whitehead et al., 2000). We also compared monthly climatological variations in ϕ and aerosol optical thickness (AOT) derived from SeaWiFS. Interestingly, we observed an inverse correlation ($r = -0.85$, $p = 0.003$, Fig. 13d), suggesting that when the aerosol thickness is at its lowest (during winter when skies are clearer) ϕ is at its highest, and during spring and summer when aeolian dust events are more frequent, ϕ is relatively low (Fig. 13d), though still considerably higher than one. If higher ϕ values in the Red Sea were linked with atmospheric aerosols, one may have expected ϕ to increase with increasing AOT. The results emphasise that seasonal variations in ϕ are consistent with knowledge on interactions between CDOM and marine physical processes (Blough, Zafriou, & Bonilla, 1993; Nelson & Siegel, 2013), and lend support to the use of the ϕ index (Morel & Gentili, 2009b) in the Red Sea.

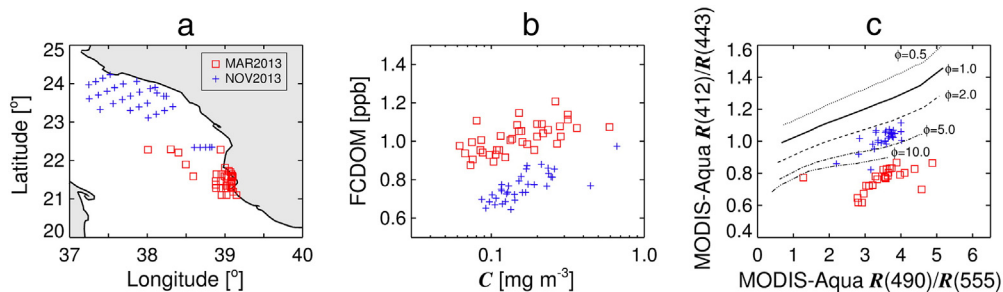


Fig. 12. (a) Shows the location of the samples used in two KAUST cruises (March and November 2013) of corresponding FCDOM and chlorophyll data. (b) Shows the relationship between FCDOM and chlorophyll for both cruises and (c) shows a plot of reflectance ratios $R(412)/R(443)$ versus $R(490)/R(555)$ for each sample, of each cruise, derived from the match-ups using a MODIS-Aqua composite averaged over the time period of each cruise.

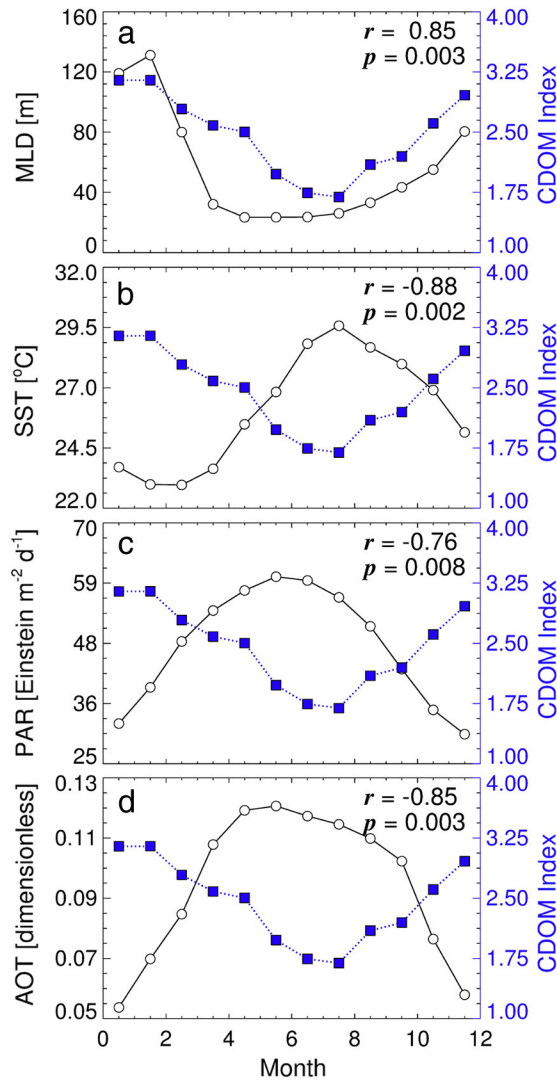


Fig. 13. Seasonal climatologies in the northern Red Sea (above 24°N) of the ϕ index (CDOM index, Morel & Gentili, 2009b) derived from the SeaWiFS sensor, the mixed layer depth (MLD) from Yao, Hoteit, Pratt, Bower, Köhl, et al. (2014), sea surface temperature (SST) derived from MODIS-Aqua, photosynthetically available radiation (PAR) derived from SeaWiFS and aerosol optical thickness (AOT) derived from SeaWiFS. Satellite data were downloaded from the NASA ocean-colour website <http://oceancolor.gsfc.nasa.gov/>.

5.1.4. Bio-optical modelling

The bio-optical framework of the Red Sea ocean-colour model can be traced back to the work of Sathyendranath et al. (2001) and Devred et al. (2006), whereby the absorption properties of two-component groups (or assemblages) were shown to vary with chlorophyll concentration. In our Red Sea ocean-colour model, all IOPs are tied to the chlorophyll biomass and dominant assemblage of phytoplankton, with the model assuming each assemblage resides in a distinctive bio-optical environment (Alvain, Loisel, & Dessailly, 2012). Two- and three-component models of phytoplankton absorption and backscattering, based on the work of Sathyendranath et al. (2001), have been developed and validated using data from different areas of the global ocean (Brewin, Dall'Olmo, et al. 2012; Brewin et al., 2011; Devred et al., 2006, 2011; Sathyendranath et al., 2001, 2004). In this study, we have simply taken this modelling framework and re-parameterised it using data in the Red Sea.

It has been shown that three-component models can be more representative of the transition in optical properties from oligotrophic to eutrophic waters when compared with two-component models (Brewin et al., 2011; Devred et al., 2011). However, a two-component model

was used in the case of the Red Sea as surface chlorophyll concentrations rarely exceed 1 mg m^{-3} (Fig. 2), such that a three-component model is unnecessary when considering the law of parsimony. However, blooms of phytoplankton with high chlorophyll may occasionally occur in the Red Sea (Genin et al., 1995), and it remains to be revealed how well the two-component model copes under such conditions. The Red Sea model is also fairly simplistic, not accounting for inelastic processes such as Raman Scattering that can impact satellite retrievals of optical constituents in oligotrophic waters (Lee et al., 2013; Sathyendranath & Platt, 1998; Westberry, Boss, & Lee, 2013). However, algorithms for deriving chlorophyll concentration (e.g. OC4 and OCI) that are empirically parameterised using *in situ* $R_{rs}(\lambda)$ and chlorophyll data include, implicitly, the Raman effects. Thus the systematic overestimation in chlorophyll observed in the Red Sea by these approaches is unlikely to be attributable to inelastic effects, although higher CDOM concentrations may depress Raman emission at bands used by OC4.

The parameterisation of any bio-optical model is ultimately related to the quality of the datasets used and their inherent uncertainties. With better quality datasets model parameterisation can be improved. The Red Sea model was parameterised using Tara data, collected during the month of January 2010, around the peak of the seasonal succession of phytoplankton (Raitos et al., 2013) (Fig. 11). Chlorophyll-specific IOPs presented in Table 1 are likely to have seasonal (Devred et al., 2006) and even inter-annual variations. We have already demonstrated that the ϕ index has a clear seasonal cycle in the northern Red Sea (Fig. 13). Retuning of OC4 and OCI algorithms for the Red Sea (OC4-RG and OCI-RG respectively) was verified using data between November and March (MIPOT, Tara and RSRC), but the performance of the algorithms during the summer (May–October) is yet to be tested. Additional data is required to evaluate the suitability of these retuned algorithms for processing summer satellite data in the Red Sea.

Algorithms such as the OC4-RG and OCI-RG proposed here are empirical in nature. The inferred relationship between chlorophyll and reflectance ratios (or differences) contains implicit dependence of the relationship on the change in phytoplankton community structure with change in chlorophyll, and on the covariance of other absorbing and scattering materials with chlorophyll. These algorithms are not designed to cope with changes in these relationships which may occur in a future climate. For instance, in recent years there have been an increasing frequency and intensity of *Noctiluca scintillans* blooms in the Indian Ocean and Arabian Sea (Gomes et al., 2014). Blooms of this species have also been observed in the Red Sea (Mohamed & Mesaad, 2007). If in the future, the phytoplankton community structure changes, or if associated variables change (e.g. CDOM and non-algal particle concentration), these alterations will interfere with the performance of empirical algorithms. On-going comparisons (and re-calibration) with *in situ* data, coupled with surveillance of other products from satellite in these regions (such as *N. scintillans*: Werdell, Roesler, & Goes, 2014), is required to monitor performance of these empirical chlorophyll algorithms.

5.1.5. Use of OC-CCI data

We chose to use OC-CCI data primarily due to improved coverage in the Red Sea region when compared with individual sensors and other merged products, so as to maximise the number of satellite and *in situ* match-ups. However, given that this merged product is relatively new, it would seem pertinent to verify whether the results found in our study using OC-CCI data are consistent with data from individual sensors. During 2010 when the Tara Oceans expedition sampled in the Red Sea, SeaWiFS, MODIS-Aqua and MERIS were all operating in parallel. Supplementary Fig. S6 shows results from a comparison of match-ups of SeaWiFS, MODIS-Aqua and MERIS derived chlorophyll, using standard band-ratio algorithms, OCI and the tuned algorithms (Tables 3 and 4), with *in situ* chlorophyll data from Tara. Consistent with OC-CCI, SeaWiFS, MODIS-Aqua and MERIS derived chlorophyll all show a systematic overestimation in chlorophyll when using the

OC4 and OCI algorithms. When using the revised algorithms (Tables 3 and 4) the systematic overestimation in chlorophyll disappears and satellite chlorophyll is in better agreement with *in situ* chlorophyll (Supplementary Fig. S6), supporting the results using OC-CCI data.

Considering OC-CCI data are merged at Level-3 processing, satellite and *in situ* match-ups in this study were conducted using daily Level-3 products, unlike standard NASA validation protocols (Bailey & Werdell, 2006). However, results from the match-ups presented here using the RSRC data (Fig. 5) resonate with the results from Brewin et al. (2013 see their Fig. 2) using Level 2 (1 km) MODIS-Aqua satellite data in-line with NASA validation protocols, supporting the use of Level-3 products for match-up analysis in this study.

5.2. Evidence for and against high CDOM absorption per unit chlorophyll in the Red Sea

Results from our analysis suggest that standard empirical ocean-colour algorithms, based either on a band-ratio (OC4) or band-difference (OCI), overestimate chlorophyll in the Red Sea, possibly due to an excess of CDOM absorption per unit chlorophyll. Knowledge on the sources and sinks of CDOM in the Red Sea remain elusive. An argument that would contradict our hypothesis is that there is very little riverine input in the Red Sea (Patzert, 1974), such that any additional CDOM is unlikely to come from terrestrial sources. A further argument against high CDOM is that the Red Sea is close to the tropics, and exposed to intense surface irradiance. Therefore, any CDOM produced would be expected to undergo rapid degradation from photo-bleaching (Nelson & Siegel, 2013; Vodacek et al., 1997). As discussed in the previous sections, there are also other reasons that may explain an overestimation in chlorophyll from standard satellite algorithms, such as uncertainty in the performance of atmospheric-correction algorithms and effects of aeolian dust deposition on the optical properties of the water (Claustre et al., 2002).

Despite these arguments, there is some evidence that may support the hypothesis of higher CDOM absorption than average in the Red Sea. Firstly, the results do not imply the Red Sea is rich in CDOM, but simply that CDOM is higher than in the average oligotrophic-mesotrophic environment (i.e. for a given chlorophyll concentration). Using satellite and *in situ* data, Morel and Gentili (2009a) found that CDOM absorption in the nearby Mediterranean Sea is twice that observed at the same latitude in the Atlantic Ocean. Nearby to the Red Sea, the Mediterranean Sea is also a semi-enclosed basin. It is possible that a refractory component of CDOM might have built up over time that is resistant to degradation from photo-bleaching. The Red Sea is also home to a vast number of coral reefs that sustain a huge amount of biological productivity and fisheries (Munday, Jones, Pratchett, & Williams, 2008). Red Sea coral reefs produce large amounts of dissolved organic matter (DOC, see van Duyl & Gast, 2001; Wild, Niggli, Naumann, & Hass, 2010; Yahel, Sharp, Marie, Häse, & Genin, 2003), and it may be that a component of this DOC is correlated with CDOM. Boss and Zaneveld (2003) observed enhanced concentrations of CDOM near coral reefs and near seagrass beds. It may be that benthic processes in basins with a high amount of area at boundaries relative to total area (such as the Red Sea and the Mediterranean), could influence CDOM more than in basins such as the Atlantic and Pacific.

Higher CDOM per unit chlorophyll may also be related to a positive relationship between enhanced metabolic processes and temperature (Taucher & Oschlies, 2011). The Red Sea is among the warmest seas on the planet, changes in microbial-loop activity with temperature (Behrenfeld, 2011) may alter production and degradation of CDOM. The Red Sea is also subject to high irradiance. In high light environments changes in the ratio of phytoplankton carbon to chlorophyll may result in modifications in the ratio of chlorophyll to CDOM, without any change in the ratio of phytoplankton carbon to CDOM. Ultimately, additional datasets are required to verify if the Red Sea has elevated CDOM absorption per unit chlorophyll.

6. Summary

Using an objective classification, and a dataset of satellite (OC-CCI products) and *in situ* chlorophyll match-up data, we tested the performance of two standard ocean-colour empirical chlorophyll algorithms (one based on a blue to green band-ratio, OC4, and the other a band-difference, OCI), a semi-analytical algorithm, and an empirical algorithm that accounts for the influence of CDOM on the remotely-sensed chlorophyll estimates. We found that the two empirical algorithms (OC4 and OCI) had the highest performance, but systematically overestimated chlorophyll when compared with the *in situ* data.

By developing a Red Sea ocean-colour model, parameterised where possible to data from the Red Sea, we adjusted the two ocean-colour empirical algorithms for chlorophyll estimation and the systematic overestimation in chlorophyll originally observed was removed. The relationships of particulate absorption and particulate backscattering with chlorophyll that are used in the Red Sea model, are similar to established global relationships, but the amount of CDOM absorption per unit chlorophyll concentration in the model is higher than standard global relationships. An enhanced amount of CDOM absorption per unit chlorophyll in the Red Sea was found to explain the overestimation in chlorophyll originally observed for the OCI and OC4 algorithms. A series of algorithms adjusted for the Red Sea have been proposed, designed for a range of ocean-colour sensors, and are now available for further testing. Given the unique and understudied marine and atmospheric environment of the region, uncertainties in the *in situ* data, and the potential influence of aeolian dust on atmospheric correction, additional information is required to scrutinise our findings.

Acknowledgements

The authors would like to thank the captains and the crews of the R/V “Aegaeo” of the Hellenic Centre for Marine Research (HCMR), and of the R/V “Oceanus” from the Woods Hole Oceanographic Institution (WHOI), who made the data collection possible. We thank in particular Leah Trafford and Amy Bower for their assistance on the cruise data used in this study. The authors are grateful to the Italian Scientific Committee for Antarctica (CSNA) for hosting the researchers and boarding the instruments in the RV *Italica*, and P. Povero and V. Saggiomo for organising the MIPOT oceanographic campaign, to Fiorani Luca for making the data available via the NASA SeaBASS website, and to NASA for archiving the MIPOT data on the SeaBASS website.

We thank staff at the Goddard Space Flight Center Ocean Ecology Laboratory for their support and J. Loftin, S. Searson, H. Le Goff, and S. Kandels for their handling of the AC-S during Tara Oceans. We thank the following people, institutions, and sponsors who made the Tara expedition possible: CNRS, EMBL, Genoscope/CEA, UPMC VIB, Stazione Zoologica Anton Dohm, UNIMIB, ANR, FWO, BIO5, Biosphere 2, agnes b., the Veolia Environment Foundation, Region Bretagne, World Courier, Cap L'Orient, the Foundation EDF Diversiterre, FRB, the Prince Albert II de Monaco Foundation, Etienne Bourgois, and the Tara Foundation teams and crew. Tara Oceans could not have happened without the support of the Tara Foundation and the Tara Consortium. This is contribution no. 27 of the Tara Oceans Expedition 2009–2012. Funding for the collection and processing of the AC-S dataset was provided by NASA Ocean Biology and Biogeochemistry Program under grants NNX11AQ14G and NNX09AU43G to the University of Maine. We also thank Alison Chase for comments on the HyperPro and HPLC data collected on the Tara cruise. We thank the editor and two anonymous reviewers for useful comments on the manuscript.

This research was supported by the King Abdullah University for Science and Technology (KAUST), Kingdom of Saudi Arabia, by a grant from the Changing Earth Science Network initiative funded by the STSE programme of the European Space Agency (ESA), by the UK National Centre for Earth Observation and by the Ocean Colour Climate Change Initiative of ESA.

- sensing single-wavelength algorithms. *International Journal of Remote Sensing*, 24, 5049–5073. <http://dx.doi.org/10.1080/0143116031000069807>.
- Sokoletsky, L., Dubinsky, Z., Shoshany, M., & Stambler, N. (2004). Single-wavelength algorithms for *in situ* or remote sensing estimation of mean pigment concentration. *International Journal of Remote Sensing*, 25, 1517–1525. <http://dx.doi.org/10.1080/01431160310001592535>.
- Stambler, N. (2005). Bio-optical properties of the northern Red Sea and the Gulf of Eilat (Aqaba) during winter 1999. *Journal of Sea Research*, 54, 186–203.
- Stambler, N. (2006). Light and picophytoplankton in the Gulf of Eilat (Aqaba). *Journal of Geophysical Research*, 111, C11009. <http://dx.doi.org/10.1029/2005JC003373>.
- Steinmetz, F., Deschamps, P.Y., & Ramon, D. (2011). Atmospheric correction in presence of sun glint: Application to MERIS. *Optics Express*, 19, 9783–9800. <http://dx.doi.org/10.1364/OE.19.009783>.
- Strickland, J. (1968). Continuous measurement of *in vivo* chlorophyll: A precautionary note. *Deep Sea Research*, 15, 225–227.
- Swan, C.M., Nelson, N.B., Siegel, D.A., & Fields, E.A. (2013). A model for remote estimation of ultraviolet absorption by chromophoric dissolved organic matter based on the global distribution of spectral slope. *Remote Sensing of Environment*, 136, 277–285.
- Taucher, J., & Oschlies, A. (2011). Can we predict the direction of marine primary production change under global warming? *Geophysical Research Letters*, 38, L02603. <http://dx.doi.org/10.1029/2010GL045934>.
- Triantafyllou, G., Yao, F., Petihakis, G., Tsiaras, K.P., Raitos, D.E., & Hoteit, I. (2014). Exploring the Red Sea seasonal ecosystem functioning using a three-dimensional biophysical model. *Journal of Geophysical Research — Oceans*, 119, 1791–1811. <http://dx.doi.org/10.1002/2013JC009641>.
- Twardowski, M.S., Boss, E., Macdonald, J.B., Pegau, W.S., Barnard, A.H., & Zaneveld, J.R.V. (2001). A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in case I and case II waters. *Journal of Geophysical Research*, 106, 14129–14142.
- van Duyl, F. C., & Gast, G.J. (2001). Linkage of small-scale spatial variations in DOC, inorganic nutrients and bacterioplankton growth with different coral reef water types. *Aquatic Microbial Ecology*, 24, 117–126.
- Vodacek, A., Blough, N.V., DeGrandpre, M.D., Peltzer, E.T., & Nelson, R.K. (1997). Seasonal variations of CDOM and DOC in the Middle Atlantic Bight: Terrestrial inputs and photooxidation. *Limnology and Oceanography*, 42, 674–686.
- 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. <http://dx.doi.org/10.1016/j.rse.2005.07.001>.
- Werdell, P.J., Bailey, S.W., Fargion, G.S., Pietras, C., Knobelspiese, K.D., Feldman, G.C., et al. (2003). Unique data repository facilitates ocean color satellite validation. *Eos, Transactions American Geophysical Union*, 84, 377. <http://dx.doi.org/10.1029/2003EO380001>.
- 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. <http://dx.doi.org/10.1016/j.mio.2013.09.001>.
- Werdell, P.J., Roesler, C., & Goes, J.I. (2014). Discrimination of phytoplankton functional groups using an ocean reflectance inversion model. *Applied Optics*, 53, 4833–4849. <http://dx.doi.org/10.1364/AO.53.004833>.
- Westberry, T.K., Boss, E., & Lee, Z. (2013). Influence of Raman scattering on ocean color inversion models. *Applied Optics*, 52, 5552–5561. <http://dx.doi.org/10.1364/AO.52.005552>.
- 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. <http://dx.doi.org/10.1364/OE.18.015419>.
- Whitehead, R.F., de Mora, S., Demers, S., Gosselin, M., Monfort, P., & Mostajir, B. (2000). Interactions of ultraviolet-B radiation, mixing, and biological activity on photobleaching of natural chromophoric dissolved organic matter: A mesocosm study. *Limnology and Oceanography*, 45, 278–291.
- Whitmire, A.L., Boss, E., Cowles, T.J., & Pegau, W.S. (2007). Spectral variability of the particulate backscattering ratio. *Optics Express*, 15, 7019–7031.
- Wild, C., Niggli, W., Naumann, M.S., & Hass, A.F. (2010). Organic matter release by Red Sea coral reef organisms-potential effects on microbial activity and *in situ* O₂ availability. *Marine Ecological Progress Series*, 411, 61–71. <http://dx.doi.org/10.3354/meps08653>.
- Xing, X., Morel, A., Claustre, H., D'Ortenzio, F., & Poteau, A. (2012). Combined processing and mutual interpretation of radiometry and fluorometry from autonomous profiling bio-argo floats: 2. Colored dissolved organic matter absorption retrieval. *Journal of Geophysical Research*, 117, C04022. <http://dx.doi.org/10.1029/2011JC007632>.
- Yahel, G., Sharp, J.H., Marie, D., Häse, C., & Genin, A. (2003). *In situ* feeding and element removal in the symbiont-bearing sponge *Theonella swinhoei*: Bulk DOC is the major source for carbon. *Limnology and Oceanography*, 48, 141–149.
- Yao, F., Hoteit, I., Pratt, L.J., Bower, A.S., Köhl, A., Gopalakrishnan, G., et al. (2014). Seasonal overturning circulation in the Red Sea: 2. Winter circulation. *Journal of Geophysical Research*, 119, 2263–2289. <http://dx.doi.org/10.1002/2013JC009004>.
- Yao, F., Hoteit, I., Pratt, L.J., Bower, A.S., Zhai, P., Köhl, A., et al. (2014). Seasonal overturning circulation in the Red Sea: 1. Model validation and summer circulation. *Journal of Geophysical Research*, 119, 2238–2262. <http://dx.doi.org/10.1002/2013JC009331>.
- Zhan, P., Subramanian, A.C., Yao, F., & Hoteit, I. (2014). Eddies in the Red Sea: A statistical and dynamical study. *Journal of Geophysical Research — Oceans*, 119, 3909–3925. <http://dx.doi.org/10.1002/2013JC009563>.
- Zhang, X., & Hu, L. (2009). Estimating scattering of pure water from density fluctuation of the refractive index. *Optics Express*, 17, 1671–1678. <http://dx.doi.org/10.1364/OE.17.001671>.
- Zhang, X., Hu, L., & He, M.X. (2009). Scattering by pure seawater: Effect of salinity. *Optics Express*, 17, 5698–5710. <http://dx.doi.org/10.1364/OE.17.005698>.