*CoastColour Round Robin data sets: a database to evaluate the performance of algorithms for the retrieval of water quality parameters in coastal waters*

Open Access Earth System Science Data

# CoastColour Round Robin data sets: a database to evaluate the performance of algorithms for the retrieval of water quality parameters in coastal waters

**B. Nechad[1], K. Ruddick[1], T. Schroeder[2], K. Oubelkheir[2], D. Blondeau-Patissier[3], N. Cherukuru[4], V. Brando[4], A. Dekker[4], L. Clementson[4], A. C. Banks[5,20], S. Maritorena[6], P. J. Werdell[7], C. Sá[8], V. Brotas[8], I. Caballero de Frutos[9], Y.-H. Ahn[10], S. Salama[11], G. Tilstone[12], V. Martinez-Vicente[12], D. Foley[13,†], M. McKibben[14], J. Nahorniak[14], T. Peterson[15], A. Siliò-Calzada[16], R. Röttgers[17], Z. Lee[18], M. Peters[19], and C. Brockmann[19]**

[1]Operational Directorate Natural Environment, Royal Belgian Institute for Natural Sciences (RBINS/ODNE), 100 Gulledelle Brussels, 1200, Belgium
[2]Commonwealth Scientific and Industrial Research Organisation (CSIRO), Land and Water, Environmental Earth Observation Program, P.O. Box 2583, Brisbane, QLD 2001, Australia
[3]Charles Darwin University, 0815 Darwin, Australia
[4]Commonwealth Scientific and Industrial Research Organisation (CSIRO), P.O. Box 1666, Canberra, ACT, Australia
[5]Hellenic Centre for Marine Research (HCMR), Institute of Oceanography, P.O. Box 2214, Heraklion 71003, Crete, Greece
[6]Earth Research Institute (ERI), University of California, Santa Barbara, CA 93106-3060, USA
[7]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[8]Marine and Environmental Sciences Centre (MARE), Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisbon, Portugal
[9]Institute of Marine Sciences of Andalucia (ICMAN-CSIC) Puerto Real-Cádiz, 11519, Spain
[10]Korea Ocean Research & Development Institute (KORDI), Ansan, P.O. Box 29, 425–600, South Korea
[11]Faculty of Geo-information Science and Earth Observation (ITC), Department of Water Resource, University of Twente, Hengelosestraat 99, 7500 AA Enschede, the Netherlands
[12]Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth PL1 3DH, UK
[13]National Oceanic and Atmospheric Administration (NOAA), Southwest Fisheries Science Center, 110 Shaffer Road, Santa Cruz, CA 95060, USA
[14]College of Earth, Ocean and Atmospheric Sciences (CEOAS), Oregon State University, Corvallis, OR, USA
[15]Center for Coastal Margin Observation and Prediction and Institute of Environmental Health, Oregon Health and Science University, 3181 SW Sam, Jackson Park Road, Portland, Oregon 97239, USA
[16]Environmental Hydraulics Institute of the University of Cantabria, Cantabria, Spain
[17]Institute of Coastal Research, Helmholtz-Zentrum Geesthacht, Centre for Materials and Coastal Research, Max-Plank-Str. 1, 21502 Geesthacht, Germany
[18]School for the Environment, University of Massachusetts Boston, Boston, MA 02125, USA
[19]Brockmann Consult, Max-Planck-Str. 2, 21502 Geesthacht, Germany
[20]European Commission – Joint Research Centre (JRC), Institute for Environment and Sustainability, Via Enrico Fermi 2749, Ispra (Va) 21027, Italy
[†]deceased

*Correspondence to:* B. Nechad (bnechad@naturalsciences.be)

**Abstract.** The use of in situ measurements is essential in the validation and evaluation of the algorithms that provide coastal water quality data products from ocean colour satellite remote sensing. Over the past decade, various types of ocean colour algorithms have been developed to deal with the optical complexity of coastal waters. Yet there is a lack of a comprehensive intercomparison due to the availability of quality checked in situ databases. The CoastColour Round Robin (CCRR) project, funded by the European Space Agency (ESA), was designed to bring together three reference data sets using these to test algorithms and to assess their accuracy for retrieving water quality parameters. This paper provides a detailed description of these reference data sets, which include the Medium Resolution Imaging Spectrometer (MERIS) level 2 match-ups, in situ reflectance measurements, and synthetic data generated by a radiative transfer model (HydroLight). These data sets, representing mainly coastal waters, are available from doi:10.1594/PANGAEA.841950.

The data sets mainly consist of 6484 marine reflectance (either multispectral or hyperspectral) associated with various geometrical (sensor viewing and solar angles) and sky conditions and water constituents: total suspended matter (TSM) and chlorophyll $a$ (CHL) concentrations, and the absorption of coloured dissolved organic matter (CDOM). Inherent optical properties are also provided in the simulated data sets (5000 simulations) and from 3054 match-up locations. The distributions of reflectance at selected MERIS bands and band ratios, CHL and TSM as a function of reflectance, from the three data sets are compared. Match-up and in situ sites where deviations occur are identified. The distributions of the three reflectance data sets are also compared to the simulated and in situ reflectances used previously by the International Ocean Colour Coordinating Group (IOCCG, 2006) for algorithm testing, showing a clear extension of the CCRR data which covers more turbid waters.

# 1   Introduction

Several studies on the intercomparison of ocean colour algorithms have been carried out to provide recommendations on appropriate methodologies and identify the domains of applicability and limitations or weaknesses of the algorithms, e.g. O'Reilly et al. (1998), Maritorena et al. (2006), Brewin et al. (2015), Odermatt et al. (2012), and Werdell et al. (2013). Except for the open ocean waters (or case 1 waters; Morel and Prieur, 1977), chlorophyll $a$ algorithm studies, no substantial consensus was achieved regarding a convergence of approaches for the retrieval of in-water properties from satellite or in situ radiometric measurements in coastal waters.

The diversity of approaches is especially high in case 2 waters (Morel and Prieur, 1977) with higher complexity of the optical properties and larger ranges of in-water constituent concentrations. To understand how these elements can affect the performance of algorithms, the CoastColour Round Robin (CCRR) project was designed (Ruddick et al., 2010). The CCRR uses a variety of reference data sets to test algorithms and compare their accuracy for retrieving water quality (WQ) parameters. These WQ parameters include chlorophyll $a$ (CHL) and total suspended matter (TSM) concentrations, inherent optical properties (IOPs), underwater light attenuation coefficients such as the diffuse attenuation of the downwelling irradiance (Kd) or the photosynthetically available radiation (PAR) with which a set of satellite data processing quality flags are associated.

Three types of data are being prepared for the CCRR: (a) match-ups, where in situ WQ is available simultaneously with a cloud-free Medium Resolution Imaging Spectrometer (MERIS) product; (b) in situ reflectances, where an in situ water-leaving reflectance measurement (denoted by RLw, which is derived from the remote-sensing reflectance, Rrs, following RLw $= \pi$ Rrs) is available simultaneously with an in situ WQ; and (c) simulated RLw for specified sets of IOPs and geometrical conditions, using HydroLight. MERIS images are also provided for the selected regions where the remote-sensing WQ algorithms are tested.

The match-ups, the in situ reflectance and the simulated data sets are presented in Sect. 2, and the variability in WQ is characterized. The data from the three data sets are intercompared in Sect. 3. This study provides documentation for the publicly available data sets (as detailed in Sect. 4) which can be used as benchmarks for ocean colour algorithm testing in coastal waters in order to ultimately improve the remote-sensing algorithms.

# 2   Data

The in situ WQ parameters provided in the match-up data set and referred to hereafter as "match-up field measurements" are described in Sect. 2.1.1. The concurrent MERIS level 2 products, reported in Sect. 2.1.2, include the MERIS reflectances and WQ, denoted respectively as L2R and L2W, and level 2 flags.

The in situ reflectance data set, described in Sect. 2.2, consists of in situ TSM and CHL measurements collected simultaneously with reflectances that cover the spectral range 440–709 nm. Inclusion of the 709 nm band in these data sets is important because it allows testing of algorithms exploiting this MERIS band, which is unique amongst any other ocean colour mission spectral specifications, operational up to 2012, e.g. for the retrieval of CHL or fluorescence line

**Table 1.** Acronyms of in situ data sources, as well as associated websites where the original data and methodologies are available.

| Acronym | Name |
|---------|------|
| CEOAS/OSU (CEOAS) | College of Earth, Ocean and Atmospheric Sciences – Oregon State university (USA) |
| CSIC | Spanish Institute for Marine Sciences (Spain) |
| CSIR | Council for Scientific and Industrial Research (South Africa) |
| CSIRO | Commonwealth Scientific and Industrial Research Organisation (Australia) |
| EMECO | European Marine ECosystem Observatory http://www.emecodata.net |
| GKSS | Centre for Materials and Coastal Research, Helmholtz-Zentrum Geesthacht (Germany) |
| HCMR | Hellenic Centre for Marine Research (Greece) |
| Ifremer | French Research Institute for Exploration of the Sea (France) http://wwz.ifremer.fr/lerpc/Activites-et-Missions/Surveillance/REPHY |
| ITC | International Institute for Geo-Information Science and Earth Observation (Netherlands) |
| KORDI | Korea Ocean Research and Development Institute (South Korea) |
| MII | Marine Institute of Ireland (Ireland) http://data.marine.ie |
| MSU | Mississippi State University (USA) |
| NOAA | National Oceanic and Atmospheric Administration (USA) |
| NOMAD | NASA bio-Optical Marine Algorithm Dataset, http://seabass.gsfc.nasa.gov |
| PML | Plymouth Marine Laboratory (UK) |
| RBINS | Royal Belgian Institute for Natural Sciences (Belgium) |
| UCSB | University of California at Santa Barbara, Earth Research Institute (USA) |
| UNICAN | Environmental Hydraulics Institute of the University of Cantabria (Spain) |

height using reflectance at band 709 nm combined with other bands in or around the phytoplankton absorption peak.

The artificial data set, based on radiative transfer simulations, is presented in Sect. 2.3. The match-up, in situ reflectance and simulated data sets come from 18 research institutes or databases (Table 1).

## 2.1 Match-up data set

The measurements in the match-up data set cover various water types from ocean and coastal regions called CoastColour sites, and consist of a collection of biogeochemical and optical measurements (inherent and apparent optical properties, hereafter referred to as IOPs and AOPs) along with the associated metadata. Only the WQ parameters for which remote-sensing algorithms are tested within the CCRR, such as CHL and TSM (see Table 2), are described in this paper, although supplementary oceanographic parameters are also included in the match-up database.

The match-up field measurements were collected at 17 CoastColour sites, selected in the framework of the CCRR (Fig. 1), where in situ WQ parameters from 2005 to 2010 were available, and measured above 5 m depth. MERIS L2R and L2W products from 2005 to 2010, derived at match-up locations, are included in the match-up data set, but only those of MERIS L2R are described in this paper.

The temporal availability of these data displayed in Fig. 2 shows unbalanced distributions over the CoastColour sites. The seasonal distribution of the match-up field measurements varies from one site to another (Fig. 3). For example, for chlorophyll $a$ measurements, 52 % of the Acadia data were collected during the period June–August, 67 %

of Chesapeake Bay data during September–November, and 100 % of Benguela data during March–May; the seasonal distribution may also vary within each site between the different WQ parameters. From all the sites, the ensemble of temperature, salinity, chlorophyll $a$, particulate organic matter (PIM), and particulate inorganic matter (POM) measurements is evenly balanced throughout the seasons. During December–February, fewer TSM, turbidity, $a$, $a_p$, $a_{phy}$, $a_d$, and $a_g$ measurements are available than during the other periods (about 13 to 18 % of the data), while the quantity of AOP data is significantly lower (2 to 9 % of the data).

### 2.1.1 Match-up field measurements

The number of stations where metadata and biogeochemical, IOP, and AOP data were collected over the CoastColour sites are reported in Table 3a and b. The availability of measurements throughout the sites varies from one parameter to another; for example, chlorophyll $a$ concentration measurements are available from 16 sites, while the scattering coefficient spectra are provided at 2 sites.

Metadata, including depth, temperature, and salinity, exceed 20 000 for each parameter, whereas the numbers of bio-geochemical data, IOPs, and AOPs are much lower: 11 208 chlorophyll $a$ concentration measurements, 538 TSM measurements, 957 reflectance spectra (the other AOP data do not reach 200 data each), and fewer than 700 IOP data (for each parameter) except for turbidity ($N = 2187$).

The number of CHL and turbidity measurements collected at the North Sea site constitute 77.0 and 99.8 % of the measurements respectively, while smaller numbers of TSM and RLw data are provided from the North Sea site: 39.4 and

**Table 2.** Metadata, IOPs, and AOPs given at wavelength λ, and biogeochemical in situ measurements available for the CoastColour sites. The two notations Chl $a$ and TChl $a$ refer to chlorophyll $a$ concentration measured by high-performance liquid chromatography (HPLC) and by fluorometry respectively.

| Metadata | Notation | Units | Concentrations | Notation | Units |
|---|---|---|---|---|---|
| Date, time | | – | Chlorophyll $a$ (fluorometry) | Chl $a$ | mg m$^{-3}$ |
| Station, cruise | | – | Total chlorophyll $a$ (HPLC) | TChl $a$ | mg m$^{-3}$ |
| File name, File_id (station) | | – | TSM | TSM | g m$^{-3}$ |
| Latitude, longitude | | degrees | Non algal particulate matter | NAP | g m$^{-3}$ |
| Wind speed | | m s$^{-1}$ | Particulate inorganic matter | PIM | g m$^{-3}$ |
| Cloud cover | | – | Particulate organic matter | POM | g m$^{-3}$ |
| Measurement depth | | m | CDOM fluorescence | CDOMf | Qse |
| Secchi depth | | m | | | |
| Water depth | | m | Flags | Notation | Units |
| Photic depth | $Z_{p\%}$ | m | General flag | Flag | – |
| Mixed layer depth | MLD | m | Location flag | Location_flag | – |
| Temperature | | °C | Time flag | Time_flag | – |
| Salinity | | psu | Chlorophyll $a$ method | Chla_flag | – |
| Provider | | – | CoastColour product | CCP_flag | – |
| IOPs | Notation | Units | AOPs | Notation | Units |
| Total absorption coefficient | $a(\lambda)$ | m$^{-1}$ | Remote-sensing reflectance | Rrs ($\lambda$) | sr$^{-1}$ |
| Particles absorption coefficient | $a_p(\lambda)$ | m$^{-1}$ | Water-leaving reflectance | RLw ($\lambda$) | – |
| NAP absorption coefficient | $a_{NAP}(\lambda)$ | m$^{-1}$ | Water-leaving radiance (or | Lw ($\lambda$) | mW cm$^{-2}$ |
| Absorption by phytoplankton | $a_{ph}(\lambda)$ | m$^{-1}$ | above-water upwelling | | μm$^{-1}$ sr$^{-1}$ |
| Absorption by detritus | $a_d(\lambda)$ | m$^{-1}$ | radiance) | | |
| CDOM absorption coefficient | $a_g(\lambda)$ | m$^{-1}$ | Above-water downwelling | Es ($\lambda$) | mW cm$^{-2}$ |
| Total (back)scattering coefficient | $b_{(b)}(\lambda)$ | m$^{-1}$ | irradiance (or incident | | μm$^{-1}$ |
| NAP scattering coefficient | $b_{NAP}(\lambda)$ | m$^{-1}$ | irradiance) | | |
| NAP backscattering coefficient | $b_{bNAP}(\lambda)$ | m$^{-1}$ | Downwelling irradiance | Ed ($\lambda$) | mW cm$^{-2}$ |
| Backscattering ratio | $b_{bp}(\lambda)/b_p(\lambda)$ | – | | | μm$^{-1}$ |
| Total beam attenuation coefficient | $c(\lambda)$ | m$^{-1}$ | Diffuse attenuation of Ed | Kd ($\lambda$) | m$^{-1}$ |
| Particles beam attenuation coefficient | $c_p(\lambda)$ | m$^{-1}$ | Diffuse attenuation of PAR | $K_{par}$ | m$^{-1}$ |
| Turbidity | | FNU, FTU | | | |

5.6 % of the total CCRR match-up field TSM and reflectance data respectively. When excluding the turbidity data, 91.6 % of the IOP measurements are contributed from the southern California (38.7 %), North Sea (22.9 %), Florida (7.6 %), and Great Barrier Reef region (7.0 %) sites.

The methods of chlorophyll $a$, TSM, IOPs, and Rrs measurements performed by each data contributor are briefly described below. Chlorophyll $a$ measurement methods by the different laboratories are summarized in Table 4.

## Chlorophyll $a$ and TSM

Chlorophyll $a$ concentrations were measured by either high-performance liquid chromatography (HPLC), fluorometry, or spectrophotometry. In the following, TChl $a$ refers to chlorophyll $a$ measurements determined by HPLC and Chl $a$ denotes chlorophyll $a$ obtained by fluorometry or spectrophotometry. TSM concentrations were collected at nine sites: the eastern Mediterranean Sea (hereafter E. Md. Sea), the

Baltic Sea and E. Md. Sea, the Great Barrier Reef region (referred to hereafter as the GBR region), the Indonesian waters, Morocco and the western Mediterranean Sea (hereafter Morocco-W. Md. Sea), the North Sea, the Red Sea, and Tasmania coastal waters.

In the CEOAS data set, 422 TChl $a$ data were measured from 2006 to 2009 at the Oregon–Washington site and 2 at the central California site. Samples were stored at $-80$ °C until HPLC analysis. The distribution of TChl $a$ measurements from Oregon–Washington is seasonally unbalanced with 8 % of the measurements collected during the period of December–February, 38 % in March–May, 50 % in June–August, and 50 % in September–November.

The CSIC data set contains 736 Chl $a$ and 667 POM measurements collected in the Gulf of Cádiz (southwest Iberian Peninsula) within the Morocco-W. Md. Sea site. The measurements were taken in the nearshore area ($< 30$ km) of the Guadalquivir estuary from 2005 to 2007, and offshore dur-
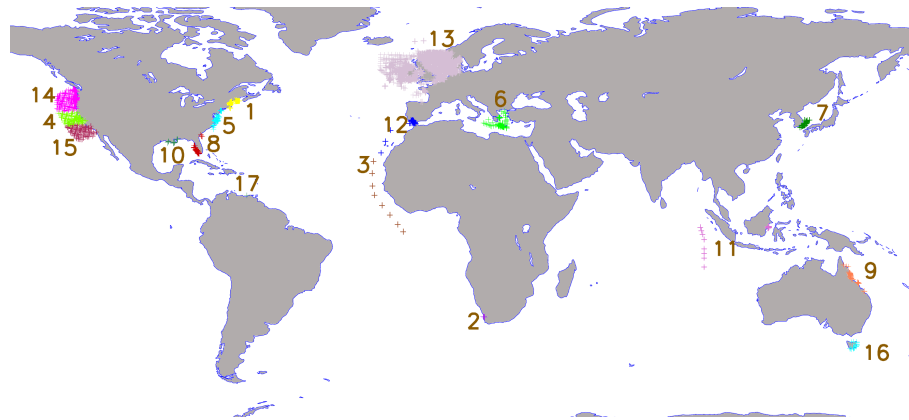
**Figure 1.** The distribution of the in situ data within the 17 CoastColour sites which are, numbered alphabetically, the coastal waters off (1) Acadia; (2) Benguela; (3) Cape Verde; (4) central California; (5) Chesapeake Bay; (6) the eastern Mediterranean Sea (referred to hereafter as E. Md. Sea); (7) the East China Sea; (8) Florida; (9) the Great Barrier Reef region (hereafter GBR region); (10) Gulf of Mexico; (11) Indonesia; (12) Morocco and western Mediterranean Sea (hereafter Morocco-W. Md. Sea); (13) the North Sea region extending to the English Channel, the Celtic and Irish seas, the Bay of Biscay, and southern Brittany (all referred to as the North Sea); (14) Oregon–Washington; (15) southern California; (16) Tasmania; and (17) Trinidad and Tobago.
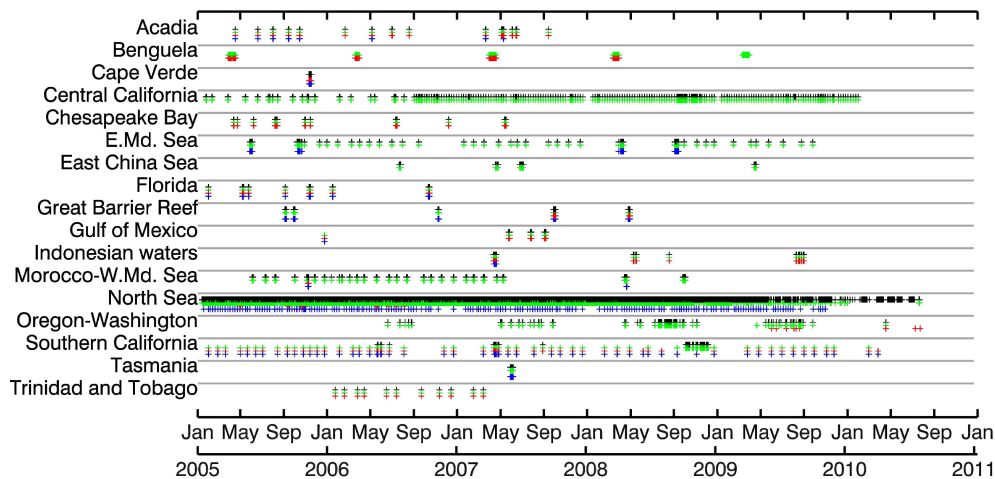


**Figure 2.** Time availability of at least one parameter available from the CoastColour sites within the match-up field measurements: metadata (black) (excluding the date, time, geographical coordinates, and data provider), biogeochemical data (green), AOPs (red), and IOPs (blue).

ing 2008 with slightly fewer measurements during the periods June–August (19 % of the data). Chlorophyll analysis was conducted by filtering samples of 500 mL through Whatman GF/F glass fibre filters (0.7 μm pore size), extracting in 90 % acetone, and measuring chlorophyll $a$ by standard fluorometric methods using a Turner Designs model 10 fluorometer following JGOFS protocols (IOC/UNESCO, 1994). TSM concentrations were measured gravimetrically on pre-weighted Whatman GF/F (0.7 μm pore size) after rinsing with distilled water, following JGOFS protocols (IOC/UNESCO, 1994). Organic matter lost on ignition was determined by reweighting the filters after 3 h in the oven at 500 °C, giving the concentrations of PIM and POM (by subtraction). TSM and PIM measurements, contaminated by salt

(filters not correctly rinsed), show low variability in TSM and PIM, with 90 % of TSM measurements comprised between 31.1 and 48.3 g m$^{-3}$. Therefore, only Chl $a$ and POM measurements are retained from the initial CSIC data set.

The CSIR chlorophyll data were collected from the Benguela coastal surface waters and measured using the standard fluorometric method of Parsons et al. (1984) with a Turner Designs 10AU fluorometer. A total of 131 Chl $a$ measurements are available from March to April for years 2005–2009.

The CSIRO data set consists of data collected at 63 stations in the GBR region from 2005 to 2008 (where 25, 19, and 55 % are available from March to May, June to August, and September to November respectively) and at 21 stations
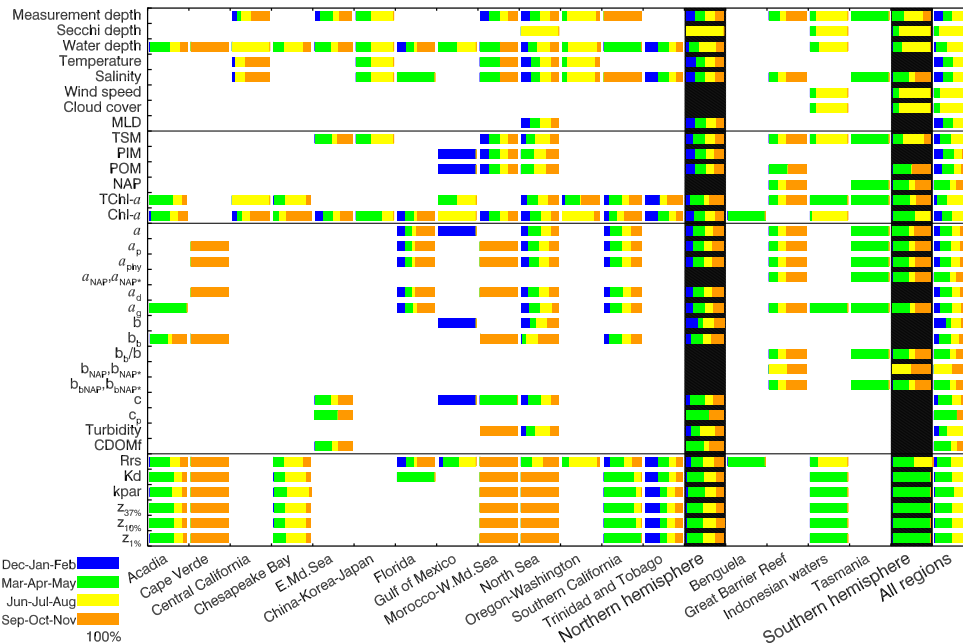
**Figure 3.** Seasonal availability of the metadata, biogeochemical, IOP, and AOP measurements from the CoastColour sites within the CCRR match-up data set.

in the Tasmanian waters in May 2007. Water samples were filtered through Whatman GF/F glass fibre filters with 0.7 μm nominal pore size and stored in liquid nitrogen until analysis by HPLC. The analyses conducted on the data set collected before July 2004 followed the method of Wright et al. (1991), while the method of Van Heukelem and Thomas (2001) was used for the subsequent campaigns (Oubelkheir et al., 2006; Blondeau-Patissier et al., 2009). For TSM analysis, the filters were pre-ashed at 450 °C, pre-washed in 100 mL of Milli-Q water, dried and pre-weighted. The samples were rinsed with 50 mL of distilled water and stored in Petri slides at 4 °C. The filters were dried at 60 °C (van der Linde, 1998).

The EMECO data set is provided by the International Council for the Exploration of the seas (ICES) and Smartbuoys data by the Centre for Environment, Fisheries and Aquaculture Science (Cefas), totaling 6274 stations with Chl *a* measurements, calibrated by HPLC. The distribution of these measurements is slightly unbalanced between the seasons (29 % of data available during March–May and 19 % in September–November).

The GKSS TSM and TChl *a* measurements were collected at 48 stations in the North Sea. TChl *a* and TSM measurements follow the protocol described in Doerffer and Schönfeld (2009). The sampling is equally distributed between the periods April–May, June–July, and September–October of years 2005–2006, with no measurements during December–February.

The HCMR data were collected at transect stations, where samples were taken in Niskin bottles from HCMR RV *Aegaeo*, in the E. Md. Sea site. For Chl *a* measurements, the

filtrations were performed using 47 mm diameter nucleopore filters consisting of Millipore® polycarbonate membrane filters, with 0.2 μm nominal pore size; Chl *a* was measured using Turner 00-AU-10 and Turner TD700 fluorometers using EPA Method 445 (Holm-Hansen et al., 1965) adapted by Arar and Collins (1992). For TSM measurements, the samples were filtered through 47 mm diameter, Isopore™ 0.45 μm polycarbonate membrane filters (Millipore®). After filtration of water samples, the filters were rinsed with Milli-Q water to remove salt. The filters were dried in the oven at 60 °C. In total, 294 Chl *a* measurements were collected from 2005 to 2009. Unbalanced percentages of Chl *a* data of 18 and 32 % are available from the periods June–August and September–November respectively. TSM measurements are available at 45 stations, sampled during years 2005 and 2008, with 47, 13, and 40 % of the data taken during the periods March–May, June–August, and September–November respectively.

The Ifremer data set consists of 975 Chl *a* measurements collected at 30 different locations within the Armorican Shelf (northwest of France), from 2005 to 2009. Data are available from the French phytoplankton surveillance network (REseau PHYtoplankton, REPHY; Gohin, 2011). Fluorometric measurements of Chl *a* were performed mostly in laboratory using a Turner C7 and C3. Over the four periods (seasons) from December–February to September–November, there are 18, 27, 32, and 23 % of the total number of Chl *a* measurements respectively.

The ITC measurements of Chl *a* and TSM were carried out in the Mahakam Delta waters from the upstream turbid

**Table 3.** Number of matchup-up field measurements provided by parameter (lines) and by site (columns).

(a) Number of metadata and biogeochemical match-up field measurements.

| CoastColour site WQ | Acadia | Benguela | Cape Verde | Central California | Chesapeake Bay | E. Md. Sea | East China Sea | Florida | GBR region | Gulf of Mexico | Indonesian waters | Morocco-W. Md. Sea | North Sea | Oregon–Washington | Southern California | Tasmania | Trinidad & Tobago | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Measurement depth | | | | 650 | | 433 | 78 | | 78 | | 119 | 738 | 27 837 | 566 | 126 | 21 | | 30 646 |
| Secchi depth | | | | | | | | | | | 119 | | 28 | | | | | 147 |
| Water depth | 76 | | 8 | 2 | 81 | 139 | 78 | 85 | | 41 | 110 | 63 | 245 | 381 | 7 | | 11 | 1327 |
| Temperature | | | | 223 | | | 77 | | | | | 63 | 25 530 | 429 | | | | 26 322 |
| Salinity | | | | 223 | | | 77 | 4 | 63 | | | 63 | 24 704 | 427 | 122 | 20 | 11 | 25 714 |
| Wind speed | | | | | | | | | | | 119 | | | | | | | 119 |
| Cloud cover | | | | | | | | | | | 113 | | | | | | | 134 |
| MLD | | | | | | | | | | | | | 124 | | | | | 124 |
| TSM | | | | | | 45 | 78 | | 63 | | 119 | | 212 | | | 21 | | 538 |
| PIM | | | | | | | | | | 6 | | 667 | 48 | | | | | 721 |
| POM | | | | | | | | 32 | | 6 | | 667 | 48 | | | | | 753 |
| NAP | | | | | | | | | 63 | | | | | | | 21 | | 84 |
| TChl a | 40 | | | 2 | 69 | | | | 63 | 41 | 4 | | 239 | | 247 | 21 | 5 | 1153 |
| Chl a | 25 | 131 | | 606 | 12 | 294 | 47 | 84 | | 6 | 96 | 736 | 7468 | 136 | 403 | | 11 | 10 055 |

(b) Number of IOP and AOP match-up field measurements.

| CoastColour site WQ | Acadia | Benguela | Cape Verde | Central California | Chesapeake Bay | E. Md. Sea | East China Sea | Florida | GBR region | Gulf of Mexico | Indonesian waters | Morocco-W. Md. Sea | North Sea | Oregon–Washington | Southern California | Tasmania | Trinidad & Tobago | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $a$ | | | | | | | | 63 | 63 | 6 | | | 117 | | 342 | 19 | | 610 |
| $a_p$ | | | 7 | | | | | 66 | 63 | | | 3 | 188 | | 346 | 21 | | 694 |
| $a_{phy}$ | | | 7 | | | | | 66 | 62 | | | 3 | 176 | | 346 | 21 | | 681 |
| $a_{NAP}, a_{NAP^*}$ | | | | | | | | | 63 | | | | | | | 21 | | 84 |
| $a_d$ | | | 7 | | | | | 66 | | | | 3 | 188 | | 347 | | | 611 |
| $a_g$ | 4 | | | | | | | 65 | 63 | | 4 | | 129 | | 342 | 19 | | 626 |
| $b$ | | | | | | | | | | 6 | | | 54 | | | | | 60 |
| $b_b$ | 23 | | 7 | | | | | | | | | 3 | 28 | | 269 | | | 330 |
| $b_b/b$ | | | | | | | | | 63 | | | | | | | 21 | | 84 |
| $b_{NAP}, b_{NAP^*}$ | | | | | | | | | 25 | | | | | | | | | 25 |
| $b_{bNAP}, b_{bNAP^*}$ | | | | | | | | | 63 | | | | | | | 21 | | 84 |
| $c$ | | | | 139 | | | | | | 6 | | 6 | 116 | | | | | 267 |
| $c_p$ | | | | 34 | | | | | | | | | | | | | | 34 |
| Turbidity | | | | | | | | | | | | 30 | 2157 | | | | | 2187 |
| CDOMf | | | | | | 132 | | | | | | | | | | | | 132 |
| Kd | 42 | | 8 | 69 | | | 4 | | | | 8 | 3 | 6 | | 16 | | 11 | 167 |
| RLw | 76 | 84 | 8 | 81 | | | | 85 | 15 | 47 | 127 | 3 | 54 | 47 | 319 | | 11 | 957 |
| kpar | 38 | | 5 | 35 | | | | | | | 8 | 3 | 4 | | 15 | | 10 | 118 |
| $z_{37\%}$ | 42 | | 8 | 69 | | | | | | | 8 | 3 | 5 | | 16 | | 10 | 161 |
| $z_{10\%}$ | 42 | | 8 | 66 | | | | | | | 8 | 3 | 6 | | 15 | | 10 | 158 |
| $z_{1\%}$ | 41 | | 8 | 61 | | | | | | | 8 | 3 | 6 | | 11 | | 10 | 148 |

**Table 4.** Instrument and methods of chlorophyll *a* measurement in the CCRR match-up data set.

| Data provider | Instrument | Filters, diameter (mm), nominal pore size (µm) | Tchl *a* measurement method (HPLC) | Chl *a* measurement method |
|---|---|---|---|---|
| CEOAS | – | Whatman GF/F, N/A, 0.7 | HPLC | – |
| CSIC | Turner Model 10 | Whatman GF/F, N/A, 0.7 | – | JGOFS protocols; IOC/UNESCO (1994) |
| CSIR | Turner 10AU | | | Parsons et al. (1984) |
| CSIRO | | GF/F, 47, 0.7 | – | – |
| | | | Wright et al. (1991), Van Heukelem and Thomas (2001) | |
| EMECO | 5LEDs (Ferrybox) | | – | In vivo fluorometry |
| GKSS | – | – | Doerffer and Schönfeld (2009) | – |
| HCMR | Turner 10AU Turner TD700 | Millipore polycarbonate membrane filters, membrane polycarbonate, 47, 0.2 | – | EPA Method 445; Holm-Hansen et al. (1965), adapted by Arar and Collins (1992) |
| Ifremer | Turner C7, C3 | – | – | Fluorometry |
| IOW | – | – | – | Fluorometry |
| ITC | | Membrane filter, 47, 0.45 | – | Spectrophotometry; Clesceri et al. (1998) |
| NOAA | – | – | – | Fluorometry |
| NOMAD | Various (see references) | | Hooker et al. (2005) | Werdell and Bailey (2005), Pegau et al. (2003) |
| PML | Hypersil 3 mm C8 Thermo Separations and Agilent | | Barlow et al. (1997); Llewellyn et al. (2005) | – |
| UCSB | Turner 10AU | | Van Heukelem and Thomas (2001) | Strickland and Parsons (1972) |
| UNICAN | Hach Lange DR-5000 | | – | Spectrophotometry; Clesceri et al. (1998) |

Mahakam River down to the clear water situated in the seaward area influenced by the Makassar Strait. From each station, two 1 L bottles of surface water samples were taken and then stored onboard in cool and dark conditions until their processing in the laboratory. TSM concentrations were determined using the gravimetric method. Water samples were filtered through previously weighted 47 mm diameter filters (Whatman GF/F filters, pore size of 0.45 µm). The filters were dried and reweighed (Clesceri et al., 1998). Chl *a* concentrations were measured using a spectrophotometer after the water samples had been filtered through 47 mm diameter filters (membrane filter, pore size of 0.45 µm) (Clesceri et al., 1998). The Chl *a* and TSM measurements cover the wet (May) and dry (August) seasons in 2008 and the dry season in August 2009, with a total of 119 stations.

The KORDI data set includes 47 Chl *a* and 78 TSM measurements collected at the East China Sea site. Samples were filtered through a 25 mm diameter GF/F glass fibre filter. Chl *a* measurements were performed through the methanol-extraction method using a PerkinElmer Lambda 19 dual-beam spectrophotometer. TSM and Chl *a* data are available

from cruises carried out during April and June 2007 and April 2009, and 31 % of TSM data are available from measurements made in July 2006. During the periods of April–May and June–July, respectively 41 and 59 % of TSM measurements are available, while 68 and 32 % of the Chl *a* data are provided for these periods.

The NOAA Chl *a* measurements were performed based on in vitro fluorescence measurements following 24 h dark period extractions in acetone. A total of 136 measurements are available from the Oregon–Washington site sampled from July to September 2008; 122 Chl *a* data from southern California acquired during the period September–November in 2008; and 606 Chl *a* data from the central California site, measured from 2005 to 2010. From the periods of September–November and June–August, respectively 52 and 30 % of the NOAA Chl *a* collection are available.

The NASA bio-Optical Marine Algorithm Dataset (NOMAD) presents a large collection of bio-optical data in ocean and coastal waters (Werdell and Bailey, 2005). The NASA SeaWiFS Bio-optical Archive and Storage System (SeaBass; Werdell et al., 2003), the source of the NOMAD data set,

includes both the HPLC and fluorometric methods. HPLC methods may have differed between laboratories in order to separate different types of pigments, which may depend on the predominant component of chlorophyll (Hooker et al., 2005). HPLC-derived TChl $a$ measurements in the NOMAD data set are the sum of monovinyl and divinyl chlorophyll $a$, plus chlorophyllide $a$, allomers, and epimers (Werdell and Bailey, 2005). The NOMAD TChl $a$ data set constitutes 24 % of the total TChl $a$ measurements gathered within the Coast-Colour match-up data set. From 2005 to 2007, 175 TChl $a$ data were collected from the six CoastColour sites – Acadia (40), Chesapeake Bay (69), Gulf of Mexico (41), Indonesian waters (4), southern California (16), and Trinidad and Tobago (5) – and 142 Chl $a$ measurements from Acadia (25 data), Chesapeake Bay (12), Florida (84), Gulf of Mexico (6), Indonesian waters (4), and Trinidad and Tobago (11).

The PML data set was collected during RV *Aegaeo* and RV *James Clark Ross* cruises in the MOS-2 and L4 areas respectively. The extraction of chlorophyll was performed in acetone including apo-carotenoate, and the separation used reversed-phase HPLC with 30 s of sonification and 5 min of centrifugation (4000 rpm) (Barlow et al., 1997). In the PML data set, divinyl-chlorophyll $a$, chlorophyllide-a and chlorophyll $a$ isomers and epimers are added to chlorophyll $a$ (Barlow et al., 1997). For TSM measurements, 2 to 4 L seawater samples were filtered in triplicates and washed with Milli-Q water. Filters were pre-ashed at 450 °C for 4 h, pre-washed in 500 mL of Milli-Q water, oven-dried at 75 °C for 24 h, and pre-weighted (van der Linde, 1998). A total of 191 pairs of Chl $a$ and TChl $a$ and 136 TSM measurements were collected by PML between 2005 and 2009. The distributions of Chl $a$, TChl $a$, and TSM measurements are overall well balanced across seasons.

The UNICAN data set includes 28 TSM and Chl $a$ measurements collected in the North Sea region (the Bay of Biscay) in July 2010. Chl $a$ was measured through a Hach Lange DR-5000 with Whatman GF/F filter following the spectrophotometric method described by Clesceri et al. (1998) (trichrometric method), using a white reference to control the quality of the measurements. TSM was estimated using a gravimetric method after filtration through GF/C glass fibre filters.

## Inherent optical properties

IOP measurements were collected at 11 sites (blue symbols in Fig. 2). The measurement methods for the total absorption coefficient, $a$; absorption by CDOM, $a_g$; absorption by particles, $a_p$; absorption by detritus, $a_d$; absorption by phytoplankton pigments, $a_{phy}$; scattering $b$ and backscattering coefficients $b_b$; total beam attenuation coefficient, $c$; and particle beam attenuation, $c_p$, are briefly described below.

For the CSIRO measurements of $a$, $a_p$, and $a_{phy}$, carried in the GBR region and Tasmania coastal waters, samples were filtered using a 25 mm Whatman GF/F filter with 0.7 μm nominal pore size and then stored in liquid nitrogen (Oubelkheir et al., 2006; Blondeau-Patissier et al., 2009). CDOM absorption was determined after filtration through polycarbonate filters (Millipore) of 0.22 μm nominal pore size, and water samples were filtered immediately after collection and stored in cool and dark conditions until analysis (Tilstone et al., 2003). The backscattering coefficients were measured using HOBI Labs HydroScat-6. The spectral dependency of the scattering coefficient was modelled as a hyperbolic function of wavelengths, using bands 412, 488, 510, 532, 555, and 650 nm (Oubelkheir et al., 2006; Blondeau-Patissier et al., 2009).

In the HCMR data set collected in the E. Md. Sea, 139 measurements of $c_p$ are provided at 470, 660, and 670 nm (available at least at one of these wavelengths), and 34 measurements of $c_p$ are given at 670 nm. The beam attenuation coefficients were measured using a 0.25 m path length transmissometer Chelsea Technologies Group Ltd Alpha Tracka II, emitting at 470 nm. The instrument was mounted on RV *Aegaeo*'s permanent CTD rosette frame for casts through the water column. The data were quality-controlled, filtered, and binned at 1 m intervals (Karageorgis et al., 2012).

MSU IOP data consist of six measurements of $a$, $b$ and $c$ coefficients collected at the Gulf of Mexico site.

The NOMAD absorption coefficients $a_p$, and $a_g$, and absorption by detritus $a_d$, were derived by spectroscopy at six CoastColour sites (Acadia, Cape Verde, Florida, Indonesian waters, Morocco-W. Md. Sea, and southern California). Note that for the Indonesian waters, only $a_g$ is provided. These data have been quality-controlled, removing unreasonable data and instrument artifacts (Werdell, 2005). The spectral backscattering coefficient provided in NOMAD data set was obtained using HOBI Labs HydroScat-2 and HydroScat-6 sensors, WET Labs ECO$b_b$ and ECOVSF sensors, and Wyatt Technology Corporation DAWN photometers. The details on $b_b$ data processing are given in Werdell (2005).

Absorption coefficient spectra were measured by PML at 5 m depth in the North Sea, using the WET Labs ac9+. As reported in Martinez-Vicente et al. (2010), the measurements were corrected to account for temperature, salinity, and scattering effects. The samples were filtered through 47 mm diameter Whatman Anopore membranes (0.2 μm pore size), using pre-ashed glassware. Absorption coefficients were determined on the spectrophotometer and a 10 cm quartz cuvette from 350 to 750 nm, relative to a bi-distilled Milli-Q reference blank. $a_g$ was calculated from the optical density and the cuvette pathlength, then the baseline offset was subtracted from $a_g$ (Groom et al., 2009). The measurement of $a_{phy}$ followed the method of Tassan and Ferrari (1995). The coefficients $a_p$ and $a_{phy}$ were measured using a PerkinElmer Lambda 2 spectrophotometer, and 25 mm GF/F filters (Tilstone et al., 2012). $a_p$ were determined before and after pigment extraction using NaClO 1 % active chloride from 350 to

**Table 5.** Number and period(s) of match-up field RLw measurements in each CoastColour site and by data provider.

| CoastColour site | Data provider | Number | Period |
|---|---|---|---|
| Acadia | NOMAD | 76 | Apr 2005 to Sep 2007 |
| Benguela | CSIR | 84 | Mar 2005 to Mar 2008 |
| Cape Verde | NOMAD | 8 | Oct 2005, Nov 2005 |
| Chesapeake Bay | NOMAD | 81 | Mar 2005, May 2007 |
| Florida | NOMAD | 85 | Jan 2005, Oct 2006 |
| Great Barrier Reef | CSIRO | 15 | Sep 2007, Apr 2008 |
| Gulf of Mexico | MSU(6) NOMAD(41) | 47 | Dec 2005 May 2007 to Jul 2007 |
| Indonesian waters | ITC(119) NOMAD(8) | 127 | May 2008, Aug 2009 Apr 2007 |
| Morocco-W. Md. Sea | NOMAD | 3 | Oct 2005 |
| North Sea | GKSS(48) NOMAD(6) | 54 | Apr 2005 to Jul 2006 Oct 2005 |
| Oregon–Washington | CEOAS | 47 | May 2009 to Jul 2010 |
| Southern California | UCSB(303) NOMAD(16) | 319 | Jan 2005 to Mar 2010 May 2006 to Aug 2007 |
| Trinidad and Tobago | NOMAD | 11 | Jan 2006 to Mar 2007 |
| All | | 957 | Jan 2005 to Jul 2010 |

750 nm. The scattering measurements were performed using an ECO VSF-3 sensor (Martinez-Vicente et al., 2010).

Backscattering coefficients provided by UCSB were estimated from profiled measurements of the total volume scattering function $\beta$ at $140°$, using a HobiLabs HydroScat-6, collected at the southern California site. These measurements were corrected for light attenuation along the photon path to the instrument detector ($\sigma$ correction of Maffione and Dana, 1997) using concurrent absorption spectra (Kostadinov et al., 2007) for measurements up to 2005, and concurrent beam attenuation and absorption modelled from the diffuse attenuation coefficient for downwelling irradiance and the irradiance reflectance (see Antoine et al., 2011, for details). A total of 269 backscattering spectra initially measured at 442, 470, 510, 589, and 671 nm were interpolated at 412, 470, 510, and 589 nm assuming a $\lambda^{-1}$ spectral dependency of the backscattering coefficient. UCSB absorption spectra up to 2005 were obtained using vertical profiles of WET Labs ac-9 measurements, after application of pure water calibration, as well as standard temperature, salinity, and scattering corrections (WET Labs ac-9 Protocol, 2003). Surface absorption values were derived from the upper 15 m absorption spectra, after filtering incomplete, negative, or extreme values; spectra were linearly interpolated at 412, 443, 490, 510, 530, 555, 620, and 665 nm (Kostadinov et al., 2007). Measurements of $a_{\mathrm{phy}}$, $a_g$, and $a_{\mathrm{d}}$ spectra were obtained using a

Shimadzu UV2401-PC spectrophotometer. CDOM samples were filtered on 0–2 µm Poretics membranes, while GF/F filters were used to retain total particulate matter for $a_{\mathrm{p}}$ measurement, corrected for pathlength effects following Guillocheau (2003). Pigment extraction was performed in 100 % methanol.

### Apparent optical properties: water-leaving reflectance

A total of 957 match-up field RLw spectra were collected at 13 CoastColour sites and provided from eight data providers, covering a variety of time periods as listed in Table 5. About 33 % of these data are provided for the southern California region, 13 % for the Indonesian coastal waters site, 9 % for the Benguela and Florida sites, and 8 % from the Acadia and Chesapeake Bay sites. Less than 19 % of the data set is provided from the rest of the CoastColour sites. Hyperspectral RLw measurements are available from the GBR region, the North Sea, and the Indonesian waters.

The instruments and methods of RLw measurements are summarized in Table 6 and briefly described below.

The CEOAS radiometric measurements in the Oregon–Washington site were performed using a Satlantic Hyper-Pro II instrument, equipped with two hyperspectral sensors to vertically profile the upwelling radiance, Lu, and downwelling irradiance, Ed, in the water column, plus a separate

**Table 6.** Instruments and methods of measurement of RLw in the CCRR match-up data set. $\theta_v$ and $\Delta\varphi$ denote respectively the sensor zenith angle and its azimuth angle relative to the sun. Ed, Lu, and Lsky denote respectively the downwelling irradiance, the upwelling radiance, and the sky radiance measured along the viewing angle $\theta_v$. The indices $+$ and $-$ refer to measurements just above and below the water surface respectively.

| Data provider | Instruments | Method | Reference |
|---|---|---|---|
| CEOAS | 3 Satlantic HyperPro | Underwater profiling of Lu−, Ed−, and above water Ed+ | http://satlantic.com/sites/default/files/documents/ProSoft-7.7-Manual.pdf |
| CSIR | 2 TriOS RAMSES | Floating buoy attached to ship, measuring Lu−, Ed+ | N/A |
| CSIRO | 1 TriOS RAMSES | Above water Lu+, Lsky, Ed+; viewing $\theta_v = 45°$, $\Delta\varphi \sim 135°$ | Tilstone et al. (2003) |
| GKSS | 3 TriOS RAMSES | Lu+, Lsky, Ed+; viewing $\theta_v = 45°$, $\Delta\varphi \sim 135°$ | N/A |
| ITC | 2 TriOS RAMSES | Lu+, Lsky, Ed+, $\theta_v = 40°$, $\Delta\varphi = 135°$ | N/A |
| MSU | N/A | N/A | N/A |
| NOMAD | Various | In-water profiling, or above-water instruments | Werdell and Bailey (2005) |
| UCSB | ASD spectrometer, Biospherical PRR-600 | Merging RLw from in-water profiling and above-water ASD reflectance | Toole et al. (2000) |

surface sensor mounted high on the ship deck that measures the above-water downwelling irradiance, Es. Processing of the collected data was performed using Satlantic ProSoft software version 8.1.3_1 (see http://satlantic.com/sites/default/files/documents/ProSoft-7.7-Manual.pdf for equations). In summary, the above-water radiance, Lw, is calculated by extrapolating the profiled Lu measurements to the subsurface (Lu(0$^-$)) and then accounting for the air–sea interface: Lw = Lu(0$^-$)$(1 - \rho)/n_w^2$, where $\rho$ is the Fresnel reflectance of the air–sea interface (set to 0.021) and $n_w = 1.345$ is the refractive index of seawater. The surface irradiance reflectance is then obtained by RLw = $\pi$ Lw / Es. Of the 137 wavelengths measured by the HyperPro II, this study presents data from 21 wavelengths covering 412 to 780 nm for RLw.

In the Benguela site, the CSIR used a Satlantic radiometer mounted on a floating buoy attached to the ship in order to measure the upwelling radiance Lu and the downwelling irradiance Ed at 0.66 m below the water surface. Lu was extrapolated to Lw by means of the upwelling diffuse attenuation coefficient, Ku, as described by Albert and Mobley (2003). RLw was estimated from Lw and Ed using a reflectance inversion algorithm optimized for local conditions.

The CSIRO RLw measurements in the GBR region were conducted under stable clear-sky conditions using one TriOS RAMSES instrument. Subsequent water-leaving radiance, Lw; sky radiance, Lsky; and Spectralon upwelling radiance, Lspec, were measured. Irradiance was calculated from Spectralon measurements according to Ed = $\pi$ Lspec C, where C is the reflectance correction factor accounting for non-perfect

Lambertian panel properties. Water-leaving reflectance was calculated according to the REVAMP protocol (Tilstone et al., 2003) by applying a sky correction factor.

The GKSS radiometric measurements were conducted onboard ferry cruises in the North Sea region, using three TriOS RAMSES radiometers that simultaneously measure Lu at 45° viewing angle, Es, and Lsky, with an azimuth angle between 130 and 140° relative to the sun. The water-leaving reflectance RLw was computed according to RLw = $\pi$(Lu − $\rho_{sky}$Lsky)/Ed, where the specular reflectance $\rho_{sky}$ is computed using the Fresnel law, as a function of the refractive index for the mean salinity along the transect.

The ITC measurements carried out in the Indonesian waters used two TriOS RAMSES spectroradiometers. The surface water upwelling and sky downwelling radiance measurements, Lu and Lsky, were measured sequentially at 40° zenith angle and at 40° nadir angle respectively. The irradiance sensor was mounted on an aluminium pole on top of the boat, pointing upward. The boat was positioned on a station to point the radiance sensor at a relative azimuth angle of 135° away from the sun. The sensors measured over the wavelength range 350–950 nm with a sampling interval of approximately 3.3 nm. The measurements were conducted under different cloud conditions. The sky radiance reflected by the water surface, $\rho_{sky}$, was estimated by assuming very small (but not zero) water-leaving reflectance in the near infrared and that $\rho_{sky}$ values were less than 0.07, which is the highest value of scattered cumulus clouds by Mobley (1999). The result of $\rho_{sky}$ values were relatively similar with $\rho_{sky}$ values given by Mobley (1999) for each cloud type condi-

www.earth-syst-sci-data.net/7/319/2015/

Earth Syst. Sci. Data, 7, 319–348, 2015

tion. The water-leaving reflectance was obtained following the equation $RLw = \pi(Lu - \rho_{sky}Lsky)/Ed$.

The MSU radiometric measurements are provided for the Gulf of Mexico in the Mississippi Sound area (around Gulfport). The reflectance spectra were measured at 14 wavelengths in the spectral range 380–780 nm.

From the NOMAD database, Lw and Es measurements were extracted for the match-up locations between 2005 and 2010 and converted to RLw spectra. Various instruments were used for the measurements of the remote-sensing reflectance, Rrs, in the NOMAD data set (Werdell and Bailey, 2005), including in-water profiling or above-water measurements. All in- and above-water data from various instruments and data providers were consistently processed to Rrs, with the methods described in Werdell and Bailey (2005).

The UCSB RLw measurements in the southern California region were obtained using above-water radiometric measurements of one Dual FieldSpec spectrometer (ASD) instrument and underwater measurements of a Biospherical Instruments (San Diego, California) profiling reflectance radiometer (PRR-600), as described by Toole et al. (2000). Sea surface radiance, Ls, at viewing zenith angle of 45°; sky radiance (which would be reflected into Ls), Lsky; and Spectralon upwelling radiance, Lspec, were measured by the FieldSpec spectrometer. The above water reflectance was estimated following Toole et al. (2000): the above-water irradiance was calculated from Spectralon measurements according to $Ed = \pi Lspec/\rho_{spec}$, where $\rho_{spec}$ is the reflectance of the plaque; the water-leaving reflectance was calculated as $RLw = \pi(Ls - \rho Lsky)/Ed - residual(750)$, where residual(750) corrects for any residual reflected sky radiance, assuming zero water-leaving radiance at 750 nm. Underwater downwelling irradiance, $Ed^-$, and upwelling radiance, $Lu^-$, were measured along vertical profiles using the Biospherical PRR-600 and then interpolated to above-water radiance and irradiance respectively, leading to a new estimate of RLw spectra, which were merged with FieldSpec reflectances (see Toole et al., 2000, for details).

### 2.1.2  MERIS data

MERIS CoastColour processing (see flow chart in Fig. 4) is applied to MERIS Level 1 Full Resolution Full Swath (FRS) to produce MERIS level 2 match-up data sets, namely MERIS water-leaving reflectance (L2R) and MERIS water quality products (L2W), over the CoastColour sites. Here, a brief description of MERIS CoastColour processing is given.

MERIS FRS products, including auxiliary data such as surface pressure, ozone, geographical location (used to identify products having an overlap with one of the test sites), viewing and sun angles, and solar flux, are processed with the Accurate MERIS Ortho-Rectified Geolocation Operational Software (AMORGOS processor, developed by ACRI-ST within ESA GlobCover project), yielding geometrically corrected MERIS child products (FSG). The L1P processor
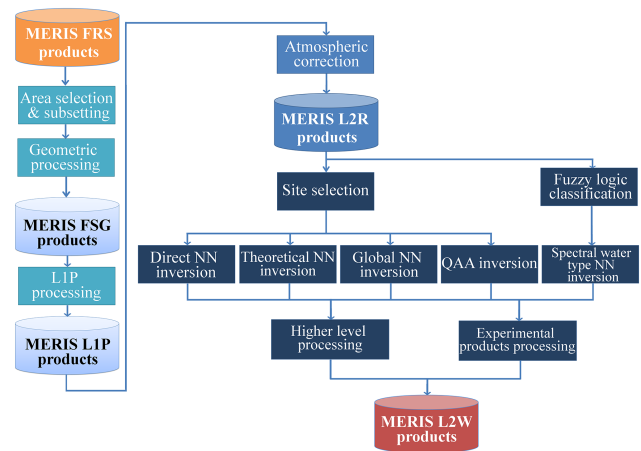


**Figure 4.** MERIS CoastColour processing.

subscenes the FSG data; applies the radiometric and smile corrections; and performs equalization following Bouvet and Ramoino (2010) and pixel classification, screening cloud pixels.

The L1P product, which contains the top of atmosphere radiance reflectance (TOA), is then atmospherically corrected to determine the water-leaving radiance reflectance, following the steps described in Doerffer (2011), which yields the L2R products. Furthermore, water pixels are classified according to their TOA reflectances and available geographical information, and L2W products are generated using various ocean colour algorithms. A complete list of the parameters contained in L2R and L2W products is given in Table 7.

Boxes of $5 \times 5$ pixels are extracted from L1P, L2R, and related L2W products at all match-up locations present for a given test site and are stored in three files associated with the site. Further processing is performed to average MERIS L2R spectra in each $5 \times 5$ box, discarding low-quality pixels (see the list of critical flags in Table 7) and yielding the mean reflectance, referred to hereafter as MERIS RLw, and its standard deviation. Other L2W and atmospheric products are also averaged over the $5 \times 5$ box (see the list in Table 8).

Finally, around each match-up location MERIS L1P, L2R, and L2W subscenes are provided in BEAM-DIMAP (".dim") format, and are associated with a KMZ file for quick visualization of area location via Google Earth.

With respect to the match-up field RLw data set, the MERIS RLw data set includes supplementary data from the following regions: the central California, E. Md. Sea and East China Sea, and Tasmania coastal waters, and extended data from Morocco-W. Md. Sea and the North Sea concurrent with extra match-up field WQ measurements (IOPs and/or biogeochemical data sets). The MERIS RLw data set is not available for all the locations of the match-up field RLw measurements (e.g. Benguela, Indonesian waters, GBR region), either because no MERIS image is available within 1 h of the match-up field measurement or because MERIS pixels

**Table 7.** The Level 2 products provided the MERIS match-up data set, as a 5 × 5 box around the locations of the match-up field measurements. The "critical" flags listed in italic font are associated with pixels being rejected (if the flags are raised) in the post-processed MERIS match-up data set.

| Navigation | Description | Units | L2R, L2W | Description | Units |
|---|---|---|---|---|---|
| ProdID | | – | reflec_x | RLw at λ (nm) | – |
| CoordID | ID of location | – | b_tsm | Scattering coefficient at 443 nm | m$^{-1}$ |
| Name | Match-up name | – | a_tot | Total absorption coefficient (443 nm) | m$^{-1}$ |
| Latitude, longitude | Geographical | degrees | | | |
| | coordinates | | Atmosphere | Description | Units |
| Date, time | | | tau_nnn | Aerosol optical thickness at λ = nnn (nm) | |
| lat_corr, lon_corr | Ortho-corrected | degrees | ang_443_865 | Aerosol Ångström coefficient | |
| | latitude/longitude | | | between 443 and 865 nm | – |
| dem_alt | DEM[a] model altitude | | | | |
| dem_rough | Roughness at sight with | | Ancillary | Description | Units |
| | intersection of line of | degrees | zonal_wind | ECMWF[c] zonal wind | m s$^{-1}$ |
| | WGS84[b] ellipsoid taken | | merid_wind | ECMWF meridional wind | m s$^{-1}$ |
| | from DEM | | glint_ratio | Glint ratio | – |
| sun_, view_zenith | Sun, view zenith angle | degrees | atm_press | ECMWF atmospheric pressure at | hPa |
| sun_, view_azimuth | Sun, view azimuth angle | degrees | | mean sea level | |
| pins | | – | ozone | ECMWF ozone concentration | DU |
| ground_control_points | | – | rel_hum | ECMWF relative humidity at 850 hPa | % |
| detector_index | Index of MERIS pixel | – | | | |

| Flags | Description | Flags | Description |
|---|---|---|---|
| *land* | Land pixel | coastline | Pixel is part of a coastline |
| water | Water pixel | cosmetic | Cosmetic flag |
| *cloud_ice* | Very high Rtoa indicating | duplicated | Pixel has been duplicated (filled in) |
| | cloud, ice, or snow pixel | f_meglint | Pixel corrected for glint |
| bright | Bright pixel | f_loinld | Low inland water flag |
| *sunglint* | Pixel affected by sun glint | f_island | Island flag |
| glint_risk | Glint correction not reliable on | f_landcons | Land product available |
| | the pixel | f_ice | Ice pixel |
| *suspect* | Suspect flag (from L1[d]) | f_cloud | IDEPIX[f] final cloud flag |
| *invalid* | Pixel is invalid | f_bright | IDEPIX bright pixel |
| solzen | High sun zenith angle | f_bright_rc | IDEPIX old bright pixel |
| ancil | Unreasonable data for ozone | f_low_p_pscatt | IDEPIX test on apparent scattering |
| | or pressure | f_low_p_p1 | IDEPIX test on P1 |
| has_flint | If the atmospheric correction | f_slope_1 | IDEPIX spectral slope test 1 flag |
| | used the flint processor | f_slope_2 | IDEPIX spectral slope test 2 flag |
| l1_flags | Level 1 classification and | f_bright_toa | IDEPIX second bright pixel test |
| | quality flag | f_high_mdsi | IDEPIX MDSI[g] above threshold |
| l1p_flags | Pixel classification flag (e.g. | f_snow_ice | IDEPIX snow/ice flag |
| | cloud screening, land, water) | agc_flags | Flag specific to the atmospheric |
| atc_oor | If RLw is out of the expected | | and flint correction |
| | range (as set in the NN[e]) | agc_land | Land pixel |
| toa_oor | Input Rtoa is out of the NN | agc_invalid | Pixel not considered for processing |
| | training range | | |
| tosa_oor | Input Rtosa is out of the NN | | |
| | training range | | |

[a] DEM refers to the digital elevation model of altitude; [b] WGS84 refers to the World Geodetic Standard 1984; [c] ECMWF is the European Centre for Medium Range Weather Forecast; [d] L1 is MERIS level 1 product; [e] NN is the atmosphere neural network algorithm; [f] IDEPIX is a generic pixel classification algorithm for optical Earth observation sensors; [g] MDSI is the MERIS differential snow index.

**Table 8.** The $5 \times 5$ box averaged L2R, L2W, and atmospheric parameters derived from the MERIS match-up data set.

| Navigation | Description | Units | L2R, L2W | Description | Units |
|---|---|---|---|---|---|
| Fid | Match-up name | – | RLw_xxx[a] | RLw at λ (nm) | – |
| Latitude, longitude | Geographical coordinates | degrees | b_tsm[a] | Scattering coefficient at 443 nm | $m^{-1}$ |
| Date, time | | | a_tot[a] | Total absorption coefficient (443nm) | $m^{-1}$ |
| sun_, view_zenith | Sun, view zenith angle | degrees | **Atmosphere** | **Description** | **Units** |
| sun_, view_azimuth | Sun, view azimuth angle | degrees | tau_nnn[a] | Aerosol optical thickness at $\lambda = nnn$ (nm) | |
| | | | ang_443_865[a] | Aerosol Ångström coefficient between 443 and and 865 nm | – |
| **Box-averaging information** | **Description** | **Units** | **Ancillary** | **Description** | **Units** |
| N(var[b]) | Number of pixels within the $5 \times 5$ box where valid var was retrieved | – | zonal_wind | ECMWF[c] zonal wind | $m\,s^{-1}$ |
| | | | merid_wind | ECMWF meridional wind | $m\,s^{-1}$ |
| | | | glint_ratio | Glint ratio | – |
| std(var[b]) | Standard deviation of var over the $N$ valid pixels in the $5 \times 5$ match-up box | var unit | atm_press | ECMWF atmospheric pressure at mean sea level | hPa |
| | | | ozone | ECMWF ozone concentration | DU |
| | | | rel_hum | ECMWF relative humidity at 850 hPa | % |

[a] Averaged over $N$ valid pixels in the $5 \times 5$ box around the match-up location. [b] The variable var refers to one of the MERIS L2 products listed under L2R, L2W, and atmosphere data types. [c] ECMWF is the European Centre for Medium Range Weather.

are flagged as cloud, land, suspect, sunglint, or invalid. After rejection of the flagged pixels, 457 MERIS RLw spectra remain from the CoastColour sites. About 80 % of these spectra are available from the North Sea region and match in situ measurements of temperature, salinity, and/or turbidity.

## 2.2 In situ reflectance data set

The in situ reflectance data set comprises a set of 336 RLw spectra sampled at nine MERIS bands from 412 to 709 nm, and collected simultaneously with CHL and/or TSM measurements at five CoastColour sites, from August 2002 to August 2009. The number of RLw data per site and per data provider, and their periods of measurement, are presented in Table 9. Part of these spectra, measured in Benguela, Indonesian waters, and the North Sea (the GKSS data set), are derived from the match-up field hyperspectral RLw data.

With respect to the match-up field data set, the in situ reflectance data set includes 266 spectra already given in the match-up field data set from the Benguela, Indonesian waters, North Sea (provided by the GKSS), and Oregon–Washington sites, plus supplementary data from the Mediterranean Sea and the North Sea (provided by RBINS; see details of measurement method hereafter) and data from Benguela covering year 2002. It excludes the entire RLw data from the Acadia, Cape Verde, Chesapeake Bay, Florida, GBR region, Gulf of Mexico, Morocco-W. Md. Sea, southern California, and Trinidad and Tobago sites, and the NOMAD RLw measurements subset collected at the North Sea and

the Indonesian waters, because no CHL and/or TSM and/or RLw spectra up to 709 nm are available. The total number of RLw spectra available within the match-up field and in situ reflectance data sets is $N = 1027$ (with no overlapping data).

The RBINS radiometric measurements were acquired in the North Sea and Mediterranean Sea using three TriOS RAMSES radiometers that simultaneously measure Es and the radiances Lw and Lsky at 40 and 140° viewing angles respectively with 135° azimuth angle relative to the sun (Ruddick et al., 2006).

The CHL data were measured by HPLC in all the sites except for measurements taken in Benguela after year 2002 (fluorometry) and in the Indonesian waters (spectrophotometry) (Table 9). The total numbers of in situ CHL and TSM data are 294 and 186 respectively.

## 2.3 Simulated data set

Radiative transfer simulations were performed with Hydro-Light version 5.0 (Mobley and Sundman, 2008), using the atmospheric, air–sea interface, and sun and viewing angle characteristics as presented in Table 10, and the specific IOPs (SIOP) for mineral particles (denoted by MP), phytoplankton, and $a_g(443)$ as given in Table 11. The SIOPs include the specific absorption coefficients for phytoplankton, $a_p^*$, and for MP, $a_{MP}^*$; the spectral slope of $a_{MP}^*$, denoted by $S_{MP}$; the specific scattering coefficient for MP, $b_{MP}^*$; the spectral variation in the beam attenuation coefficient for phytoplankton,

**Table 9.** The number and period(s) of measurement of in situ RLw and TSM and/or CHL concentrations, collected at each CoastColour site within the in situ reflectance data set. The methods for CHL and TSM measurements are also provided.

| CoastColour site (data provider) | Number of RLw spectra CHL, TSM | Period | CHL method | TSM method |
|---|---|---|---|---|
| Benguela (CSIR) | 135, 135, 0 | Aug 2002 to Mar 2008 | year 2002: HPLC; other years: fluorometric | – |
| Indonesian waters (ITC) | 119, 92, 119 | May 2008, Aug 2009 | Spectrophotometry | Gravimetric, GF/F |
| Mediterranean Sea (RBINS) | 7, 7, 7 | Mar 2009 | HPLC, 90 % acetone, cell homogenizer | Gravimetric, GF/F |
| North Sea (GKSS) | 48, 48, 48 | Apr 2005 to Jul 2006 | HPLC | Gravimetric, GF/F |
| North Sea (RBINS) | 12, 12, 12 | Apr 2006 to Jun 2009 | HPLC, 90 % acetone, cell homogenizer | Gravimetric, GF/F |
| Oregon– Washington (CEOAS) | 15, 15, 0 | May 2009 to Aug 2009 | HPLC | – |
| Total | 321 | Aug 2002 to Aug 2009 | – | – |

**Table 10.** Atmospheric, air–sea interface, and solar and viewing geometry specifications in CCRRv1.

| Parameter | Values |
|---|---|
| Sun angles | Zenith: 0, 40, and 60°; azimuth: 0° |
| Viewing angles | Zenith: 0°; azimuth: 90° |
| Surface wind speed | $5\,\mathrm{m\,s^{-1}}$ |
| Cloud fraction | 0 |
| Sky radiance distribution | Semi-empirical sky model; Harrison and Coombes (1988) |
| Direct and diffuse sky irradiances | Semi-empirical sky model RADTRAN |

$\gamma_{\mathrm{CHL}}$, and for MP, $\gamma_{c_{\mathrm{MP}}}$; and the spectral slope of CDOM absorption, $S_{\mathrm{CDOM}}$.

This simulated data set is denoted as "CCRRv1" to facilitate comparison with future versions, e.g. with variability in the specific inherent optical properties.

A total of 5000 triplets of CHL and MP concentrations and $a_g(443)$ were generated according to the following:

– A random number function modelling a log-normal probability density function was used for CHL.

– The associated MP and $a_g(443)$ values were also generated by a random number function but constrained to yield reasonable covariation of the triad, comparable to that reported by Babin et al. (2003b) from in situ measurements in coastal European waters.

Figure 5a and b show the distributions of the simulated MP and $a_g(443)$ vs. CHL concentrations and their co-variations.

Based on these concentrations and SIOP models, a set of hyperspectral (2.5 nm resolution) data were generated, including the total absorption $a$, scattering $b$, and backscattering $b_b$ coefficients; the phytoplankton absorption coefficient, $a_{\mathrm{phy}}$; and the ratio $b_b/a + b_b$. For each in-water con-

tent (5000 cases) and sun angle (3 cases), HydroLight computed RLw and the diffuse downwelling irradiance attenuation spectra, Kd, as well as the photosynthetically available radiation, PAR. The spectra were further spectrally subsampled to (a) MERIS band central wavelengths (412.5, 442.5, 490, 510, 560, 620, 665, 681.25, 708.75, 753.75, 761.875, 865, 885, and 900 nm), (b) MODIS bands (412, 443, 469, 488, 531, 547, 645, 667, 678, 748, 859, and 869 nm) and (c) SeaWiFS bands (412, 443, 490, 510, 555, 670, 765, and 865 nm). In the following, only spectra generated at MERIS bands are presented.

## 3 Results and discussion

The distributions of water depth, temperature and salinity, CHL and TSM concentrations, IOPs, and AOPs are presented in Sects. 3.1–3.6, followed by the analysis of the covariation between CHL and TSM and bio-optical relationships existing in the CCRR data sets (Sect. 3.7).

The distributions of CHL, TSM, and IOPs in the match-up field data set and the in situ reflectance data set are related to the AOPs measured throughout the CoastColour sites. The similarities/differences in these relationships characteristic

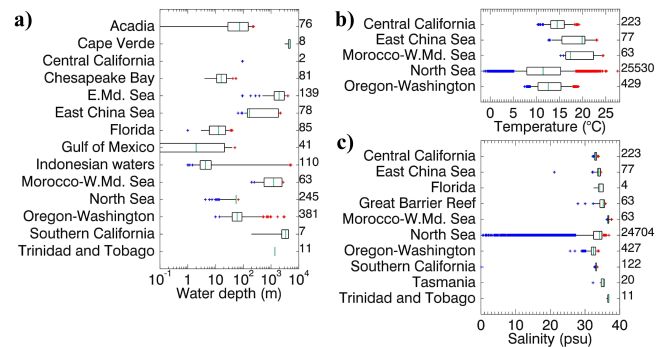**Table 11.** The inherent optical properties as established in CCRRv1.

| Parameter and value | Description | Reference |
|---|---|---|
| $c_{\text{phy}}(660\,\text{nm}) = 0.407\,\text{CHL}^{0.795}$ | Phytoplankton beam attenuation coefficient at 660 nm | Loisel and Morel (1998) |
| $\gamma_{\text{CHL}>2} = 0$<br>$\gamma_{\text{CHL}\leq2} = 0.5\log_{10}(\text{CHL}) - 0.3$ | Spectral variation in $c_{\text{phy}}$ (power law exponent) | Morel et al. (2002) |
| $\beta_{\text{phy}}(\lambda)$ : Fournier–Forand | Phytoplankton scattering phase function with $b_{b\text{phy}}/b_{\text{phy}} = 0.006$ | Similar to Morel et al. (2002) |
| $a_{\text{p}}^*(\lambda) = A(\lambda)\text{CHL}^{B(\lambda)}$ | Phytoplankton specific absorption coefficient | Bricaud et al. (1998) |
| $b_{\text{MP}}^*(555\,\text{nm}) = 0.51\,\text{m}^2\text{g}^{-1}$ | Specific scattering coefficient for MP | Babin et al. (2003a) |
| $\beta_{\text{MP}}(\lambda)$ : Petzold | MP scattering phase function | Mobley (1994) |
| $a_{\text{MP}}^*(443\,\text{nm}) = 0.04\,\text{m}^2\text{g}^{-1}$ | Specific absorption coefficient for MP | Babin et al. (2003b) |
| $S_{\text{MP}} = -0.0123\,\text{nm}^{-1}$ | Spectral slope of $a_{\text{MP}}^*$ (exponential) | Babin et al. (2003b) |
| $\gamma_{c\text{MP}} = -0.3749$ | Spectral variation in the beam attenuation coefficient for MP (power law), giving $b_{\text{p}}^{715}/b_{\text{p}}^{555} = 0.925$ | In agreement with Babin et al. (2003a) |
| $S_{\text{CDOM}} = -0.0176\,\text{nm}^{-1}$ | Spectral slope of $a_g$ (exponential) | Babin et al. (2003b) |
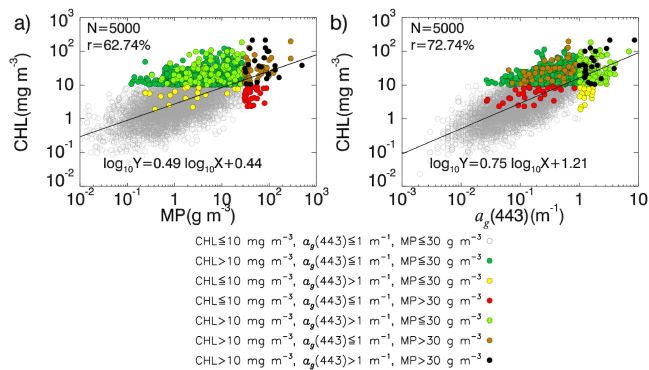


**Figure 5.** The simulated **(a)** MP and **(b)** $a_g(443)$ vs. the simulated CHL concentrations, in the CCRRv1. The colours represent the ranges of MP, CHL, and $a_g(443)$ as reported in the key above.



**Figure 6.** The distribution of **(a)** water depth (m), **(b)** temperature (°C), and **(c)** salinity (psu) as given in the in situ data set at all available depths. The black boxes delimit the 25th and 75th percentiles of the data and the black horizontal lines show the extension of up to the 5th and 95th percentiles. The green line represents the median value and the blue (red) "+" the minimum (maximum) plot values below (above) the 5th (95th) percentile. The number of measurements taken at each test site is reported on the right axis of the graph. The scale is logarithmic for the water depth.

of these sites may shed light on the common (universal) bio-optical relationships and/or emphasize some more regional features, which is of interest for remote-sensing algorithm development and validation. The bio-optical relationships within the match-up field and in situ data sets are also compared to the models, as well as to the ranges of TSM, CHL, and CDOM concentrations assumed in the simulated CCRRv1.

## 3.1　Water depth, temperature, and salinity

The CoastColour sites are characterized by different distributions of water depth, temperature, and salinity (Fig. 6). The median water depth varies from 2 m in the Gulf of Mexico to more than 1000 m in the Morocco-W. Md. Sea, Trinidad and Tobago, E. Md. Sea, southern California, and Cape Verde sites (Fig. 6a). The sea surface temperature in the North Sea ranges from −0.6 to 26 °C, encompassing the ranges of tem-

perature reported at the four other sites (Fig. 6b), probably due to the quasi-continuous sampling in the North Sea throughout the cold and warm seasons (Fig. 2). The frequent sampling of salinity in the North Sea across seasons is exhibited in the large range of this measurement (0.5–37 psu). About 82 % of salinity data measured in the CoastColour sites exceed 32 psu (Fig. 6c).

　　Note, however, that these distributions may not represent all the conditions within which the entire in situ measurements were collected, since the time windows of the metadata (excluding the date, time, and geographic coordinates) do not always cover those of the measurement of the biogeochemical data, IOPs, and AOPs (Figs. 2 and 3).
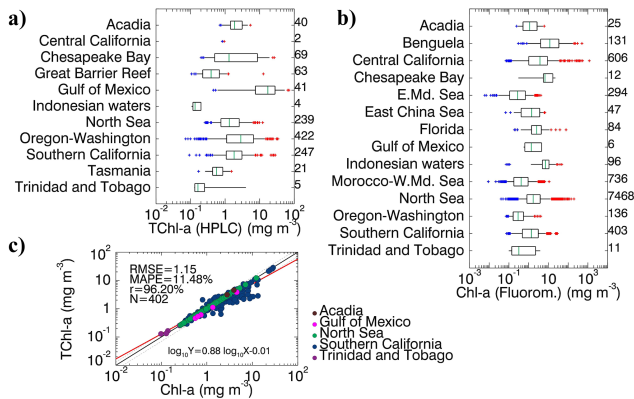
**Figure 7.** The distribution of **(a)** TChl $a$ and **(b)** Chl $a$ concentrations (in mg m$^{-3}$) as given in the in situ data set at all measurement depths, and **(c)** Chl $a$ vs. TChl $a$. The number of measurements taken at each test site is reported on the right axis of the graph. The graphical convention in panels **(a)** and **(b)** is identical to Fig. 6. In panel **(c)** the solid line represents the 1 : 1 ratio, the dashed lines $\pm 30\,\%$, and the red line the linear regression fitting the log-transformed TChl $a$ and Chl $a$ measurements.

## 3.2 Chlorophyll $a$ concentration

TChl $a$ (HPLC method) and Chl $a$ (fluorometric method) span from 0 to extremely high values ($> 1000\,\text{mg m}^{-3}$) in the central California site (Fig. 7a, b). TChl $a$ values vary by about 2 orders of magnitude in most of the sites. The low number of TChl $a$ measurements ($\leq 5$) in the data from the Indonesian waters and Trinidad and Tobago sites may explain the reduced variability observed there. With the higher number (temporal and spatial coverage) of Chl $a$ measurements, larger ranges of variability are found in the measurements from the Indonesian waters and about 7 orders of magnitude from the measurements taken in the central California site. For most of the sites, Chl $a$ varies at least 3 orders of magnitude.

Chlorophyll $a$ concentrations (either Chl $a$ or TChl $a$) exhibit median values less than 1 mg m$^{-3}$ from the E. Md. Sea, GBR region, Morocco-W. Md. Sea, Tasmania, and Trinidad and Tobago sites. Some of these sites have been extensively studied and characterized as ultra- to oligotrophic (CHL $\leq 1\,\text{mg m}^{-3}$) or mesotrophic to eutrophic waters:

– The eastern Mediterranean Sea is oligotrophic due to nutrient limitations. CHL ranges from $\sim 0.02\,\text{mg m}^{-3}$ in the Cyprus eddy to 0.3 mg m$^{-3}$ during the winter bloom (Groom et al., 2005). Similar ranges of CHL were reported in the ultra-oligotrophic eddies of the western Mediterranean Sea (Loisel et al., 2011).

– In the GBR region the water composition is largely influenced by the land use in the adjacent catchments (Schaffelke et al., 2012). Chlorophyll $a$ concentrations are generally low, with median values ranging from

0.1 mg m$^{-3}$ inshore to 0.25 mg m$^{-3}$ offshore along a cross-shelf gradient (Brodie et al., 2007).

– The eastern Atlantic off the Morocco coast is characterized by nutrient-rich waters (Freudenthal et al., 2002) and by the upwelling regime from April to September. Based on a single vertical profile in the chlorophyll maximum layer off the Moroccan coast in September 1999, the average CHL was estimated at about 1.4 mg m$^{-3}$ (Dolan et al., 2002), while Oubelkheir et al. (2005) found that surface CHL ranged from 0.01 to 3.75 mg m$^{-3}$ during the same cruise; these reported maxima values lie at the upper end (between the 75th and 95th percentiles) of the CCRR match-up field data range collected at 5 m depth.

– In the data from the central California site, the variations in Chl $a$ are primarily determined by sea surface temperature and wind-driven coastal upwelling loading nutrient-rich waters (Chavez et al., 2002). This site exhibits the widest range of CHL variability ($> 6$ orders of magnitude).

In the data from the Acadia, East China Sea, Florida, North Sea, Indonesian waters, Oregon–Washington, and southern California sites, the median CHL ranges from 1 to 10 mg m$^{-3}$. For the Benguela, Chesapeake Bay, and Gulf of Mexico sites, the concentration of chlorophyll $a$ exceeds 10 mg m$^{-3}$. It may exceed 50 mg m$^{-3}$ during algal blooms in the Benguela upwelling system (Probyn, 1985) and reach very high values (CHL $> 500\,\text{mg m}^{-3}$) during a dinoflagellate bloom of *Ceratium balechii* (Pitcher and Probyn, 2011).

The data from Oregon–Washington encompass a wide range of temporal and spatial variability. TChl $a$, collected between April and September during years 2006 to 2010, varies over 3 orders of magnitude, up to 33 mg m$^{-3}$ with a median value of 2.9 mg m$^{-3}$, while Chl $a$ spans from 0.07 to 4 mg m$^{-3}$ during the period July–September 2008 with a median value of 0.3 mg m$^{-3}$. This is due to the productive upwelling season and the low-productivity downwelling season, more productive areas onshore, and less productivity near Oregon than to the north, close to Washington and in the Columbia River plume. It is also possible that variability in the data set is due to slight differences in sampling protocols between the laboratory groups although this would likely be minimal.

In Chesapeake Bay, a distribution similar to the match-up data was described in Tzortziou et al. (2007) based on measurements performed in 2001 where the mean CHL value was about 15 mg m$^{-3}$, and higher CHL values up to 74 mg m$^{-3}$ occurred during spring and summer periods.

Overall, the chlorophyll $a$ match-up data set collected for the CCRR exercise are representative of the distributions reported in the literature. Moreover, the measured Chl $a$ and TChl $a$ in the CoastColour sites show a high correlation ($r = 96.2\,\%$, $N = 402$) with mean absolute percentage error
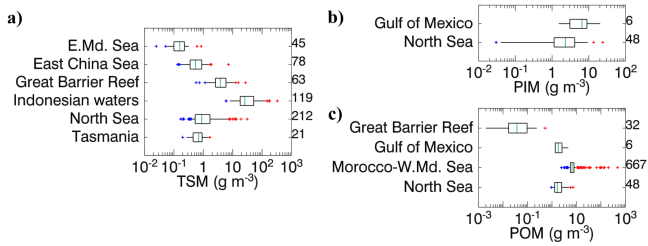
**Figure 8.** The distribution of **(a)** TSM $(\mathrm{g\,m^{-3}})$, **(b)** PIM $(\mathrm{g\,m^{-3}})$ and **(c)** POM $(\mathrm{g\,m^{-3}})$ as given in the in situ data set at all measurement depths. The number of measurements taken at each test site is reported on the right axis of the graphs. The graphical convention is identical to Fig. 6.

(MAPE) equal to 11.5 % (Fig. 7c). Most of the discrepancies between TChl $a$ and Chl $a$ are noticed in measurements from the southern California site. When this site is excluded, a significantly lower MAPE is obtained for the seven sites (MAPE = 3.6 % with correlation $r = 99.8$ %).

## 3.3 TSM, turbidity, Kd, and Kpar

The distributions of TSM are reported in Fig. 8a. PIM and POM concentrations, measured over two and four CoastColour sites respectively, and their distributions are indicated in Fig. 8b and c.

The measurements from the E. Md. Sea show the lowest TSM concentrations (TSM $< 1\,\mathrm{g\,m^{-3}}$, $N = 45$), whereas the region of Indonesian waters exhibits the highest values (median TSM $> 20\,\mathrm{g\,m^{-3}}$, $N = 119$). In Tasmania, TSM varies between 0.1 and $2\,\mathrm{g\,m^{-3}}$ (Cherukuru et al., 2014). The median TSM concentrations observed from the E. Md. Sea, East China Sea, Tasmania, North Sea, GBR region, and Indonesian waters sites are 0.2, 0.6, 0.7, 0.9, 3.8, and $26\,\mathrm{g\,m^{-3}}$ respectively.

Turbidity measurements are provided at two sites (see Fig. 9a). The distribution of turbidity matches that of TSM over the North Sea – likely due to significantly overlapping periods where TSM and turbidity measurements were collected (see the green and red colours in Fig. 2).

The ranges of Kd (443) and Kpar measurements (Fig. 9c, d) show similar differences amongst the Acadia, Cape Verde, Chesapeake Bay, Indonesian waters, Morocco-W. Md. Sea, North Sea, southern California, and Trinidad and Tobago sites: the highest mean values are observed in Acadia and Chesapeake Bay (corresponding to the lowest mean values of $Z_{1\%} < 20\,\mathrm{m}$; see Fig. 9b), and the lowest in Cape Verde, Morocco-W. Md. Sea, and the Indonesian waters, which correspond to the highest mean values of $Z_{1\%} > 60\,\mathrm{m}$ found at these three sites.

Kd (and Kpar) values are lower in the Indonesian waters than in the North Sea site. However, the Secchi disk data sets for these two sites, larger than the Kd (and Kpar) data set,
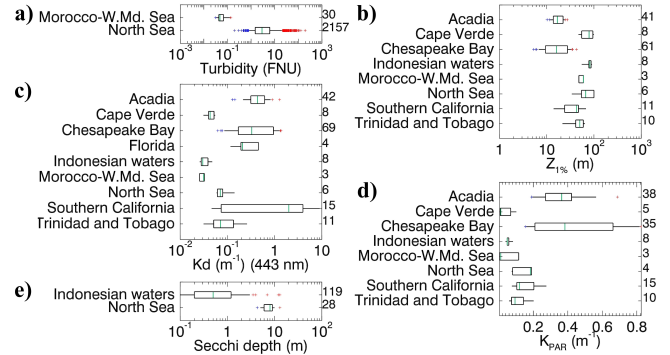


**Figure 9.** The distribution of **(a)** turbidity (FNU for the North Sea and FTU for Morocco-W. Md. Sea), **(b)** the photic depth $Z_{1\%}$ (m), **(c)** Kd at 443 nm $(\mathrm{m^{-1}})$, **(d)** Kpar $(\mathrm{m^{-1}})$, and **(e)** Secchi depth (m). The scale is logarithmic for turbidity and Kd, and linear elsewhere. The number of measurements taken at each test site is reported on the right axis of the graph. The graphical convention is identical to Fig. 6.
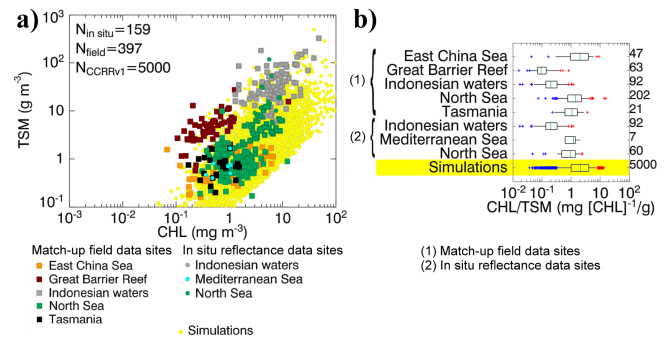


**Figure 10.** The distribution of **(a)** CHL concentrations $(\mathrm{mg\,m^{-3}})$ vs. TSM concentrations $(\mathrm{g\,m^{-3}})$ from the in situ reflectance data set (in the three sites as indicated in the key) plotted as filled circles, and from the match-up data set, including Chl $a$ and TChl $a$ (where available), and the associated match-up field TSM concentrations (in the six sites indicated in the key) and plotted as filled squares, both superimposed on the simulated data (yellow circles). **(b)** CHL / TSM ratio $(\mathrm{mg\,[CHL]^{-1}\,g^{-1}})$ from the match-up, in situ reflectance and simulated data sets. The graphical convention in panel **(b)** is identical to Fig. 6; the yellow colour distinguishes the simulated data set from the in situ measurements.

suggest a higher water clarity in the North Sea than in the Indonesian waters (Fig. 9e).

## 3.4 CHL vs. TSM

The co-variation of CHL with TSM from the in situ reflectance data set at 159 locations (where both CHL and TSM are available) is compared to the co-variation of CHL with TSM from the match-up field data set at 1062 locations. Both co-variations can be visually compared to that of CHL vs. TSM from the simulated data set (Fig. 10a). The distribution
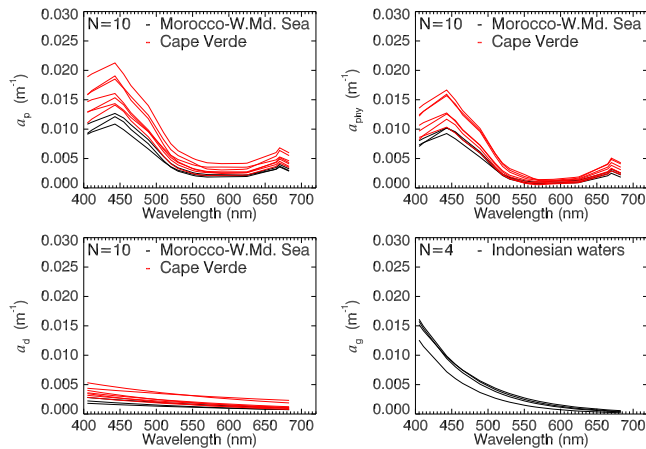
**Figure 11.** The match-up field absorption spectra provided from Morocco-W. Md. Sea, Cape Verde, and Indonesian waters.



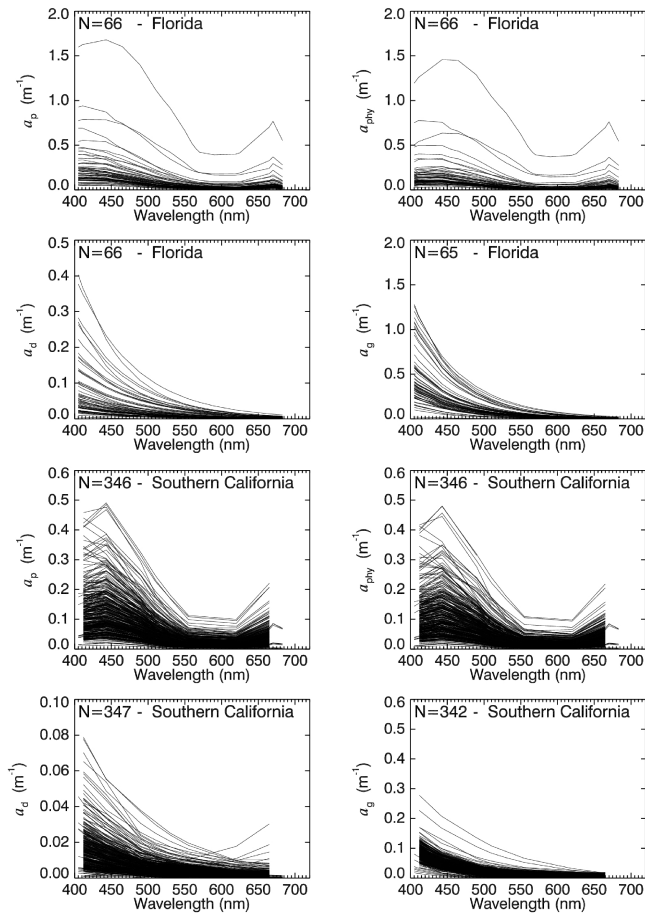**Figure 12.** The match-up field absorption spectra provided from the North Sea site.



**Figure 13.** The match-up field absorption spectra provided from the GBR region and Acadia (upper panel) and Tasmania (bottom) sites.

of the ratio CHL / TSM is shown for the match-up field measurements (Fig. 10b).

The co-variations of CHL and TSM are generally consistent for the majority of in situ and match-up test sites, showing a general tendency of CHL increasing with TSM, as reported in Babin et al. (2003b, their Fig. 2). The simulated data fit better the distributions of CHL and TSM collected in the North Sea, since their models adopted the distributions documented in Babin et al. (2003b), based on measurements taken in European coastal waters including the North Sea.

As previously reported in Sects. 3.2 and 3.3, various CHL and TSM ranges are observed in the match-up field measurements throughout the CoastColour sites (the GBR region, the North Sea, and Tasmania coastal waters). The in situ reflectance data set showed differences in CHL and TSM ranges between the Indonesian waters, the North Sea, and the Mediterranean Sea (Fig. 10).

The simulated data encompass all the ranges covered by the in situ CHL and TSM (from the in situ reflectance data set), and partially the ranges of the match-up field data: excluding few measurements collected in the Indonesian waters and GBR region sites associated with very low CHL / TSM ratio (see Fig. 10b).

Large variability in the ratio CHL / TSM from the match-up field measurements is noticeable amongst the six Coast-Colour sites, spanning over 3 orders of magnitude (see Fig. 10b). The GBR region and Indonesian coastal waters sites present the lowest median value of the ratio CHL / TSM (from 0.1 to 0.2 mg $[CHL]^{-1} g^{-1}$), being approximately 10 times lower than the median magnitudes measured in the North Sea and Tasmania (around 1.4 and 1.1 mg $[CHL]^{-1} g^{-1}$ respectively). The East China Sea site exhibits the highest median value of CHL / TSM of 2 mg $[CHL]^{-1} g^{-1}$.

Identical CHL and TSM data (92 pairs) from the Indonesian waters are available both in the in situ reflectance
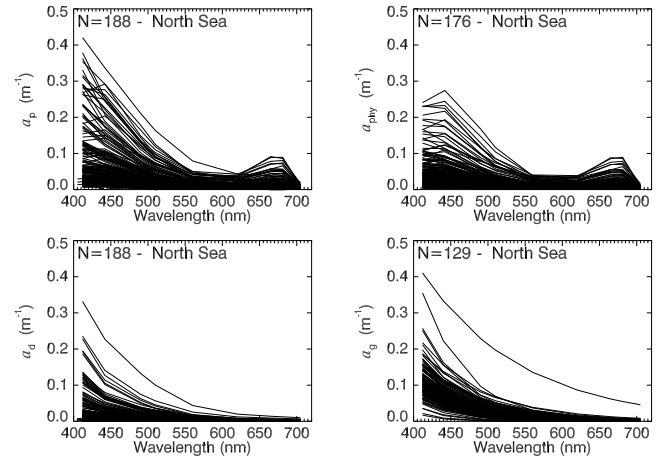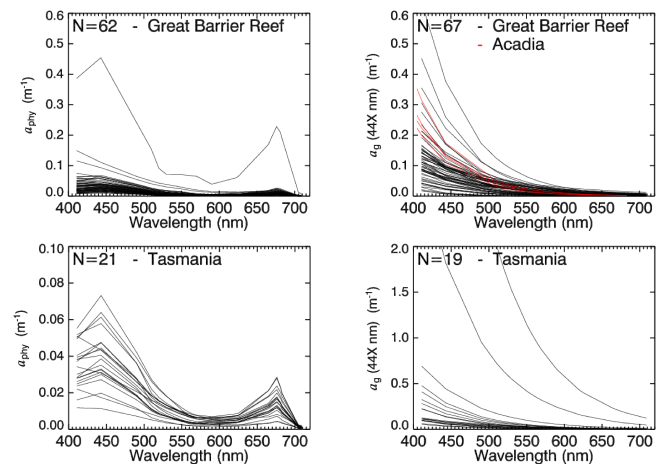
data set and the match-up field measurements, giving identical distributions of CHL vs. TSM and CHL / TSM ratio. From the Mediterranean Sea site, the in situ reflectance data set collected during March 2009 shows a median value of CHL / TSM of 1.2 mg $[CHL]^{-1} g^{-1}$ (only seven data). From the North Sea site, the distribution of CHL / TSM ratios in the in situ reflectance data set (60 data) is slightly shifted towards lower values relative to the ratios estimated from the match-up field measurements (202 data).
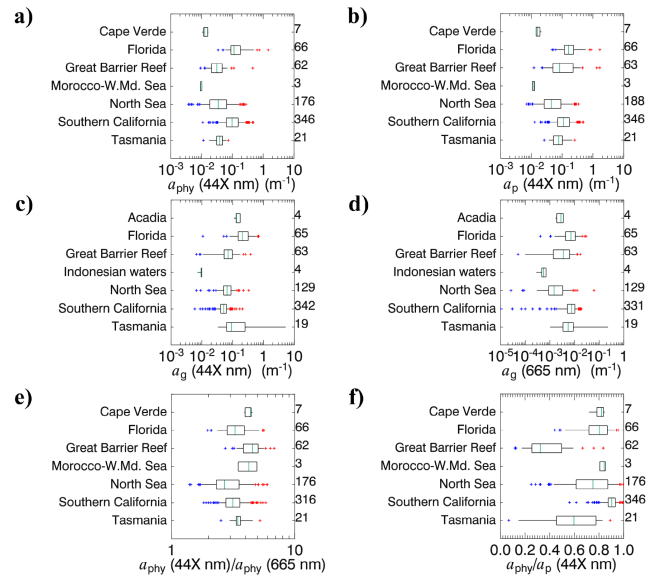
## 3.5 Inherent optical properties

The match-up field absorption coefficient spectra can be classified into four groups, starting from the sites where the lowest amplitudes around 443 nm are observed, to the highest amplitudes: (a) Morocco-W. Md. Sea, Cape Verde, and the Indonesian waters sites (Fig. 11); (b) the North Sea (Fig. 12);

**Figure 14.** The match-up field absorption spectra provided from the Florida (upper panel) and southern California (bottom) sites.



**Figure 15.** The distributions of (**a**) $a_{phy}(44X)$, (**b**) $a_p(44X)$, (**c**) $a_g(44X)$, (**d**) $a_g(665)$, (**e**) the ratios $a_{phy}(44X)/a_{phy}(665)$, and (**f**) $a_{phy}/a_p(443)$ measured at the CoastColour sites. When coefficients at wavelength 443 nm are missing, they are replaced by data at 440 or 442 nm. The graphical convention is identical to Fig. 6.

(c) the GBR region, Acadia, and Tasmania sites (Fig. 13); and (d) the southern California and Florida sites (Fig. 14). All the absorption coefficient spectra exhibit a large variability at shorter wavelengths (around 443 nm, denoted by 44X to refer to 440, 442, or 443 nm) and for $a_p$ and $a_{phy}$ around the phytoplankton absorption peak at 665 nm.

The median values for the available $a_{phy}(44X)$ data span between 0.01 m$^{-1}$ in the Morocco-W. Md. Sea and Cape Verde sites and 0.1 m$^{-1}$ in the Florida site (Fig. 15a). The median values encountered in the GBR region, North Sea, and Tasmania are between 0.031 and 0.039 m$^{-1}$. Note that the median concentrations of chlorophyll $a$ from the GBR region and Tasmania sites are between 0.4 and 0.6 mg m$^{-3}$, which is 2 to 3 times lower than in the North Sea (1.3 mg m$^{-3}$). This indicates that, on average (for the available measurements sampled), the chlorophyll-specific absorption coefficients around 443 nm from the North Sea are lower than from the other two sites (as a comparison see spectra in Tilstone et al., 2012, for the North Sea and Blondeau-Patissier et al., 2009, for Australia). The different periods of sampling throughout the seasons for each site

(Fig. 3) may partly explain this general discrepancy between the $a_{phy}^*(44X)$ data, which can be highly impacted during algal bloom events.

For $a_p(44X)$ data, the median values are 0.012 and 0.016 m$^{-1}$ in Morocco-W. Md. Sea and Cape Verde sites respectively (Fig. 15b). Noticeably higher median values are observed from the North Sea data (0.044 m$^{-1}$), Tasmania (0.078 m$^{-1}$), and the GBR region (0.083 m$^{-1}$), and exceed 0.1 m$^{-1}$ in the Florida and southern California sites.

The coefficients $a_g$ taken around 443 nm and at 665 nm span over 3 and 4 orders of magnitude respectively throughout the CoastColour sites (Fig. 15c, d). The Florida measurements exhibit the highest median values around 443 nm, exceeding 0.2 m$^{-1}$ with a high median value of 0.006 m$^{-1}$ at 665 nm. Conversely, the southern California measurements show the highest median value of $a_g$ around 665 nm exceeding 0.007 m$^{-1}$, with a significantly low median value of 0.05 m$^{-1}$ at 443 nm. Note the overall similar distributions of $a_g(443)$ from the GBR region and the North Sea sites (median values around 0.07 m$^{-1}$), while $a_g(665)$ shows a significant shift towards higher values in the data from the GBR region. The Tasmania data contain two extreme spectra of $a_g$ (see Fig. 13), which slightly increases the median value for this site up to 0.09 m$^{-1}$, above the value observed in the GBR region.

The ratio $a_{phy}(44X)/a_{phy}(665)$ shows the lowest median value of 2.7 in the North Sea and the highest value of 4.6 in the GBR region (Fig. 15e), which is inversely related to the distribution of CHL / TSM: the highest median
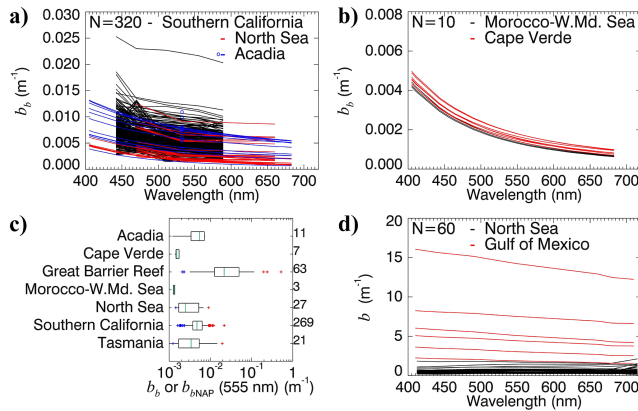
**Figure 16.** The spectra of **(a)** $b$ (m$^{-1}$) measured in the Gulf of Mexico and North Sea sites and **(b)** $b_b$ (m$^{-1}$) measured in five Coast-Colour sites (note that the coefficients from the southern California were limited to the spectral range 442–589 nm).



**Figure 17.** The distributions of **(a)** $b_b/b$, **(b)** $b_{bNAP}^*$ (m$^2$ g$^{-1}$) and **(c)** $c$ (660) (m$^{-1}$). Note the different scaling used for these plots. The graphical convention in Fig. 6 is used.

value of CHL / TSM of 1.5 (mg [CHL]$^{-1}$ g$^{-1}$) is observed in the North Sea (Fig. 10b), while a lower value of $\sim 0.1$ (mg [CHL]$^{-1}$ g$^{-1}$) is found in the GBR region.

The distributions of $a_{phy}/a_p(440)$ in the GBR region, North Sea, and Tasmania sites (see Fig. 15f) nearly follow the distributions of the associated CHL / TSM, marked by the lowest median values in the GBR region, Tasmania, and the North Sea sites (0.32, 0.60 and 0.75 respectively), and higher values ($> 0.80$) in the Cape Verde, Florida, Morocco-W. Md. Sea, and southern California sites. The large variability in CHL / TSM in the North Sea and GBR region (spanning over 3 and 2 orders of magnitude; see Fig. 8) can be related to the high variability in $a_{phy}/a_p(440)$ (about 10-fold magnitudes). From the Morocco-W. Md. Sea site, the number of $a_{phy}/a_p(440)$ data is too low ($N = 3$) compared to the number of CHL / TSM measurements ($N = 665$), yielding a mismatch between both distributions.

The southern California measurements of $b_b$ show the highest variability in shapes and amplitudes, with values at 555 nm spanning from 0.0016 to 0.0216 m$^{-1}$, encompassing the ranges of $b_b$ measurements from the Acadia and North Sea sites (Fig. 16a). From the Cape Verde and Morocco-W. Md. Sea sites, only 10 $b_b$ spectra are available, lying at the bottom limit of $b_b$ measurements from the three previous sites (Fig. 16b). The noticeable shift between the ranges of $b_b$ measured in the Acadia and Cape Verde sites may partly explain the shift between Kd (or Kpar) in Acadia and Cape Verde (Fig. 9a).

The distributions of the total backscattering coefficients from Acadia, Cape Verde, Morocco-W. Md. Sea, North Sea, and southern California, and of non-algal particles' backscattering coefficients collected at the GBR region and Tasmania sites at 555 nm, are presented in Fig. 16c. Quite similar median values of $b_b(555)$ are observed in Acadia, southern
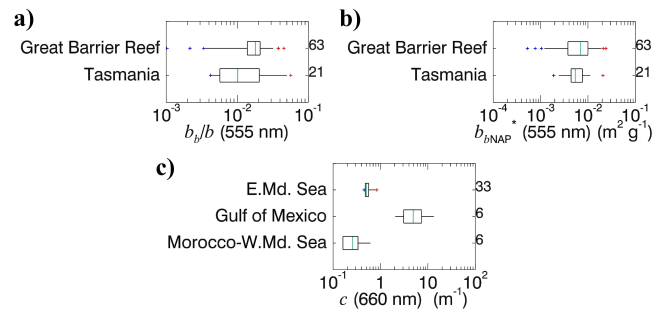
California, and Tasmania coastal waters, being respectively 0.0041, 0.0040, and 0.0034 m$^{-1}$ (Fig. 16c).

In the GBR region, the coefficient $b_{bNAP}(555)$ spans over 3 orders of magnitude around the highest median value 0.021 m$^{-1}$. The distributions of $b_{bNAP}(555)$ coefficients and TSM (see Fig. 8) differ notably between Tasmania and GBR region.

Lower $b_b$ values are found in the Cape Verde and Morocco-W. Md. Sea sites, where only a few backscattering measurements ($< 10$) are available, showing a limited variability.

The total scattering coefficients provided in the North Sea and Gulf of Mexico exhibit high relative variability in the two sites, with the highest amplitudes measured in the Gulf of Mexico (Fig. 16d).

The scattering to backscattering ratio $b_b/b$ at 555 nm and the mass-specific non-algal particulate backscattering, $b_{bNAP}^*$, at 555 nm are available exclusively from the Tasmanian and GBR coastal waters (Fig. 17a, b). Most of $b_b/b(555)$ values from the GBR region lie above the 75th percentile of $b_b/b$ measurements in Tasmania, their respective median values being 0.02 and 0.01 (Fig. 17a). Although different distributions are described by $b_{bNAP}(555)$ coefficients and TSM from the GBR region and Tasmania sites, the range of $b_{bNAP}^*(555)$ observed from the Tasmania site is within that observed from the GBR region (Fig. 17b), which spans from $5 \times 10^{-4}$ to $5 \times 10^{-2}$ m$^2$ g$^{-1}$. Generally similar median values are found in the GBR region and Tasmania site: 0.0053 and 0.0075 m$^2$ g$^{-1}$ respectively.

The beam attenuation coefficients measured in the Gulf of Mexico cover the spectral range 410 to 710 nm; those measured in the E. Md. Sea are given at 470, 660, and 670 nm; and those in Morocco-W. Md. Sea are at 660 nm (only coefficients at 660 nm are reported here, Fig. 17c). The coefficients $c(660)$ span over 3 orders of magnitude, ranging from 0.04 to 0.9 m$^{-1}$ in Morocco-W. Md. Sea and the E. Md. Sea (with median values being respectively 0.12 and 0.46 m$^{-1}$) and from 2 to 13 m$^{-1}$ in the Gulf of Mexico (with a median value equal to 4.91 m$^{-1}$).
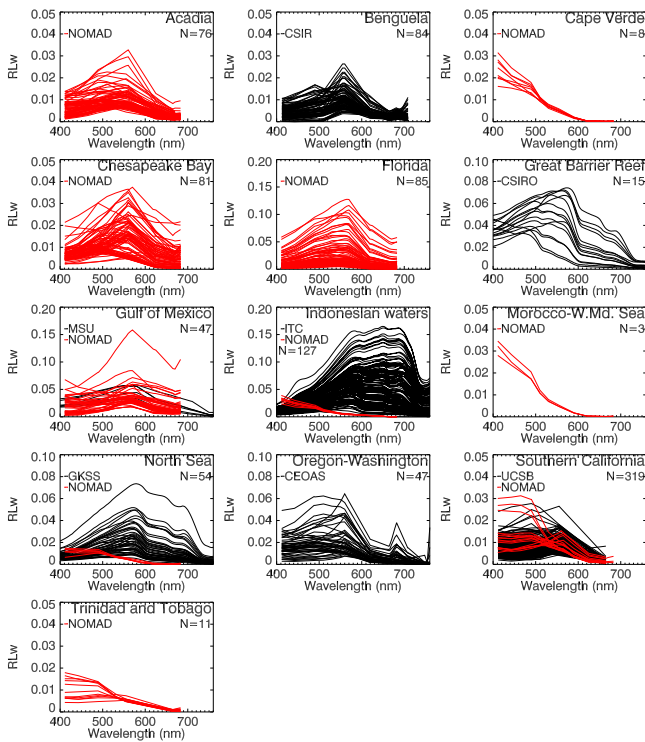
**Figure 18.** Match-up field RLw provided from 13 CoastColour sites by the eight data providers indicated in the figures. Note the different scales used for the sites.

Concurrent measurements of $a_g(443)$, TSM, and CHL available at the GBR region, Tasmania, and North Sea sites, and of $a_g(443)$, $a_{phy}(443)$, and $a_d(443)$ measured at the GBR region, Tasmania, southern California, and Florida sites, show large variability in water optical properties and biogeochemical parameters, covering the case 1 and case 2 waters (see ternary plots in the Supplement). Note, however, that these plots do not provide an accurate overview of the water masses sampled for this study.

## 3.6 Water-leaving radiance reflectance

The match-up field reflectance measurements, the MERIS RLw (both in the CCRR match-up data set), and the in situ reflectance spectra (in the CCRR in situ reflectance data set) are presented successively in Figs. 18–20. Note that the percentage of RLw data per site available from the in situ measurements (i.e. the match-up field and the in situ reflectance data sets) is different from that of the MERIS RLw data set: 31, 13, and 12 % of the in situ RLw data are provided from the southern California, Benguela, and Indonesian waters respectively; about 8 % from the Chesapeake Bay and Florida sites; and less than 7 % of RLw data from the North Sea site, while 80 % of MERIS RLw measurements are provided for the North Sea.
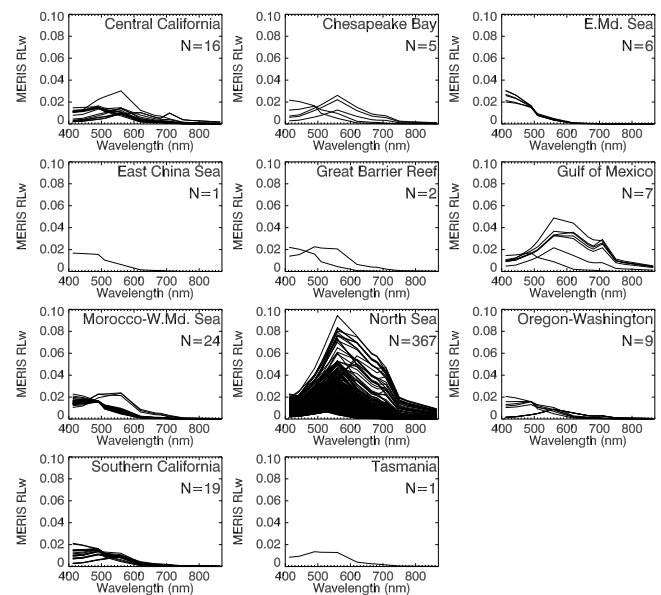


**Figure 19.** MERIS RLw provided in the CCRR match-up data set for the 11 CoastColour sites.
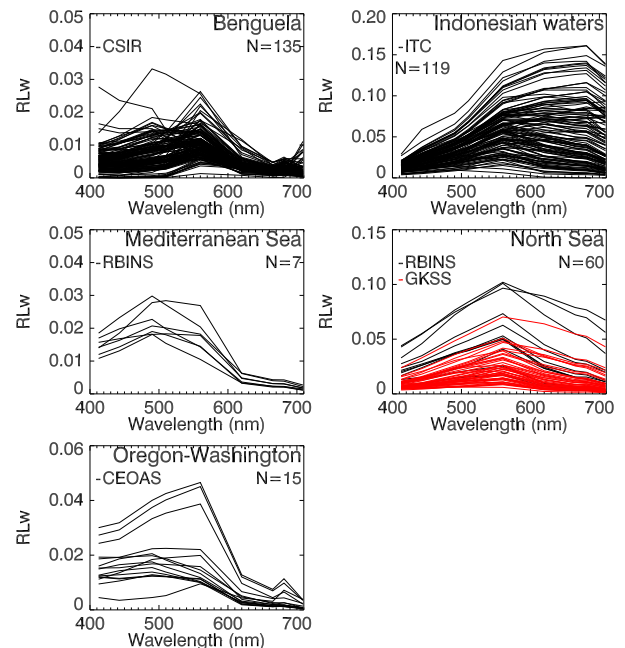


**Figure 20.** The in situ reflectance spectra provided from five Coast-Colour sites. Note the different scales used for the sites. The spectra from the North Sea site coloured in red are provided by GKSS.

The match-up field RLw measurements from the southern California, Morocco-W. Md. Sea, and Benguela sites present the generally lowest amplitudes amongst the CoastColour sites, where more than 75 % of the RLw values at 555 nm, RLw(555), are less than 0.01, and only 6 % of the collected spectra have RLw(555) > 0.02. For the southern California

Earth Syst. Sci. Data, 7, 319–348, 2015

www.earth-syst-sci-data.net/7/319/2015/

site, this is in agreement with the extremely high absorption coefficients reported earlier (see Fig. 14); from the Morocco-W. Md. Sea site, the three low match-up field reflectance spectra observed during October 2005 can be associated with the three relatively low absorption spectra measured during the same period (Fig. 11); and for the eutrophic waters of the Benguela site, these low reflectances can be explained by the high phytoplankton absorption and possibly high detrital and/or CDOM absorption. Note that MERIS RLw spectra for the southern California and Morocco-W. Md. Sea sites show ranges of RLw amplitudes comparable to those of the match-up field spectra.

The MERIS RLw measurements from the E. Md. Sea site show low RLw(555) < 0.01 and higher values at 412 nm ranging from 0.015 to 0.03 and inversely the in situ RLw measurements are slightly higher, up to 0.027, with relatively lower RLw at 412 nm < 0.019. These generally low reflectance values observed by MERIS and in situ are related to the clear oligotrophic waters of the Mediterranean Sea (95 % of Chl $a$ data are less than 2 mg m$^{-3}$, with a median value of 0.3 mg m$^{-3}$; Fig. 7). The difference in the spectral shapes between MERIS and in situ measurements can be explained partly by the different periods of observations which were conducted in March 2009 for the in situ data: while only two measurements were available from MERIS in March 2008, all the other measurements were collected during September 2008 and May and October 2005, outside of the blooming period (Barale et al., 2008).

For the Chesapeake Bay, both the 5 MERIS RLw and 81 match-up field RLw spectra exhibit values less than 0.04 at all wavelengths, with 27 % of match-up field RLw(555) higher than 0.02.

From the North Sea site, most of the match-up field (Fig. 18), MERIS RLw (Fig. 19), and in situ spectra (Fig. 20) show a peak around 550–570 nm, not exceeding 0.05, with 50 % of the RLw(555) above 0.02. Lower reflectances are measured at shorter wavelengths (< 450 nm) associated with the higher CDOM and particles absorption in this spectral range (Fig. 12).

The match-up field RLw(555) measurements from the GBR region lie in a range comparable to that observed from the North Sea, but with a significantly different distribution: 86 % of RLw(555) measurements exceed 0.02. With respect to the North Sea RLw spectra, the spectral shapes and magnitudes of RLw from the GBR region are also markedly different in the blue spectral range. This may be attributed to the notable difference between the spectral shapes and magnitudes of phytoplankton absorption coefficients measured in the North Sea site (Fig. 12) and in the GBR region (Fig. 13), where the concentration of CHL (Chl $a$ or TChl $a$) is generally 4 times lower than in the North Sea (Fig. 7).

Amongst the 47 match-up field reflectance measurements collected at the Gulf of Mexico site, 95 % of RLw measurements around 555 nm are higher than 0.02, 34 % range from 0.03 to 0.05, and one extreme value (> 0.15) is reported. The
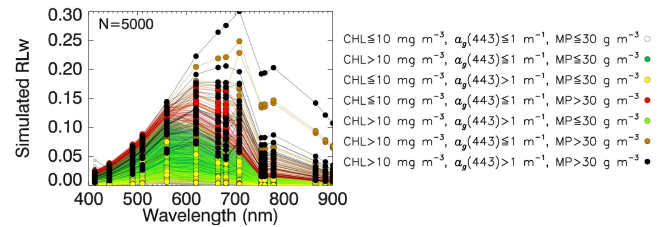


**Figure 21.** The simulated reflectance spectra in the CCRRv1. The colours represent the ranges of MP, CHL, and $a_g(443)$ as reported in the key above.

peak of chlorophyll absorption is noticeable around 665 nm on the NOMAD reflectance spectra, which can be related to the generally high Chl $a$ measurements (75 % of TChl $a$ are higher than 7.6 mg m$^{-3}$ and the median value is 17 mg m$^{-3}$; see Fig. 7).

The 127 match-up field reflectance spectra collected from the Indonesian waters exhibit the highest variability in the amplitudes in the red and near-infrared spectra range, with some values exceeding 0.1 around 700 nm, likely due to high TSM concentrations as shown in Fig. 8 (TSM may exceed 100 g m$^{-3}$, with a median value about 25 g m$^{-3}$). At 555 nm, 96 % of the RLw data are above 0.02. Most of the spectra show a minimum around the chlorophyll absorption peak, which can be related to Chl $a$ distribution in these waters with a median value about 7 mg m$^{-3}$ (Fig. 7).

The 47 Oregon–Washington match-up field spectra exhibit a high variability in reflectance at 412 nm, with values ranging from $3 \times 10^{-3}$ to $5 \times 10^{-2}$, with varying spectral shapes depicting the high spatial and time variability in phytoplankton concentrations noted earlier for that site (Sect. 3.2).

The simulated reflectance spectra in the CCRRv1 data set are presented in Fig. 21, and related to the ranges of the simulated CHL, MP, and $a_g(443)$ via their colours (as indicated in the key).

The comparison between the CCRRv1 and the IOCCG Algorithm Working Group simulated data (IOCCG, 2006) indicates that the ranges of the total absorption coefficient at 440 nm, $a(440)$, and the remote-sensing reflectance, Rrs(440), in CCRR are globally within those of IOCCG (Fig. 22a), with a few points of higher total absorption coefficient (maximum in CCRR is $a(440) = 23.6$ m$^{-1}$). While more variability in the reflectance of CCRR for the mid- and high ranges of absorption is noted, the ranges of the reflectance band ratios 410 : 440 and 490 : 555 of CCRR are within those of the IOCCG data (Fig. 22b). The large variability in the CCRR reflectance is mainly due to the extended ranges of MP and CHL towards higher concentrations, yielding extended ranges of particle backscattering.

The distribution of reflectance products from the three CCRR data sets is examined through the following:

– RLw band ratio 490 : 555 vs. RLw band ratio 412 : 443 from reflectance measurements in the match-up field

www.earth-syst-sci-data.net/7/319/2015/

Earth Syst. Sci. Data, 7, 319–348, 2015
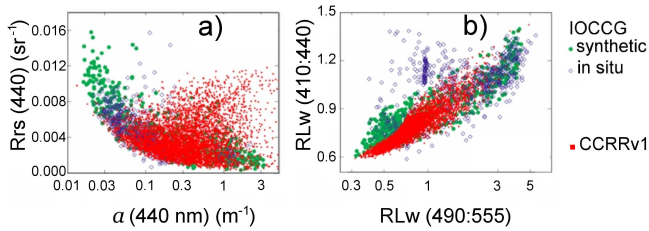
**Figure 22.** Comparison between IOCCG (green, reproduced from Fig. 2.3 in IOCCG, 2006), and CCRRv1 (red) simulated data sets. **(a)** Variations of the remote-sensing reflectance Rrs (440) with $a$(440), **(b)** variations in Rrs band ratio $410:440$ with respect to Rrs band ratio $490:555$. The blue diamonds represent the NOMAD (Werdell and Bailey, 2005) subset of in situ data extracted from the SeaBASS data set and used in the algorithm testing in IOCCG (2006).



**Figure 23.** RLw band ratio $490:555$ vs. RLw band ratio $412:443$ **(a)** and vs. RLw band ratio $709:665$ **(b)** within the match-up field data set, RLw band ratio $490:560$ vs. RLw band ratio $412:443$ **(c)** and vs. RLw band ratio $709:665$ **(d)** within the MERIS RLw products of the match-up data set, and in the in situ reflectance data set **(e, f)**. The yellow circles represent the simulated data set.

data set (Fig. 23a) and vs. RLw band ratio $709:665$ (Fig. 23b). Note that since most of the match-up field measurements contain RLw at 555 nm, that band is chosen instead of MERIS band 560 nm, where only few data are available.

– RLw band ratio $490:560$ vs. RLw band ratio $412:443$ (Fig. 23c) and vs. RLw band ratio $709:665$ (Fig. 23d) from the MERIS RLw products of the match-up data set.

– RLw band ratio $490:560$ vs. RLw band ratio $412:443$ and vs. RLw band ratio $709:665$ from the in situ reflectance data set (Fig. 23e, f).

From the match-up field data set, fewer measurements are available at band 709 nm (only 312 points). There is a general consistency in the distribution of RLw band ratios $709:665$ (respectively $412:443$) vs. $490:560$ from the three data sets except for the in situ reflectances measured in the Benguela waters, which exhibits a high ratio of RLw(412) / RLw(443) > 1 in the lower range of reflectance band ratio $490:560$, likely due to the hypertrophic nature of these waters.

Apart from the extreme ranges of reflectance ratios collected from the Benguela site, the large scatter of points observed in the in situ reflectance data set from the Gulf of Mexico, the Mediterranean Sea, and the North Sea (Fig. 23e and f) is drastically reduced in the match-up data set (that is, MERIS RLw) as shown in Fig. 23c and d. The distribution of the ratio RLw(709) / RLw(665) derived from the simulated data set better reproduces the ranges covered by the in situ data set (> 70 %) and by the match-up data set (> 95 %) than the distribution of the simulated RLw(412) / RLw(443). This is mainly attributable to the fact that the reduced variability in phytoplankton and CDOM inherent optical properties modelled in the simulated data set does not represent the large natural variability in these IOPs, which greatly affect RLw
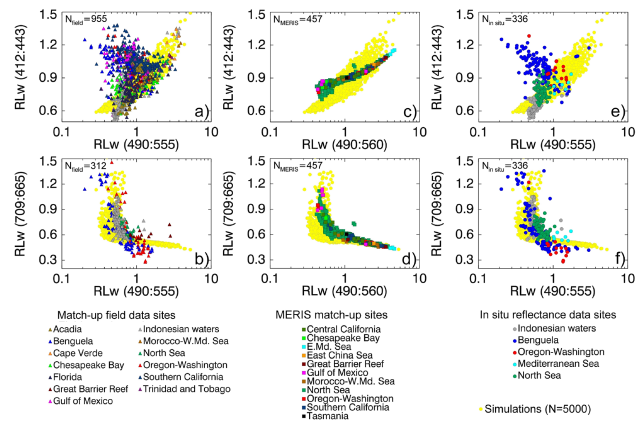
particularly at shorter wavelengths with the effect lessening at longer wavelengths.

## 3.7 Bio-optical relationships

For the comparison of MERIS and the match-up field data, only concurrent data (i.e. within ±1 h of MERIS overpass) are considered. Moreover, only match-up field data measured at depths less than 2 m are taken into account since in situ data collected at larger depths are not correlated with the surface remote-sensing signal (e.g. in the case of stratified waters).

The number of match-up field TSM and TChl $a$ data measured above 2 m depth, concurrent with reflectance measurements, is 48 and 322 respectively. The number of concurrent in situ IOPs and RLw measurements is low (three from the North Sea and Morocco-W. Md. Sea sites, four from Acadia and the Indonesian waters, six from the Gulf of Mexico, and seven from the Cape Verde site), except from the southern California and Florida sites (313 and 66 data respectively). Furthermore, no IOP parameter is available from all these sites (see Table 3b). In the following, the analyses are focused on the distributions of CHL, TSM, and reflectance data and their relationships within the match-up, in situ, and simulated data sets.

### 3.7.1 CHL vs. RLw

An overview of the optical conditions and CHL ranges covered by the in situ measurements is given by Fig. 24.

Figure 24 presents a scatter plot of CHL vs. RLw band ratio $709:665$ which shows that the highest CHL concentrations are exhibited during phytoplankton blooms in the
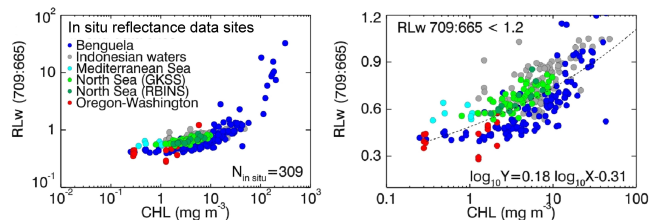
**Figure 24.** From the CCRR in situ reflectance data set: CHL $(mg\,m^{-3})$ vs. RLw band ratio 709 : 665 with a close-up on the lower range of RLw 709 : 665 < 1.2 presented in the left figure.
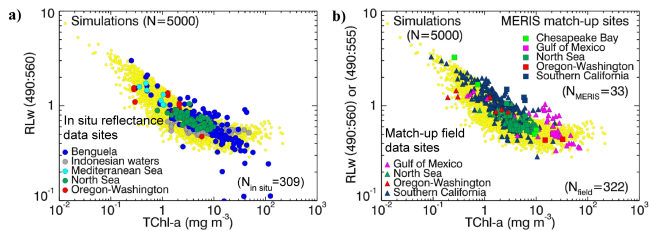


**Figure 25.** Reflectance band ratio 490 : 560 vs. CHL concentrations $(mg\,m^{-3})$ from **(a)** the in situ reflectance data set and from **(b)** MERIS RLw products and match-up field RLw measurements (using band 555 nm instead of 560 nm) with the associated match-up field TChl $a$ concentrations, both superimposed on the simulated reflectance band ratio 490 : 560 vs. the simulated CHL (yellow circles). Match-up field TChl $a$ data are restricted to measurements collected within 1 h after/before the time of MERIS overpass, and to the maximum measurement depth of 2 m.

Benguela waters, where the RLw band ratio 709 : 665 is the most sensitive to CHL variations. For CHL < 10 $mg\,m^{-3}$, the measurements from the Oregon–Washington and Benguela sites contain globally lower values of RLw band ratio than in the other sites (left graph in Fig. 24).

The relationship between CHL and reflectance band ratio 490 : 560 is quite consistent throughout the CCRR data sets (Fig. 25), except for the in situ measurements from the Benguela site (blue filled circles in Fig. 25a). This is due to the very high CHL > 100 $mg\,m^{-3}$ present in the Benguela site, as previously noted in Fig. 24c, as well as the match-up field data set in Fig. 7b associated with low RLw ratios outlying the rest of data, and spanning from 0.2 down to 0.02.

### 3.7.2 TSM vs. RLw

Figure 26 shows the distribution of RLw at 620 nm as a function of TSM concentrations, plotted using different colours for each of the CoastColour sites: Indonesian waters, Mediterranean Sea, and the North Sea (light- and dark-green colours are used for the North Sea region to distinguish the data provider). Linear regression is applied and the associated equations are reported in the same figure. The slopes of the regression lines range from 0.48 to 0.81. The associated goodness of fit coefficients are 39, 59, 89, and 67 % for
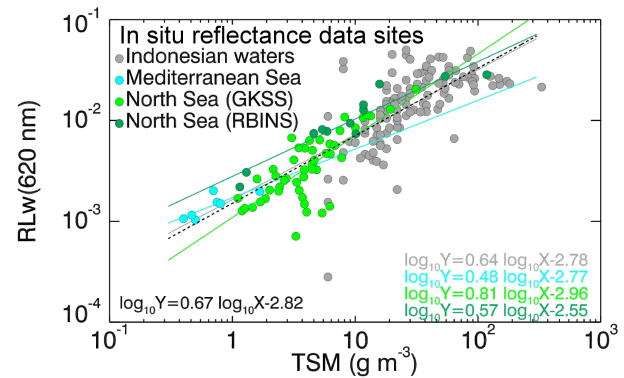


**Figure 26.** TSM $(g\,m^{-3})$ vs. RLw (620 nm) from the CCRR in situ reflectance data set and their associated regression lines.

the Indonesian, North Sea (GKSS), North Sea (RBINS), and Mediterranean Sea measurements respectively. In the GKSS data, most of the scatter occurs at low TSM ranges < 7 $g\,m^{-3}$, whereas the measurements from the Indonesian site, taken mainly in highly turbid waters (average and median values being 41.6 and 26 $g\,m^{-3}$ respectively), show a global scatter. This scatter can be due to a high variability in the specific inherent optical properties of particles, caused by varying particles size, and/or composition within the sites. The scatter may also be impacted by mismatches of RLw and/or TSM measurements in water with high spatiotemporal variability.

The regression line fitting all the data is given by $\log_{10}(\mathrm{RLw}(620)) = 0.67\log_{10}(\mathrm{TSM}) - 2.82$ (shown by a dashed black line in Fig. 26) with $R^2 = 69\%$.

The distribution of the reflectance as a function of TSM concentrations in the three CCRR data sets is presented by the scatter plots of TSM vs. the reflectance at 665 nm (Fig. 27a, b) and of TSM vs. the reflectance band ratio 665 : 490 (Fig. 27c, d). The simulated TSM vs. RLw(665) scatter plots follow the trend of the match-ups and in situ data sets, with a scatter indicating either a variable particulate mass-specific backscattering coefficient different from that assumed in the simulations (Table 11) or measurement errors.

When using the reflectance band ratio, there is a significantly larger scatter of the simulated data due to the effects of CDOM and phytoplankton absorption affecting the reflectance at 490 nm, whereas this scatter is less noticeable in the in situ or match-up data of the five regions (Fig. 27c, d), which indicates less variability in CDOM absorption coefficient and phytoplankton concentrations in these measurements than in the modelled data set.

The general shift between the in situ reflectance data set and match-up data (Fig. 27a, b) is not noticeable in the reflectance ratios (Fig. 27c, d), suggesting similar 665 : 490 absorption coefficient ratios in the two data sets, at least in the North Sea regions; at lower TSM range (< 10 $g\,m^{-3}$), the reduced scatter (in the reflectance band ratios vs. TSM, com-
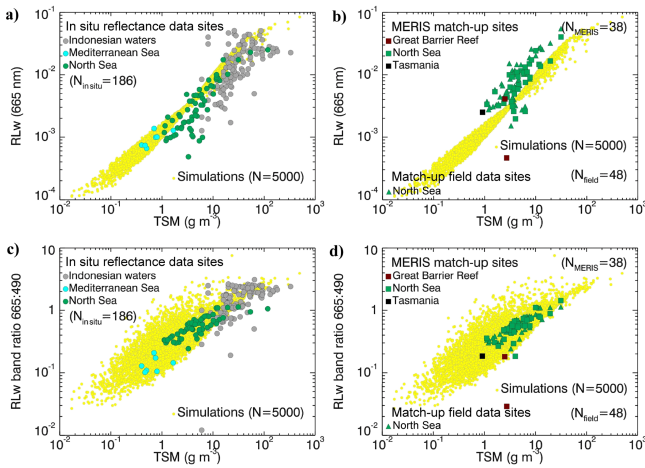
**Figure 27.** The relationships TSM vs. RLw(665) and TSM vs. RLw band ratio 665 : 490 in the in situ reflectance data set (non-yellow circles) plotted respectively in panels **(a)** and **(c)**, in the MERIS and match-up field data sets (squares and triangles) respectively in panels **(b)** and **(d)**, and in the simulated data set (yellow circles) in panels **(a–d)**.

pared to RLw(665) vs. TSM) could also be due to a removal of spectrally white errors.

The distribution of the ranges of RLw(665) and RLw(665) / RLw(490) in terms of TSM and within the sites – Indonesian waters, the Mediterranean Sea, and the North Sea – is consistent with the distribution of CHL vs. TSM (Fig. 10) in these sites, especially at low TSM $< 10\,\mathrm{g\,m^{-3}}$, where CHL is highly correlated with RLw(665).

## 4   Data repository

The match-up, in situ reflectance and simulated data sets are accessible from the PANGAEA website at http://doi.pangaea.de/10.1594/PANGAEA.841950. A description of files format and access follows.

The match-up field data for a site "SiteX" are stored in the compressed file at http://hs.pangaea.de/model/ccrr/Matchup_Dataset.zip, under directory /Matchup_Dataset/FieldData, in CSV files named following the classification of the parameters given in Table 2:

1. SiteX_metadata.csv, including the metadata and flags;

2. SiteX_biogeochem.csv, including the concentrations of the biogeochemical measurements;

3. SiteX_iops.csv, with the inherent optical properties;

4. SiteX_aops.csv, which includes the apparent optical properties.

For the North Sea region, two files are provided, having the names "North_Sea" and "North_Sea_Emeco", related to the

origin of the data: the North_Sea_Emeco data were downloaded from the EMECO website and North_Sea data from other data providers.

MERIS match-up products derived at a $5 \times 5$ pixel box around the locations of the match-up field measurements are provided as CSV files, with the headers as listed in Table 7, and stored in the compressed file http://hs.pangaea.de/model/ccrr/Matchup_Dataset.zip under directory /Matchup_Dataset/MERIS_5x5_L2R. The MERIS match-up products averaged at each location (from the $5 \times 5$ pixel box; see Sect. 2.1.2 for details) are stored in CSV files, with the headers listed in Table 8, and made available at directory /Matchup_Dataset/MERIS_average_L2R.

The in situ reflectance data are given in one CSV file, listing for each data provider (in the first column) the sample number, date, start and end of measurement time, latitude, longitude, the site identification number, the name of the location, the RLw spectra (nine columns for the nine MERIS selected bands), and CHL and TSM concentrations. These data are stored under the compressed directory InSituReflectance_Dataset.zip, accessible from the web address http://hs.pangaea.de/model/ccrr/InSituReflectance_Dataset.zip.

The simulated data are written in ASCII file format, and saved under the directory http://hs.pangaea.de/model/ccrr/Simulated_Dataset.zip. The concentrations of CHL, MP, and $a_g(443)$ are given in a separate file (named "Conc.txt"), where the simulation numbers going from 1 to 5000 are listed in the first column. Each entry (e.g. each simulation number or line) is associated with a given combination of CHL, MP, and $a_g(443)$.

The IOPs modelled for each entry, being the total absorption, $a_{\mathrm{tot}}$, scattering, $b_{\mathrm{tot}}$, and backscattering, $b_{\mathrm{btot}}$ coefficients excluding the pure water contributions; the absorption by phytoplankton pigments, $a_{\mathrm{phy}}$; and the ratio of the total backscattering coefficient to the sum of the total absorption and backscattering coefficients $b_{\mathrm{b}}/(a + b_{\mathrm{b}})$ – are provided in ASCII files called SPC_Atot.dat, SPC_Btot.dat, SPC_BBtot.data, SPC_Aphy.data and SPC_BBoABB.dat respectively, where SPC = "hyper" refers to hyperspectral input (from 350 to 900 nm, with a 5 nm step) and SPC = "maqua", "meris", or "swifs" to multispectral input at the band-centred wavelengths of the three sensors: MODIS-Aqua, MERIS, and SeaWiFS respectively.

The simulations generated hyperspectral and multispectral outputs specified in the prefix of the output filename. The three sun zenith angles ($x = 0$, 40, and 60°), assumed successively for the set of the 5000 simulations are given in the output AOPs file names as suffixes "_sza$x$.dat". Separate files are provided for RLw and Kd, stored as SPC_RLw_sza$x$.dat and SPC_Kd_sza$x$.dat respectively. The column entry in the spectral data files gives the wavelength (in nm), and the line entry gives the simulation number.

The simulated data also include the photosynthetically available radiation PAR$_{\mathrm{Ed}}$ and PAR$_{\mathrm{Eo}}$, defined as the integration over 400 to 700 nm of the spectral downwelling

irradiance Ed and of the scalar irradiance Eo respectively. $PAR_{Ed}$ and $PAR_{Eo}$ are profiled from 0 m above the water surface down to 80 m depth, at 27 depths listed along the columns (the line entry is related to the simulation number). The euphotic depths, $Zeu_{Ed}$ and $Zeu_{Eo}$, defined as the depths where $PAR_{Ed}$ and $PAR_{Eo}$ have 1 % of their respective values at the water surface, are provided in the files called Zeu_from_PAR_Ed_sza$x$.dat and Zeu_from_PAR_Eo_sza$x$.dat respectively.

The concentrations, IOP and AOP spectral data, and PAR and Zeu data files include headers to facilitate reading the data. The IOPs and concentration files are stored under the subfolder "Input IOPs Concentrations" and the simulated RLw, Kd, $PAR_{Ed}$, $PAR_{Eo}$, $Zeu_{Ed}$, and $Zeu_{Eo}$ under "Output AOPs".

# 5 Conclusion

The CCRR match-up, in situ, and simulated data sets form a large database covering a wide range of water types, from oligotrophic to hypertrophic, and from clear to very turbid waters with a high diversity of IOPs.

The data sets contain 336 in situ reflectance spectra (covering the spectral range 412 to 709 nm) from five CoastColour sites, 957 match-up field reflectance spectra from 13 sites and 457 MERIS RLw spectra from 11 sites which show global consistency over the match-up sites, despite the absence of harmonized protocols used for RLw measurements by the different laboratories. In total, 80 % of the MERIS RLw measurements are provided from the North Sea, matching various in situ water quality parameters collected at this site. This is balanced by the distribution of RLw measurements, throughout the CoastColour sites, given in the match-up field and in situ reflectance data sets where fewer than 5 % of RLw spectra are available from the North Sea, while 23 % of the match-up field IOP data (excluding turbidity) are provided for that site.

The high-quality reflectance data sets along with the biogeochemical (CHL, TSM) and inherent optical properties provided at 17 CoastColour sites, covering the period 2005 to 2010, are fully documented and made available publicly for use in ocean colour algorithm testing.

The simulated data set includes 5000 reflectance spectra and the associated concentrations and inherent optical properties of chlorophyll $a$, mineral particles, and CDOM. The simulated reflectance data have been compared to the in situ and match-up reflectance data, showing a global consistency and giving clues for the discrepancies noticed (e.g. variable inherent optical properties, measurement uncertainties).

The strengths and weaknesses of each individual data set are recognized; for example, an in situ measurement represents "sea truth" better than simulated data but is subject to measurement uncertainty and represents only a small volume. Testing of an algorithm on all three data sets using re-

spectively MERIS, in situ, and simulated RLw input has significantly added value: evaluating the robustness of an ocean colour algorithm against remote-sensing measurements uncertainties; identifying its domain of validity (e.g. the detection or saturation limits); and testing its performance on various regions, days and daytimes, sea and sky conditions, etc. Such exercises may point out the disadvantage/advantage of using an algorithm for a regional or global application.

Oceanographic databases have been built during the last few years and made available to the scientific community (e.g. open ocean phytoplankton data by O'Brien et al., 2013, and Buitenhuis et al., 2013), facilitating the sharing of data and stimulates collaboration between the research institutes. In this paper, the first public optical–biogeochemical database was established representing the core of an open resource dedicated to case 2 remote-sensing data validation and algorithm testing. With joint efforts from the research centres and laboratories, this database may be updated with extended in situ data for the existing sites and for new regions in coastal and inland waters, with extra information (e.g. data quality flags), and with artificial data sets covering extra ranges of optical properties (e.g. extremely absorbing waters, extremely turbid waters) and/or underlying new bio-optical models.

**The Supplement related to this article is available online at doi:10.5194/essd-7-319-2015-supplement.**

www.earth-syst-sci-data.net/7/319/2015/

Earth Syst. Sci. Data, 7, 319–348, 2015

influence of the Psittalia sewage treatment plant. Ministry of
Environment of the Hellenic Government.

– Young Je Park for providing the KORDI data set.

– Syarif Budhiman for processing the ITC data set.

– Guy Westbrook for providing the MII data set.

– Bryan A. Franz for processing the NOMAD data set.

– Francis Gohin, Catherine Belin, and Alain Lefèbvre for the
  Ifremer (REPHY phytoplankton network) data set.

– Griet Neukermans for collecting and processing the RBINS
  data set.

– Xabier Guinda, Beatriz Echavarri, Isabel Santamaría, and
  Pablo Ruíz for the UNICAN data collection and processing.

– Elisabete Mota (formerly in the Center of Oceanography,
  FCUL) and Katharina Poser (formerly in Brockman Consult)
  are thanked for their participation in the collection and organi-
  zation of the in situ data sets.

David Foley actively contributed to this paper through the NOAA
data set, before he passed away in December 2013. May this work
be a way to remember him for his rigorousness and generosity.

Special thanks to Frank Müller-Karger and two anonymous
reviewers for their very appreciated suggestions and comments, and
to the Topical Editor François Schmitt for his assistance throughout
the reviewing process.

Edited by: F. Schmitt

# References

Albert, A. and Mobley, C. D.: An analytical model for subsurface
    irradiance and remote sensing reflectance in deep and shallow
    Case 2 waters, Opt. Express, 11, 2873–2890, 2003.

Antoine, D., Siegel, D. A., Kostadinov, T. S., Maritorena, S.,
    Nelson, N. B., Gentili, B., Vellucci, V., and Guillecheau, N.:
    Variability of optical particle backscattering in contrasting bio-
    optical oceanic regimes, Limnol. Oceanogr., 56, 3, 955–973,
    doi:10.4319/lo.2011.56.3.0955, 2011.

Arar, E. J. and Collins, G. B.: Method 445.0 – In vitro determina-
    tion of chlorophyll a and pheophytin a in marine and freshwater
    algae by fluorescence, National Exposure Research Laboratory,
    Office of Research and Development, U.S. Environmental Pro-
    tection Agency, Cincinnati, Ohio, USA, 1992.

Babin, M., Morel, A., Fournier-Sicre, V., Fell, F., and Stramski, D.:
    Light scattering properties of marine particles in coastal and open
    ocean waters as related to the particle mass concentration, Lim-
    nol. Oceanogr., 28, 843–859, 2003a.

Babin, M., Stramski, D., Ferrari, G. M., Claustre, H., Bricaud, A.,
    Obolensky, G., and Hoepffner, N.: Variations in the light absorp-
    tion coefficients of phytoplankton, nonalgal particles and dis-
    solved organic matter in coastal waters around Europe, J. Geo-
    phys. Res., 108, 3211, doi:10.1029/2001JC000882, 2003b.

Barale, V., Jaquet, J.-M., and Ndiayé, M.: Algal blooming patterns
    and anomalies in the Mediterranean Sea as derived from the Sea-
    WiFS data set (1998–2003), Remote Sens. Environ., 112, 3300–
    3313, 2008.

Barlow, R. G., Cummings, D. G., and Gibb, S. W.: Improved resolu-
    tion of mono and divinyl chlorophylls a and and zeaxanthin and
    lutein in phytoplankton extracts using reverse phase C-8 HPLC,
    Mar. Ecol.-Prog. Ser., 161, 303–307, 1997.

Blondeau-Patissier, D., Brando, V. E., Oubelkheir, K., Dekker, A.
    G., Clementson, L. A., and Daniel, P.: Bio-optical variability of
    the absorption and scattering properties of the Queensland in-
    shore and reef waters, Australia, J. Geophys. Res.-Oceans, 114,
    C5, doi:10.1029/2008JC005039, 2009.

Bouvet, M. and Ramoino, F.: Equalization of MERIS L1b prod-
    ucts from the 2nd reprocessing, European Space Agency, Tech-
    nical Report TEC-EEP/2009.521/MB, Issue 1, Revision 3,
    08/12/2009, 77 p., published by ESTEC, Noordwijk, the Nether-
    lands, 2010.

Brewin, R. J. W., Sathyendranath, S., Mueller, D., Brockmann, C.,
    Deschamps, P.-Y., Devred, E., Doerffer, R., Fomferra, N., Franz,
    B., Grant, M., Groom, S., Horseman, A., Hu, C., Krasemann, H.,
    Lee, Z., Maritorena, S., Mélin, F., Peters, M., Platt, T., Regner, P.,
    Smyth, T., Steinmetz, F., Swinton, J., Werdell, J., and White, G.
    N.: The Ocean Colour Climate Change Initiative: III. A round-
    robin comparison on in-water bio-optical algorithms, Rem. Sens.
    Environ., 162, 271–294, 2015.

Bricaud, A., Morel, A., Babin, M., Allali, K., and Claustre, H.: Vari-
    ations of light absorption by suspended particles with chlorophyll
    a concentration in oceanic (Case 1) waters: Analysis and impli-
    cations for bio-optical models, J. Geophys. Res., 103, 31033–
    31044, 1998.

Brodie, J., De'ath, G., Devlin, M., Furnas, M. J., and Wright, M.:
    Spatial and temporal patterns of near surface chlorophyll a in the
    Great Barrier Reef lagoon, Mar. Freshwater Res., 58, 342–353,
    2007.

Buitenhuis, E. T., Vogt, M., Moriarty, R., Bednaršek, N., Doney, S.
    C., Leblanc, K., Le Quéré, C., Luo, Y.-W., O'Brien, C., O'Brien,
    T., Peloquin, J., Schiebel, R., and Swan, C.: MAREDAT: towards
    a world atlas of MARine Ecosystem DATa, Earth Syst. Sci. Data,
    5, 227–239, doi:10.5194/essd-5-227-2013, 2013.

Chavez, F. P., Pennington, J. T., Castro, C. G., Ryan, J. P., Michisaki,
    R. P., Schlining, B., Walz, P., Buck, K. R., McFadyen, A.,
    and Collins, C. A.: Biological and chemical consequences of
    the 1997–1998 El Niño in central California waters, Prog.
    Oceanogr., 54, 205–232, doi:10.1016/S0079-6611(02)00050-2,
    2002.

Cherukuru, N., Brando, V. E., Schroeder, T., Clementson, L. A., and
    Dekker, A. G.: Influence of river discharge and ocean currents on
    coastal optical properties, Cont. Shelf Res., 84, 188–203, 2014.

Clesceri, L. S., Greenberg, A. E., and Eaton, A. D.: Standard
    methods for the examination of water and wastewater, 20th edi-
    tion, American Public Health Association, 1325 pp., Washington
    D.C., 1998.

Doerffer, R.: Protocols for the validation of MERIS water prod-
    ucts, ESA/GKSS Doc. No. PO-TN-MEL-GS-0043, 46 p.,
    available as Appendix at: http://envisat.esa.int/workshops/mavt_
    2003, MERIS protocols, Issue 1.3.5, 2002.

Doerffer, R.: Alternative Atmospheric Correction Procedure for
    Case 2 Water Remote Sensing using MERIS, 2011.

Doerffer, R. and Schönfeld, W.: Validation transect between Cux-
    haven and Helgoland: GKSS Technical Note, 18 February 2009,
    Helmholtz-Centrum Geesthacht, Geesthacht, 2009.

Dolan, J. R., Claustre, H., Carlotti, F., Plounevez, S., and Moutin, T.: Microzooplankton diversity: relationships of tintinnid ciliates with resources, competitors and predators from the Atlantic Coast of Morocco to the Eastern Mediterranean, Deep-Sea Res. Pt. I, 49, 1217–1232, 2002.

Freudenthal, T., Meggers, H., Henderiks, J., Kuhlmann, H., Moreno, A., and Wefer, G.: Upwelling intensity and filament activity off Morocco during the last 250,000 years, Deep-Sea Res. Pt. II, 49, 3655–3674, 2002.

Gohin, F.: Annual cycles of chlorophyll $a$, non-algal suspended particulate matter, and turbidity observed from space and in-situ in coastal waters, Ocean Sci., 7, 705–732, doi:10.5194/os-7-705-2011, 2011.

Groom, S. B., Herut, B., Brenner, S., Zodiatis, G., Psarra, S., Kress, N., Krom, M. D., Law, C. S., and Dracopoulos, P.: Satellite-derived spatial and temporal biological variability in the Cyprus Eddy, Deep-Sea Res. Pt. II, 52, 2990–3010, 2005.

Groom, S. B., Martinez-Vicente, V., Fishwick, J. R., Tilstone, G., Moore, G., Smyth, T. J., and Harbour, D.: The Western English Channel observatory: Optical characteristics of station 4, J. Marine Syst., 77, 278–295, 2009.

Guillocheau, N.: $\beta$-Correction Experiment Report, ICESS, University of California, Santa Barbara, CA, USA, 2003.

Harrison, A. W. and Coombes, C. A.: An opaque cloud cover model of sky short wavelength radiance, Sol. Energy, 41, 387–392, 1988.

Holm-Hansen, O., Lorenzen, C. J., Holmes, R. W., and Strickland, J. D. H.: Fluorometric determination of chlorophyll, Journal du Conseil Permanent International pour l'Exploration de la Mer, 30, 33–15, 1965.

Hooker, S. B., Van Heukelem, L., Thomas, C. S., Claustre, H., Ras, J., Barlow, R. G., Sessions, H., L., S., Perl, J., Trees, C. C., Stuart, V., Head, E., Clementson, L. A., Fishwick, J. R., Llewellyn, C., and Aiken, J.: The Second SeaWiFS HPLC Analysis Round-Robin Experiment (SeaHARRE-2), NASA Aeronautics and Space Information, Goddard Space Flight Center, Greenbelt, Maryland, USA, 2005.

IOCCG: Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications, edited by: Lee, Z.-P., Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada, pp. 126, 2006.

IOC/UNESCO: Protocols for Joint Global Ocean Carbon Flux Study (JGOFS) Core Measurements, Paris, 1994.

Karageorgis, A. P., Georgopoulos, D., Kanellopoulos, T. D., Mikkelsen, O. A., Pagou, K., Kontoyiannis, H., Pavlidou, A., and Anagnostou, C.: Spatial and seasonal variability of particulate matter optical and size properties in the Eastern Mediterranean Sea, J. Marine Syst., 105–108, 123–134, 2012.

Kostadinov, T. S., Siegel, D. A., Maritorena, S., and Guillocheau, N.: Ocean color observations and modeling for an optically complex site: Santa Barbara Channel, California, USA, J. Geophys. Res., 112, C07011, doi:10.1029/2006JC003526, 2007.

Llewellyn, C. A., Fishwick, J. R., and Blackford, J. C.: Phytoplankton community assemblage in the English Channel: a comparison using chlorophyll a derived from HPLC-CHEMTAX and carbon derived from microscopy cell counts, J. Plankton Res., 27, 103–119, 2005.

Loisel, H. and Morel, A.: Light scattering and chlorophyll concentration in Case 1 waters: a re-examination., Limnol. Oceanogr. (Methods), 43, 847–857, 1998.

Loisel, H., Vantrepotte, V., Norkvist, K., Mériaux, X., Kheireddine, M., Ras, J., Pujo-Pay, M., Combet, Y., Leblanc, K., Dall'Olmo, G., Mauriac, R., Dessailly, D., and Moutin, T.: Characterization of the bio-optical anomaly and diurnal variability of particulate matter, as seen from scattering and backscattering coefficients, in ultra-oligotrophic eddies of the Mediterranean Sea, Biogeosciences, 8, 3295–3317, doi:10.5194/bg-8-3295-2011, 2011.

Maffione, R. A. and Dana, D. R.: Instruments and methods for measuring the backward-scattering coefficient of ocean waters, Appl. Optics, 36, 6057–6067, 1997.

Maritorena, S., Lee, Z.-P., Du, K.-P., Loisel, H., Doerffer, R., Roesler, C., Tanaka, A., Babin, M., and Kopelevich, O. V.: Syntheric and in situ data sets for algorithm testing. Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications, IOCCG, Dartmouth, Canada, 73–79, 20065, 13–18, 2006.

Martinez-Vicente, V., Land, P. E., Tilstone, G. H., Widdicombe, C., and Fishwick, J. R.: Particulate scattering and backscattering related to water constituents and seasonal changes in the Western English Channel, J. Plankton Res., 32, 603–619, 2010.

Mobley, C. D.: Light and water: radiative transfer in natural waters, Academic Press, London, UK, 1994.

Mobley, C. D.: Estimation of the remote sensing reflectance from above-surface measurements, Appl. Optics, 38, 7442-7455, 1999.

Mobley, C. D. and Sundman, L. K.: HYDROLIGHT 5/ECOLIGHT 5 Technical Documentation, Sequoia Scientific, Inc., 2700 Richards Road, Suite 107, Bellevue, WA 98005, 2008.

Morel, A. and Prieur, L.: Analysis of variations in ocean color, Limnol. Oceanogr., 22, 709–722, 1977.

Morel, A., Antoine, D., and Gentili, B.: Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function, Appl. Optics, 41, 6289–6306, 2002.

O'Brien, C. J., Peloquin, J. A., Vogt, M., Heinle, M., Gruber, N., Ajani, P., Andruleit, H., Arístegui, J., Beaufort, L., Estrada, M., Karentz, D., Kopczyńska, E., Lee, R., Poulton, A. J., Pritchard, T., and Widdicombe, C.: Global marine plankton functional type biomass distributions: coccolithophores, Earth Syst. Sci. Data, 5, 259–276, doi:10.5194/essd-5-259-2013, 2013.

Odermatt, D., Gitelson, A. A., Brando, V. E., and Schaepman, M.: Review of constituent retrieval in optically deep and complex waters from satellite imagery, Remote Sens. Environ., 118, 116–126, 2012.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., Kahru, M., and McClain, C. R.: Ocean color chlorophyll algorithms for SeaWiFS, J. Geophys. Res., 103, 24937–24953, 1998.

Oubelkheir, K., Claustre, H., Sciandra, A., and Babin, M.: Bio-optical and biogeochemical properties of different trophic regimes in oceanic waters, Limnol. Oceanogr., 50, 1795–1809, 2005.

Oubelkheir, K., Clementson, L. A., Webster, I., Ford, P., Dekker, A. G., Radke, L., and Daniel, P.: Using inherent optical properties to investigate biogeochemical dynamics in a tropical

www.earth-syst-sci-data.net/7/319/2015/

Earth Syst. Sci. Data, 7, 319–348, 2015

macrotidal coastal system, J. Geophys. Res.-Oceans, 111, C7, doi:10.1029/2005JC003113, 2006.

Parsons, T. R., Maita, Y., and Lalli, C. M.: A manual of Chemical and Biological Methods for seawater Analysis, Pergamon Press, New York, USA, 173 pp., 1984.

Pegau, W. S., Zaneveld, J. R. V., Mitchell, B. G., Mueller, J. L., Kahru, M., Wieland, J., and Stramska, M.: Inherent Optical Properties: Instruments, Characterizations, Field Measurements and Data Analysis Protocols, volume IV, in: Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, 76 pp., 2003.

Pitcher, G. C. and Probyn, T. A.: Anoxia in southern Benguela during the autumn of 2009 and its linkage to a bloom of the dinoflagellate Ceratium balechii, Harmful Algae, 11, 23–32, 2011.

Probyn, T. A.: Nitrogen uptake by size-fractionated phytoplankton populations in the southern Benguela upwelling system, Mar. Ecol.-Prog. Ser., 22, 249–258, 1985.

Ruddick, K., Brockmann, C., Doerffer, R., Lee, Z., Brotas, V., Fomfera, N., Groom, S. B., Krasemann, H., Martinez-Vicente, V., Sá, C., Santer, R., Sathyendranath, S., Stelzer, K., and Pinnock, S.: The CoastColour project regional algorithm round robin exercise, in: Remote Sensing of the Coastal Ocean, Land and Atmosphere Environment, SPIE, Incheon, Republic of Korea, 7858–7807, 2010.

Ruddick, K. G., De Cauwer, V., Park, Y., and Moore, G.: Seaborne measurements of near infrared water-leaving reflectance – the similarity spectrum for turbid waters, Limnol. Oceanogr., 51, 1167–1179, 2006.

Schaffelke, B., Carleton, J., Skuza, M., Zagorskis, I., and Furnas, M. J.: Water quality in the inshore Great Barrier Reef lagoon: Implications for long-term monitoring and management, Mar. Pollut. Bull., 65, 249–260, 2012.

Strickland, J. D. H. and Parsons, T. R.: A practical handbook of the sea water analysis, Fisheries Research Board of Canada, Ottawa, Bulletin 167, 311 p., 1972.

Tassan, S. and Ferrari, G. M.: An alternative approach to absorption measurements of aquatic particles retained on filters, Limnol. Oceanogr., 40, 1358–1368, 1995.

Tilstone, G., Moore, G., Sorensen, K., Doerffer, R., Rottgers, R., Ruddick, K. G., Pasterkamp, R., and Jorgensen, P. V.: Regional Validation of MERIS Chlorophyll products in North Sea coastal waters: REVAMP protocols, ENVISAT validation workshop European Space Agency, available as Appendix at: https://earth.esa.int/workshops/mavt_2003/, Frascatti, Italy, 20–24 October 2003.

Tilstone, G., Peters, S., van Der Woerd, H., Eleveld, M., Ruddick, K., Schönfeld, W., Krasemann, H., Martinez-Vicente, V., Blondeau-Patissier, D., Jorgensen, P. V., Rottgers, R., and Sorensen, K.: Variability in specific-absorption properties and their use in Ocean Colour Algorithms for MERIS in North Sea coastal waters, Remote Sens. Environ., 118, 320–338, 2012.

Toole, D. A., Siegel, D. A., Wenzies, D. W, Neumann, M. J., and Smith, R. C.: Remote-sensing reflectance determinations in the coastal ocean environment: impact of instrumental characteristics and environmental variability, Appl. Optics, 39, 3, 456–469, 2000.

Tzortziou, M., Subramaniam, A., Herman, J. R., Gallegos, C. L., Neale, P. J., and Harding, L. W. J.: Remote sensing reflectance and inherent optical properties in the mid Chesapeake Bay, Estuarine, Coastal and Shelf Science, 72, 16–32, 2007.

van der Linde, D. W.: Protocol for the determination of total suspended matter in oceans and coastal zones, Technical Note I.98.182, European Commission Joint Research Centre, Ispra, Italy, 1998.

Van Heukelem, L. and Thomas, C. S.: Computer-assisted high-performance liquid chromatography method development with applications to the isolation and analysis of phytoplankton pigments, J. Chromatogr. A, 910, 31–49, 2001.

Werdell, P. J.: An evaluation of inherent optical property data for inclusion in the NASA Bio-optical Marine Algorithm Data Set, NASA Ocean Biology Processing Group paper, NASA Goddard Space Flight Cent., Greenbelt, Md., available at: http://seabass.gsfc.nasa.gov/seabass/data/werdell_nomad_iop_qc.pdf, 2005.

Werdell, P. J. and Bailey, S. W.: An improved in situ bio-optical data set for ocean color algorithm development and satellite data product validation, Remote Sens. Environ., 98, 122–140, 2005.

Werdell, P. J., Bailey, S. W., Fargion, G. S., Pietras, K., Knobelspiesse, K. D., Feldman, G. C., and McClain, C.: Unique data repository facilitates ocean color satellite validation, EOS, Transactions, American Geophysical Union, 84, 377–392, 2003.

Werdell, P. J., Franz, B. A., Bailey, S. W., Feldman, G. C., Boss, E., Brando, V. E., Dowell, M. D., Hirata, T., Lavender, S. J., Lee, Z., Loisel, H., Maritorena, S., Mélin, F., Moore, T. S., Smyth, T. J., Antoine, D., Devred, E., Hembise Fanton d'Andon, O., and Mangin, A.: Generalized ocean color inversion model for retrieving marine inherent optical properties, Appl. Optics, 52, 2019–2037, 2013.

WET Labs ac-9 Protocol: Revision H, WET Labs, Inc., Philomath, OR, USA, 42 pp., 2003.

Wright, S. W., Jeffrey, S. W., Mantoura, R. F. C., Llewellyn, C. A., Bjornland, T., and Repeta, D. J.: Improved HPLC method for the analysis of chlorophylls and carotenoids from marine phytoplankton, Mar. Ecol.-Prog. Ser., 77, 183–196, 1991.