Marine Chemistry xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Marine Chemistry

journal homepage: www.elsevier.com/locate/marchem

Spectral absorption by marine chromophoric dissolved organic matter: Laboratory determination and piecewise regression modeling

M. Guillermina Ruiz[a,b,*], Vivian Lutz[a,b], Robert Frouin[c]

[a] Instituto Nacional de Investigación y Desarrollo Pesquero, Paseo Victoria Ocampo n°1, B7602HSA Mar del Plata, Buenos Aires, Argentina
[b] Instituto de Investigaciones Marinas y Costeras-CONICET, Mar del Plata, Buenos Aires, Argentina
[c] Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, United States

ARTICLE INFO

Keywords:
Chromophoric dissolved organic matter
Filter types
Piecewise regression analysis
South Atlantic
South Pacific

ABSTRACT

Chromophoric dissolved organic matter (CDOM) is an important light-absorbing component of seawater. Yet spectrophotometric determinations of CDOM absorption from existing laboratory methods differ substantially. Since CDOM absorption in the visible usually remains below the detection limit of traditional spectrophotometers, its spectral shape has been modeled from the ultra-violet, by applying a single exponential model (SEM) from which a unique parameter, the spectral slope S, is derived. The usefulness of SEM and S is controversial, due to the lack of agreement on the fitting procedures and the poor ability of the SEM to fit equally well all CDOM absorption spectra. In view of this, empirical factors affecting the measurement of CDOM absorption coefficient by spectrophotometry were tested. No differences in CDOM spectra obtained by filtration through 0.2 µm membrane or 0.7 µm GFF filters were found for either high (Case II) or low (Case I) CDOM content situations. Two spectral shape groups were distinguished after applying a multivariate approach to 145 spectra from the South Atlantic, Strait of Magallanes, and South Pacific. The two groups were associated mainly with coastal and oceanic waters. A segmented regression model (SRM) with two free breakpoints better represented the CDOM absorption spectra than a SEM. The SRM fitted both CDOM spectral shape groups with accuracy. This concatenated exponential model is useful for understanding CDOM dynamics and developing improved satellite ocean-color algorithms.

1. Introduction

Chromophoric dissolved organic matter (CDOM) is one of the optical components of seawater that absorb sunlight, reducing photosynthetically available radiation for phytoplankton growth (Højerslev, 1980; Mei et al., 2010). In coastal areas, the discharge of terrestrial degradation products in riverine and estuarine waters could play a dominant role in CDOM composition (Hernes and Benner, 2003). However, in the open ocean, the main source is in situ microbial production (Lalli and Parsons, 1997; Nelson and Siegel, 2013). In the upper layer it is mostly produced from decaying materials of marine origin (e.g., dead plankton) (Nelson et al., 1998), while in the deep ocean, microbial DOM processing plays a substantial role. (Catalá et al., 2015; Jørgensen et al., 2014; Kirk, 2011; Romera-Castillo et al., 2010).

CDOM absorption spectra have been described as featureless, typically decreasing with increasing wavelength from the ultraviolet (UV) to the visible (VIS) in an almost exponential fashion. The ubiquitous CDOM can dominate light absorption in some coastal areas of the ocean, especially at blue wavelengths in coincidence with the chlorophyll absorption peak close to 440 nm. This has been a source of impairment for the accuracy of band–ratio algorithms to retrieve chlorophyll absorption remotely (IOCCG, 2000; Mannino et al., 2014, Siegel et al., 2005). A wide disparity is observed in the laboratory methods used to determine the CDOM spectral absorption coefficient, a_{CDOM} (λ), [m^{-1}] (Andrew et al., 2013; Clark et al., 2008; D'Sa et al., 2006; Kowalczuk et al., 2006; Lorenzoni et al., 2011; Lutz et al., 2006; Toming et al., 2009). This is not a trivial issue since CDOM absorption is operationally defined by the filtrate that passes through a small pore size filter and absorbs light in the visible and ultraviolet (Blough and Del Vecchio, 2002).

The most common method employed to measure a_{CDOM} (λ) involves reading absorption spectra in a dual beam spectrophotometer with a 10 cm path length quartz cell against a pure water reference. Disparities in the literature regarding the pre-filtration, type of filtration system, filter pore size, type of ultrapure water, the usage of a "real blank" (pure water treated as a sample), or time span between sample acquisition, filtration and scanning can be easily observed. For example, several systems that produce ultrapure water are available in the

* Corresponding author at: Instituto Nacional de Investigación y Desarrollo Pesquero, Paseo Victoria Ocampo nº1, B7602HSA, Mar del Plata, Buenos Aires, Argentina.
E-mail addresses: mgruiz@inidep.edu.ar (M.G. Ruiz), vlutz@inidep.edu.ar (V. Lutz), rfrouin@ucsd.edu (R. Frouin).

http://dx.doi.org/10.1016/j.marchem.2017.03.012
Received 25 August 2016; Received in revised form 30 March 2017; Accepted 31 March 2017

market, although not all are of the same quality. This introduces variability on a global scale since each laboratory would be subtracting a different spectrum of "ultrapure" water. Moreover, some authors use 0.2 μm membrane filters while others use 0.7 μm glass fiber filters in closure experiments of optical properties, acknowledging that the contribution to absorption of the CDOM fraction comprised between 0.2 and 0.7 μm might be lost (Pope et al., 2000).

Most dual-beam spectrophotometers set with a 10 cm cell have a detection limit at optical density of 0.005 (equivalent to 0.115 m^{-1} absorption coefficient); hence, values below this threshold, which can be usually found especially in oligotrophic waters with low CDOM, cannot be determined with accuracy. According to Lambert-Beer law, absorption increases with increasing optical path length for a given solution. Thus more sensitive instruments with longer optical path lengths and higher accuracy have been developed and are already in use, such as the liquid waveguide systems (D'Sa et al., 1999; Miller et al., 2002) or the PSICAM (Röttgers and Doerffer, 2007). However, these instruments are still not routinely used and require CDOM spectrophotometer measurements for validation and data comparison. Moreover, many of the existing CDOM datasets have been acquired using bench-spectrophotometers, which will probably be in use for some time. Hence an up-to-date of the best practice method using this configuration to obtain $a_{CDOM}(\lambda)$ is required.

CDOM absorption spectra have been commonly characterized with a single exponential model (SEM) of the form:

$$a_{CDOM}(\lambda) = a_{CDOM}(\lambda_0) * \exp[-S(\lambda - \lambda_0)] \quad (1)$$

where $a_{CDOM}(\lambda)$ is expressed in (m^{-1}) and from which a single parameter, the spectral slope S (nm^{-1}), is derived (Bricaud et al., 1981; Carder et al., 1989; Højerslev, 1980; Jerlov, 1957). Various authors have demonstrated the utility of S to study different aspects of CDOM dynamics. For instance, Carder et al. (1989) showed that the $S_{350–500}$ qualitatively describes the ratio of fulvic to humic acids in a sample. Higher S values have been mostly reported for oceanic environments with lower $a_{CDOM}(\lambda)$, while lower S values are common in coastal, estuarine zones with higher $a_{CDOM}(\lambda)$ (Blough and Del Vecchio, 2002). Furthermore, the S slope is also an input in bio-optical algorithms developed to retrieve CDOM absorption or closely related products from ocean color satellite data (Magnuson et al., 2004; Maritorena et al., 2002). Slope values reported in the literature vary widely (e.g. from 11 to 25 × 10^{-3} nm^{-1}, Blough and Del Vecchio, 2002, Table 1). This reported S variability is impacted by the modeling procedures (Helms et al., 2008; Højerslev and Aas, 2001) and depends on the wavelength range and type of fitting method (linear or non-linear), making comparisons of different works almost impossible (Blough and Del Vecchio, 2002; Nelson and Coble, 2009). This range is as high as the variability found in the S values after adjusting a SEM in different wavelength ranges to a given CDOM spectrum (Table S1). It has been recently demonstrated that the SEM is not a good model to describe the spectral shape and predict $a_{CDOM}(\lambda)$ (Twardowski et al., 2004). In our case, by fitting the SEM through Eq. 1, it overlooks subtle structural and compositional features in CDOM spectral shape. Alternative methods have been proposed to characterize CDOM spectra adjusting S in several short wavelength ranges yielding promising results (Fichot and Benner, 2012; Helms et al., 2008; Loiselle et al., 2009; Mannino et al., 2014; Sarpal et al., 1995).

Considering the aforementioned disparities two main objectives are pursued in this work. The first objective consists in testing and finding the best practice for some aspects of laboratory analysis concerning $a_{CDOM}(\lambda)$ determination by traditional bench spectrophotometry, i.e., the purity of the reference water, the need for a "real blank", and the type of filter used (0.2 μm membrane or 0.7 μm glass fiber). The second objective is to verify the hypothesis of the existence of different CDOM spectral shapes associated to different environmental conditions, and to propose a model to fit CDOM spectra with better accuracy than the SEM.

Table 1
Linear regression parameters, and their standard errors, obtained after plotting the decimal logarithm of n = 10 absorption spectra between 255 and 300 nm for the different experimental groups (see below). The residual standard error (RSE) is considered as a measure of the non-explained variability of each group.

	Intercept	Slope	Standard error of the intercept	Standard error of the slope	Residual standard error
Pure water type experiment					
AOF	− 0.7882	− 0.0278	0.0053	2e-04	0.0616
AOFNUC	− 0.6235	− 0.0284	0.0115	4e-04	0.1163
AOFGFF	− 0.6652	− 0.0274	0.0079	3e-04	0.0761
Type of filter experiment					
Case II – NUC	− 1.8561	− 0.0198	0.0037	1e-04	0.0372
Case II – GFF	− 1.8404	− 0.0197	0.0035	1e-04	0.0354
Case I – NUC	− 3.2679	− 0.0220	0.0066	2e-04	0.0670
Case I – GFF	− 3.2612	− 0.0224	0.0096	3e-04	0.0977

AOF: 2-day aged osmosis water; AOFNUC: AOF water filtered through membrane filter; AOFGFF: AOF water filtered through glass fiber filters. Case II – NUC: CDOM sample from a "Case II" water sample filtered through membrane filters. Case II – GFF: CDOM sample from a "Case II" water sample filtered through glass fiber filters. Case I – NUC: CDOM sample from a "Case I" water sample filtered through membrane filters. Case II – GFF: CDOM sample from a "Case II" water sample filtered through glass fiber filters.

2. Materials and methods

2.1. Laboratory analysis of CDOM samples

First, three types of ultra-pure water were compared. The first type, OFw, was obtained from a purification system (Gota a Gota, S.R.L.) producing ultra-pure water free of organic matter from a source of reverse osmosis water. The second type, AOF, consisted of OFw aged for two days in acid-washed glass bottles. The third type was a database of Milli-Q water absorption spectra obtained onboard R/V Melville during the MV1102 cruise (see Section 2.2). At our facility it was observed that historical "real blanks" had lower absorption values in the UV than the freshly acquired OFw. Real blanks differ from fresh OFw in that they are treated as a sample (i.e., taken to the cruise and then filtered), with at least two days between their collection and the acquisition of their spectra. Therefore, the effect of aging OFw water on its absorption spectra was investigated. Note that only AOF was used as a source of pure water in the subsequent filter type experiments.

Two types of samples representing "Case I" and "Case II" waters were used for this laboratory experiment following the classical optical characterization of waters by Gordon and Morel (1983) and Morel and Prieur (1977). "Case I" waters refer to those waters in which phytoplankton and other materials of autochthonous generation are the principal agents responsible for variations in optical properties of the water. The "Case I" sample was collected onboard BIP E. Holmberg during the EH-02/13 cruise (May 2013) nearby Islas Georgias del Sur (54°11′S, 37°30′W). It is worth mentioning that since this sample was stored in a cool room in the dark for more than 6 months, it was used here as an example of oligotrophic waters, but not as a true measurement for the area. Instead, "Case II waters" refer to those which are influenced by phytoplankton and also by other substances that vary independently of phytoplankton, notably inorganic particles in suspension and other alloctonous materials. The "Case II" sample was acquired with a bucket at the "Muelle de Pescadores" pier in Mar del Plata (38°00′S, 57°33′W, between April and July 2014) and transferred into acid-cleaned glass bottles with Teflon lids and kept at 4 °C until analysis (within 3 h from collection).

To test the filter types, GFF fiber filters (Whatman, hereafter referred to as "GFF") with a nominal pore size of 0.7 μm were

combusted at 450 °C during 3.5 h to oxidize any organic matter and then washed with 200 ml of AOF water to avoid glass fiber particles passing to the filtrate. Membrane filters with a 0.2 μm pore size (Poretics, hereafter referred to as "NUC") were rinsed in 10% HCl during 15 min and then soaked in AOF water until use following Mitchell et al. (2003). One filter per sample was used in all cases. All glassware was rinsed in 10% HCl for at least 2 h, washed 6 times with reverse osmosis water and another 3 times with AOF water. All openings of the glassware were carefully covered with aluminum foil until use to prevent contamination. This procedure was always performed the day before the experiments.

CDOM and pure water spectra were acquired employing a Shimadzu 2401 UV–VIS spectrophotometer using a 10 cm quartz cell. Before reading a batch of samples, the cell was thoroughly washed twice with AOF water, 1 time with high grade pure ethanol and another 6 times with AOF water. All readings were done setting 1 nm sampling interval, medium scan speed, and 2 nm slit wide over the range 250–750 nm, after letting the instrument lamps warm up for 30 min. All the samples were allowed to reach room temperature and scanned against air after setting an air-air baseline (0.000 ± 0.001 absorbance units). CDOM absorption coefficient was calculated using the equation:

$$a_{CDOM}(\lambda) = 2.303 * [A_{CDOM}(\lambda) - A_{REAL\ BLANK}(\lambda)]/L \quad (2)$$

where L is the optic path length in meters, $A_{REAL\ BLANK}$ is the average of n replicates of the absorption spectra of filtered AOF water (calculated as follows: for each replicate, its own average absorption value between 460 and 470 nm was subtracted, and then the average of these replicates was taken as the real blank spectrum), and A_{CDOM} is the average CDOM absorption spectrum of n replicates, after subtracting to each replicate the absorption between 460 and 470 nm (same range as for pure water spectra) and then the average absorption between 590 and 600 nm to correct the apparent CDOM offset (Mitchell et al., 2003). A smoothing moving average (7 nm) was applied to each $a_{CDOM}(\lambda)$ spectrum in order to minimize noise.

Data analysis was performed using UVProbe 2.21 and free software *R* (R Core Team, 2015), with codes developed ad hoc. The regression slope after plotting the logarithm 10 of the n a_{CDOM} spectra between 255 and 300 nm was used to compare spectra of different treatment groups in the type of pure water and type of filter experiments (with AOF, AOFNUC and AOFGFF and Case II-NUC, Case II-GFF, Case I-NUC and Case II-GFF as treatment groups respectively). The residual standard error (RSE) was used as a measure of the non-explained variability of each group.

2.2. Modeling of CDOM field samples

For the second objective, a CDOM data set collected during the R/V Melville MV1102 cruise was used (see cruise track in Fig. 1). The ship left Cape Town on the 20th of February 2011 and arrived in Valparaiso on the 14th of March 2011, sailing across the South Atlantic Ocean, the Strait of Magallanes, and along the Pacific Chilean coast during the austral late summer. An extensive description of the oceanographic regimes can be found elsewhere (Rudorff, 2013). Briefly, the MV1102 cruise covered six biogeochemical provinces (Longhurst, 2007) that differed in terms of water masses and bio-optical characteristics (see Fig. 4.1 in Rudorff, 2013). A few stations were made in a highly productive region in the south of the Benguela Current Coastal Province followed by the South Atlantic Gyral Province, finding more stable, clear and oligotrophic waters. R/V Melville continued sailing in the Southwest direction through the South Subtropical Convergence Province, a relatively more dynamic region with intrusion of Sub-Antarctic surface waters, which promoted a strong chlorophyll gradient. The route proceeded westward along the Subantartic Water Ring Province, which is a high latitude zone, marked by a strong seasonality of mixing-stratification processes driven by wind stress and solar irradiance. Upon reaching 50°S the ship entered the Southwest Atlantic Shelves Province, sailing again in waters of high productivity due to the presence of the Malvinas Current. She then passed through the Strait of Magallanes, which connects the South Atlantic with the South Pacific oceans. The hydrography of the straits is highly complex (Lutz et al., 2016; Panella et al., 1991); the water circulation in this channel is mainly regulated by strong tides, with a mixture of Sub-Antartic and Pacific Waters, continental runoff and glacial-fluvial waters, favoring the re-suspension of the fine sediments especially in shallow sectors. Finally, R/V Melville sailed across the Humboldt Current Coastal Province from the southern fjords up to 36°S in central Chile, again finding high productive waters with high biomass levels driven by the Humboldt Current.

CDOM samples were filtered through acid-washed 0.2 μm membrane filters (Nuclepore) and scanned aboard just after sampling in a Cary-50 UV–VIS-NIR spectrophotometer between 250 and 750 nm and corrected for the absorption by the "real blank" (Rudorff, 2013). As mentioned above, a correction for the apparent CDOM offset and a smoothing moving average (7 nm) was applied to each a_{CDOM} spectrum. A total of 145 spectra were included in this study, after 3 outliers were excluded from the original set of samples (the first one due to a bump suspected to be produced by mycosporine like amino-acids, MAAs (Carignan and Carreto, 2013; Subramaniam et al., 1999) or by nitrate (Catalá et al., 2016); the second one due to having negative values that prevented the calculation of its geometric mean; and the third one due to suspected contamination).

The existence of different spectral shapes was addressed with cluster analysis (CA) on the normalized, natural log-scaled CDOM spectra using the absorption in the 275–400 nm range as the only variable to discriminate among groups. Normalization was performed to eliminate the absorption magnitude with the aim of exploring the underlying shapes. Assuming that CDOM spectra can be modeled like a concatenation of exponentials, it can be demonstrated that the distance between two spectra can be calculated as the square Euclidean distance between the spectra normalized by their geometric mean and transformed into logarithmic scale (Suppl. info. S1). First, a hierarchical cluster (square Euclidean distances, Wards method) was performed, and second, after determining the number of groups from this analysis (k = 2 groups), a non-hierarchical cluster (K-means) was performed to separate spectra in groups, minimizing the variance within groups and maximizing it between groups.

The segmented regression model (SRM) described by Muggeo (2003) was applied using the R *segmented* package (Muggeo, 2008) to model CDOM spectra. This is a regression model where the relationships between the response and the explanatory variables are piecewise linear, represented by at least two straight lines connected at unknown (unspecified) breakpoints (hereafter, BKs). First, a SEM obtained by linear fit was applied to each natural log-linearized CDOM spectrum between 275 and 400 nm; then the model was updated by the *segmented ()* function adding the segmented relationships. The upper wavelength limit (400 nm) was chosen to be a good and conservative wavelength at which the optical density of our dataset was above the detection limit of the instrument.

Linear discriminant analyses (LDA) were performed to distinguish the observed spectral shape groups. The BKs and slopes derived from the SRM models with 1 and 2 breakpoints were used as input for the first LDA, while five environmental variables were considered as input for the second. The selected environmental variables were: sea surface temperature (SST), sea surface salinity (SSS), bottom depth (BD), distance to the nearest coast (DNC), and fluorescence intensity (Fl). Each of these variables was standardized so that it had zero mean and within-groups variance of 1 to compare the discriminating power of predicting variables. DNC, Fl and BD were transformed with the log(x) function to reach multivariate normal distribution. The environmental variables were selected because they are correlated to CDOM. In the first place a major source of CDOM is land, therefore a relationship to DNC and BD is expected; while Fl would indicate a source of CDOM by phytoplankton (mainly in open ocean). In turn, these sources might

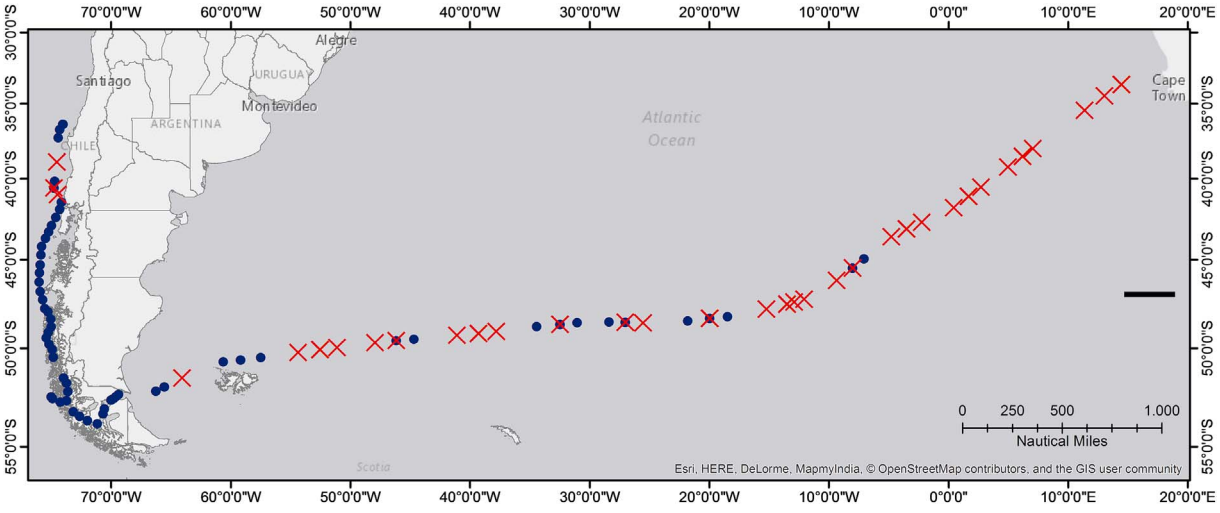

Fig. 1. Map of the study region showing the MV1102 ship track and the location of the CDOM samples classified according to their spectral shape into Group 1 (blue circles) and Group 2 (red cross). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

determine the chemical nature of CDOM and hence the shape of its absorption spectrum. SST and SSS are relevant to characterize water masses. Besides, all these variables could be obtained through remote sensors (although in the present study we used in situ data).

Several environmental variables were recorded during all sampling stations. Methods for the retrieval of sea surface salinity, sea surface temperature, fluorescence, and bottom depth are described elsewhere (Rudorff, 2013; Rudorff et al., 2014). The distance of each station location to the nearest coast was estimated using NASA "Distance to Nearest Coastline: 0.01-Degree Grid" product (NASA, 2009) by assigning to each station the distance value of the closest point in NASA's grid. Minimum distance to NASA's grid points took into account the ellipsoidal shape of the Earth (Bivand et al., 2013; Pebesma and Bivand, 2005). Although for some sampling stations the distance assigned was slightly larger than the true distance to the coast, it was a good approximation due to the random distribution of the distances between the stations and the points in the grid and the consistency of the results with the vessel trajectory.

All statistical analyses where performed in R version 3.1.2 (R Core Team, 2015). CA was performed using the *hclust()* and *kmeans()* functions from the*base* R package and the LDA using the *lda()* function of the *MASS* package.

3. Results

3.1. Impact of reference pure water and filter type

The effect of aging OFw water on its absorption spectra was investigated. This procedure decreased water absorption in the UV (i.e., made it look closer to the commonly used Milli-Q water spectrum) (Fig. 2), probably due to oxidation and volatilization of traces of organic compounds remaining in the purified water. For all the subsequent experiments, OFw aged for two days (AOF) was used as a source of ultra-pure water.

In order to determine the best reference for CDOM spectra acquisition, we addressed the spectral differences between unfiltered and filtered AOF water, either through glass fiber filters (GFF) or 0.2 µm membrane filters (NUC). First, we observed that rinsing combusted GFF filters with at least 200 ml of the purest water available minimizes the introduction of glass fibers into the filtrate (Fig. S1). Although subtracting a blank of the same kind (i.e., AOF water filtered through not rinsed combusted GFF) cancelled this effect, we preferred to rinse with 200 ml AOF water all combusted GFF filters for the subsequent experiments. The mean absorption spectra of unfiltered AOF water had

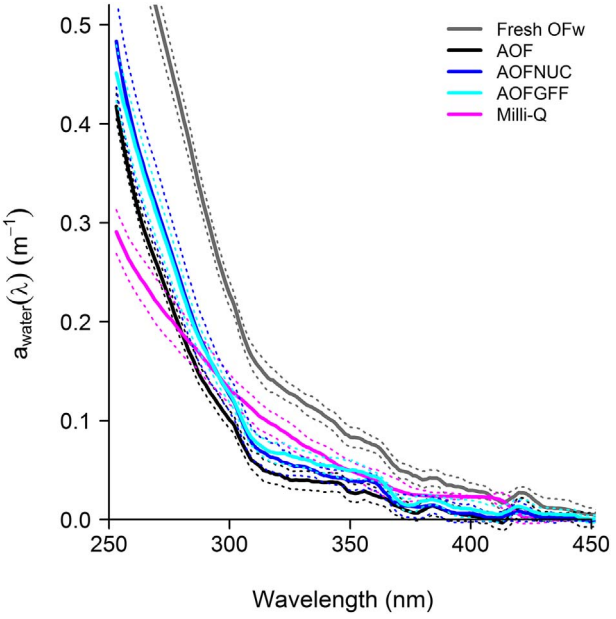

Fig. 2. Absorption coefficient spectra of ultrapure water obtained from different ultrapure systems. Fresh osmosis water (OFw, grey line), 2-days aged osmosis water (AOF, black line), Milli-Q water (magenta line), AOF water filtered through membrane filter (AOFNUC, blue line) and through glass fiber filters (AOFGFF, light blue). Solid lines represent the mean spectrum of 10 replicates (except for MilliQ water, n = 145) and dotted lines ± 1 SD. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

slightly lower absorption in the UV region than the mean absorption spectra of filtered AOF water, either through NUC (AOFNUC) or GFF (AOFGFF) filters. Visual inspection suggested that the averaged spectrum of AOFNUC was identical to that of AOFGFF (Fig. 2). To compare spectra of different treatment groups, we estimated the slope after plotting the logarithm 10 of the n a_{WATER} spectra between 255 and 300 nm (Table 1). The standard error of the slopes of water filtered through thoroughly rinsed GFF filters was similar to that of water filtered through cleaned membrane filters, and to that of unfiltered pure water.

We further addressed the effect of the filter type (GFF or NUC) on the CDOM spectra acquisition of a low CDOM content "Case I" and a high CDOM content "Case II" water sample to assess possible adsorption of organic matter on GFF filters. The intercepts and slopes fitted by linear regression were identical for both groups; no differences were

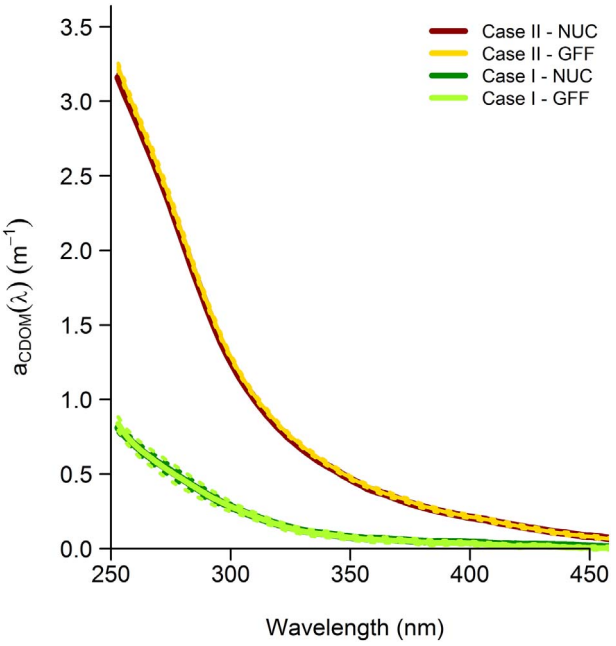

Fig. 3. CDOM spectral absorption coefficients of Case II and Case I water samples obtained after filtration through membrane filters (NUC) and glass fiber filters (GFF). Solid lines represent the mean spectrum of 10 replicates and dotted lines represent ± 1 SD. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

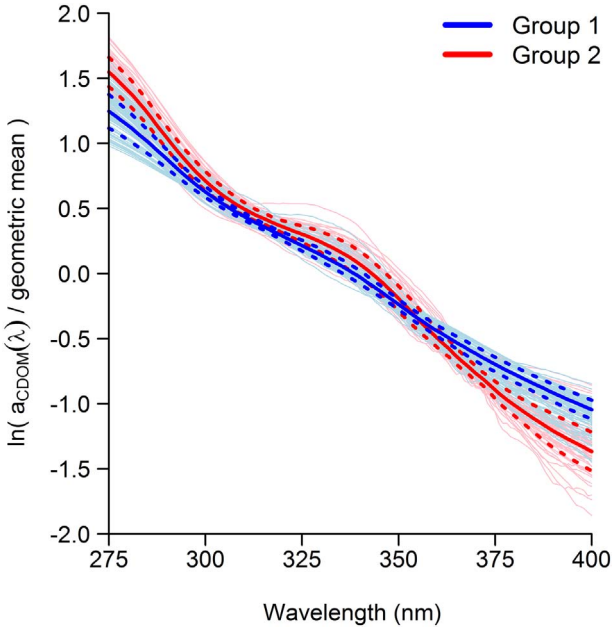

Fig. 4. Centroids of the two CDOM spectral shape groups obtained by a non-hierarchical cluster analysis (K-means). Spectra were normalized by their geometric mean and linearized into logarithmic scale within the 275–400 nm range before running the cluster. Solid lines represent the centroids (mean spectrum) and dotted lines ± 1 SD; thin lines in the back correspond to each individual CDOM sample belonging to spectral shape Group 1 (blue, n = 84) or Group 2 (red, n = 61). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

observed in the average spectrum obtained using NUC or GFF filters, for the two water types (Fig. 3). Difference in non-explained variability indicated by the RSE of the linear regression between Case II – NUC and Case II – GFF groups was negligible and small between Case I-NUC and Case I-GFF (Table 1).

3.2. Identification of CDOM spectral shape groups

In the first place we fitted the SEM by lineal fit in different wavelength ranges and compared their goodness of fit. We observed that the equation that best fits a spectrum with high absorption values did not fit equally well a spectrum with low absorption values, as it was shown by the sum of squared errors of the fitted model (Table S1). These observations were in agreement with other published results (Twardowski et al., 2004). Thus, we assumed that not all CDOM absorption spectra behave as a single decreasing exponential, but that some of them are better described by a concatenation of exponentials (Fig. S2). We hypothesized that it is possible to classify CDOM spectra only according to their shape and regardless of the magnitude of the absorption coefficients. To test this hypothesis, a cluster analysis (CA) was performed on a set of $n = 145$ normalized, log-scaled CDOM spectra. The result of a hierarchical cluster showed a dendrogram tree with 2 notable branches (Fig. S3). Hence, a non-hierarchical cluster (*K-means*) was performed to separate spectra in $k = 2$ groups. This CA rendered one group with 84 and another with 61 spectra (almost identical to the hierarchical cluster). The *centroids* (i.e., average spectrum) of each cluster group are displayed in Fig. 4. The spatial distribution of the spectral shape groups predicted by the CA is displayed in Fig. 1.

3.3. Modeling CDOM spectra by a concatenation of exponentials

The question addressed was whether it is possible to model all CDOM spectra equally well with a unique model, regardless of their shape. We applied the piecewise or segmented regression model (SRM) described by Muggeo (2003). We compared the results of fitting to the MV1102 CDOM dataset the models M0 (SRM with no BKs and 1 slope, equivalent to the SEM), M1 (SRM with 1 BK and 2 slopes) and M2 (SRM with 2 BKs and 3 slopes). The seed values for the BKs were 345 nm for M1 and 295 and 345 nm for M2. The equations describing M0, M1, and M2 are as follows:

M0:

$$\ln\ (a_{CDOM}(\lambda)) = -S(\lambda - \lambda_0) + \ln(a_{CDOM}(\lambda_0)) \quad (3)$$

M1:

$$\begin{aligned} \ln(a_{CDOM}(\lambda)) &= -S_1(\lambda - \lambda_0) + \ln(a_{CDOM}(\lambda_0)) \quad \text{if} \quad \lambda < BK1 \\ \ln(a_{CDOM}(\lambda)) &= -S_2(\lambda - \lambda_0) + \ln(a_{CDOM}(\lambda_0)) \quad \text{if} \quad \lambda \geq BK1 \end{aligned} \quad (4)$$

M2:

$$\begin{aligned} \ln(a_{CDOM}(\lambda)) &= -S_1(\lambda - \lambda_0) + \ln(a_{CDOM}(\lambda_0)) \quad \text{if } \lambda < BK1 \\ \ln(a_{CDOM}(\lambda)) &= -S_2(\lambda - \lambda_0) + \ln(a_{CDOM}(\lambda_0)) \quad \text{if } BK1 \geq \lambda < BK2 \\ \ln(a_{CDOM}(\lambda)) &= -S_3(\lambda - \lambda_0) + \ln(a_{CDOM}(\lambda_0)) \quad \text{if } \lambda \geq BK2 \end{aligned} \quad (5)$$

In all cases, λ_0 is a lambda of reference. Then, for either of the models M0–M2, the original absorption spectrum in m^{-1} can be obtained with the exponential function of the modeled data. Fig. 5 displays the residuals against the wavelength for M0, M1 and M2 models for the two groups of spectral shapes derived from the CA. The goodness of fit improves with the number of BKs, as indicated by the smaller values of the residuals, AIC and adjusted R^2 (Table S2). Model 2 fit spectra of both shape groups with equal goodness of fit, whereas M1 performed better for Group 1 than for Group 2. Model 0 always produced larger residuals than the two segmented models, whatever the spectral shape. Three examples of CDOM spectra fitted by the SRM models M0, M1, and M2 are displayed in Fig. 6.

The distribution of the retrieved parameters from M1 (Fig. 7, A) and from M2 (Fig. 7, B) provides insights into the underlying CDOM spectral shapes. Model 1 does not reveal any clear position for the BK, according to the broad distribution of the boxes for either group; yet it yields significant differences in both slope means between groups (Table 2). In contrast, M2 reveals a subtle feature in CDOM absorption coefficient spectra: both groups present a BK close to 302 nm, as denoted by the narrow distribution of the BK1 boxes. In addition, BK2 is close to

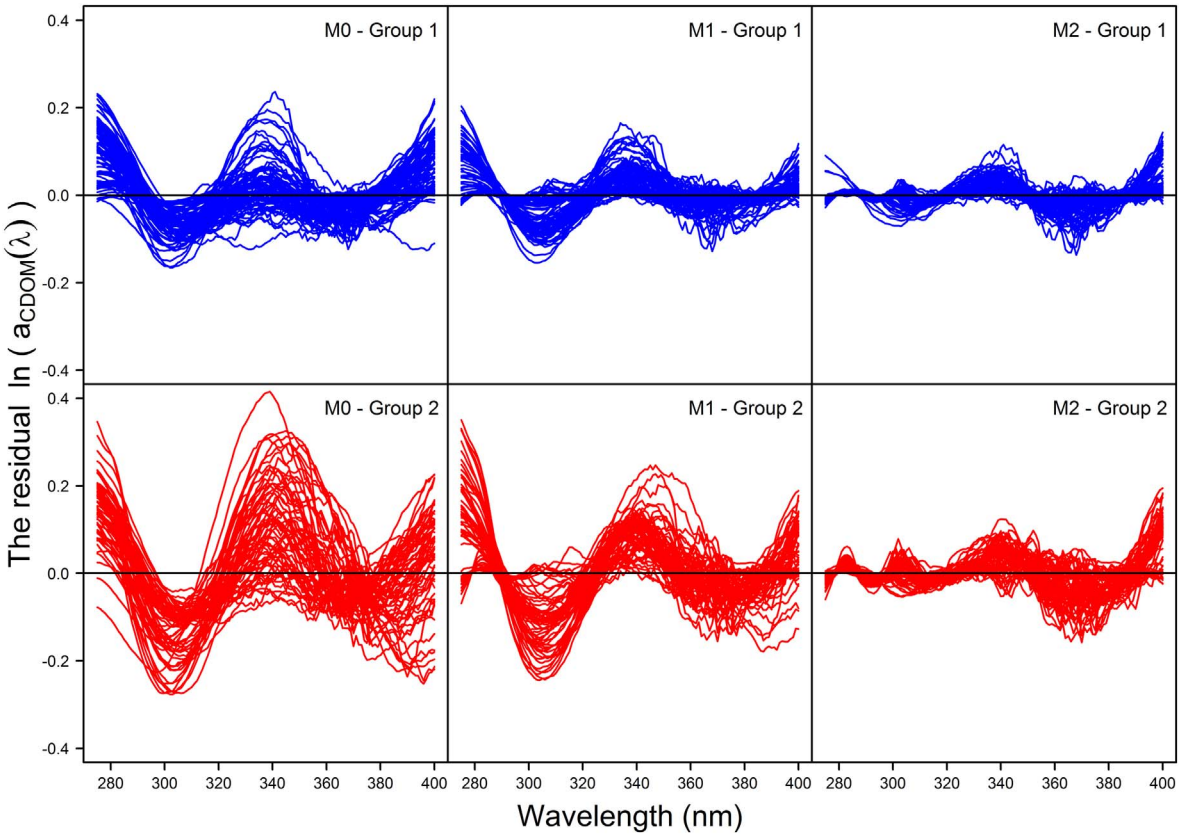

Fig. 5. Spectral distribution of residuals from the fitted segmented regression models M0, M1 and M2. Upper panel: spectral shape Group 1 (n = 84); lower panel: spectral shape Group 2 (n = 61).Note that models were fitted in the logarithmic scale, hence residuals are dimensionless.

343 nm in Group 2, while more variable in Group 1. The mean and standard deviation values of the breakpoints and slopes are summarized in Table 2 for both groups and models.

In order to evaluate the usefulness of the segmented model to distinguish the observed spectral shape groups, we performed a linear discriminant analysis (LDA) using as predictor variables the slopes and BKs retrieved from M1 and M2. Significant differences in the group means were observed for all the predictor variables except for the first breakpoint BK1 in M1 and M2 after analysis of variance test (i.e. ANOVA) (Table 3). Thus, BK1 was excluded from the LDA based on M2 parameters as predictor variables. The standardized linear discriminant coefficients indicated that for M1, the variable with the highest discriminating power was slope 2, followed by the BK1, while in M2 slope 3 and slope 2 were the best predictors (Table 3). The distribution of the discriminant function scores indicated that there was almost null degree of overlapping (Fig. 8) and the cross-classification showed that overall the 90.34% and 93.11% of the spectra were correctly classified by each of the discriminant analysis respectively (Table 4). For classification purposes, a discriminant function based on the unstandardized linear coefficient was derived:

$$D_{M2} = (-0.0016 \times BK2) + (131.28 \times Slope\ 1) + (126.35 \times Slope\ 2) + (202.00 \times Slope\ 3)$$
$$\text{Cut-off} = 0.2229 \quad (6)$$

3.4. Association of CDOM spectral shapes with environmental conditions

Finally, a LDA on a set of selected environmental variables was performed to assess whether an environmental configuration was associated to each of the observed CDOM spectral shape groups. Significant differences in the group means were observed for all the predictor variables except for SST by ANOVA test (Table 5), hence it was excluded from the LDA. The standardized linear discriminant coefficients indicated that the most powerful discriminating variable was the logDNC, followed by logFl, with logBD and SSS being the poorest predictors (Table 5). The distribution of the discriminant function scores shown in Fig. 9 revealed a small degree of overlapping. The cross-validated classification showed that overall the 82.80% were correctly classified (21 samples of Group 1 were classified as Group 2 and 4 of Group 2 as Group 1, Table 6). For classification purposes, a discriminant function based on unstandardized linear coefficients and the environmental variables was derived:

$$D_{environment} = (-0.221 \times SSS) + (-0.396 \times logFl) + (0.261 \times logBD) + (0.934 \times logDNC)$$
$$\text{Cut-off} = 0.1429 \quad (7)$$

4. Discussion

4.1. Reference pure water

One of the most common procedures to get a CDOM absorption spectrum involves filtering a seawater sample, transfer it to the quartz cuvette, and scan it versus a filtered pure water reference (i.e., a real blank) (Mitchell et al., 2003). Using a real blank as a reference provides the opportunity to cancel the absorption of any leachate from the filters. However, the source of pure water differs from one laboratory to another, which means that a different pure water spectrum is being subtracted at each facility. Furthermore, the absorption spectra of any of the commercially available pure water differ from the spectra reported by Pope and Fry (1997) for 380–700 nm and by Sogandares and Fry (1997) for 350–380 nm. This combined spectrum has been

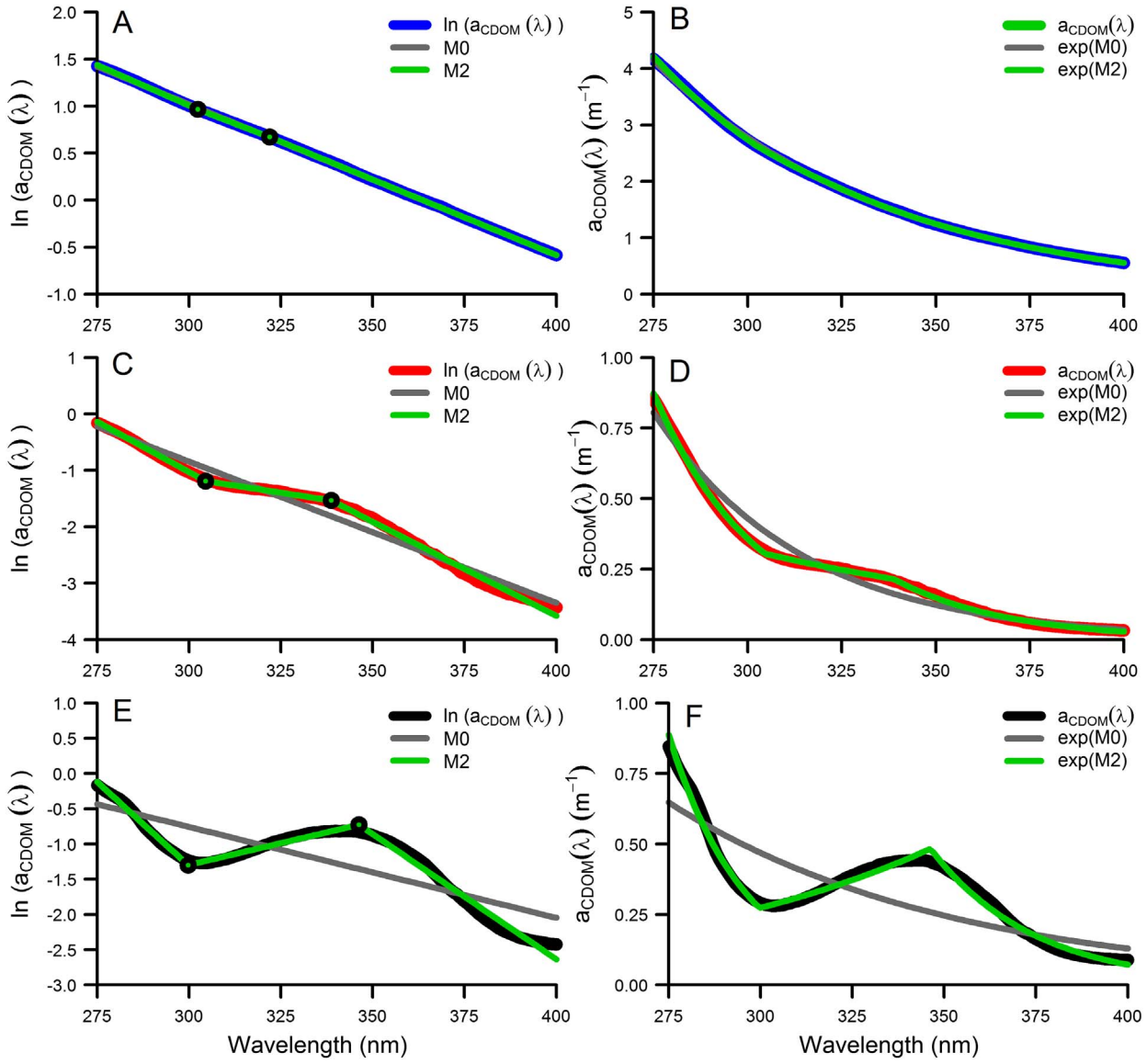

Fig. 6. CDOM spectra in logarithmic and absorption coefficient scales (left and right, respectively) fitted by the SRM M0 and M2. Examples representing spectral shape Group 1 (A and B), Group 2 (C and D), and an unusual CDOM spectrum with a bump excluded from the CA (E and F) are displayed. Dots indicate the position of the BKs found by the SRM M2. Both models were fitted in the 275–400 nm range. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

generally considered to be the "ideal" water spectrum for optical studies and ocean color processing. Since these studies were performed, absorption spectra of "pure" water were derived from optical measurements of the "clearest" seawater (Lee et al., 2015; Morel et al., 2007), suggesting that pure water absorption coefficients could be even lower than the aforementioned "ideal" spectrum. Recently, pure water absorption coefficients have been determined in the 250–550 nm UV-region (Mason et al., 2016), revealing even lower values in the 350–400 nm range. Nevertheless, for the sake of argument in the absence of a community-recommended UV absorption by pure water, we propose that it would be possible to theoretically refer all CDOM spectra to this "ideal" water spectrum, eliminating a source of variation between laboratories. This could be done as follows, i.e., by adding the difference between the spectrum of the purest available water (PW_a) – used as a reference in a given facility (e.g. Milli-Q, osmosis, etc.) and the "ideal" spectrum (PW_i):

$$Sa = PW_i + DM + Fi \quad (9)$$

$$PW_a = PW_i + Im \quad (10)$$

$$RB = PW_a + Fi = PW_i + Im + Fi \quad (11)$$

We assume that in seawater (SW) salts have no absorption and DM and PM are dissolved and particulate matter respectively (Eq. 8). After filtering SW (Eq. 9), PM is retained in the filter but the CDOM sample (Sa) may now contain any leachate from the filter (Fi). PW_a is the purest water available and "Im" is any impurity from the purification system (Eq. 10). The real blank (RB) may also carry the Fi component (Eq. 11). If in the spectrophotometer the CDOM sample is scanned against the RB (Eq. 10), then:

Hence, what we usually consider to be the "dissolved matter absorption spectrum" is actually the difference between the dissolved matter and the impurities from the purification system. Eq. 10 can be written as follows (Eq. 10′):

$$Im = PW_a - PW_i \quad (10')$$

The impurities (Im) absorption spectrum is the difference between PW_a and PW_i spectrum. Then to really obtain the DM spectrum it is necessary to add the Im spectrum to the spectrum obtained reading the Sa against a RB (Eq. 12′). This could be achieved by acquiring the spectrum of the PW_a against air and subtracting the PW_i spectrum (which is deducted replacing Eq. 10′ into Eq. 12′):

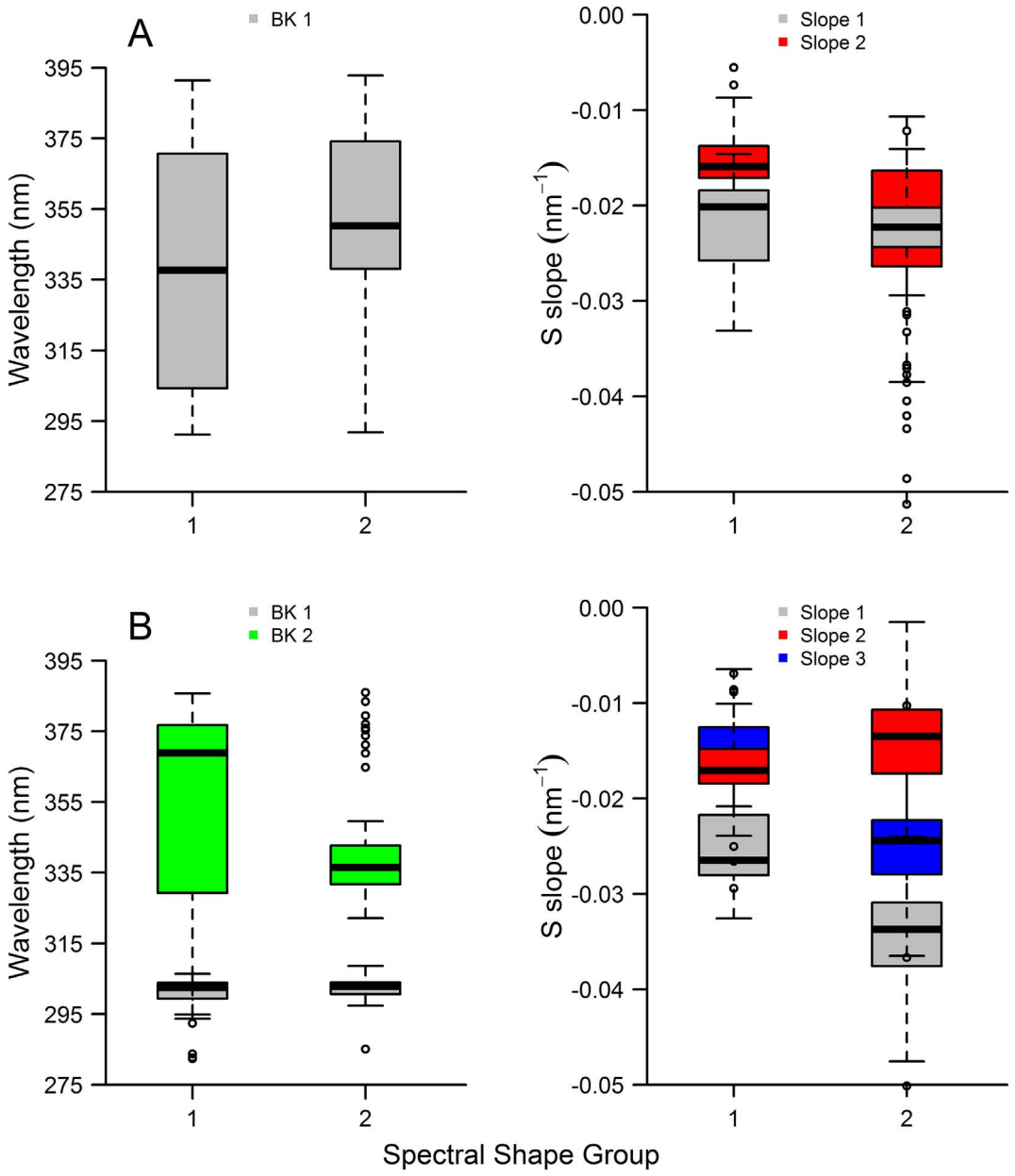

Fig. 7. Boxplot of the parameters retrieved after applying the segmented regression models (SRM) to a set of 145 CDOM spectra obtained from natural samples. A: SRM M1 (1 break point and 2 slopes); B: SRM M2 (2 breakpoints and 3 slopes). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$Sa - RB + Im = DM \quad (12')$$

$$Sa - RB + PW_a - PW_i = DM \quad (10') + (12')$$

In this way, variability from the purification systems would cancel out and all CDOM spectra would be referred to the same "ideal" water spectrum. This procedure requires obtaining the absorption spectra of the PW_a, which is not exactly the same as the RB spectrum. While waiting for a consensus "ideal" UV water spectrum, it is possible to use

Table 2
Comparison of the mean and standard deviation of the slopes and breakpoints retrieved after applying two different segmented regression models (M1 and M2) to the two spectral shape groups (Group 1 and Group 2) obtained from the cluster analysis. Model M1 has 1 BK and 2 slopes; model M2 has 2 BKs and 3 slopes.

	Model 1 (M1)				Model 2 (M2)			
	Group 1		Group 2		Group 1		Group 2	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
BK1 (nm)	338	33.9	346	31.5	302	10.5	302	3.3
BK2 (nm)	–	–	–	–	355	25.6	343	16.6
Slope 1 (nm^{-1})	− 0.0216	0.0047	− 0.0247	0.0082	− 0.0252	0.0045	− 0.0347	0.0049
Slope 2 (nm^{-1})	− 0.0153	0.0031	− 0.0219	0.0070	− 0.0167	0.0035	− 0.0139	0.0053
Slope 3 (nm^{-1})	–	–	–	–	− 0.0151	0.0034	− 0.0244	0.0062

Table 3
P values and *F* statistics of the ANOVA applied to the regression model parameters (BKs and Slopes) obtained for spectral shape Group 1 and Group 2 previous to the linear discriminant analysis (LDA). Standardized and unstandardized coefficients obtained by the LDA using as predictor variables BKs and Slopes. Results are presented for M1 and M2 segmented regression models.

	Model 1 (M1)				Model 2 (M2)			
	Group 1		Group 2		Group 1		Group 2	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
BK1 (nm)	338	33.9	346	31.5	302	10.5	302	3.3
BK2 (nm)	–	–	–	–	355	25.6	343	16.6
Slope 1 (nm^{-1})	− 0.0216	0.0047	− 0.0247	0.0082	− 0.0252	0.0045	− 0.0347	0.0049
Slope 2 (nm^{-1})	− 0.0153	0.0031	− 0.0219	0.0070	− 0.0167	0.0035	− 0.0139	0.0053
Slope 3 (nm^{-1})	–	–	–	–	− 0.0151	0.0034	− 0.0244	0.0062

the above correction to refer all CDOM spectra to a reference pure water spectrum in order to maintain coherence between laboratories. Scanning CDOM samples and pure water spectrum against air should perform this.

4.2. Filtration procedures

Nearly identical CDOM spectra were obtained using either polycarbonate 0.2 μm membrane or 0.7 μm GFF filters (Fig. 2). Even though the nominal pore size differs, in average, these two types of filters allow the passage of the same type of molecules. There are several explanations for this observation. One explanation is that the fraction between 0.2 and 0.7 μm has a minor contribution to light absorption in the open ocean (Nelson et al., 1998). Electron microscopy, however, has revealed that membrane filters can have pores larger than 0.2 μm diameter as a consequence of the perforation process, which allows the overlapping of holes, creating pores twice or even three times larger than intended (Negri, pers. comm.). Nayar and Chou (2003) reported that the average pore size of a GFF filter is reduced in combusted filters due to contraction of the matrix of fibers. Furthermore, it is not unreasonable to expect that after filtering a few milliliters of sea water filters get clogged due to the formation of a tight network of soluble molecules in the sample. The importance of rinsing combusted GFF filters previous to sample collection has been already documented (Mitchell et al., 2003). We found that rinsing filters with 200 ml of pure water was enough to elute most loose glass fibers that might affect the absorption readings. Our results showed that pure water and CDOM samples filtered through GFF filters treated in this way were comparable to those obtained using membrane filters.

4.3. CDOM spectral shape

Our results demonstrated statistically by a multivariate analysis that at least two spectral forms of CDOM can be distinguished in the 275–400 nm range (Fig. 4), which in turn could be related to “coastal” and “oceanic” conditions (Fig. 1). Here we used a segmented regression model (SRM) to fit all spectral shapes of CDOM. This approach has the advantage of having higher accuracy than the SEM as shown by the residuals of the regressions (Fig. 5).

The SRM (Muggeo, 2003) has been applied successfully to find breakpoints (also known as joint points or change points) in several types of biological data like genomic sequences or shark maturity stages (Muggeo and Adelfio, 2011; Segura et al., 2013). It is easily run through a statistical package in open free software (Muggeo, 2008). Although the number of BKs must be provided a priori, the piecewise regression has the benefit of being able to statistically estimate the wavelength of the BKs. Nevertheless, it is important to bear in mind that in some cases the results were sensitive to the seed values provided for the BKs within the given wavelength range used for fitting M2.

We observed that the goodness of fit increases with the number of BKs. We explored the segmented models M1 (1 breakpoint and 2 slopes) and M2 (2 breakpoints and 3 slopes) within a narrow wavelength range (275–400 nm), similar to the one used by Helms et al. (2008), in a set of 145 CDOM spectra from a wide range of oceanographic regimes.

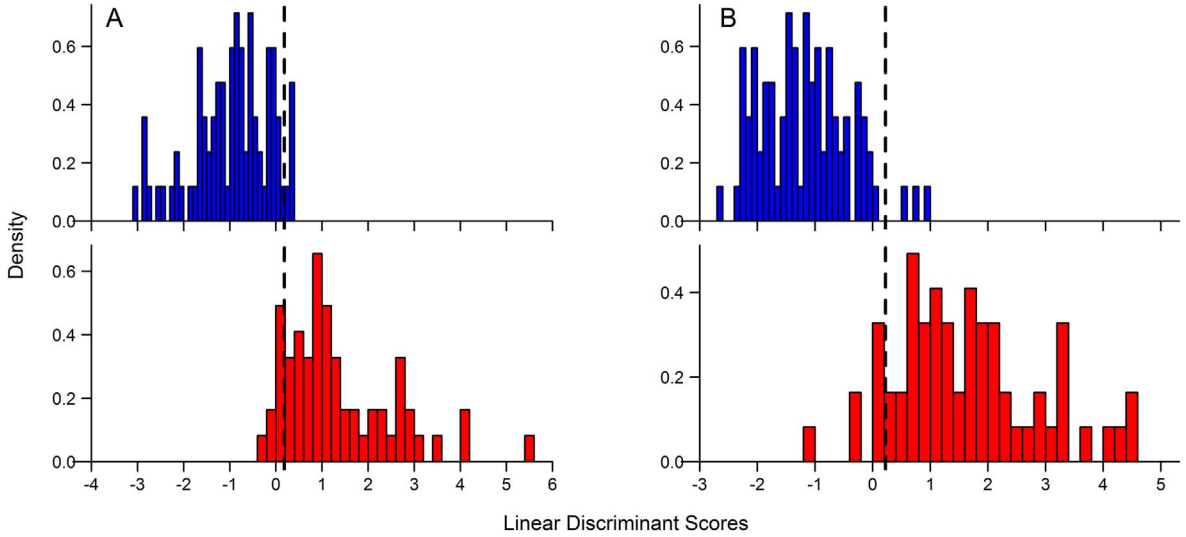

Fig. 8. Distributions of linear discriminant scores of the LDA based on the segmented regression parameters as predictor variables. A: LDA based on BKs and slopes after applying M1; B: LDA based on BKs and slopes after applying M2. The distributions show minimal degree of overlapping between spectral shape Groups 1 (blue, upper panel) and 2 (red, lower panel) in both cases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Group membership predicted by the linear discriminant analysis (LDA) using as predictor variables the breakpoints and slopes derived after applying the segmented regression models M1 and M2.

	LDA on segmented model parameters M1				LDA on segmented model parameters M2			
	Group 1	Group 2	Total	Error (%)	Group 1	Group 2	Total	Error (%)
Group 1	79	5	84	5.95	81	3	84	3.57
Group 2	9	52	61	14.75	7	54	61	4.83
Total	88	57	145	9.65	88	57	145	6.89

Table 5
Linear discriminant analysis (LDA) using as predictor variables a set of environmental variables (SST, SSS, logFl, logBD, and logDNC). ANOVA's p values and F statistics correspond to the prior group means for the set of variables.

	ANOVA		LDA coefficients	
	F	p value	Standardized	Unstandardized
SST	1.280	0.2597	–	–
SSS	37.202	$9.441\ e^{-9}$	− 0.2217	− 0.1015
logFl	15.875	0.0001	− 0.3963	− 0.4824
logBD	60.653	$1.261\ e^{-12}$	− 0.2608	0.2069
logDNC	81.829	$9.704\ e^{-16}$	0.9340	0.5416

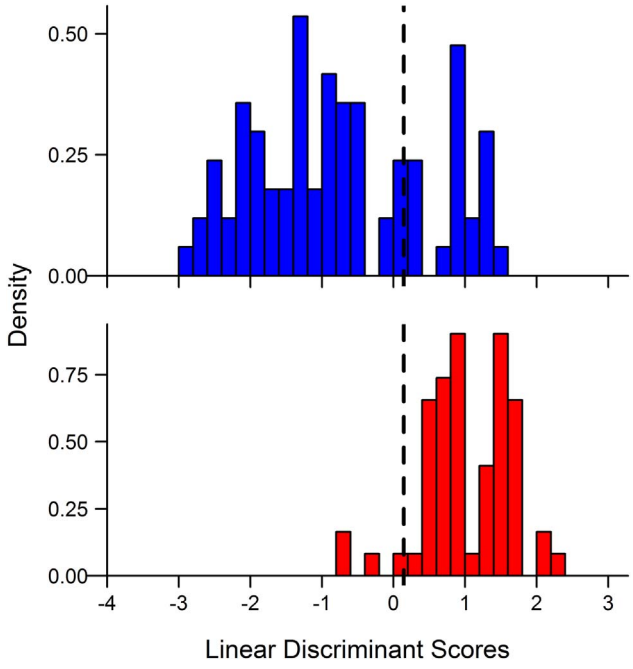

Fig. 9. Linear discriminant scores distributions after LDA based on five environmental variables as predictors. The distribution shows a small degree of overlapping between spectral shape Groups 1 (blue, upper panel) and 2 (red, lower panel). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
Group membership predicted by linear discriminant analysis based on environmental information.

	LDA on environmental variables			
	Group 1	Group 2	Total	Error (%)
Group 1	63	21	84	25.00
Group 2	4	57	61	6.55
Total	67	78	145	17.24

Besides, the upper wavelength limit (400 nm) was chosen to be a good and conservative wavelength at which the optical density of our dataset was above the detection limit of the instrument.

Other authors already stated the segmented aspect of some CDOM spectra. Sarpal et al. (1995) sampled oceanic and coastal sites in the Bellinghausen Sea and Antarctic Peninsula for CDOM and fitted two slopes ($S_{260–330}$ and $S_{330–410}$) in the log-linearized scale with accuracy higher than one. Both their oceanic and coastal spectra look like the spectra characterized in Group 2 and associated with "oceanic" environments (indeed the authors noticed the high transparency of Antarctic waters). Helms et al. (2008) proposed using the ratio of the slopes fitted between $S_{275–295}$ and $S_{350–400}$ (the S_R) to characterize CDOM spectra. They found that S_R increases with salinity, which suggests different spectral shapes. They also reported that $S_{275–295}$ and S_R shifts to lower values with increasing molecular weights of the organic matter, and $S_{350–400}$ increased after photo bleaching CDOM samples. Another complex method involving several parameters was proposed by Massicotte and Markager (2016). It is based on a Gaussian decomposition of the absorption spectra, which gives robust estimates of S and also models specific chromophores present in complex spectra from oceanic water samples. Other studies have looked at variations in CDOM shapes according to water masses (Nelson et al., 2010), including the transformation of recalcitrant DOM at depth (Catalá et al., 2015).

Most of the spectra classified as Group 1 were sampled in the Strait of Magallanes and along the coast of Chile, where high levels of CDOM are expected due to the terrestrial runoff from the surrounding fjords, shallow depths, and strong tides, on the one hand, and the presence of an eastern boundary upwelling system, on the other hand. Instead, most of the spectra classified as Group 2 were sampled in the open ocean, where lower levels of CDOM absorption typical of oligotrophic environments are expected. The SRM with 2 breakpoints, i.e., M2, fitted spectra of both shape groups with similar accuracy and provided the best fit compared to M1 and M0 (Fig. 5). Moreover, an unusual CDOM spectrum obtained from a sample collected near Cape Town showing a bump by mycosporine like amino-acids (MAAs) or by nitrate (Catalá et al., 2016) was modeled by M2 in the 275–400 nm range with significantly better accuracy than M1 and M0 (Fig. 7). This case was excluded from the CA because it was separated in a group on its own. The SRM M2 revealed subtle characteristics in the form of CDOM absorption spectra, which questions the idea that all CDOM spectra resemble a simple exponential. The position of BK1 was constantly close to 302 nm in both spectral shape groups. While the position of BK2 was somehow trivial in Group 1, it was quite close to 343 nm in Group 2, coincident with the previously reported inflection point in the second derivative of natural-log CDOM spectra (Helms and Mopper, 2006).

The value of Slope 3 (which is estimated by M2 in a narrow wavelength range between BK2 and 400 nm) could be used to estimate CDOM in the VIS where measured optical density values remain below the detection limit of traditional spectrophotometers (Fig. S4). This provides an objective solution to the unresolved problem of choosing a wavelength range for fitting a single exponential to model absorption in the VIS as a function of absorption in the UV region, which would facilitate comparing values among regions and dates.

The spectral shape is a feature defined by the chemical nature of the

compounds in a CDOM sample. That is, although our results show a strong association between the spectral forms and environments that we denominated “coastal” and “oceanic”, it must be emphasized that spectral shapes are not direct consequences of the geographical location alone. The absorption spectra encompass the sum of absorption by the mix of chemical components. Some works (Röttgers and Koch, 2012; Catalá et al., 2016) have investigated variations in CDOM spectra from open ocean waters in relation to the chemical composition, resulting in the identification of specific chromophores absorbing at different wavelengths (e.g., citochrome *c* at 415 nm, and nitrate at 302 nm). Here, the CA classified a small number of spectra sampled in the open South Atlantic Ocean as Group 1, and spectra of both spectral forms were found at five stations. At those stations the sampling was carried out at the surface and at different depths. It was noticeable that surface samples were classified as Group 2 (“oceanic”), while deep samples as Group 1 (“coastal”). Previous studies showed that the *S* slope shifted to lower values with depth due to lack of photo-degradation (Sarpal et al., 1995; Twardowski and Donaghay, 2002). This phenomenon is expected for this high latitude zone and season where these stations were located, which is marked by a strong seasonality of mixing-stratification processes driven by wind stress and solar irradiance (Longhurst, 2007). Thus, high slope values as those reported for Group 2 are expected for the upper layer samples. Another plausible explanation for the classification of these CDOM samples as Group 1 despite having been collected in the open ocean may be related to the vicinity of these sampling stations to Sea Mountains (Rudorff, 2013). These elevations of the seafloor are known to induce eddy upwelling cells of nutrient-rich deep waters that sustain local phytoplankton growth (Borrione and Schlitzer, 2012). Likewise, spectra close to Malvinas archipelago were classified as Group 1, which may be due to the presence of terrestrial runoff from the islands and a shallower bottom depth. Therefore, the denomination “terrestrial” rather than “coastal” may also be appropriate for Group 1.

An important benefit of the linear discriminant analysis (LDA) is the ability to predict group membership for classifying new cases given the identifying features. Thus, after fitting M2 within 275–400 nm, the BKs and slopes could be used to classify CDOM spectra into spectral shape “Group 1 - coastal” and “Group 2 - oceanic” using the discriminant function (Eq. 6) based on the unstandardized linear discriminant coefficients and a simple classification rule (above or below the reported cut-off). Similarly, environmental data could be used with Eq. 7 to predict the probable spectral shape of CDOM absorption spectra in the absence of samples.

CDOM is an important contributor to ocean color that does not always covariate with chlorophyll-a concentration. Improving the acquisition method of $a_{CDOM}(\lambda)$ is a first essential step to facilite the inter-comparison of in situ measurements performed at different laboratories. The SRM presented here seems to be a good tool to reconstruct the whole spectrum of $a_{CDOM}(\lambda)$ at all wavelengths (Fig. S4). This approach would allow retrieving a more reliable value of $a_{CDOM}(440)$, a variable often used in bio-optical studies, and characterizing CDOM absorption spectra in a simple way. The development of satellite applications of the SRM M2 parameters, which relate spectral shapes to environmental conditions, would contribute to improve ocean color products and understanding of CDOM dynamics in marine ecosystems.

5. Conclusions

From the assessment of factors affecting the laboratory analysis of CDOM, the main result was that nearly identical CDOM spectra were obtained from replicate samples using either membrane (0.2 μm) or GFF filters (0.7 μm nominal), which may be due to several possible factors, i.e., minor absorption of the fraction between 0.2 and 0.7 μm (Nelson et al., 1998) and pore contraction of GFF filters after combustion (Nayar and Chou, 2003). Scanning spectra of CDOM samples, real blanks (filtered pure water), and the water used at a given laboratory against air is a practice that provides the possibility of referencing CDOM absorption spectra to the same “ideal” water spectrum, eliminating possible sources of variations.

A more accurate alternative to the single decreasing exponential model, i.e., a concatenated exponential model, was proposed to describe CDOM spectra. It was first demonstrated statistically from a multivariate approach using only information provided by the $a_{CDOM}(\lambda)$ in the range 275–400 nm that at least 2 spectral shapes generally exist. Spectra classified in Group 1 showed a clear breakpoint close to 302 nm while spectra in Group 2 showed a breakpoint close to 302 nm and another close to 343 nm. These results showed that the single exponential model traditionally used to describe CDOM spectra lacks the ability to reveal these spectral features, in agreement with previous studies. The concatenated exponential model with 2 breakpoints (Muggeo, 2008; Muggeo, 2003), determined by piecewise regression, fitted with high accuracy the spectral shape of both groups. This was true even for special cases such as spectra having a pronounced peak in the UV (probably due to MAAs or nitrate). Moreover, it was possible to associate different spectral shapes with “coastal” and “oceanic” environmental conditions. Using the Slope 3 derived from the SRM with 2 breakpoints provides an objective way to estimate CDOM absorption in the VIS ($a_{CDOM}(440)$ is usually involved in satellite algorithms). Detailed examination of the location of the breakpoints and the spectral slopes in the various segments, including their ratios, made in relation to CDOM composition measurements (not available in this study), may provide further information about CDOM origin and dominant processes (photo degradation, microbial alteration), a perspective for future work. Additional research on the application of the segmented regression in CDOM analysis may lead to the development of improved ocean color algorithms.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.marchem.2017.03.012.

Acknowledgments

The R/V Melville crew and colleagues on board are gratefully acknowledged for their help with data collection and processing. The National Science Foundation and The Scripps institution of Oceanography provided ship time, logistics, and technical support for the MV1102 cruise. The National Space and Aeronautics Administration supported Robert Frouin under various grants. This work was partially funded by the Instituto Nacional de Investigación y Desarrollo Pesquero (INIDEP), the Consejo Nacional de Investigaciones Científicas y Técnicas through a scholarship for M.G. Ruiz, and the Inter-American Institute for Global Change Research (grant CRN3094). We thank Daniel Hernández, INIDEP, for his meticulous advice on statistics, Lucrecia Allega, INIDEP, for assistance with the map, and the reviewers, for helpful comments. This work is part of M. Guillermina Ruiz's PhD thesis. This is INIDEP contribution n° 2047.

References

Andrew, A.A., Del Vecchio, R., Subramaniam, A., Blough, N.V., 2013. Chromophoric dissolved organic matter (CDOM) in the Equatorial Atlantic Ocean: optical properties and their relation to CDOM structure and source. Mar. Chem. 148 (0), 33–43.

Bivand, R.S., Pebesma, E., Gómez-Rubio, V., 2013. Applied spatial data analysis with R. In: Use R!. Springer-Verlag New York, New York.

Blough, N.V., Del Vecchio, R., 2002. Chromophoric DOM in the coastal environment. In: Hansell, D.A., Carlson, C.A. (Eds.), Biogeochemistry of Marine Dissolved Organic Matter. Elsevier Science, San Diego, pp. 509–546.

Borrione, I., Schlitzer, R., 2012. Distribution and recurrence of phytoplankton blooms around South Georgia, Southern Ocean. Biogeosci. Discuss. 9 (8), 10087–10120.

Bricaud, A., Morel, A., Prieur, L., 1981. Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains. Limnol. Oceanogr. 26 (1), 43–53.

Carder, K.L., Steward, R.G., Harvey, G.R., Ortner, P.B., 1989. Marine humic and fulvic acids: their effects on remote sensing of ocean chlorophyll. Limnol. Oceanogr. 34 (1), 68–81.

Carignan, M., Carreto, J.I., 2013. Characterization of mycosporine-serine-glycine methyl

ester, a major mycosporine-like amino acid from dinoflagellates: a mass spectrometry study. J. Phycol. 49 (4), 680–689.
Catalá, T.S., et al., 2015. Water mass age and aging driving chromophoric dissolved organic matter in the dark global ocean. Glob. Biogeochem. Cycles 29 (7), 917–934.
Catalá, T.S., et al., 2016. Chromophoric signatures of microbial by-products in the dark ocean. Geophys. Res. Lett. 43 (14), 7639–7648.
Clark, C.D., Litz, L.P., Grant, S.B., 2008. Salt marshes as a source of chromophoric dissolved organic matter (CDOM) to Southern California coastal waters. Limnol. Oceanogr. 53 (5), 1923–1933.
D'Sa, E.J., Miller, R.L., Del Castillo, C., 2006. Bio-optical properties and ocean color algorithms for coastal waters influenced by the Mississippi River during a cold front. Appl. Opt. 45 (28), 7410–7428.
D'Sa, E.J., Steward, R.G., Vodacek, A., Blough, N.V., Phinney, D., 1999. Determining optical absorption of colored dissolved organic matter in seawater with a liquid capillary waveguide. Limnol. Oceanogr. 44 (4), 1142–1148.
Fichot, C.G., Benner, R., 2012. The spectral slope coefficient of chromophoric dissolved organic matter (S275–295) as a tracer of terrigenous dissolved organic carbon in river-influenced ocean margins. Limnol. Oceanogr. 57 (5), 1453–1466.
Gordon, H.R., Morel, A.Y., 1983. Remote assessment of ocean color for interpretation of satellite visible imagery. A review. In: Barber, R.T., Mooers, C.N.K., Bowman, M.J., Zeitzschel, B. (Eds.), Lecture Notes on Coastal and Estuarine Studies. Springer-Verlag, New York.
Helms, J.R., Mopper, K., 2006. Spectral shape as an indicator of molecular weight in chromophoric dissolved organic matter. In: Old Dominion University theses: Chemistry: 2006, .
Helms, J.R., et al., 2008. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. Limnol. Oceanogr. 53 (3), 955–969.
Hernes, P.J., Benner, R.C., 2003. Photochemical and microbial degradation of dissolved lignin phenols: implications for the fate of terrigenous dissolved organic matter in marine environments. J. Geophys. Res. Oceans 108 (C9), 3291.
Højerslev, N.K., 1980. On the origin of yellow substance in the marine environment. In: Oceanogr. Rep. 42. University Copenhagen, pp. 1–35.
Højerslev, N.K., Aas, E., 2001. Spectral light absorption by yellow substance in the Kattegat-Skagerrak area. Oceanologia 43 (1), 39–60.
IOCCG, 2000. Remote Sensing of Ocean Colour in Coastal, and Optically-Complex, Waters. Reports of the International Ocean-Colour Coordinating Group Vol. 3 (Dartmouth, Canada).
Jerlov, N.G., 1957. A transparency-meter for ocean water. Tellus 9 (2), 229–233.
Jørgensen, L., Stedmon, C.A., Granskog, M.A., Middelboe, M.C.G.L., 2014. Tracing the long-term microbial production of recalcitrant fluorescent dissolved organic matter in seawater. Geophys. Res. Lett. 41 (7), 2481–2488.
Kirk, J.T.O., 2011. Light and Photosynthesis in Aquatic Ecosystems. Cambridge University Press, New York (662 pp.).
Kowalczuk, P., Stedmon, A., C. and Markager, S., 2006. Modeling absorption by CDOM in the Baltic Sea from season, salinity and chlorophyll. Mar. Chem. 101 (1–2), 1–11.
Lalli, C.M., Parsons, T.R., 1997. Biological Oceanography: An Introduction. Butterworth-Heinemann Ltd, Oxford England, pp. 314.
Lee, Z., et al., 2015. Hyperspectral absorption coefficient of "pure" seawater in the range of 350–550 nm inverted from remote sensing reflectance. Appl. Opt. 54 (3), 546–558.
Loiselle, S.A., et al., 2009. Optical characterization of chromophoric dissolved organic matter using wavelength distribution of absorption spectral slopes. Limnol. Oceanogr. 54 (2), 590–597.
Longhurst, A.R., 2007. Chapter 2 - biogeographic partition of the ocean. In: Ecological Geography of the Sea, second ed. Academic Press, Burlington, pp. 19–34.
Lorenzoni, L., et al., 2011. Bio-optical characteristics of Cariaco Basin (Caribbean Sea) waters. Cont. Shelf Res. 31 (6), 582–593.
Lutz, V., et al., 2016. Bio-optical characteristics along the Straits of Magallanes. Cont. Shelf Res. 119, 56–67.
Lutz, V.A., Subramaniam, A., Negri, R.M., Silva, R.I., Carreto, J.I., 2006. Annual variations in bio-optical properties at the 'Estación Permanente de Estudios Ambientales (EPEA)' coastal station, Argentina. Cont. Shelf Res. 26 (10), 1093–1112.
Magnuson, A., Harding, L.W.J., Mallonee, M.E., Adolf, J., 2004. Bio-optical model for Chesapeake Bay and the Middle Atlantic Bight. Estuar. Coast. Shelf Sci. 61, 403–424.
Mannino, A., Novak, M.G., Hooker, S.B., Hyde, K., Aurin, D., 2014. Algorithm development and validation of CDOM properties for estuarine and continental shelf waters along the northeastern U.S. coast. Remote Sens. Environ. 152, 576–602.
Mason, J.D., Cone, M.T., Frye, E.S., 2016. Ultraviolet (250–550 nm) absorption spectrum of pure water. Appl. Opt. 55, 7163–7172.
Maritorena, S., Siegel, D.A., Peterson, A.R., 2002. Optimization of a semianalytical ocean color model for global-scale applications. Appl. Opt. 41, 2705–2714.
Massicotte, P., Markager, S., 2016. Using a Gaussian decomposition approach to model absorption spectra of chromophoric dissolved organic matter. Mar. Chem. 180, 24–32.
Mei, Z.-P., et al., 2010. Modeling the timing of spring phytoplankton bloom and biological production of the Gulf of St. Lawrence (Canada): effects of colored dissolved organic matter and temperature. Cont. Shelf Res. 30 (19), 2027–2042.
Miller, R.L., Belz, M., Castillo, C.D., Trzaska, R., 2002. Determining CDOM absorption spectra in diverse coastal environments using a multiple pathlength, liquid core waveguide system. Cont. Shelf Res. 22 (9), 1301–1310.
Mitchell, B.G., Kahru, M., Wieland, J., Stramska, M., 2003. Determination of spectral absorption coefficients of particles, dissolved material and phytoplankton for discrete water samples. In: Mueller, J.L., Fargion, G.S., McClain, C.R. (Eds.), Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4. 20771. NASA Goddard Space Flight Center, NASA/TM - 2003, Greenbelt, Maryland, pp. 39–56.
Morel, A., et al., 2007. Optical properties of the "clearest" natural waters. Limnol. Oceanogr. 52 (1), 217–229.
Morel, A., Prieur, L., 1977. Analysis of variations in ocean color. Limnol. Oceanogr. 22 (4), 709–722.
Muggeo, V., 2008. Segmented: an {R} package to fit regression models with broken-line relationships. R News 8 (1), 20–25.
Muggeo, V.M., Adelfio, G., 2011. Efficient change point detection for genomic sequences of continuous measurements. Bioinformatics 27 (2), 161–166.
Muggeo, V.M.R., 2003. Estimating regression models with unknown break-points. Stat. Med. 22 (19), 3055–3071.
NASA, 2009. Distance to the Nearest Coast. NASA Ocean biology DAAC. (released June 2009, http://oceancolor.gsfc.nasa.gov/DOCS/DistFromCoast/, accessed April 22 2015).
Nayar, S., Chou, L.M., 2003. Relative efficiencies of different filters in retaining phytoplankton for pigment and productivity studies. Estuar. Coast. Shelf Sci. 58 (2), 241–248.
Nelson, N., Siegel, D.A., Michaels, A.F., 1998. Seasonal dynamics of colored dissolved material in the Sargasso Sea. Deep-Sea Res. II 45, 931–957.
Nelson, N.B., Coble, P.G., 2009. Optical analysis of chromophoric dissolved organic matter. In: Wurl, O. (Ed.), Practical Guidelines for the Analysis of Seawater. CRC Press, pp. 87.
Nelson, N.B., Siegel, D.A., Carlson, C.A., Swan, C.M.C.L., 2010. Tracing global biogeochemical cycles and meridional overturning circulation using chromophoric dissolved organic matter. Geophys. Res. Lett. 37, L03610. http://dx.doi.org/10.1029/2009GL042325.
Nelson, N.B., Siegel, D.A., 2013. The global distribution and dynamics of chromophoric dissolved organic matter. Annu. Rev. Mar. Sci. 5, 447–476.
Panella, A., et al., 1991. A preliminary contribution to understanding the hydrological characteristics of the Straits of Magellan: Austral Spring 1989. Boll. Ocean. Teor. Appl. 9 (2–3), 106–126.
Pebesma, E.J., Bivand, R.S., 2005. Classes and methods for spatial data in R. R News 5 (5), 9–13.
Pope, R.M., Fry, E.S., 1997. Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements. Appl. Opt. 36 (33), 8710–8723.
Pope, R.M., Weidemann, A.D., Fry, E.S., 2000. Integrating Cavity Absorption Meter measurements of dissolved substances and suspended particles in ocean water. Dyn. Atmos. Oceans 31 (1–4), 307–320.
R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Romera-Castillo, C., Sarmento, H., Álvarez-Salgado, X.A., Gasol, J.M., Marraséa, C., 2010. Production of chromophoric dissolved organic matter by marine phytoplankton. Limnol. Oceanogr. 55 (1), 446–454.
Röttgers, R., Doerffer, R., 2007. Measurements of optical absorption by chromophoric dissolved organic matter using a point-source integrating-cavity absorption meter. Limnol. Oceanogr. Methods 5 (May), 126–135.
Röttgers, R., Koch, B.P., 2012. Spectroscopic detection of a ubiquitous dissolved pigment degradation product in subsurface waters of the global ocean. Biogeosciences 9 (7), 2585–2596.
Rudorff, N.D., 2013. Ocean Colour Variability in the Southern Atlantic and Southeastern Pacific. (PhD Thesis, INPE, São José dos Campos, 336 pp.). http://mtc-m16d.sid.inpe.br/col/sid.inpe.br/mtc-m19/2013/09.17.14.26/doc/publicacao.pdf.
Rudorff, N.D., et al., 2014. Ocean-color radiometry across the Southern Atlantic and Southeastern Pacific: accuracy and remote sensing implications. Remote Sens. Environ. 149, 13–32.
Sarpal, R.S., Mopper, K., Keiber, D.J., 1995. Absorbance properties of dissolved organic matter in Antartic sea water. Antarct. J. 30, 139–140.
Segura, A.M., Milessi, A.C., Vögler, R., Galván-Magaña, F., Muggeo, V., 2013. The determination of maturity stages in male elasmobranchs (Chondrichthyes) using a segmented regression of clasper length on total length. Can. J. Fish. Aquat. Sci. 70 (6), 830–833.
Siegel, D.A., Maritorena, S., Nelson, N.B., Behrenfeld, M.J., McClain, C.R., 2005. Colored dissolved organic matter and its influence on the satellite-based characterization of the ocean biosphere. Geophys. Res. Lett. 30 (20), L20605.
Sogandares, F.M., Fry, E.S., 1997. Absorption spectrum (340–640 nm) of pure water. I. Photothermal measurements. Appl. Opt. 36 (33), 8699–8709.
Subramaniam, A., Carpenter, E.J., Karentz, D., Falkowski, P.G., 1999. Optical properties of the marine diazotrophic cyanobacteria *Trichodesmium* spp.; I -absorption and spectral photosynthetic characteristics. Limnol. Oceanogr. 44 (3), 608–617.
Toming, K., Arst, H., Paavel, B., Laas, A., Nõges, T., 2009. Spatial and temporal variations in coloured dissolved organic matter in large and shallow Estonian waterbodies. Boreal Environ. Res. 14 (6), 959–970.
Twardowski, M.S., Boss, E., Sullivan, J.M., Donaghay, P.L., 2004. Modeling the spectral shape of absorption by chromophoric dissolved organic matter. Mar. Chem. 89 (1–4), 69–88.
Twardowski, M.S., Donaghay, P.L., 2002. Photobleaching of aquatic dissolved materials: Absorption removal, spectral alteration, and their interrelationship. J. Geophys. Res. Oceans 107 (C8), 6-1–6-12.