The COASTℓOOC dataset

Philippe Massicotte[1], Marcel Babin[4, 1], Frank Fell[2], Vincent Fournier-Sicre[3, 5], and David Doxaran[4]

[1]*Takuvik International Research Laboratory (IRL 3376) ULaval - CNRS, Biology department, Laval University, Quebec G1V 06A, Canada
[2]Informus GmbH, 13187 Berlin, Germany
[3]ACRI-ST, 260 Route du Pin Montard, 06904 Sophia-Antipolis, France
[4]Laboratoire d'Océanographie de Villefranche, CNRS / Sorbonne Université, 06230 Villefranche-sur-mer, France
[5]*European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), EUMETSAT-Allee 1, 64295 Darmstadt, Germany

Correspondence: Marcel Babin (marcel.babin@takuvik.ulaval.ca)

Abstract. Coastal Surveillance Through Observation of Ocean Color (COASTℓOOC) oceanographic expeditions were conducted in 1997 and 1998 to examine the relationship between the optical properties of seawater and related biological and chemical properties across the coastal-to-open ocean gradient in various European seas. A total of 379 stations were visited along the coasts of the Gulf of Lion in the Mediterranean Sea ($n = 61$), Adriatic Sea ($n = 39$), Baltic Sea ($n = 57$), North
5 Sea ($n = 99$), English Channel ($n = 85$) and Atlantic Ocean ($n = 38$). Particular emphasis was dedicated to the collection of a comprehensive set of apparent (AOPs) and inherent (IOPs) optical properties to support the development of ocean color remote sensing algorithms. The data were collected in situ using traditional ship-based sampling, but also from a helicopter, which is a very efficient means for that type of coastal sampling. The dataset collected during the COASTℓOOC campaigns is unique in that it is fully consistent in terms of operators, protocols, and instrumentation. This rich and historical dataset is
10 still today frequently requested and used by other researchers. Therefore, we present the result of an effort to compile and standardize a dataset which will facilitate its reuse in future development and evaluation of new bio-optical models adapted for optically-complex waters. The dataset is available at https://doi.org/10.17882/93570 (Massicotte et al., 2023).

1 Introduction

Since the launch of the Coastal Zone Color Scanner (CZCS) by NASA in 1978, ocean color remote sensing has been used to
15 monitor the state of, and the changes in global marine ecosystems both in time and space. In open oceans, the main component that affects the variations in the inherent (IOPs) and apparent (AOPs) optical properties of seawater is phytoplankton, which is usually represented by the concentration of chlorophyll *a* (Morel and Prieur, 1977). Such open ocean waters are traditionally termed Case 1 waters. Many simple spectral band ratio algorithms have been developed to link changes in remotely-sensed ocean color (OC), measured as reflectance, to variations in chlorophyll *a* concentration (see O'Reilly and Werdell 2019 for
20 an extensive evaluation of OC band ratio algorithms). Because these algorithms perform well, a plethora of studies have

---

*Current address
*Current address

been conducted based on those, notably about phytoplankton phenology (e.g., Vargas et al. 2009) and phytoplankton primary production (see Carr et al. 2006 and references therein).

Space-borne monitoring of aquatic ecosystems is more challenging for optically complex Case 2 waters (Morel and Prieur, 1977) often found in coastal areas. In contrast to Case 1 waters, the optical properties of Case 2 waters are determined by several types of constituents occurring in individually highly variable concentrations. Rivers draining large catchment areas deliver important quantities of optically significant substances such as chromophoric dissolved organic matter (CDOM) and suspended particulate matter (SPM) to the coastal waters (Hedges et al., 1997; Cole et al., 2007; Massicotte et al., 2017). In shallow areas, bottom reflectance and resuspension of sediments can also alter the signal of water-leaving reflectance (Lee et al., 1998). Because CDOM and SPM do not necessarily covary with chlorophyll a and can mask the presense of phytoplankton (Sathyendranath, 2000), bio-optical algorithms developed for Case 1 waters are generally not appropriate for optically-complex Case 2 waters (Gordon and Morel, 1983). Hence, concomitant in situ measurements of optically active components (CDOM, SPM, chlorophyll a) and related radiometric quanfities (AOPs and IOPs) are, to this day, critically needed to develop and improve bio-optical algorithms for coastal areas.

The objective of the Coastal Surveillance Through Observation of Ocean Color (COASTℓOOC) oceanographic expeditions was to acquire a comprehensive set of AOPs, IOPs, and concentrations of optically significant components along the European coasts to help the development of new remote sensing algorithms in optically complex waters. Back in 1997 and 1998, the COASTℓOOC campaigns were among the very first concerted efforts to fill the knowledge gaps in coastal marine optics by establishing a large, consistent, and comprehensive optical dataset for coastal waters. A total of 379 stations were visited along the coasts of the Mediterranean, Adriatic, Baltic and North Sea, in the English Channel, and in the Atlantic Ocean, mostly in coastal, but also in open oceanic waters. While this unique set of data has been used or referenced over the years in numerous peer-reviewed publications (see list in Appendix A), a reference final dataset, with well-documented quality controls, has never been published. Even though the COASTℓOOC campaigns were carried out more than 20 years ago, the rich and historical dataset that has been collected still has great potential to contribute to the development and evaluation of new bio-optical models adapted for optically-complex waters. The goal of this paper is to formally present a quality-controlled and final reference version of the COASTℓOOC dataset, archived in a public repository where users can download it readily.

2 Study area and sampling overview

2.1 Study area and general sampling strategy

During the six COASTℓOOC campaigns, further referred to as C1 to C6, a total of 379 locations were visited (Fig. 1). The stations were located along the coasts of the Gulf of Lion in the Mediterranean Sea ($n = 61$), Adriatic Sea ($n = 39$), Baltic Sea ($n = 57$), North Sea ($n = 99$), English Channel ($n = 85$) and Atlantic Ocean ($n = 38$). Within each area, the stations were generally distributed along across-shore or along-shore transects to capture the land-to-sea gradients and document river plumes (Fig. 1B). COASTℓOOC campaigns were either ship-based (C1, C4, C5) or used a helicopter as a sampling platform (C2, C3, C6) and took place between 1997-04-02 and 1998-09-25 (Fig. 2A). Compared to traditional ship-based sampling, the

helicopter platform allowed efficient stations close to the coast in areas difficult to access by boat (some samples were collected in waters as shallow as 1 m, Babin 2003). Combining both ship and airborne sampling approaches allowed covering the whole inshore to open-ocean aquatic continuum. The bathymetry (2021, 2021) varied greatly across the stations, where it averaged ca. 10 meters in the Adriatic Sea and ca. 800 meters in the Case 1 Atlantic Ocean (Fig. 2B).

2.2 Water sampling and processing

On board the ship, water samples were analyzed and/or stored immediately after collection. When using a helicopter as a sampling platform, water samples were kept in polyethylene containers for no longer than two hours after sampling (see Babin 2003 for further details). Filtration was conducted at low vacuum onto 25 mm glass fiber filters (Whatman, GF/F) to collect particles for subsequent analyses, as described below. Three separate subsamples (up to two liters each) were filtered for each sample with the following purposes. To determine the concentration of suspended particulate matter (SPM; g m^{-3}), the particles collected onto a pre-weighed GF/F filter were dried and stored in a freezer at -80 °C until the dry weight was determined less than two months later in the laboratory (Van Der Linde, 1998). For pigment analysis, the filter with collected particles was inserted into a cryotube and kept in liquid nitrogen until analysis less than three months later. High-performance liquid chromatography (HPLC) was used as described by Vidussi et al. 1996 to determine liposoluble pigment concentrations. Total chlorophyll *a* is here defined as the sum of chlorophyll *a*, divinyl-chlorophyll *a*, chlorophyll *a* isomer and epimer, chlorophyllids a, and phaeopigments. The concentrations of particulate organic carbon and nitrogen (POC and PON, respectively) were determined using a Carlo Erba NCS 2500 elemental analyzer as described in Ferrari et al. 2003. These determinations were made on particles collected on a GF/F filter precombusted at 450 °C for 2 h and stored in the freezer at -80 °C until analysis less than two months later. Further details concerning laboratory analyses can be found in Babin 2003; Ferrari 2000; Ferrari et al. 2003. To determine the concentration of dissolved organic matter (DOC, M), seawater samples were filtered on a 0.22 mm Millipore membrane. The final filtrate was transferred to a 100-mL amber-glass bottle, stored in the refrigerator and analyzed within three weeks. DOC was determined by high-temperature catalytic oxidation using a Carlo Erba 480 analyzer. The methodological details can be found in Ferrari 2000.

2.3 Optical measurements

At each station, an optical package was deployed in the water column from the surface, down to one to two meters above the ocean bottom in shallow waters, and to maximum depths of ca. 100 m in deep waters. The package included a CTD (conductivity, temperature, and depth) as well as an array of optical instruments to simultaneously measure beam attenuation, absorption and scattering coefficients, as well as irradiance in the water column. The Modular Data and Power System (MODAPS) data-gathering system from Wetlabs Inc., USA was used to combine the time-stamped data from the different instruments to produce a data matrix with a common depth grid. For further information, please refer to the COASTℓOOC final report (Aiken et al., 2000).

A specific aspect of COASTℓOOC is the use of a helicopter as a sampling platform, which has enabled high-frequency visiting of sites in very shallow waters (see section 2.1). An additional advantage is due to the fact that helicopter-based optical

measurements are not influenced by the sampling platform itself, unlike ship-based measurements unavoidably affected by the presence of the ship hull, most prominently close to the surface. A drawback of this approach is the difficulty to operate a helicopter stationary at a constant height, especially at low or gusty wind conditions: while the former requires substantial flight heights to avoid the helicopter downdraft impacting the optical measurements, rendering sampling less efficient, the latter can result in unwanted vertical movements of the instrument package in water on the order of 1-2 m s^{-1}. To ensure the availability of data of sufficient quality close to the surface, several consecutive up- and downcasts were therefore performed at each "helicopter" site. Among the measured vertical profiles at each site, the most suitable casts for depth merging and surface extrapolation (which are not necessarily the same, see the end of section 2.3.3) were selected using a combination of objective criteria and visual inspection.

2.3.1 Irradiance measurements

The SeaWiFS Profiling Multichannel Radiometer (SPMR, Satlantic Inc, Canada) was used to measure downward (E_d) and upward (E_u) spectral irradiances (W m^{-2} µm^{-1}) in the water column. Irradiance was measured at 13 wavelengths matching the MERIS channels of direct relevance for ocean observations (411, 443, 456, 490, 509, 531, 559, 619, 664, 683, 706, 779, 866 nm, except for COASTℓOOC 1 operating at 590 nm instead of 619 nm), ranging from the blue part of the spectrum to the near-infrared at an acquisition rate of 6 Hz. The actual wavelengths differ slightly between upward and downward observations.

The SPMR was fixed to the IOP instrument frame, except for the COASTℓOOC 5 mission, where it was used in free-falling mode. For the ship-based measurements, concomitant observations of the in-air downward spectral irradiance at the sea surface were performed to account for changes in the incoming solar radiation during the in situ profiling (e.g., due to clouds). This approach was not achievable for the helicopter-based measurements, since there was no suitable place to mount the radiometer. For these campaigns, measurements of E_d taken by the profiler while in the air (i.e., before and after the in-water measurements) were used to assess E_d on the sea surface. In order to minimize the time lag between the in-water measurements and the E_d values on the sea surface, our processing code automatically chose the air measurements closest in time. Under stable atmospheric conditions, no significant error is expected to result from this approach, while it will contribute to increased variability in the retrieved parameters when sampling under varying cloud cover. However, the practical impact of this is limited due to the short duration of helicopter-based sampling, typically lasting between one and three minutes per profile.

2.3.2 Water-leaving reflectance

To calculate the water-leaving reflectance, accurate measurements of both E_d and E_u just below the water surface ($z = 0^-$) are needed. In case no spectrally matching in-air E_d reference data are available (as is the case for the helicopter-based campaigns), this was done by extrapolating both $E_d(z)$ and $E_u(z)$ vertical profiles toward the sea surface by fitting an exponential model for the depth dependence of the measured spectral irradiance values (equation 1):

$$\Gamma(\lambda, z) = \Gamma(\lambda, 0^-)e^{-K_\Gamma(\lambda)z} \tag{1}$$

where $\Gamma(\lambda, z)$ is either E_d or E_u measured at wavelength λ and at depth z, $\Gamma(\lambda, 0^-)$ is E_d or E_u estimated just below the air-water interface and K_Γ is the estimated diffuse attenuation coefficient for downward (K_{Ed}, m^{-1}) or upward (K_{Eu}, m^{-1}) irradiance. Note that a correction to account for instrument self-shading of the upwelling irradiance was not applied. We assume the impact of this omission is relatively small except for highly absorbing waters (depending on constituent concentration and/or observation wavelength), where it may lead to an underestimation of near-surface upwelling irradiances.

In case concomitant spectrally matching in-air E_d (0$^+$) reference data are available (as is the case for most ship-based campaigns), E_d (0$^-$) is calculated from E_d (0$^+$) by applying a constant factor of 0.943 to account for losses of the downwelling irradiance (direct + diffuse) due to Fresnel reflection at the rough sea surface. The derived E_u (0$^-$) and E_d (0$^-$) values were thereafter used to calculate the water reflectance R (0^+) just below the sea surface:

$$R(0^-) = \frac{E_u(0^-)}{E_d(0^-)} \tag{2}$$

Vertical irradiance profiles in the red and near-infrared channels at wavelengths beyond 600 nm were fitted using a sum of two exponential functions of which one represents the diffuse attenuation due to absorption and elastic scattering for the channel under consideration, while the other accounts for Raman scattering from the corresponding (shorter) excitation wavelengths significantly contributing to the upward light field (e.g., Sugihara et al. 1984). Not doing so, i.e., using just a single exponential function, leads to substantial underestimation of the extrapolated subsurface values due to significant Raman contribution from excitation wavelengths with significantly lower K values (see Appendix 1). Figure A1 showing the near-surface depth dependence of E_u at 779 nm demonstrates why the two-function approach is required. Below a depth of ca. 1.5 m, the signal is strongly dominated by Raman scattering and is therefore characterized by a diffuse attenuation coefficient very close to that of the corresponding excitation wavelengths.

2.3.3 Irradiance depth merging

Depth-merged irradiance profiles were derived for all sites by applying a standardized approach. The underlying strategy for depth merging was to provide users with the information they need to do their analyses while modifying the original data as little as possible. Processing was therefore limited to procedures that could be applied equally to all COASTℓOOC campaigns, whether helicopter- or ship-based. For example, variations of the incoming solar radiation were not considered in the depth merging process, since such observations were not available for the helicopter-based campaigns. However, if measured while operating from research vessels, such information is included in the depth-merged profile data to allow users doing their own subsequent analyses.

Typically, several up- and downcast vertical profiles were performed, especially when using the helicopter as a sampling platform, to increase the potential availability of high-quality observations. The actual merging process was executed as follows.

First, any observation not meeting both of the two following conditions was discarded: (1) SPMR instrument tilt smaller than 20° and (2) SPMR vertical speed between 0.1 and 2.0 m s^{-1}. To automatically identify the best-suited cast for profile generation, the observations meeting the above two conditions were then distributed cast-wise over 0.2 m wide vertical bins, starting at zero depth. Finally, the cast with the highest number of bins containing at least one valid observation was selected for depth profile generation, if it fulfilled two further conditions: (3) at least 2 m vertical distance between the highest and the lowest bin and (4) at least 50% of the bins containing at least one observation. No depth merged data is provided for stations where there is no cast meeting all of the above four conditions. Combining several casts taken at one station has not proven successful due to potentially large discontinuities in the aggregated values at depth levels where the number of available casts changes, especially in unstable atmospheric conditions with the corresponding short-term irradiance fluctuations.

Actual depth merging was done by taking the median of all observations of a particular irradiance parameter within each depth bin. Spectral irradiance observations were only considered for values above 0.01 W m^{-2} μm^{-1} to reduce the impact of residual dark current and radiometric noise. Finally, an attempt was made to identify and mask out remaining outliers in the depth-merged irradiance profiles. This was done by applying Grubb's test at a 5% significance level (Guthrie, 2012) to the residual of a second-order polynomial fit of the logarithmic-transformed spectral irradiance observations for a 2 m depth window, sliding down in steps of 1 m. While this procedure has successfully identified individual spikes, some obvious outliers extending over several depth bins remain in the merged dataset, and users will need to make their own judgment on how to identify and remove those.

Examples of depth-merged irradiance profiles are shown in Fig. 3 for one clear-water site in the NE Atlantic (Fig. 3A) and one CDOM-rich site in the Baltic Sea (Fig. 3B). Figure 3A features a double deep chlorophyll maximum with the corresponding irradiance increase due to chlorophyll fluorescence at water depths between ca. 35 m and 45 m, most prominently visible at 683 nm, but also affecting the neighboring channels at 664 nm and 706 nm. Diffuse attenuation in the blue to green wavelengths is correspondingly enhanced as reflected by the change in slope at a depth of ca. 34 m. Fig. 3B features Baltic Sea waters rich in CDOM, resulting in strong absorption in the UV and blue parts of the spectrum. The reduced decrease of the upwelling irradiance at 559 nm towards larger depths is likely caused by the reflection of downwelling irradiance from the ocean bottom, while chlorophyll fluorescence is likely the reason for the relatively enhanced (change of slope) upwelling irradiance at 683 nm for depths below ca. 7 m.

Note that extrapolated surface values and depth-merged profiles for a specific site are not necessarily derived from the same observations. This is especially true for helicopter-based operations typically comprising several up- and downcast close to the water surface to increase the likelihood of having sufficient observations for the surface extrapolation, while usually only one or two deep-water casts were made to obtain observations at greater depth. The best-suited observations for surface extrapolation, therefore, do not necessarily correspond to the cast used for depth merging.

2.3.4 Profiles of the diffuse attenuation coefficient

Downward (K_{Ed}, m^{-1}) and upward (K_{Eu}, m^{-1}) spectral diffuse attenuation coefficients have been derived from the depth-merged irradiance data by fitting an exponential function through the irradiance observations within a 2 m depth window, sliding down

in steps of 0.2 m. The calculation was only performed if at least five observations were available for fitting within the depth window. The choice of the window width was based on a compromise between two conflicting requirements: on the one hand, it should be small enough to allow resolution of relatively narrow vertical features, such as a deep chlorophyll maximum. On the other hand, it should be large enough to allow for an accurate estimation of K. The window width of 2 m chosen herein favors vertical resolution over accuracy: while even relatively narrow vertical features remain distinguishable in the derived their own K profiles, these are on the other hand affected by remaining outliers, instrument tilt variations, or instrumental noise. Obviously, users may use depth-merged irradiance profiles to derive K profiles or depth-integrated K values better meeting individual requirements.

2.3.5 Chromophoric dissolved organic matter (laboratory measurements)

Measurements of chromophoric dissolved organic matter (CDOM) absorption (a_{CDOM}, m^{-1}) were performed in the laboratory using a spectrophotometer (Perkin Elmer, Lambda 12) on water samples (10 cm cuvette path length) filtered on a 0.22 μm Millipore membrane pre-rinsed with 50 ml of Milli-Q water (Babin, 2003). a_{CDOM} spectra were measured between either 300-750 nm or 350-750 nm at 1-nm increment.

2.3.6 Particulate, non-algal and phytoplankton absorption (laboratory measurements)

Water samples were filtered onto 25 mm glass fiber filters (Whatman, GF/F) at a low vacuum before absorption measurement (Babin, 2003). Total particulate absorption (a_P, m^{-1}) was measured on particles retained on the filters between 380 and 750 nm at 1-nm increment using a spectrophotometer equipped with a 60 mm integrating sphere (Perkin Elmer, Lambda 19), and following the transmittance-reflectance (TR) method (Tassan and Ferrari, 1995, 1998). Afterward, pigments were removed from the particles with sodium-hypochloride to measure non-algal (or non-pigmented) absorption (a_{NAP}, m^{-1}). Finally, phytoplankton absorption (a_ϕ, m^{-1}) was retrieved by subtracting a_{NAP} from a_P. a_{CDOM} and a_{NAP} absorption spectra were fitted according to the following equation (Jerlov, 1968; Bricaud et al., 1981):

$$a_\tau(\lambda) = a_\tau(\lambda_0)e^{-S_\tau(\lambda-\lambda_0)} + k \qquad (3)$$

Where a_τ is the absorption coefficient, either a_{CDOM} or a_{NAP} (m^{-1}), λ is the wavelength (nm), λ_0 is a reference wavelength (443 nm), S_τ is the spectral slope (nm^{-1}) that describes the approximate exponential decrease in absorption with increasing wavelength and k a background constant (m^{-1}) accounting for scatter in the cuvette and drift of the instrument. a_{CDOM} spectra were fitted using observations below 700 nm, whereas a_{NAP} fits were performed between 380 and 730 nm, excluding spectral regions between 400-480 nm and 620-710 nm to avoid possible residual pigment absorption (Babin, 2003). a_{CDOM} and a_{NAP} were baseline corrected by subtracting the background parameter (k) derived from the following equation:

$$\hat{a}_\tau(\lambda) = a_\tau(\lambda) - k \qquad (4)$$

Finally, the spectral background of a_P spectra (i.e., the average absorption measured between 746 and 750 nm) was added to a_{NAP}. The underlying rationale is that absorption by phytoplankton in the near-infrared (NIR) is null, and that absorption measured in the NIR for total particles using the TR method is real (Tassan and Ferrari, 2003) and belongs exclusively to non-algal particles. The absorption signal measured for total particles retained on untreated filters is assumed more reliable than that measurement on bleached filters.

2.3.7 Total absorption and scattering (in situ measurements)

In situ vertical profiles of absorption (a, m^{-1}) and beam attenuation (c, m^{-1}) were acquired at nine wavelengths (412, 440, 488, 510, 555, 630, 532 or 650 depending on the instrument configuration, 676, and 715 nm) using a flow-through in situ absorbance-attenuance meter (AC9, Wetlabs). As the measurements are referenced to pure (Milli-Q) water, the obtained absorption and attenuation coefficients exclude the contribution of water. To correct absorption measurements for incomplete recovery of scattered light, $a(715)$ was subtracted from $a(\lambda \leq 715)$. Scattering coefficients (b, m^{-1}), were calculated by subtracting absorption from attenuation. As the contribution of molecular (water) scattering is excluded (see above), this coefficient essentially corresponds to particle scattering and is hereafter denoted b_P (m^{-1}).

The such derived values for a, b_p, and c were subsequently averaged over the first attenuation length (1 / K_{Ed}) to provide values representative of the surface layer. Due to practical difficulties when operating the AC-9 from a helicopter (e.g., trapped air in the instrument), the availability of a, b_p, and c values is limited for helicopter-based campaigns, especially for COASTlOOC 3 (Adriatic Sea). Surface-to-bottom depth merging of the full AC-9 profiles has not been attempted so far, but this is envisaged for a potential future reprocessing of the COASTℓOOC dataset (see Section 5).

3 Data quality control and data processing

Different general quality control procedures were adopted to ensure the integrity of the data. First, the raw data were visualized and screened to eliminate errors both originating from the measurement devices, including sensors (systematic or random), and errors inherent to measurement procedures and methods. Statistical metrics such as average, standard deviation, and range were computed to detect and remove anomalous values in the data. Then, data were checked for duplicates and remaining outliers. The complete list of variables is presented in Table 1.

4 Data description (an overview)

4.1 Spatial variability along the coastal-ocean gradient

The COASTℓOOC sampling strategy was primarily designed to capture the bio-optical gradient across the sampled ecosystems and along transects from the coast towards the open ocean within each area (Fig. 1). The following sections present an overview of a few selected variables measured in the different areas. An extensive and detailed explanatory visualization analysis is presented in the final report of the COASTℓOOC project (Aiken et al., 2000).

4.1.1 chlorophyll *a* and particulate organic carbon

As observed in Fig. 4, both total chlorophyll *a* and particulate organic carbon (POC) concentrations varied markedly across the sampled areas, reflecting the natural gradient captured by the sampling strategy. Across all stations, total chlorophyll *a* ranged between 0.05 and 29 mg m^{-3}. The median chlorophyll *a* ranged between ca. 0.15 mg m^{-3} in the Atlantic Ocean and 5.9 mg m^{-3} in the Baltic Sea (Fig. 4A). Based on the definition proposed by Antoine et al. 1996, the sampled stations were representative of a wide range of trophic status: oligotrophic ($n = 17$, chlorophyll $a \leq 0.1$ mg m^{-3}), mesotrophic ($n = 106$, $0.1 \leq$ chlorophyll $a \leq 1$ mg m^{-3}) and eutrophic ($n = 245$, chlorophyll $a > 1$ mg m^{-3}). A similar shore-to-sea gradient pattern could be observed for POC, with median values varying between ca. 0.07 g m^{-3} and 0.8 g m^{-3} in the Mediterranean Sea (case 2) and Baltic Sea, respectively (Fig. 4B, range between 0.02 and 2.5 g m^{-3}).

4.1.2 Inherent optical properties (IOPs)

Measured absorption spectra in each area for total particulate (a_P), non-algal particles (a_{NAP}), phytoplankton (a_ϕ), and chromophoric dissolved organic matter (a_{CDOM}) are presented in Fig. 5. Phytoplankton maximum absorption peaks at around 440 nm and 675 nm are easily distinguishable in Fig. 5C and were found to be highly correlated with the concentration of total chlorophyll *a* (Pearson's $r > 0.90$). Across all the areas, $a_{CDOM}(350)$, a proxy for dissolved carbon concentration, varied between 0.03 and 3.66 m^{-1}. These values fall within the ranges of the mean values reported globally in the ocean (0.14 m^{-1}), coastal (1.82 m^{-1}), and estuarine (4.11 m^{-1}) ecosystems (Massicotte et al., 2017). Overall, the highest absorption by dissolved organic matter was observed in the Baltic Sea (Fig. 5D). Sampled stations in this area were located West of the Oder River plume, which drains important quantities of humic substance from its catchment area. In contrast, the lowest CDOM absorption was observed in the Atlantic Ocean, where stations were located away from land and thus less influenced by terrestrial inputs (Fig. 1). The spatial variability of CDOM is further emphasized in Fig. 5E, where $a_{CDOM}(350)$ decreased by a factor of four between the two most distant points (approximately 40 km) of the westernmost transect sampled in the North Sea (see Fig. 1B).

Light scattering by suspended particles in the water column is a driver of reflectance variability and is often used by remote sensing applications to discriminate between Case 1 and Case 2 waters (Sathyendranath, 2000; Morel and Bélanger, 2006). The particulate scattering coefficient at 440 nm, $b_p(440)$ ranged between 0.05 m^{-1} and 35.8 m^{-1} (Fig. 6A). The median values varied by almost two orders of magnitude between the Atlantic Ocean and the Adriatic Sea (Fig. 6A). Furthermore, $b_p(440)$ showed more or less the same spatial pattern as the downward irradiance attenuation coefficient at 443 nm, $K_{Ed}(443)$, with median $b_p(440)$ values varying between 0.05 m^{-1} and 1.21 m^{-1} in the Atlantic Ocean and the North Sea respectively (Fig. 6B, Pearson's $r = 0.76$).

4.1.3 Apparent optical properties (AOPs)

Subsurface reflectance $R(0^-)$ for all COASTℓOOC sites is shown in Fig. 7, reflecting the fundamental differences in the bio-optical properties of the different areas visited. For example, due to enhanced CDOM absorption, reflectance in the blue part of

the spectrum is generally low at the Baltic Sea sites, while it is higher in the Case-1 waters encountered in the Atlantic Ocean or the Mediterranean Sea. In the red and NIR, very high reflectance values are observed in the sediment-rich Case-2 waters of the North Sea and the English Channel.

Sample profiles of the diffuse attenuation coefficient $K_{Ed}(490)$ at a vertical resolution between 2.0 m close to the surface and 5.0 m underneath are shown in Fig. 8 for two transition zones ranging from turbid waters close to the coast to clearer waters offshore. Fig. 8A shows $K_{Ed}(490)$ values along a transect taken in the Gulf of Lion on 30. Sept. 1997 from site C4014 located ca. 5 km to the west of the mouth of the main Rhône branch (Grand Rhône) to site C4019 some 20 km further south in clear Mediterranean waters. While $K_{Ed}(490)$ at the offshore sites C4017 to C4019 adopts low values between ca. 0.02 and 0.05 m^{-1} reflecting clear-water conditions, sites C4014 and C4015 are strongly influenced by the Rhône river plume, resulting in $K_{Ed}(490)$ reaching values of up to 0.8 m^{-1} in the surface layer. The turbid Rhône freshwater is floating on top of denser and clearer Mediterranean seawater, resulting in a strong stratification. Fig. 8B shows a similar transect taken in the North Sea on 12. Sept. 1998 extending from site C6063 close to the western shore of Texel island to site C6069 located ca. 22 km further west in the open North Sea. Again, there is a distance-to-coast-dependent decrease of $K_{Ed}(490)$ values likely related to reduced turbidity, from ca. 0.8 to 1.0 m^{-1} close to the coast down to values of ca. 0.2 to 0.3 m^{-1} offshore. In contrast to the transect in the Gulf of Lion, no obvious stratification is observed, $K_{Ed}(490)$ values are rather homogeneously distributed over the entire water column, pointing to effective mixing.

4.2 Co-variability across bio-optical measurements

Few research papers have previously shown how optically significant components and AOPs/IOPs measured during the COASTℓOOC missions co-varied (Ferrari, 2000; Ferrari et al., 2003; Babin, 2003; Babin et al., 2003). In this section, a brief overview of selected pairwise relationships is presented. Total chlorophyll a co-varies with POC (Fig. 4) and the relationship was highly variable (Fig. 9A). Whereas the global Pearson's correlation is 0.79 ($n = 299$), it ranged between 0.36 in the North Sea ($n = 88$) to 0.78 in the Adriatic Sea ($n = 37$) showing that optically significant components do not necessarily covary altogether. Likewise, the global Pearson's correlation between a_ϕ and total chlorophyll a (Fig. 9B) was relatively high ($r = 0.90$, $n = 338$). The lowest and highest correlations were observed in the North Sea ($r = 0.69$, $n = 88$) and the Adriatic Sea ($r = 0.91$, $n = 38$), respectively. POC concentration is well known to be an important driver of both IOPs and AOPs in aquatic ecosystems (Stramski et al., 2008; Cetinić et al., 2012). Unsurprisingly, positive correlations were observed between POC and $K_{Ed}(443)$ (Fig. 9C, $r = 0.81$, $n = 204$) and $b_p(440)$ (Fig. 9D, $r = 0.67$, $n = 165$).

Based on the size-reactivity continuum model proposed by Benner and Amon 2015, physicochemical and photochemical processes are shaping the size distribution of organic matter along the aquatic continuum, and determine the contrasted intrinsic nature of the particles for each type of ecosystem. This is likely one key reason explaining why the observed relationships are quite variable across the different sampled ecosystems (Fig. 9). For example, in coastal areas (e.g., Baltic Sea, Fig. 1), the POC content is generally influenced by large humic organic particles drained by rivers from the surrounding watershed (Babin et al., 2003). Consequently, in these areas, POC concentration (Fig. 4B), absorption (Fig. 5), and downward attenuation coefficients

(Fig. 9C) were higher than in other sampled areas. In contrast, the stations that were farther away from the coast (such as the Atlantic Sea, Fig. 1) had lower POC concentrations (Fig. 4B) and absorption (Fig. 5).

5 Code and data availability

The COASTℓOOC data is provided as a collection of comma-separated values (CSV) files that are regrouping measurements
associated with each measurement. The processed and tidied version of the data is hosted at SEANOE (SEA scieNtific Open data Edition) under the CC-BY license (https://www.seanoe.org/data/00824/93570/, Massicotte et al. 2023). To aid the user to merge these files, there is a lookup table file designated station list (https://www.seanoe.org/data/00824/93570/data/100311.csv) that can serve as a table to join the data together based on the station's unique identifier. Table 1 shows the complete list of available measurements. Please note that the actual filenames (e.g., /100311.csv) are assigned by SEANOE during the upload
process and cannot be altered. All statistical analyses were performed in R version 4.3.0 (R Core Team, 2023). The code used to produce the figures and the analysis presented in this paper is available under the GNU GPLv3 license (10.5281/zenodo.8091717, Massicotte 2023). Irradiance depth merging was done with an IDL code available under the Creative Commons Attribution 3.0 Germany license (10.5281/zenodo.8096682). Note that there are plans to update and further improve the current data set. More data will likely be recovered from the backup archives, most prominently AC-9 measurements to derive
in situ vertical profiles of a, b_p and c. Intended improvements to radiometer processing comprise (not exhaustively) a better dark current correction, the application of a self-shading correction, a more accurate treatment of the air/sea transmission of downward irradiance (E_d), and a closer look at uncertainties. Any update of the COASTℓOOC dataset will be made publicly available through SEANOE.

6 Conclusions

The consolidated, quality-controlled, data collected during the COASTℓOOC oceanographic expeditions offer, even if collected more than twenty years ago, many possibilities to better understand the bio-optical dynamics along the land-to-sea gradient where optically-complex water drained from watersheds mix with seawater. To this date, nearly 40 peer-reviewed studies have made use of the COASTℓOOC data set to develop remote sensing algorithms and to study the optical properties of seawater across the coastal-to-open ocean gradient (see Table A1). This now consolidated and easily accessible data set will
facilitate future development and evaluation of new bio-optical models adapted for optically-complex waters. In this paper, only a subset of variables has been presented. The reader is referred to Table 1 for a complete list and description of variables collected during the COASTℓOOC expeditions.
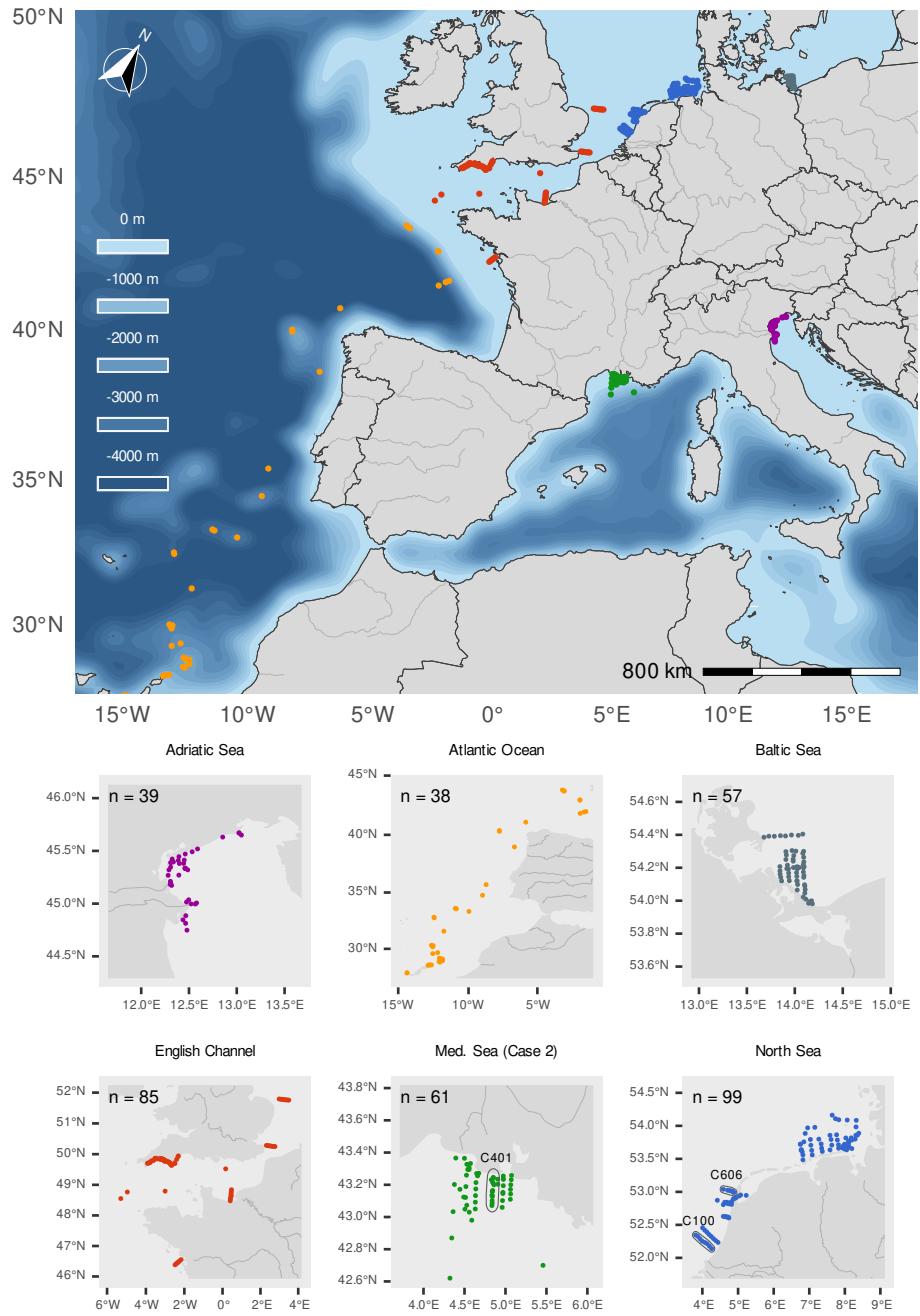
Figure 1. Map of the sampling stations. The three black outlines in the bottom panels (Mediterranean Sea and North Sea) identify the stations and the transects used in Fig. 5E (C100) as well as for the examples of vertical profiles of $K_{\mathrm{Ed}}(490)$ in Fig. 8 (C401 and C606).
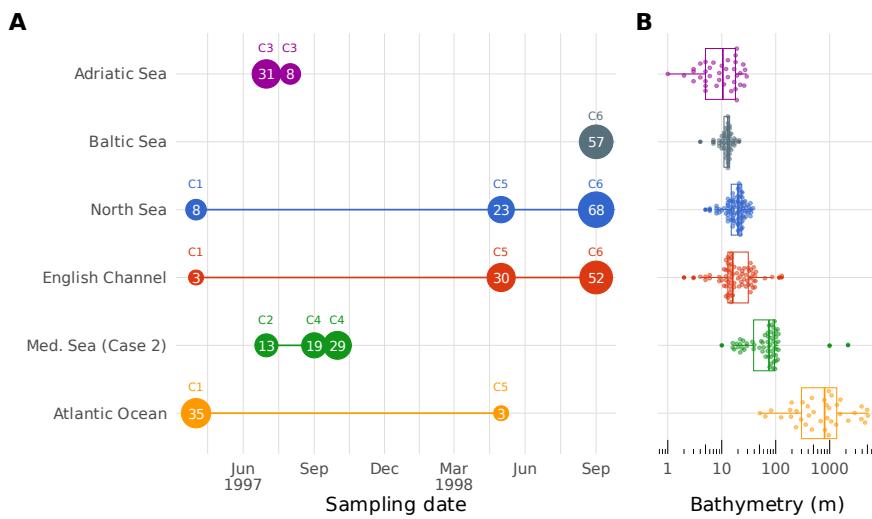
Figure 2. (**A**) Overview of the temporal sampling for the six areas. The numbers in the circles indicate the number of visited stations each month. (**B**) Boxplot showing the bathymetry at the sampling locations by area. The labels on top of each circle identify the COASTℓOOC campaigns (C1 to C6).
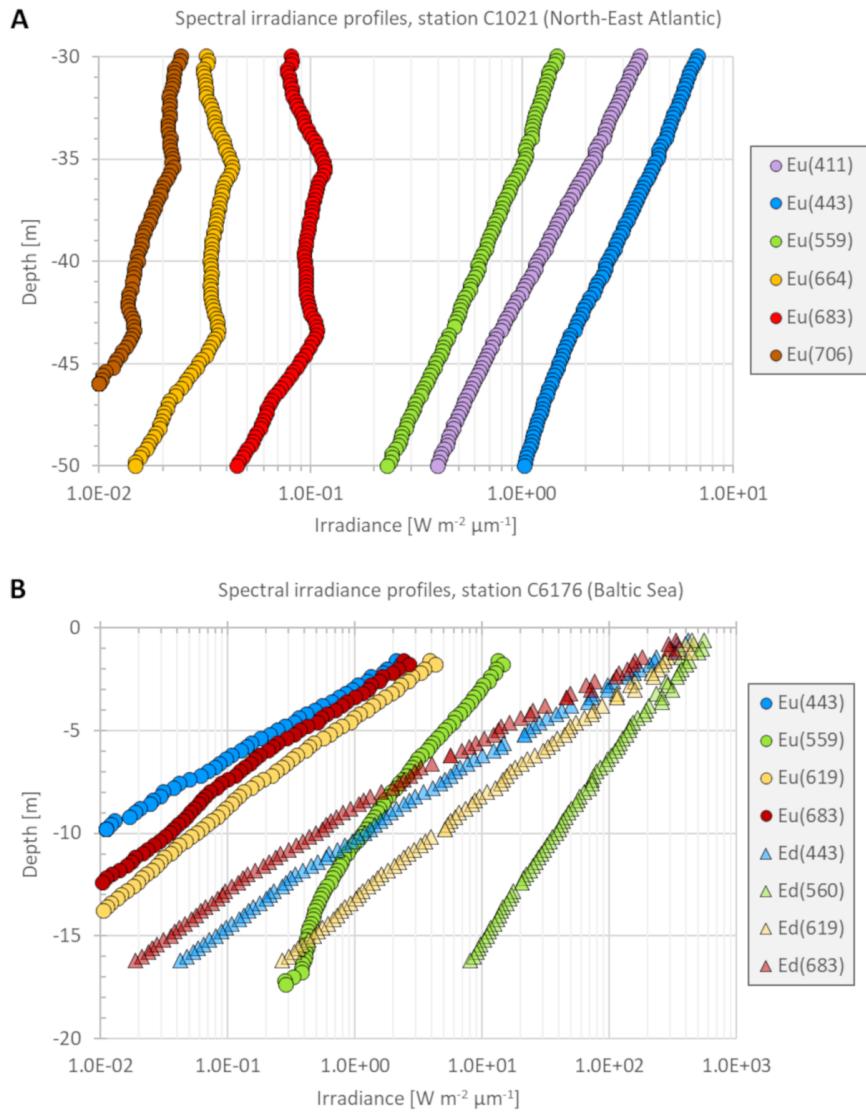
Figure 3. (A) Vertical profile of the upwelling irradiance taken on 11 April 1997 at station C1021 located in the North-East Atlantic offshore Northern Portugal. (B) Vertical profiles of the upwelling and downwelling irradiance taken on 25 September 1998 at station C6176 located in the Baltic Sea offshore Usedom Island in North-East Germany. See text for further detail.
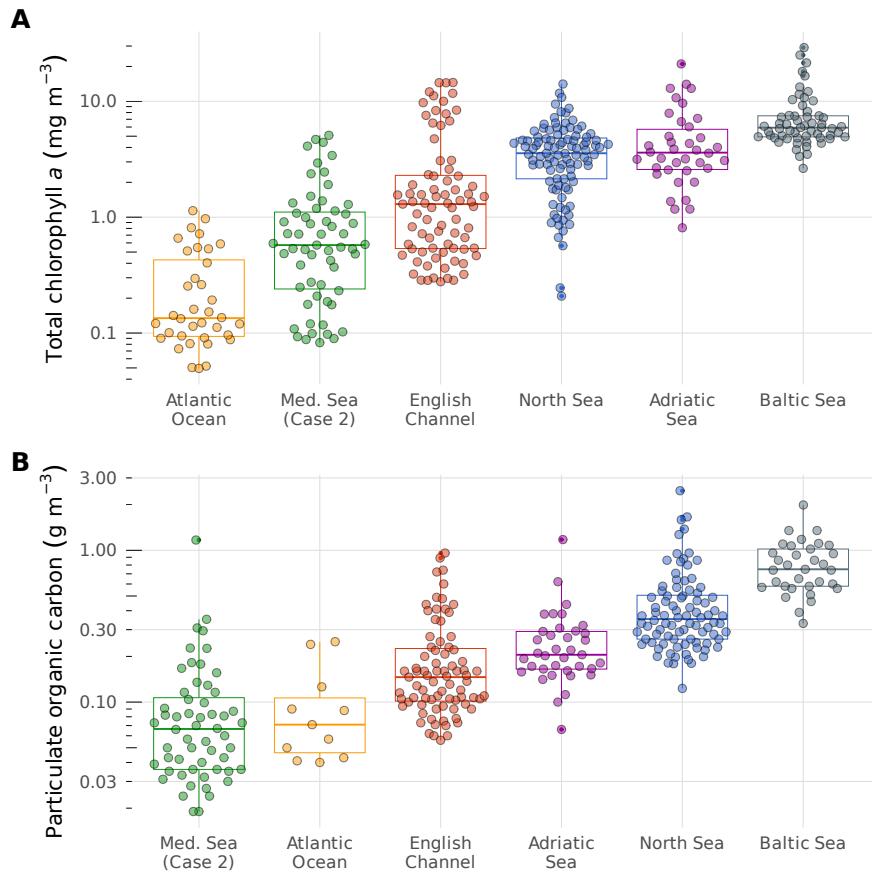
Figure 4. (**A**) Total chlorophyll *a* and (**B**) particulate organic carbon across the sampled areas.
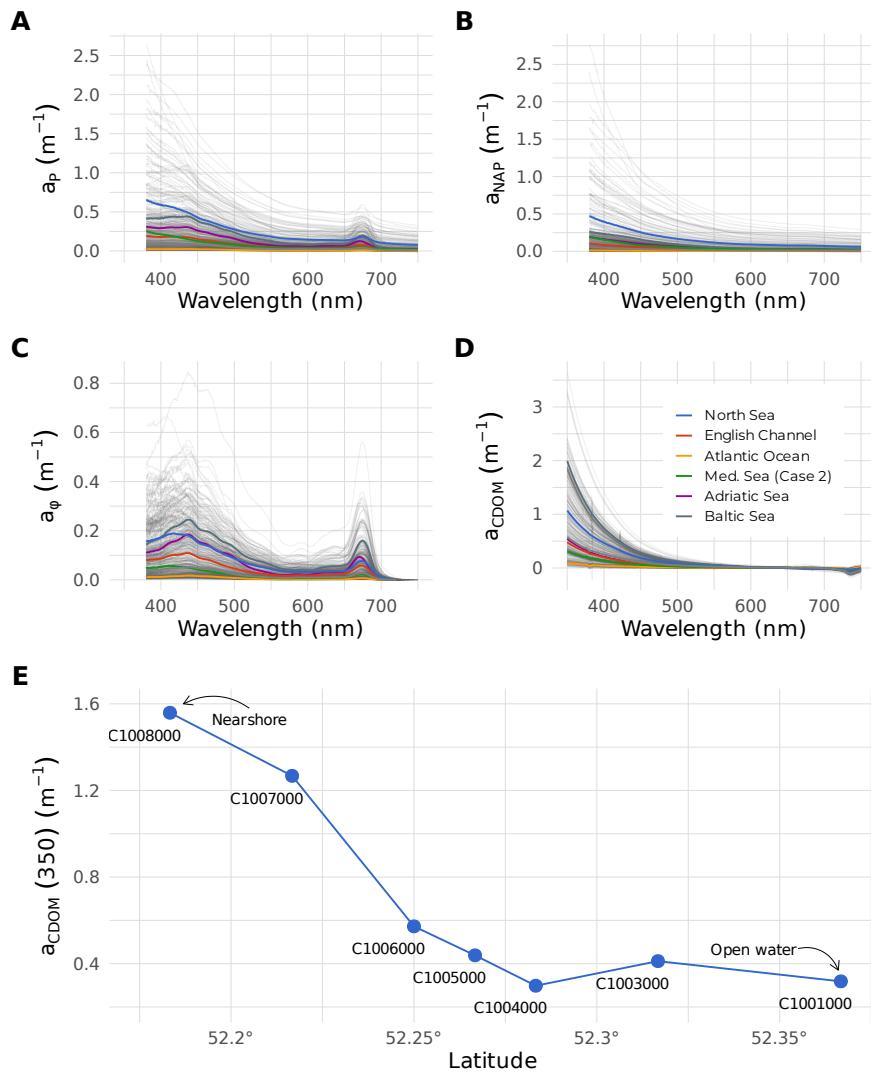
Figure 5. (**A**) Average total particulate (a_p), (**B**) non-algal (a_NAP), (**C**) phytoplankton (a_ϕ) and (**D**) chromophoric dissolved organic matter (a_CDOM) absorption spectra in each area. Light gray lines in panels A-D show individual absorption spectra measured across the sampled areas. (**E**) $a_\mathrm{CDOM}(350)$ along the westernmost transect in the North Sea (see Fig. 1B).
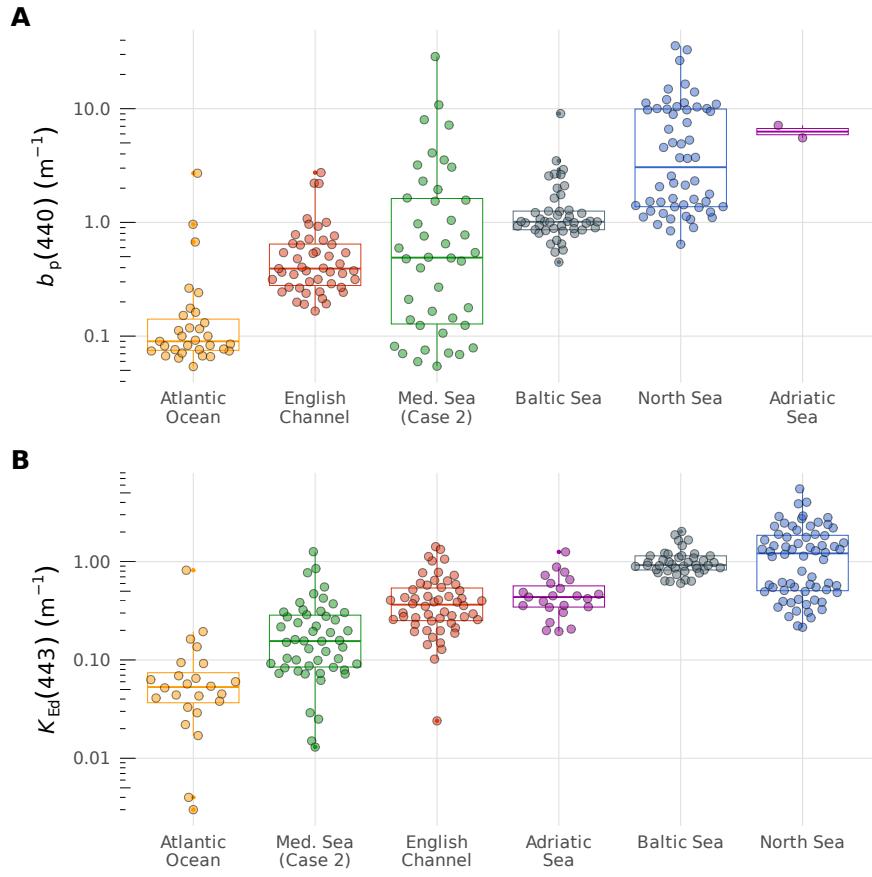
Figure 6. (**A**) Particulate scattering coefficient at 440 nm ($b_b(440)$) and (**B**) attenuation coefficient for downward irradiance at 443 nm ($K_{Ed}(443)$) across the sampled areas.
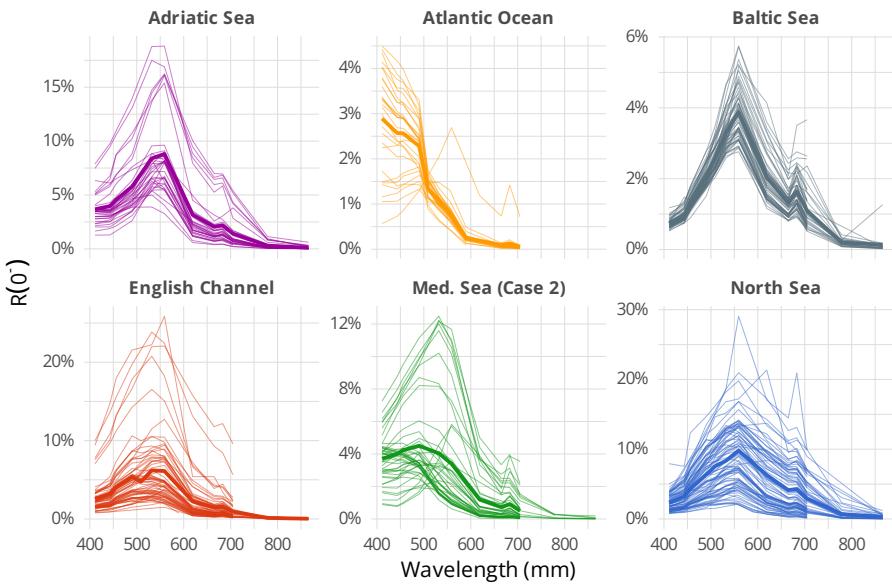
Figure 7. Subsurface reflectance $R\left(0^{-}\right)$ across the sampled areas. The thick lines represent the regional averages. Note that $R\left(0^{-}\right)$ at wavelengths above 705 nm could only be derived from helicopter-based measurements. Please also note the y-axes are adjusted to the data presented within each subplot.
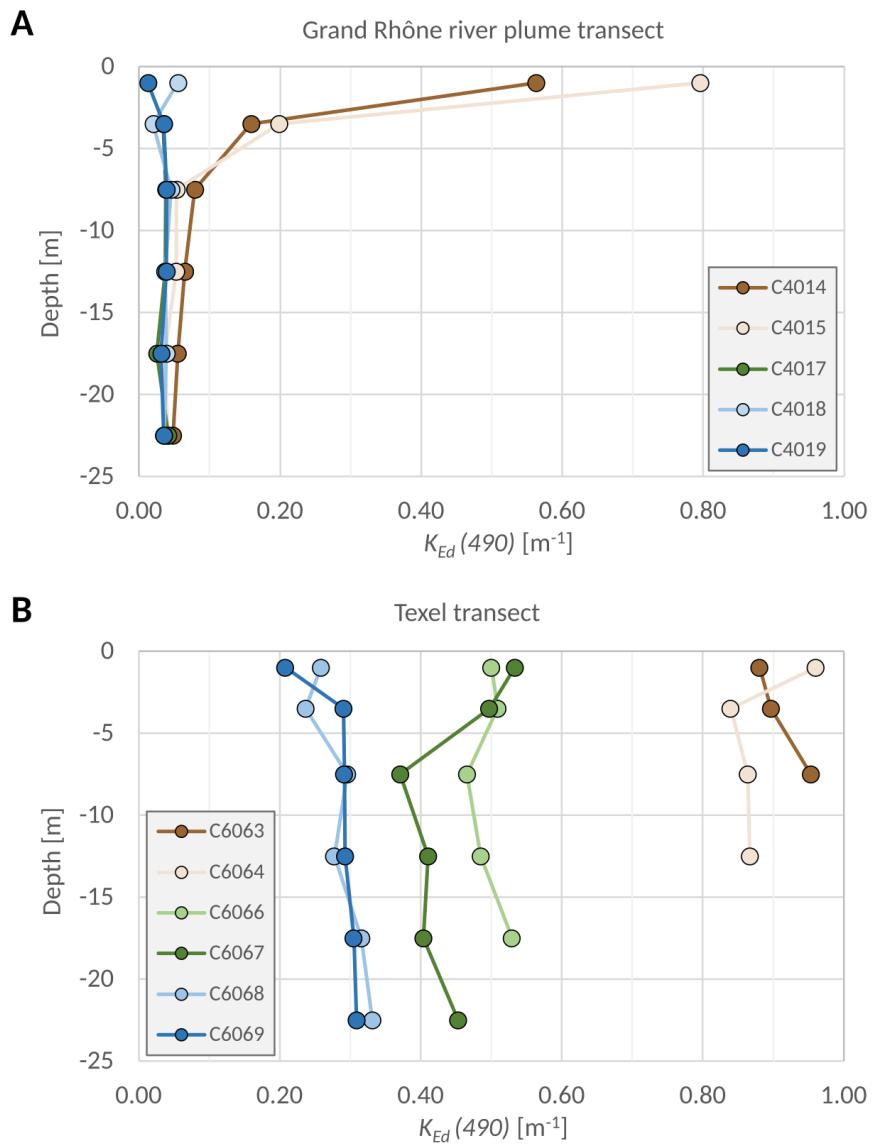
Figure 8. (A) $K_{Ed}(490)$ profiles taken in the Lion's Gulf on 30. Sept. 1997 along a transect extending from site C4014 located ca. 5 km to the west of the mouth of the main Rhône branch (Grand Rhône) to site C4019 located some 20 km further south in clear Mediterranean waters. **(B)** $K_{Ed}(490)$ profiles taken in the North Sea on 12. Sept 1998 along a transect extending from site C6063 close to the western shore of the island of Texel to site C6069 located ca. 22 km further west in the open North Sea.
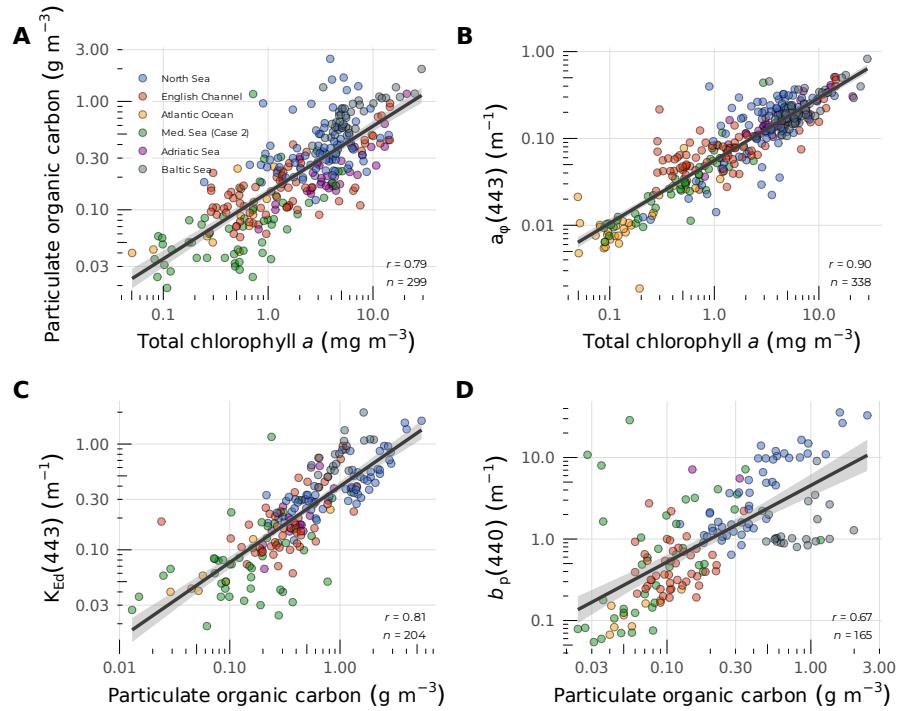
Figure 9. Scatterplots showing relationships among different selected variables. (**A**) Particulate organic carbon (POC) and (**B**) phytoplankton absorption at 443 nm ($a_\phi(443)$) against total chlorophyll *a*. (**C**) Downward irradiance at 443 nm ($K_{Ed}(443)$) and (**D**) particulate scattering at 440 nm ($b_b(440)$) against particulate organic carbon. The red lines show the linear relationships between the variables. The shaded gray areas represent the 95% confidence intervals around the fitted models.

Table 1: List of measured parameters

Variable	Units	Source file	Description
wavelength	nm	100303.csv	
a_p_m1	m^{-1}	100303.csv	Total particulate absorption
a_nap_m1	m^{-1}	100303.csv	Non-algal absorption
a_nap_adjusted_m1	m^{-1}	100303.csv	Non-algal absorption adjusted so that baseline background is equal to that of ap
a_cdom_m1	m^{-1}	100303.csv	Chromophoric dissolved organic matter absorption
a_cdom_adjusted_m1	m^{-1}	100303.csv	Chromophoric dissolved organic matter absorption with background baseline removed
a_phy_m1	m^{-1}	100303.csv	Phytoplankton absorption

Table 1: List of measured parameters *(continued)*

Variable	Units	Source file	Description
background_a_p_m1	m^{-1}	100303.csv	Baseline background of total particulate absorption
background_a_cdom_m1	m^{-1}	100303.csv	Baseline background of chromophoric dissolved organig matter absorption
background_a_nap_m1	m^{-1}	100303.csv	Baseline background of non-algal absorption
a_m1	m^{-1}	100304.csv	Total non-water absorption coefficient
c_m1	m^{-1}	100304.csv	Total non-water attenuation coefficient
bp_m1	m^{-1}	100304.csv	Particle scattering coefficient
longitude	Degree decimal	100305.csv	Longitude of the pixel used to extract the bathymetry
latitude	Degree decimal	100305.csv	Latitude of the pixel used to extract the bathymetry
bathymetry_m	m	100305.csv	Bathymetry depth at the sampled stations
eu_w_m2_um	W m-2 μm-1	100307.csv	Upward irradiance just beneath the water surface (Eu0-)
ed_w_m2_um	W m-2 μm-1	100307.csv	Downward irradiance just beneath the water surface (Ed0-)
k_eu_m1	m^{-1}	100307.csv	Attenuation coefficient for upward irradiance (just beneath the water surface)
k_ed_m1	m^{-1}	100307.csv	Attenuation coefficient for downward irradiance (just beneath the water surface)
measured_reflectance_percent	Percent	100309.csv	Surface water reflectance
s_cdom_nm1	nm^{-1}	100310.csv	Spectral slope that describes the approximate exponential decrease in aCDOM
s_nap_nm1	nm^{-1}	100310.csv	Spectral slope that describes the approximate exponential decrease in aNAP
a_cdom443_m1	m^{-1}	100310.csv	Chromophoric dissolved organic matter absorption at 443 nm
a_nap443_m1	m^{-1}	100310.csv	Non-algal absorption at 443 nm
station	free text	100311.csv	Unique ID of the sampled station. Can be used as unique key to merge the data across other files.
date	YYYY-MM-DD	100311.csv	Date at which the measurement was made
depth_m	m	100311.csv	Depth at which the measurement was made
longitude	Degree decimal	100311.csv	Longitude of the sampling station
latitude	Degree decimal	100311.csv	Latitude of the sampling station

Variable	Units	Source file	Description
area	free text	100311.csv	Region where the measurement was made. One of: (1) North Sea, (2) English Channel, (3) Atlantic Ocean, (4) Med. Sea (Case 2), (5) Adriatic Sea, (6) Baltic Sea
system	free text	100311.csv	Dominant river system influencing the site
gmt_time	Hours decimal	100311.csv	Representative time of observations at site in Greenwich Mean Time (UTC)
solar_zenith_angle	degree	100311.csv	Solar zenith angle
chlorophyll_a_mg_m3	mg m^{-3}	100308.csv	Chlorophyll-a
chlorophyll_b_mg_m3	mg m^{-3}	100308.csv	Chlorophyll-b
chlorophyll_c_mg_m3	mg m^{-3}	100308.csv	Chlorophyll-c
pheopigment_mg_m3	mg m^{-3}	100308.csv	Pheopigment
fucoxanthin_mg_m3	mg m^{-3}	100308.csv	Fucoxanthin
hexanoyloxyfucoxanthin_19_mg_m3	mg m^{-3}	100308.csv	Hexanoyloxyfucoxanthin-19
butanoyloxyfucoxanthin_19_mg_m3	mg m^{-3}	100308.csv	Butanoyloxyfucoxanthin-19
alloxanthin_mg_m3	mg m^{-3}	100308.csv	Alloxanthin
zeaxanthin_mg_m3	mg m^{-3}	100308.csv	Zeaxanthin
prasixanthin_mg_m3	mg m^{-3}	100308.csv	Prasixanthin
neoxanthin_mg_m3	mg m^{-3}	100308.csv	Neoxanthin
violaxanthin_mg_m3	mg m^{-3}	100308.csv	Violaxanthin
diatoxanthin_mg_m3	mg m^{-3}	100308.csv	Diatoxanthin
diadinoxanthin_mg_m3	mg m^{-3}	100308.csv	Diadinoxanthin
peridinin_mg_m3	mg m^{-3}	100308.csv	Peridinin
carotene_mg_m3	mg m^{-3}	100308.csv	Carotene
lutein_mg_m3	mg m^{-3}	100308.csv	Lutein
suspended_particulate_matter_g_m3	g m^{-3}	100306.csv	Suspended particulate matter
particulate_organic_nitrogen_g_m3	g m^{-3}	100306.csv	Particulate organic nitrogen
total_particulate_carbon_g_m3	g m^{-3}	100306.csv	Total particulate carbon
particulate_organic_carbon_g_m3	g m^{-3}	100306.csv	Particulate organic carbon
dissolved_organic_carbon_g_m3	g m^{-3}	100306.csv	Dissolved organic carbon

Variable	Units	Source file	Description
Cast	1	100312.csv	Processed cast number
Depth	m	100312.csv	Depth of vertical bin, e.g. -1.00 representing the depth bin [-0.90, -1.10 m]
TmpWat	Degree celsius	100312.csv	Water temperature
Cond	ms cm^{-1}	100312.csv	Conductivity
Salin	PSU	100312.csv	Salinity
SigmaT	1	100312.csv	Quantity (no units) to specify the density of sea water
TiProf	Degree	100312.csv	Tilt of profiling radiometer
TiRef	Degree	100312.csv	Tilt of reference radiometer
VSpeed	m s^{-1}	100312.csv	Vertical speed
Altim	m	100312.csv	Altimeter sounding of distance from the ocean ground
N_OBS	1	100312.csv	Number of observations within depth bin
EUnnn	W m-2 μm-1	100312.csv	In-water upwelling irradiance at wavelength nnn
EDnnn	W m-2 μm-1	100312.csv	In-water downwelling irradiance at wavelength nnn
ERnnn	W m-2 μm-1	100312.csv	In-air downwelling irradiance at wavelength nnn
KUnnn	m^{-1}	100312.csv	Diffuse attenuation at wavelength nnn calculated from the upwelling irradiance
KDnnn	m^{-1}	100312.csv	Diffuse attenuation at wavelength nnn calculated from the downwelling irradiance
PAR_ABS	μmol m^{-2} s^{-1}	100312.csv	Phytosynthetically Active Radiation (PAR)
PAR%SRF	Percent	100312.csv	PAR at depth z relative to PAR on the sea surface
K_PAR	m^{-1}	100312.csv	Diffuse attenuation for PAR

Appendix A

A1
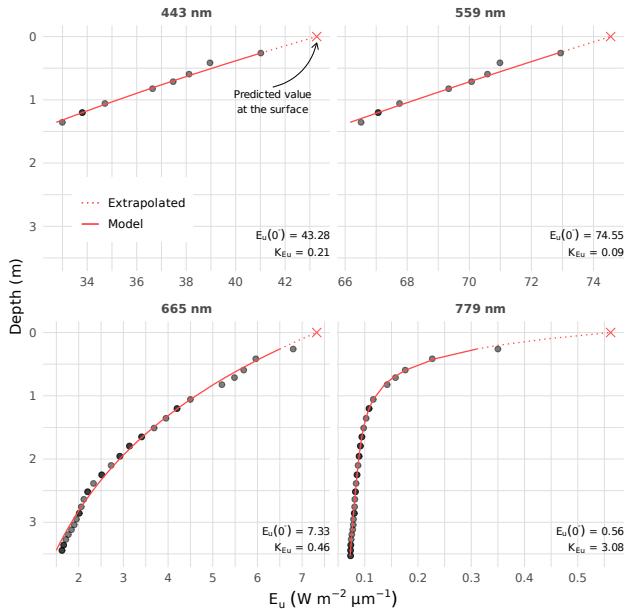
Figure A1. Example of extrapolation of the upward irradiance towards the sea surface (station C3004000). The irradiance values just below the sea surface (red crosses, $E_u\,(0^-)$) and the corresponding diffuse attenuation coefficient for the topmost water layer are specified in the figures. The full lines represent the interpolated values whereas the dashed lines show extrapolation values up to the sea surface.

Table A1: Scientific articles in peer-reviewed journals using or referencing COASTLOOC

Publications
Babin, M., Stramski, D., Ferrari, G. M., Claustre, H., Bricaud, A., Obolensky, G., & Hoepffner, N. (2003). Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. Journal of Geophysical Research: Oceans, 108(C7).
Begouen Demeaux, C., & Boss, E. (2022). Validation of Remote-Sensing Algorithms for Diffuse Attenuation of Downward Irradiance Using BGC-Argo Floats. Remote Sens. 2022, 14, 4500.
Belanger, S., Babin, M., & Larouche, P. (2008). An empirical ocean color algorithm for estimating the contribution of chromophoric dissolved organic matter to total light absorption in optically complex waters. Journal of Geophysical Research: Oceans, 113(C4).
Beltrán-Abaunza, J. M., Kratzer, S., & Brockmann, C. (2014). Evaluation of MERIS products from Baltic Sea coastal waters rich in CDOM. Ocean Science, 10(3), 377-396.
Blix, K., Li, J., Massicotte, P., & Matsuoka, A. (2019). Developing a new machine-learning algorithm for estimating chlorophyll-a concentration in optically complex waters: A case study for high northern latitude waters by using Sentinel 3 OLCI. Remote Sensing, 11(18), 2076.
Caillault, K., Roupioz, L., & Viallefont-Robinet, F. (2021). Modelling of the optical signature of oil slicks at sea for the analysis of multi-and hyperspectral VNIR-SWIR images. Optics Express, 29(12), 18224-18242.

Publications

Chami, M., & Platel, M. D. (2007). Sensitivity of the retrieval of the inherent optical properties of marine particles in coastal waters to the directional variations and the polarization of the reflectance. Journal of Geophysical Research: Oceans, 112(C5).

Claustre, H., Fell, F., Oubelkheir, K., Prieur, L., Sciandra, A., Gentili, B., & Babin, M. (2000). Continuous monitoring of surface optical properties across a geostrophic front: Biogeochemical inferences. Limnology and Oceanography, 45(2), 309-321.

D'Alimonte, D., Zibordi, G., Kajiyama, T., & Berthon, J. F. (2014). Comparison between MERIS and regional high-level products in European seas. Remote sensing of environment, 140, 378-395.

Defoin-Platel, M., & Chami, M. (2007). How ambiguous is the inverse problem of ocean color in coastal waters?. Journal of Geophysical Research: Oceans, 112(C3).

Doerffer, R., & Schiller, H. (2007). The MERIS Case 2 water algorithm. International Journal of Remote Sensing, 28(3-4), 517-535.

Doron, M., Babin, M., Mangin, A., & Hembise, O. (2007). Estimation of light penetration, and horizontal and vertical visibility in oceanic and coastal waters from surface reflectance. Journal of Geophysical Research: Oceans, 112(C6).

Doron, M., Babin, M., Hembise, O., Mangin, A., & Garnesson, P. (2011). Ocean transparency from space: Validation of algorithms estimating Secchi depth using MERIS, MODIS and SeaWiFS data. Remote Sensing of Environment, 115(12), 2986-3001.

Dransfeld, S., Tatnall, A. R., Robinson, I. S., & Mobley, C. D. (2005). Prioritizing ocean colour channels by neural network input reflectance perturbation. International Journal of Remote Sensing, 26(5), 1043-1048.

Ferrari, G. M. (2000). The relationship between chromophoric dissolved organic matter and dissolved organic carbon in the European Atlantic coastal area and in the West Mediterranean Sea (Gulf of Lions). Marine Chemistry, 70(4), 339-357.

Groom, S., Martinez-Vicente, V., Fishwick, J., Tilstone, G., Moore, G., Smyth, T., & Harbour, D. (2009). The western English Channel observatory: Optical characteristics of station L4. Journal of Marine Systems, 77(3), 278-295.

Jamet, C., Loisel, H., & Dessailly, D. (2012). Retrieval of the spectral diffuse attenuation coefficient Kd (lambda) in open and coastal ocean waters using a neural network inversion. Journal of Geophysical Research: Oceans, 117(C10).

Kratzer, S., & Moore, G. (2018). Inherent optical properties of the baltic sea in comparison to other seas and oceans. Remote Sensing, 10(3), 418.

Loisel, H., Stramski, D., Mitchell, B. G., Fell, F., Fournier-Sicre, V., Lemasle, B., & Babin, M. (2001). Comparison of the ocean inherent optical properties obtained from measurements and inverse modeling. Applied Optics, 40(15), 2384-2397.

Loisel, H., Vantrepotte, V., Ouillon, S., Ngoc, D. D., Herrmann, M., Tran, V., ... & Van Nguyen, T. (2017). Assessment and analysis of the chlorophyll-a concentration variability over the Vietnamese coastal waters from the MERIS ocean color sensor (2002–2012). Remote sensing of environment, 190, 217-232.

Loisel, H., Stramski, D., Dessailly, D., Jamet, C., Li, L., & Reynolds, R. A. (2018). An inverse model for estimating the optical absorption and backscattering coefficients of seawater from remote-sensing reflectance over a broad range of oceanic and coastal marine environments. Journal of Geophysical Research: Oceans, 123(3), 2141-2171.

Matsuoka, A., Hill, V., Huot, Y., Babin, M., & Bricaud, A. (2011). Seasonal variability in the light absorption properties of western Arctic waters: Parameterization of the individual components of absorption for ocean color applications. Journal of Geophysical Research: Oceans, 116(C2).

Publications

Matsuoka, A., Babin, M., Doxaran, D., Hooker, S. B., Mitchell, B. G., Bélanger, S., & Bricaud, A. (2014). A synthesis of light absorption properties of the Arctic Ocean: application to semi-analytical estimates of dissolved organic carbon concentrations from space. Biogeosciences, 11(12), 3131-3147.

Morel, A., & Bélanger, S. (2006). Improved detection of turbid waters from ocean color sensors information. Remote Sensing of Environment, 102(3-4), 237-249.

Neukermans, G., Loisel, H., Mériaux, X., Astoreca, R., & McKee, D. (2012). In situ variability of mass-specific beam attenuation and backscattering of marine particles with respect to particle size, density, and composition. Limnology and oceanography, 57(1), 124-144.

Oubelkheir, K., Claustre, H., Bricaud, A., & Babin, M. (2007). Partitioning total spectral absorption in phytoplankton and colored detrital material contributions. Limnology and Oceanography: Methods, 5(11), 384-395.

Schroeder, T., Behnert, I., Schaale, M., Fischer, J., & Doerffer, R. (2007). Atmospheric correction algorithm for MERIS above case-2 waters. International Journal of Remote Sensing, 28(7), 1469-1486.

Schroeder, T., Schaale, M., Lovell, J., & Blondeau-Patissier, D. (2022). An ensemble neural network atmospheric correction for Sentinel-3 OLCI over coastal waters providing inherent model uncertainty estimation and sensor noise propagation. Remote Sensing of Environment, 270, 112848.

Shahraiyni, T. H., Schaale, M., Fell, F., Fischer, J., Preusker, R., Vatandoust, M., ... & Tavakoli, A. (2007). Application of the Active Learning Method for the estimation of geophysical variables in the Caspian Sea from satellite ocean colour observations. International Journal of Remote Sensing, 28(20), 4677-4683.

Tassan, S., Ferrari, G. M., Bricaud, A., & Babin, M. (2000). Variability of the amplification factor of light absorption by filter-retained aquatic particles in the coastal environment. Journal of Plankton Research, 22(4), 659-668.

Tassan, S., & Ferrari, G. M. (2002). A sensitivity analysis of the 'Transmittance–Reflectance' method for measuring light absorption by aquatic particles. Journal of Plankton Research, 24(8), 757-774.

Wei, J., Wang, M., Jiang, L., Yu, X., Mikelsons, K., & Shen, F. (2021). Global estimation of suspended particulate matter from satellite ocean color imagery. Journal of Geophysical Research: Oceans, 126(8), e2021JC017303.

Yu, X., Salama, M. S., Shen, F., & Verhoef, W. (2016). Retrieval of the diffuse attenuation coefficient from GOCI images using the 2SeaColor model: A case study in the Yangtze Estuary. Remote Sensing of Environment, 175, 109-119.

Yu, X., Lee, Z., Shen, F., Wang, M., Wei, J., Jiang, L., & Shang, Z. (2019). An empirical algorithm to seamlessly retrieve the concentration of suspended particulate matter from water color across ocean to turbid river mouths. Remote Sensing of Environment, 235, 111491.

Zhang, T., Fell, F., Liu, Z. S., Preusker, R., Fischer, J., & He, M. X. (2003). Evaluating the performance of artificial neural network techniques for pigment retrieval from ocean color in Case I waters. Journal of Geophysical Research: Oceans, 108(C9).

Zhang, T., & Fell, F. (2004). An approach to improving the retrieval accuracy of oceanic constituents in Case II waters. Journal of Ocean University of China, 3(2), 220-224.

Zhang, T., & Fell, F. (2007). An empirical algorithm for determining the diffuse attenuation coefficient Kd in clear and turbid waters from spectral remote sensing reflectance. Limnology and Oceanography: Methods, 5(12), 457-462.

Zheng, G., & Stramski, D. (2013). A model based on stacked-constraints approach for partitioning the light absorption coefficient of seawater into phytoplankton and non-phytoplankton components. Journal of Geophysical Research: Oceans, 118(4), 2155-2174.

335 *Author contributions.* Conceptualization: MB, Formal analysis: FF, DD, VSF, MB. Investigation: MB, FF, VFS. Software: DD (AC-9) FF (radiometry), VFS (data acquisition and pre-processing). Writing: PM with support from all co-authors.

Competing interests. The authors declare no competing interests.

Acknowledgements. Many people contributed to the generation of the COASTℓOOC database as described herein, be it by contributing to project implementation, supporting field campaigns, performing lab work, or providing scientific guidance. In this respect, we would like 340 especially to thank James Aiken, Jean-François Berthon, Annick Bricaud, Hervé Claustre, Roland Doerffer, Giovanni Massimo Ferrari, Hans Hakvoort, Geir Johnsen, Khadija Oubelkheir, Matt Pinkerton, Rainer Reuter, Marcel Wernand and Grigor Obolensky. We are also grateful to the crews of the ships (NO Thetis-2, RV Poseidon, RV Victor Hensen) and helicopter (Commerc'Air SA) we used for collecting the field data.

References

345   2021, G. B. C. G.: The GEBCO_2021 Grid - a Continuous Terrain Model of the Global Oceans and Land., https://doi.org/10.5285/C6612CBE-50B3-0CFF-E053-6C86ABC09F8F, 2021.

Aiken, J., Antoine, D., Babin, M., Barth, H., Bricaud, A., Chauton, M., Claustre, H., Doerffer, R., Dowell, M., Fell, F., Ferrari, M., Fischer, J., Fournier-Sicre, Vincent, Hakvoort, H., Hoepffner, N., Johnsen, G., Montagner, F., Moore, G., Morel, A., Obolensky, G., Olbert, C., Pinkerton, M., Reuter, R., Sakshaug, E., and Wernand, M.: COASTLOOC (COAstal Surveillance Through Observation of Ocean Colour)

350   Final Report, Tech. rep., Zenodo, https://doi.org/10.5281/ZENODO.7428384, 2000.

Antoine, D., André, J.-M., and Morel, A.: Oceanic Primary Production: 2. Estimation at Global Scale from Satellite (Coastal Zone Color Scanner) Chlorophyll, Global Biogeochemical Cycles, 10, 57–69, https://doi.org/10.1029/95GB02832, 1996.

Babin, M.: Variations in the Light Absorption Coefficients of Phytoplankton, Nonalgal Particles, and Dissolved Organic Matter in Coastal Waters around Europe, Journal of Geophysical Research, 108, 3211, https://doi.org/10.1029/2001JC000882, 2003.

355   Babin, M., Morel, A., Fournier-Sicre, V., Fell, F., and Stramski, D.: Light Scattering Properties of Marine Particles in Coastal and Open Ocean Waters Asrelated to the Particle Mass Concentration, Limnology and Oceanography, 48, 843–859, https://doi.org/10.4319/lo.2003.48.2.0843, 2003.

Benner, R. and Amon, R. M.: The Size-Reactivity Continuum of Major Bioelements in the Ocean, Annual Review of Marine Science, 7, 185–205, https://doi.org/10.1146/annurev-marine-010213-135126, 2015.

360   Bricaud, A., Morel, A., and Prieur, L.: Absorption by Dissolved Organic Matter of the Sea (Yellow Substance) in the UV and Visible Domains, Limnology and Oceanography, 26, 43–53, https://doi.org/10.4319/lo.1981.26.1.0043, 1981.

Carr, M.-E., Friedrichs, M. A., Schmeltz, M., Noguchi Aita, M., Antoine, D., Arrigo, K. R., Asanuma, I., Aumont, O., Barber, R., Behrenfeld, M., Bidigare, R., Buitenhuis, E. T., Campbell, J., Ciotti, A., Dierssen, H., Dowell, M., Dunne, J., Esaias, W., Gentili, B., Gregg, W., Groom, S., Hoepffner, N., Ishizaka, J., Kameda, T., Le Quéré, C., Lohrenz, S., Marra, J., Mélin, F., Moore, K., Morel, A.,

365   Reddy, T. E., Ryan, J., Scardi, M., Smyth, T., Turpie, K., Tilstone, G., Waters, K., and Yamanaka, Y.: A Comparison of Global Estimates of Marine Primary Production from Ocean Color, Deep Sea Research Part II: Topical Studies in Oceanography, 53, 741–770, https://doi.org/10.1016/j.dsr2.2006.01.028, 2006.

Cetinić, I., Perry, M. J., Briggs, N. T., Kallin, E., D'Asaro, E. A., and Lee, C. M.: Particulate Organic Carbon and Inherent Optical Properties during 2008 North Atlantic Bloom Experiment: POC AND OPTICS-NAB08, Journal of Geophysical Research: Oceans, 117, n/a–n/a,

370   https://doi.org/10.1029/2011JC007771, 2012.

Cole, J. J., Prairie, Y. T., Caraco, N. F., McDowell, W. H., Tranvik, L. J., Striegl, R. G., Duarte, C. M., Kortelainen, P., Downing, J. A., Middelburg, J. J., and Melack, J.: Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget, Ecosystems, 10, 172–185, https://doi.org/10.1007/s10021-006-9013-8, 2007.

Ferrari, G. M.: The Relationship between Chromophoric Dissolved Organic Matter and Dissolved Organic Carbon in the European At-

375   lantic Coastal Area and in the West Mediterranean Sea (Gulf of Lions), Marine Chemistry, 70, 339–357, https://doi.org/10.1016/S0304-4203(00)00036-0, 2000.

Ferrari, G. M., Bo, F. G., and Babin, M.: Geo-Chemical and Optical Characterizations of Suspended Matter in European Coastal Waters, Estuarine, Coastal and Shelf Science, 57, 17–24, https://doi.org/10.1016/S0272-7714(02)00314-1, 2003.

Gordon, H. R. and Morel, A. Y.: Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery: A Review, Springer US,

380   New York, NY, 1983.

Guthrie, W. F.: NIST/SEMATECH e-Handbook of Statistical Methods (NIST Handbook 151), https://doi.org/10.18434/M32189, 2012.

Hedges, J., Keil, R., and Benner, R.: What Happens to Terrestrial Organic Matter in the Ocean?, Organic Geochemistry, 27, 195–212, https://doi.org/10.1016/S0146-6380(97)00066-1, 1997.

Jerlov, N.: Optical Oceanography, Elsevier Publishing Company, New York, 1968.

385 Lee, Z., Carder, K. L., Mobley, C. D., Steward, R. G., and Patch, J. S.: Hyperspectral Remote Sensing for Shallow Waters I A Semianalytical Model, Applied Optics, 37, 6329, https://doi.org/10.1364/AO.37.006329, 1998.

Massicotte, P.: PMassicotte/Coastlooc_data_paper: V1.0.0, Zenodo, https://doi.org/10.5281/ZENODO.7708653, 2023.

Massicotte, P., Asmala, E., Stedmon, C., and Markager, S.: Global Distribution of Dissolved Organic Matter along the Aquatic Continuum: Across Rivers, Lakes and Oceans, Science of The Total Environment, 609, 180–191, https://doi.org/10.1016/j.scitotenv.2017.07.076, 2017.

390 Massicotte, P., Babin, M., Fell, F., Fournier-Sicre, V., and Doxaran, D.: The COASTLOOC Project Dataset, https://doi.org/10.17882/93570, 2023.

Morel, A. and Bélanger, S.: Improved Detection of Turbid Waters from Ocean Color Sensors Information, Remote Sensing of Environment, 102, 237–249, https://doi.org/10.1016/j.rse.2006.01.022, 2006.

Morel, A. and Prieur, L.: Analysis of Variations in Ocean Color: Ocean Color Analysis, Limnology and Oceanography, 22, 709–722,
395 https://doi.org/10.4319/lo.1977.22.4.0709, 1977.

O'Reilly, J. E. and Werdell, P. J.: Chlorophyll Algorithms for Ocean Color Sensors - OC4, OC5 & OC6, Remote Sensing of Environment, 229, 32–47, https://doi.org/10.1016/j.rse.2019.04.021, 2019.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2023.

Sathyendranath, S.: Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters, Reports of the International Ocean-
400 Colour Coordinating Group, No. 3, IOCCG, Dartmouth, Canada, 2000.

Stramski, D., Reynolds, R. A., Babin, M., Kaczmarek, S., Lewis, M. R., Röttgers, R., Sciandra, A., Stramska, M., Twardowski, M. S., Franz, B. A., and Claustre, H.: Relationships between the Surface Concentration of Particulate Organic Carbon and Optical Properties in the Eastern South Pacific and Eastern Atlantic Oceans, Biogeosciences, 5, 171–201, https://doi.org/10.5194/bg-5-171-2008, 2008.

Sugihara, S., Kishino, M., and Okami, N.: Contribution of Raman Scattering to Upward Irradiance in the Sea, Journal of the Oceanographical
405 Society of Japan, 40, 397–404, https://doi.org/10.1007/BF02303065, 1984.

Tassan, S. and Ferrari, G. M.: An Alternative Approach to Absorption Measurements of Aquatic Particles Retained on Filters, Limnology and Oceanography, 40, 1358–1368, https://doi.org/10.4319/lo.1995.40.8.1358, 1995.

Tassan, S. and Ferrari, G. M.: Measurement of Light Absorption by Aquatic Particles Retained on Filters: Determination of the Optical Pathlength Amplification by the 'Transmittance-Reflectance' Method, Journal of Plankton Research, 20, 1699–1709,
410 https://doi.org/10.1093/plankt/20.9.1699, 1998.

Tassan, S. and Ferrari, G. M.: Variability of Light Absorption by Aquatic Particles in the Near-Infrared Spectral Region, Applied Optics, 42, 4802, https://doi.org/10.1364/AO.42.004802, 2003.

Van Der Linde, D.: Protocol for Determination of Total Suspended Matter in Oceans and Coastal Zones, Tech. Rep. I.98, 1998.

Vargas, M., Brown, C. W., and Sapiano, M. R. P.: Phenology of Marine Phytoplankton from Satellite Ocean Color Measurements, Geophys-
415 ical Research Letters, 36, L01 608, https://doi.org/10.1029/2008GL036006, 2009.

Vidussi, F., Claustre, H., Bustillos-Guzmàn, J., Cailliau, C., and Marty, J.-C.: Determination of Chlorophylls and Carotenoids of Marine Phytoplankton: Separation of Chlorophyll a from Divinylchlorophyll a and Zeaxanthin from Lutein, Journal of Plankton Research, 18, 2377–2382, https://doi.org/10.1093/plankt/18.12.2377, 1996.