A model for remote estimation of ultraviolet absorption by chromophoric dissolved organic matter based on the global distribution of spectral slope

Chantal M. Swan, Norman B. Nelson, David A. Siegel, Erik A. Fields

Abstract

Absorption of ultraviolet radiation (UV, 280–400 nm) by chromophoric dissolved organic matter (CDOM) precedes a host of light-sensitized surface ocean processes relevant to global climate. These include photo-biogeochemical cycling of organic material, release of sulfur and carbon-containing gases to the atmosphere, and the photoprotection of marine microorganisms. Synoptic CDOM absorption data in the UV is highly desired yet difficult to estimate by satellite methods as the atmosphere interferes with direct detection of water-leaving UV radiance. The absorption spectrum of CDOM is typically modeled as an exponential function in which a spectral slope parameter, $S$, describes the rate of decrease in absorption with increase in wavelength. Significant functional relationships are observed in aquatic environments between $S$ and the CDOM absorption coefficient at 443 nm, $a_{CDOM}(443)$. In this paper, we use a large, systematic dataset of spectrosopic CDOM measurements from the U.S. CO$_2$/CLIVAR Repeat Hydrography Survey to examine the relationship between $S$ and $a_{CDOM}(443)$ as a means to model $a_{CDOM}($λ$)$ in the UV from ocean color. Our resultant model predicts $a_{CDOM}($λ$)$ at wavelengths from 325 to 412 nm from the absorption coefficient of colored dissolved and detrital materials (CDM) at 443 nm, $a_{CDM}(443)$, retrieved by an existing semi-analytical ocean color algorithm. Expected agreement (near 1:1) with the training dataset was achieved ($r^2 = 0.71–0.85$, $p = 0$, $n = 127$). Considering inherent satellite data uncertainties as well as the model’s limitations in regions with potential terrestrial influence, good correspondence between modeled and in situ values was observed during independent validation with open ocean CDOM data, such as from BIOSOPE ($r^2 = 0.77–0.85$, $p < 0.05$, $n = 29$). The model has immediate application in global scale assessments of photochemical rate processes and CDOM cycling in the open ocean due to its simplicity and optimization using a large base of field data (>7500 samples) from diverse Case I waters.

1. Introduction

Chromophoric dissolved organic matter (CDOM) plays a crucial role in marine systems via its absorption of ultraviolet radiation (UV, 280–400 nm). Over broad areas of the upper ocean, CDOM can account for as much as 95% of total non-water UV absorption (Johannessen et al., 2003; Nelson et al., 1998; Zepp et al., 2007). UV absorption by CDOM sensitizes a suite of photochemical reactions at the sea surface including dimethylsulfide photolysis (Kieber et al., 1996; Toole et al., 2006); CO$_2$, CO, and COS release (Mopper & Kieber, 2002; Stubbins et al., 2006; Zafiriou et al., 2008); and transformation of microbial bioavailability of organic matter (Benner & Biddanda, 1998). The quantity of UV absorbed by oceanic CDOM is thus an essential factor in the study of such reactions and their level of impact on biogeochemical cycles and climate (Reader & Miller, 2011; Whitehead & de Mora, 2000; Zepp et al., 2007).

To make global-scale assessments of photochemical processes, optically data are essential and primarily acquired through remote sensing. Ocean color algorithms (Carder et al., 1999; Lee et al., 2002; Mannino et al., 2008; Maritorena et al., 2002, 2010; Morel & Gentili, 2009) have proven to be critical for determining visible light absorption by CDOM and its impact on the accuracy of chlorophyll estimates from space (Carder et al., 1989; Nelson & Siegel, 2013; Siegel et al., 2002, 2005a, 2005b, 2013). Recently published empirical models have made use of remotely sensed visible reflectance ratios (or multiple linear regression thereof) to predict the irradiance attenuation coefficient, $k_d($λ$)$, in the UV, which when combined with knowledge of the relation between $k_d($λ$)$ and the CDOM absorption coefficient, $a_{CDOM}($λ$)$, can be used to extrapolate $a_{CDOM}($λ$)$ in the UV assuming a single spectral slope value (Fichot et al., 2008; Johannessen et al., 2003). The training datasets used to develop even the most sophisticated working algorithms of this sort (SeaUV and SeaUV$_C$ algorithms, Fichot & Miller, 2010; Fichot...
et al., 2008) tend to favor the U.S. coastal water bodies and North Atlantic Ocean. The authors themselves note the need for datasets inclusive of a greater variety of optical domains to improve parameterizations, especially due to assumptions inherent when using empirical relationships between optical properties (Fichot et al., 2008). Larger field datasets and additional approaches for estimating UV absorption by CDOM that are optimized to handle both productive pelagic regimes as well as the vast oligotrophic deserts of the open ocean, the subtropical gyres, are thus desirable. The general need to incorporate field data from diverse oceanographic provinces in bio-optical model development serves as motivation for the present study. Here we employ a large, systematic set of global field observations of CDOM in a novel yet simple approach toward predicting $\alpha_{\text{CDOM}}(\lambda)$ in the UV from the remotely sensed absorption coefficient of colored dissolved and detrital materials (CDM) at 443 nm, $\alpha_{\text{CDM}}(443)$, which is the quantity retrieved by existing ocean color algorithms.

Spectral light absorption by CDOM in the ocean, which can be measured either by in situ instruments or from discrete water samples via absorption spectroscopy, is frequently modeled as an exponential function over a short wavelength interval (Eq. (1)) (Bricaud et al., 1981):

$$a_{\text{CDOM}}(\lambda) = a_{\text{CDOM}}(\lambda_o) \exp^{-S(\lambda - \lambda_o)}$$  \hspace{1cm} (1)

where $\lambda$ and $\lambda_o$ are a desired known wavelength and reference wavelength, respectively, and $S$ is a spectral slope parameter modeled through either a linear fit to the log-transformed spectrum, or a non-linear fit to the curve of the spectrum over the spectral range considered (Stedmon et al., 2000; Twardowski et al., 2004).

The selection of wavelength interval as well as the fitting routine used in computing $S$ can greatly affect the interpretation of data, which has often complicated comparison of results across CDOM studies (Del Vecchio & Blough, 2004; Stedmon et al., 2000; Twardowski & Donahgay, 2002; Twardowski et al., 2004). On the other hand, the option of wavelength interval selection for computing $S$ has resulted in a variety of investigations into CDOM composition and dynamics using the spectral slope parameter. Several permutations, including ratios of $S$ determined for progressive wavelength intervals, as well as the functional relationship of $S$ to the magnitude of CDOM absorption, have been proposed as investigative tools for understanding both source and diagenetic state of CDOM in aquatic environments (Del Castillo & Cole, 2000; Helms et al., 2005; Kitidis et al., 2006; Loiselle et al., 2009a, 2009b; Nelson et al., 2007, 2010; Stedmon & Markager, 2003; Weishaar et al., 2003; Zhang et al., 2007). An inverse relationship between $S$ and $\alpha_{\text{CDOM}}(\lambda)$ at a given wavelength has been documented within a number of studies (Bricaud et al., 2010; Del Castillo & Cole, 2000; Rochelle-Newall et al., 2004; Stedmon & Markager, 2001, 2003; Stedmon et al., 2000; Vodacek et al., 1997; Yacobi et al., 2003). This relationship has been used to diagnose the relative contributions of terrestrial versus marine organic matter, as well as the effect of end-member mixing on the optical properties of coastal waters (Del Vecchio & Blough, 2004; Stedmon & Markager, 2001, 2003).

In the open ocean, increases in $S$ are generally an indication of the process of solar photobleaching, as greater proportional absorption loss relative to initial occurs with increase in wavelength across the UV-visible spectrum (Del Vecchio & Blough, 2004; Swan et al., 2012; Twardowski & Donahgay, 2002; Váňáhoňá & Wetzel, 2004; Vodacek et al., 1997). Photobleaching regulates the global surface distribution of CDOM as modulated by autochthonous production pathways and vertical mixing of CDOM. Climatology of $\alpha_{\text{DMF}}(443)$ in the surface ocean (Fig. 1) from SeaWiFS reveals distinct patterns that support this, reflecting the relative impact of solar bleaching of CDOM with latitude, and vertical input of elevated CDOM from subthermocline waters (Nelson & Siegel, 2013; Siegel et al., 2002, 2013). A stark contrast is observed, for example, between the belt of elevated CDOM abundance in the eastern equatorial Pacific upwelling region relative to the CDOM-depleted waters that have longer surface residence times, such as the subtropical gyres of the major ocean basins (Fig. 1). This suggests that throughout most of the open ocean where terrestrial sources of organic material have little influence, photobleaching is the primary control on the relationship between $S$ and $\alpha_{\text{CDOM}}(\lambda)$. The univocal nature of the relationship provides a potential means for extrapolating satellite-derived $\alpha_{\text{CDOM}}(443)$ into the UV, assuming the absence of riverine influence (Stedmon & Markager, 2001).

In this study, we develop a model based on the relation between $S$ and $\alpha_{\text{CDOM}}(443)$ from thousands of field measurements of CDOM absorption spectra from a wide range of oceanic biomes, including temperate shelf seas, waters in proximity to the Antarctic ice shelf, the oligotrophic subtropical gyres in both hemispheres, productive subarctic gyres and active upwelling zones at the equator. Model coefficients are then tuned using input of $\alpha_{\text{CDOM}}(443)$ from the semi-analytical GSM algorithm (Siegel et al., 2002) in order to estimate CDOM absorption over 325–412 nm from ocean color. In the following sections we describe the development, validation and interpretation of the model using archival data from SeaWiFS and MODIS Aqua, and indicate its potential applications.

2. Data and methods

2.1. CLIVAR data

Global field measurements of CDOM absorption used in model development and validation were acquired through spectroscopy of filtered seawater samples collected along selected transects as part of the U.S. CO2/CLIVAR Repeat Hydrography Survey from 2003 through 2008. The station locations of data from transects used in the present study are superimposed on the surface map of $\alpha_{\text{CDOM}}(443)$ (Fig. 1) and listed along with dates of collection in Tables 1 and 2. The model development dataset consisted of 8 CLIVAR transects totaling 7611 discrete measurements of CDOM absorption throughout the global ocean in regions with contrasting trophic conditions, and from depths ranging from the surface to 5500 m deep.

The A20 and A22 sections sampled waters of the Sargasso Sea, Gulf Stream, U.S. continental shelf and sections of the Caribbean Sea, as well as the North Atlantic subtropical mode water (relatively low in CDOM) and North Atlantic Deep Water masses (Nelson et al., 2007). The P16S and P16N transects collectively covered the meridional North Pacific along 210° longitude, sampling the productive subarctic gyre and coastal waters near Alaska, the eastern equatorial Pacific upwelling zone characterized by elevated CDOM, the subtropical gyres in both hemispheres that are relatively low in CDOM, and Antarctic cumpolar and bottom waters (Swan et al., 2009). The IBS and I9N transects spanned the meridional Indian Ocean along approximately 90° longitude sampling Southern Ocean waters close to the ice shelf, then northward into the Indian Ocean equatorial regime as influenced by the Indonesian throughflow, and up into the relatively fresh and warm waters of the Bay of Bengal. The I6 transect sampled the polar front of the Indian sector of the Southern Ocean as well as the relatively fresh and highly productive surface waters of the Agulhas current system flowing around South Africa. Finally, the P18 transect further sampled waters of the eastern Pacific, including the hyper-oligotrophic sub-tropical gyre in that region. The model development dataset derives from a single methodology, comprising the first large systematic dataset of CDOM that captures the range of variability in CDOM spectral signature throughout the world ocean, thus optimizing the potential for global applicability of the model.

A subset of CLIVAR data consisting of 3 transects from the surface ocean spanning both hemispheres (CLIVAR A16N, A16S and P2) was reserved for independent validation of the model. The A16N cruise sampled waters of the North Atlantic, including subarctic freshwater lenses originating from Arctic waters flowing through the Denmark Strait, as well as waters in the vicinity of the North Atlantic Deep
CLIVAR campaigns were treated according to this protocol. For most of the campaigns, samples were analyzed shipboard immediately following filtration, whereas for several of the campaigns (I6S, A16S, A16N and P2) samples were filtered and refrigerated at 4 °C for later on-shore analysis within a few months from the date of collection. Laboratory tests show no appreciable alteration of the CDOM absorption coefficient in filtered water samples from the open ocean stored at 4 °C for up to 1 year (Swan et al., 2012).

The exponential decrease in CDOM absorption with increase in wavelength (Eq. (1)) permits estimation of the spectral slope parameter. We used a least-squares non-linear curve fitting approach to the CDOM spectrum over the wavelength interval 320–400 nm, and refer to this parameter herein as $S_{320-400}$ to emphasize its derivation from the spectrophotometric CDOM absorption measurement. We selected this spectral range for the fit because many of our data were sampled from oligotrophic waters that comprised vast areas of the open ocean, where absorption values in the visible were too low to provide consistent fits. The use of a narrow spectral range for computing the spectral slope of CDOM not only improves the resultant coefficients at 325 nm that were less than the instrumental detection limit of 0.005 m$^{-1}$, and thus potential artifacts of the $S_{320-400}$ calculation that arise below this absorption value. The 325 nm absorption criterion was chosen because the best signal-to-noise ratio for CDOM absorption is achieved at this wavelength using UltraPath™.

Table 1

<table>
<thead>
<tr>
<th>Cruise transect</th>
<th>Region</th>
<th>Latitude range</th>
<th>Longitude range</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIVAR P16S</td>
<td>South Pacific/Southern Ocean</td>
<td>71°S–17°S</td>
<td>210°</td>
<td>Jan.–Feb. 2005</td>
</tr>
<tr>
<td>CLIVAR I9N</td>
<td>Equatorial/North Indian Ocean</td>
<td>28°S–18°N</td>
<td>85°–95°</td>
<td>Mar.–May. 2007</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Cruise transect</th>
<th>Region</th>
<th>Latitude range</th>
<th>Longitude range</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIVAR A16S</td>
<td>South Atlantic</td>
<td>2°S–60°S</td>
<td>325°–335°</td>
<td>Jan.–Feb. 2005</td>
</tr>
<tr>
<td>BIOSOPE</td>
<td>South Pacific</td>
<td>30°N</td>
<td>133°–242°</td>
<td>Jul.–Aug. 2004</td>
</tr>
</tbody>
</table>

Water formation region. The A16S transect sampled the offshore waters of Brazil, the South Atlantic subtropical gyre, as well as the subantarctic frontal zone of the Atlantic (~45°S). The P2 cruise transected the zonal North Pacific, including the subtropical gyre and low salinity tongue that characterizes North Pacific Intermediate Water (Talley, 1997).

Datasets used in model development. U.S. CO2/CLIVAR Repeat Hydrography Survey and BIOSOPE Project (Claustre et al., 2008) transect coverage and dates of spectrophotometric CDOM measurements.

Fig. 1. Surface distribution of absorption by colored dissolved and detrital materials, $a_{\text{CDM}}(443, \text{m}^{-1})$, from the GSM algorithm using SeaWiFS climatology (1997–2005) (Siegel et al., 2002). Yellow markers = field data used in model development; white markers = field data used for independent model validations (See Tables 1 and 2 for locations and dates of field campaigns). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

---

2.2. BIOSOPE data

We used data from the Biogeochemistry and Optics South Pacific Experiment (BIOSOPE, Claustre et al., 2008) to independently validate the performance of the model in predicting CDOM absorption coefficients in the UV. The acquisition of BIOSOPE data, funded through grants from the Centre National de Recherche Scientifique and Institut National des Sciences de l’Univers, was possible through the following website: http://www.obs-vlfr.fr/proof/vt/op/ec/biosope/bio.htm.

The BIOSOPE cruise transected the South Pacific subtropical gyre, which is presumed to host the most optically clear waters in the global ocean (Morel et al., 2007a, 2007b; Tedetti et al., 2007). The cruise track extended from the mesotrophic waters surrounding the Marquesas Islands, through the hyper-oligotrophic subtropical gyre, and into the highly productive waters of the coastal Chilean upwelling zone. Samples were filtered through 0.2 μm membranes, and analyzed on an UltraPath ™ instrument similar to that used for the CLIVAR samples as described above, with the slight methodological difference that NaCl solution was used to correct for refractive index offsets instead of the artificial seawater (Bricaud et al., 2010). Surface (5 m) values of CDOM absorption in the 325–412 nm range were extracted from the BIOSOPE database for the validation.

2.3. Ocean color-derived $\alpha_{CDOM}(443)$ data

The semi-analytical GSM algorithm uses constant spectral slopes for the inherent optical properties of chlorophyll, CDOM and particulate backscattering in order to parse component absorption from the normalized water-leaving radiance spectrum provided by ocean color sensors (Maritorena et al., 2010). For model testing and implementation, we utilized $\alpha_{CDOM}(443)$ data of various spatial and temporal resolutions computed by GSM. For the validation, 8-day mean $\alpha_{CDOM}$ values were computed using 1–9 pixels from the neighborhood of each station location. Data from the SeaWiFS reprocessing (R2009.1) were used everywhere in the present study except with field data collected in 2008 (CLIVAR transects I6S and P18) when SeaWiFS was not operational. For those samples, MODIS data were used with GSM. SeaWiFS climatology over 10 years (1997–2007) was used to model the spatial and time-latitude distributions of zonal mean CDOM presented.

2.4. Model development

As a first step, $\alpha_{CDOM}(443, m^{-1})$ was plotted against $S_{320-400}$ (nm$^{-1}$) for all 7611 data points from the Pacific, Atlantic, Indian and Southern Ocean transects within the training dataset (Fig. 2A). $S_{320-400}$ values in our study ranged from 0.0076 to 0.0433 nm$^{-1}$ with a mean of 0.0184 (SD = 0.004 nm$^{-1}$), which fell within range of oceanic S values reported over similar wavelength intervals by Babin et al. (2003) for European coastal waters and the eastern equatorial Atlantic (mean $S_{320-500} = 0.0176$, SD = 0.002 nm$^{-1}$), by Kitidis et al. (2006) for surface waters of the Atlantic (mean $S_{320-350} = 0.026$, SD = 0.006 nm$^{-1}$), and by Bricaud et al. (2010) for the southeast Pacific (range of $S_{320-500} = 0.007$–0.032 nm$^{-1}$).

A least-squares exponential fit to the scatter plot of $S_{320-400}$ (nm$^{-1}$) versus $\alpha_{CDOM}(443, m^{-1})$ was performed (Fig. 2A, gray dashed line). An added constant within the exponential function was logical as the trend in the field data suggests that there is a minimum threshold that spectral slopes do not fall below, which in turn suggests that the CDOM spectrum shall always have a relatively exponential shape hence non-zero value for $S_{320-400}$ regardless of absorption magnitude. The resulting equation from this original regression is:

$$S_{320-400} = 0.013 + 0.017 \exp^{-75.18 \alpha_{CDOM}(443)}.$$  \hspace{1cm} (2)

We also tested the use of a hyperbolic curve to describe $S_{320-400}$ as a function of $\alpha_{CDOM}(443)$ (Stedmon & Markager, 2001; Twardowski et al., 2004), but found that an exponential function provided a better fit to the data. Furthermore, a hyperbolic function implies that at very low absorption, spectral slopes approaches infinity, which is implausible as there is also likely a natural maximum threshold for $S_{320-400}$ in natural waters.

The original fit between $S_{320-400}$ (nm$^{-1}$) and $\alpha_{CDOM}(443, m^{-1})$ is good ($r^2 = .73$, p-value = 0, n = 7611) (see Fig. 2A inset), which indicates the potential utility of the functional form for predicting spectral slope values from remotely sensed $\alpha_{CDOM}(443)$, which in turn could be used to extrapolate CDOM absorption to the UV using the standard Eq. (1). However, this line of reasoning has the implicit assumption that at 443 nm, CDOM is CDOM. Previous experiments on North Atlantic samples have demonstrated that the particulate detrital component of CDOM absorption does make a 5–20% contribution to absorption at 443 nm, which the satellite cannot distinguish from the CDOM component of the retrieved signal (Nelson et al., 1998; Siegel et al., 2002). This indicated a need for model tuning since the desired variable for estimations of CDOM-driven photochemical rate processes is absorption strictly due to CDOM (Reader & Miller, 2011).
To evaluate the initial model, surface field data (<7 m) within the training dataset (Table 1) were matched with 8-day mean GSM output of \(a_{CDOM}(443)\) from the 3 × 3 pixel neighborhood of the sampling station, resulting in 288 surface matchups. Mean \(a_{CDOM}(443)\) values computed from less than 7 pixels of satellite data were rejected to ensure quality. 10 surface samples from the Orinoco River outflow and North Atlantic continental shelf along the A22 transect represented terrestrially influenced, and thus more optically complex (Case II) waters (Blough et al., 1993; Del Vecchio & Subramaniam, 2004; Nelson et al., 2007). These 10 samples were excluded from the analysis because their \(a_{CDOM}(443)\) levels (~0.72 m\(^{-1}\)) were significant outliers as the GSM is optimized for Case I conditions. These criteria brought down the total number of validation points from the training dataset to 127. A preliminary linear regression of observed versus predicted \(a_{CDOM}(680)\) over 325–412 nm (using the original fit (Eq. (2)) in conjunction with Eq. (1) to estimate CDOM absorption) revealed offsets (y-intercepts) from the regression that were in the range of 0.002 to 0.066 m\(^{-1}\) (not shown). These relatively high offsets were further indication that a better model supplying a CDOM-to-CDM approximation was needed.

To improve on our approach, the original model coefficients from Eq. (2) provided a basis for model optimization, wherein they were used as initial guesses in a non-linear minimization scheme to solve for parameters \(\sigma, A, B, C,\) and \(D\) using a model of the form:

\[
\sigma = B + C \exp(\frac{a_{CDOM}(443)}{0.005})
\]

(3)

\[
a_{CDOM}(443) = A a_{CDOM}(443) \exp(\sigma(\lambda - 443)).
\]

(4)

The derived parameter, \(\sigma\) (nm\(^{-1}\)), is exponentially related to \(a_{CDOM}(443)\) (Fig. 2B) by a similar function as \(S_{320–400}\) is related to \(a_{CDOM}(443)\) within the field data (Fig. 2A). Parameters \(A, B, C,\) and \(D\) are constants independent of wavelength that were fit using a cost function that minimized the square and standard deviation of the difference between measured and modeled \(a_{CDOM}(680)\) over the wavelengths 325–412 nm. Eqs. (5) and (6) display the solutions for the model constants. The fidelity of the fits with the training data set was markedly improved by this approach (Table 3; Fig. 3). The resultant model (paired equations) predicts \(\sigma\) from the input of \(a_{CDOM}(443)\), which is then used for estimating \(a_{CDOM}(\lambda)\) over the wavelength range 325–412 nm:

\[
\sigma = 0.02 + 0.01 \exp(-0.66 \times a_{CDOM}(443))
\]

(5)

\[
a_{CDOM}(\lambda) = 0.46 a_{CDOM}(443) \exp(-\sigma(\lambda - 443)).
\]

(6)

To ascertain the exact meaning of this parameter we performed an additional optimization by using a cost function that minimized the square and standard deviation of the difference between measured and modeled \(a_{CDOM}(680)\) over the wavelengths 325–412 nm (using the original fit (Eq. (2)) in conjunction with Eq. (1) to estimate CDOM absorption) revealed offsets (y-intercepts) from the regression that were in the range of 0.002 to 0.066 m\(^{-1}\) (not shown). These relatively high offsets were further indication that a better model supplying a CDOM-to-CDM approximation was needed.

To improve on our approach, the original model coefficients from Eq. (2) provided a basis for model optimization, wherein they were used as initial guesses in a non-linear minimization scheme to solve for parameters \(\sigma, A, B, C,\) and \(D\) using a model of the form:

\[
\sigma = B + C \exp(\frac{a_{CDOM}(443)}{0.005})
\]

(3)

\[
a_{CDOM}(443) = A a_{CDOM}(443) \exp(\sigma(\lambda - 443)).
\]

(4)

The derived parameter, \(\sigma\) (nm\(^{-1}\)), is exponentially related to \(a_{CDOM}(443)\) (Fig. 2B) by a similar function as \(S_{320–400}\) is related to \(a_{CDOM}(443)\) within the field data (Fig. 2A). Parameters \(A, B, C,\) and \(D\) are constants independent of wavelength that were fit using a cost function that minimized the square and standard deviation of the difference between measured and modeled \(a_{CDOM}(680)\) over the wavelengths 325–412 nm. Eqs. (5) and (6) display the solutions for the model constants. The fidelity of the fits with the training data set was markedly improved by this approach (Table 3; Fig. 3). The resultant model (paired equations) predicts \(\sigma\) from the input of \(a_{CDOM}(443)\), which is then used for estimating \(a_{CDOM}(\lambda)\) over the wavelength range 325–412 nm:

\[
\sigma = 0.02 + 0.01 \exp(-0.66 \times a_{CDOM}(443))
\]

(5)

\[
a_{CDOM}(\lambda) = 0.46 a_{CDOM}(443) \exp(-\sigma(\lambda - 443)).
\]

(6)

The \(\sigma\) parameter characterizes the spectral relationship between CDM and CDOM absorption in the model. Though it assumes a similar function as the spectral slope parameter of CDOM in relation to the absorption coefficient (Fig. 2B), it should not be interpreted as a spectral slope of CDM or CDOM, particularly given that values of \(\sigma\) and \(S_{320–400}\) in our datasets are not highly correlated (\(r^2 = 0.24\); \(p < 0.001\); \(n = 127\)). The unitless \(A\) parameter serves as a CDM-to-CDM scaling factor, although it is difficult to ascertain the exact meaning of this parameter as it works in concert with \(\sigma\) to compute \(a_{CDOM}(\lambda)\) from \(a_{CDOM}(443)\) over the spectral range of model application. The A value of 0.46 (Eq. (6)) should not be interpreted to imply, for example, that nearly half of CDM is detritus since previous reports indicate that less than 20% of the CDM signal of open ocean samples is detrital absorption (Nelson et al., 1998; Swan et al., 2009).

3. Results and discussion

3.1. Model validation

The model was optimized to predict CDOM absorption coefficients in excellent agreement with measured surface values from the global CLIVAR training dataset (Table 3; Fig. 3). Regression slopes are near unity, y-intercepts fall below the largest inherent uncertainty (0.005 m\(^{-1}\)) in the CDOM absorption coefficient for 325–412 nm as measured using UltraPath™, and the correlation is very strong between the measured and modeled data, with \(r^2 = 0.85, 0.85, 0.81,\) and 0.71 (\(p\)-value < 0.05, \(n = 127\)) for CDOM absorption at 325, 340, 380, and 412 nm, respectively.

The independent CLIVAR validation data subset consisted of 28 surface ocean samples. The matchups between measured and modeled \(a_{CDOM}(\lambda)\) using this independent subset are good (\(r^2 = 0.65–0.77\); \(p\)-value < 0.05; Table 3; Fig. 4); however, slopes of the regressions at the UV wavelengths approach values of 1.6 (Table 3), thus the model under predicts the observed CLIVAR data in the UV (Fig. 4). Overall, the percent relative error from validation with the CLIVAR subset is 21%. Two data points (seen in Fig. 4 as measured \(a_{CDOM}(325)\) values above 0.18 m\(^{-1}\)) significantly impact the model predictions. If the two points are excluded from the analysis, regression slopes (1.26–1.40) and offsets (−0.009–0.003) are reduced at each wavelength (regression not shown). Since the model is originally based on the mean relationship between the spectral slope parameter and the CDOM absorption coefficient, this result roughly suggests that these two samples have a higher than expected spectral slope value given their CDOM absorption values. The samples originated from the San Diego near-coastal zone sampled along P2 where it is possible that allochthonous sources of organic material may have contributed to the CDOM pool, resulting in spectral properties different than the rest of the transect samples; however, at present we do not have supporting data to confirm this.

Similar observations were made when evaluating the model against data from the South Pacific. The BIOSOPE Project dataset consisted of 29 points for the independent validation (Table 3; Fig. 4). The regression of observed versus modeled BIOSOPE data yields high \(r^2\) values ranging from 0.77 to 0.85 (\(p < 0.05\)); however, regression slopes fall below the 1:1 line (Fig. 4). Overall, percent relative error in the BIOSOPE predictions at each wavelength averages 31%. The over prediction of the BIOSOPE data appears to be driven by 8 data points, also identifiable by \(a_{CDOM}(325)\) values above 0.18 m\(^{-1}\) (Fig. 4a), as near 1:1 matchups are observed (1.06–1.09) and offsets significantly reduced (0.001–0.003) if the points are excluded from the regression (not shown). These 8 samples originated from longitudes east of 286° within the eutrophic coastal Chilean upwelling region where a strong surface salinity gradient indicates the potential for riverine influences on the CDOM composition of the samples, and a range of low spectral slope values for CDOM has been previously observed (Bricaud et al., 2010). With respect to the global mean trend between \(a_{CDOM}\) and spectral slope (Fig. 2), the over prediction roughly suggests that spectral slope is
lower in these samples relative to expected given their absorption values. As with the CLIVAR data subset however, we do not have supporting information to validate whether certain samples from BIOSOPE had traces of terrestrial organics from near-continental waters. Since the CDOM absorption values (0.18–0.3 m$^{-1}$ at 325 nm) for all validation points from CLIVAR and BIOSOPE are well within the observed range for the global open ocean, and given that spectral slope varies by at least an order of magnitude in the open ocean (Fig. 2A; see also Bricaud et al., 2010; Swan et al., 2009), we have no basis on which to flag any of the points as outliers. Future application of the model may benefit from a priori knowledge or data identifying regions of possible terrestrial organic influence on the local CDOM pool. Certainly where Case II waters mix with the open ocean, other algorithms (e.g. Fichot & Miller, 2010; Fichot et al., 2008) may be more appropriately applied.

![Fig. 3. Dashed line indicates regression of observed CDOM absorption with modeled CDOM absorption for wavelengths 325, 340, 380 and 412 nm using the CLIVAR training dataset. Solid line indicates the 1:1 line.](image)

![Fig. 4. Regression of observed CDOM absorption with modeled CDOM absorption for wavelengths 325, 340, 380 and 412 nm using the CLIVAR validation (dashed line) and BIOSOPE validation (dotted line) datasets, respectively. Solid line indicates the 1:1 line.](image)
Additional validation with data from the Bermuda Atlantic Time-series Study (BATS; 1994–2006) highlighted a few considerations when implementing the model on the regional scale. Most surface \(a_{\text{CDOM}}(325)\) values from BATS fell within the range of 0.04–0.17 m\(^{-1}\), as compared with the training dataset in which \(a_{\text{CDOM}}(325)\) ranges three orders of magnitude up to 0.4 m\(^{-1}\). Model predictions of BATS measurements were poor \((r^2 = 0.04, p = 0.17, n = 49; \text{data not shown})\), which we attribute to the lack of dynamic range in the BATS dataset. Our model is based on a training dataset with a broad range shown), which we attribute to the lack of dynamic range in the BATS environments on seasonal time scales (Steinberg et al., 2004), and may possess CDOM of a different spectral quality and slope value due to the different phytoplankton and microbial alteration produces lower slope values \((\text{Del Vecchio \\& Blough, 2002; Helms et al., 2008; Nelson et al., 2004; Váhatalo \\& Wetzel, 2004; Whitehead et al., 2000})\), recent reports suggest secondary biological controls on the spectral quality of CDOM. For example, Kitidis et al. (2006) observed distinct end-member values of spectral slope in phytoplankton-produced CDOM within the deep chlorophyll maximum and in microbially produced CDOM in surface waters of the Atlantic, suggesting that slope may be useful as a quality index of autochthonous CDOM, even in the open ocean where absorption is relatively low. Nelson et al. (2004) demonstrated that CDOM produced by microbial cultures amended with fresh algal exudate had a relatively higher spectral slope \((S_{320-400} = 0.025 \text{ m}^{-1})\) than CDOM in the upper mixed layer of the Sargasso Sea \((0.019 \text{ m}^{-1})\), suggesting that the source of microbial substrate in the open-ocean also impacts CDOM spectral properties.

An inherent assumption when modeling the relationship between spectral slope and the CDOM absorption coefficient for the global ocean \((\text{Fig. 2A})\) is that the \(S_{320-400}\) value of ‘new’ (non-light exposed) microbially produced CDOM within the ocean interior inhabits a relatively constant and narrow range \((\text{Whitehead \\& De Mora, 2000; Nelson et al., 2004})\), and that light exposure and physical processes act on this CDOM pool altering spectral slope in a consistent manner \((\text{Nelson et al., 2007; Swan et al., 2009})\). However, in the high range of \(\text{in situ} \ a_{\text{CDOM}}(443)\) values \((>0.2 \text{ m}^{-1})\) presented herein \((\text{Fig. 2A})\), we observe roughly two trends in the \(S_{320-400}\) versus CDOM relationship. One trend is comprised largely of data from the North Pacific, which have higher \(S_{320-400}\) values for a given \(a_{\text{CDOM}}(443)\) value than data originating from the equatorial regions of the Pacific and Indian Oceans that comprise the other trend \((\text{lower} \ S_{320-400}\) values for the same \(a_{\text{CDOM}}(443)\) value). The North Pacific has been characterized by high seasonality in primary productivity and POC fluxes \((\text{Buesseler et al., 2007})\), and may possess CDOM of a different spectral quality and slope value due to the different phytoplankton and microbial speciation than CDOM found in the picoplankton-dominated equatorial regimes \((\text{Li, 2002; Szeto et al., 2011; Zhang et al., 2007})\). Such taxa-specific signals in the CDOM absorption spectrum may not necessarily be diluted from the background signal in open ocean environments on seasonal time scales \((\text{Steinberg et al., 2004})\), and thus may explain the subtle latitudinal differences that we observe in the global \(S_{320-400}\) versus \(a_{\text{CDOM}}(443)\) relationship \((\text{Fig. 2A})\). The biological origin and spectral properties of the oceanic CDOM pool are an interesting topic for future investigation, as it remains to be determined if the spectral signature of CDOM is capable of diagnosing smaller scale processes and ecosystem-level dynamics affecting organic matter quality in the global ocean.

3.3. Model application

We applied the model to SeaWiFS-based \((1997–2007)\) climatological means of \(a_{\text{CDOM}}(443)\) from GSM to produce maps of the surface distribution of modeled \(a_{\text{CDOM}}(325)\) \((\text{Fig. 5A})\) and the zonal mean annual cycle of \(a_{\text{CDOM}}(325)\) \((\text{Fig. 5B})\). These distributions provide an indication of the magnitude of modeled CDOM absorption in the UV across the global ocean, and an example of model implementation that is readily achieved. Use of the model with shorter time series \((e.g., \text{monthly})\) \(a_{\text{CDOM}}(443)\) data may be readily applied in rate calculations of photochemical processes. Coupling temporal modeled \(a_{\text{CDOM}}(\lambda)\) values in the UV with quantum yield estimates and UV irradiance data \((\text{such as from the Total Ozone Mapping Spectrometer})\) enables the potential estimation of CDOM photobleaching rates on basin to global scales \((\text{Del Vecchio \\& Blough, 2004; Swan et al., 2012})\). Rates for various photoproduction and destruction processes, which are sensitized by CDOM and impact elemental cycling in the surface ocean, may also be quantified in this manner and potentially integrated into biogeochemical models characterizing mixed layer dynamics \((\text{Fichot \\& Miller, 2010; Millet et al., 2010; Mopper \\& Kieber, 2002; Toole et al., 2006; Zafiriou et al., 2008})\) \((\text{Fig. 5})\).

4. Conclusion

We presented a model for estimating UV absorption by CDOM in the surface ocean that is optimized for synoptic scale use with global ocean color data. The model was developed for Case 1 waters from a simple construct and large training dataset consisting of thousands of systematically measured CDOM absorption spectra sampled from a multitude of trophic regimes and optical provinces in the world ocean. The global applicability of the model naturally raises its overall uncertainty as validation with regional field datasets reveals. In waters where terrestrial proximity or other known second-order processes may influence the spectral character of organic matter, confidence in the model predictions likely decreases. Even in waters remote from terrestrial influence, there exists the potential for isolated cases \((e.g., \text{strong off-shore bloom conditions})\) where other algorithms such as the SeaUV/UVf may be more applicable for predicting CDOM absorption due to their basis on field data from optically complex waters \((\text{Fichot et al., 2008; Reader \\& Miller, 2011})\).

Overall the model performs well during independent validations when factoring in the potential sampling of Case II waters, as well as the uncertainties inherent in satellite-retrieved reflectances and methodological inconsistencies in CDOM filtration and spectroscopic analysis. There is a clear need for additional validation and potential model improvement as more field data becomes available. The model is based on CDOM absorption values from all depths of the ocean; therefore, it encapsulates the dynamic range in CDOM spectral properties resulting from physical and biogeochemical processes in the ocean interior, as well as the mixing and light-driven dynamics in the surface ocean. While radiometric profile data is equally necessary for ground-truthing, our approach highlights the value of water sample data in today’s algorithm development, particularly as more sensitive and rapid detection methods for CDOM absorption continuously evolve \((\text{Flage et al., 2009})\). Furthermore, the datasets and approach detailed herein offer a basis for establishing regionally constrained models of the same kind.

The impact of CDOM on upper ocean processes and ocean color remote sensing is now widely acknowledged, making new information about the sinks, sources and cycling pathways of CDOM a valuable target for future studies. The model presented here will provide
additional service to those exploring the role of CDOM in the surface ocean, its effects on radiative transfer, and the impact of CDOM on biogeochemical cycles via UV photochemistry. The model will be particularly useful in estimating photochemical fluxes across wide swaths of the surface ocean, and could potentially aid in better representing the contribution of CDOM to total absorption in future incarnations of marine bio-optical models.

Acknowledgements

This work was funded through grants from NSF Chemical Oceanography and NASA Biology and Biogeochemistry Programs to D. Siegel, N. Nelson and C. Carlson, and by the NASA ESS Fellowship Program for C. Swan. The authors are grateful for critical feedback from three anonymous referees. The authors acknowledge the organizing committees and chief scientists of the U.S. CO2/CLIVAR Repeat Hydrography Program, as well as the captains and crew of the participating research vessels: Revelle, Melville, Knorr, Brown, and Thompson. Dave Menzies, Elisa Haledow, Stuart Goldberg, Dennis Hansell, Jon Klabam, Meredith Meyers, Meg Murphy, Alexey Mishonov, Stacy Brown, Wenhai Chen, Stan Hooker and Mary Russ are also gratefully acknowledged for their efforts in data collection during the CLIVAR campaigns. Hervé Claustre provided access to the BIOSOPE database. Processed satellite data were obtained from the Ocean Color MEAsURES project at UCSC (http://wiki.iceose.ucsd.edu/Measures/Products).

References


