Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Spectral range within global $a_{CDOM}(440)$ algorithms for oceanic, coastal, and inland waters with application to airborne measurements

Henry F. Houskeeper [a,b,*], Stanford B. Hooker [c], Raphael M. Kudela [b]

[a] Department of Geography, University of California, Los Angeles 315 Portola Plaza, Los Angeles, CA 90095, USA
[b] Ocean Sciences Department, University of California, Santa Cruz 1156 High Street, Santa Cruz, CA 95064, USA
[c] NASA Goddard Space Flight Center, Ocean Ecology Laboratory/Code 616.2, Greenbelt, MD 20771, USA

ARTICLE INFO

Keywords:
CDOM
Airborne
Remote sensing
Open Ocean
Coastal zone
Inland waters
Spectral range
Radiometry

ABSTRACT

The optically active component of dissolved organic material in aquatic ecosystems, or colored dissolved organic matter (CDOM), is represented by the coefficient of absorption due to the dissolved aquatic constituents at 440 nm, $a_{CDOM}(440)$. Remote sensing of $a_{CDOM}(440)$ enables characterization of ecosystem processes and aids in retrieval of chlorophyll *a*, a proxy for phytoplankton biomass. Spectrally adjacent band-ratio domains, e.g., blue to green, have previously been applied for remote sensing of $a_{CDOM}(440)$ in coastal and oceanic waters with similar results compared to more complex semi-analytical algorithms. Estimation of $a_{CDOM}(440)$ from ratios of the most spectrally separated ocean color wavebands (end members), e.g., ultraviolet (UV) to near-infrared (NIR), termed end-member analysis (EMA), has previously been shown to increase the accuracy of global $a_{CDOM}(440)$ retrievals from in-water observations of diffuse attenuation and to enable a unified algorithmic perspective without requiring regional adjustment of internal bio-optical parameters. EMA of above-water observations is evaluated herein, with a focus on coastal and inland waters in which increasing optical complexity and likelihood of bottom reflectance challenge the oceanic algorithms developed for deep and optically simple (case-1) waters. Analysis herein of three independent, in situ, bio-optical datasets indicates significant correlation between $a_{CDOM}(440)$ and end-member band ratios (next-generation 320 and 780 nm or legacy 412 and 670 nm ratios) with a coefficient of determination, R^2, of 0.87 (log-scale) or higher based on a dataset spanning the dynamic range of global, conservative water bodies. For applicable wavelengths, EMA algorithms are shown to agree with case-1 relationships and to produce consistent log-scale uncertainties across more than three orders of magnitude in $a_{CDOM}(440)$ values (0.001–2.305 m^{-1}). EMA using UV and NIR wavelengths (320 and 780 nm) is applied to low-altitude airborne observations and satisfies 25% uncertainty based on unbiased percent differences (UPDs) within each of three dissimilar match-up sites ranging in $a_{CDOM}(440)$ from 0.02–0.57 m^{-1}. Results demonstrate that EMA is a useful and robust approach for the remote sensing of $a_{CDOM}(440)$ in coastal and inland waters, which are generally shallower, contain more optically complex environments, and span a greater range in $a_{CDOM}(440)$ than oceanic waters.

1. Introduction

Aquatic dissolved organic matter (DOM) modifies the vertical deposition of solar heat to surface waters, provides carbon storage on ephemeral to geological timescales, and fuels the microbial pump that recycles nutrients to maintain global phytoplankton stocks (Pegau, 2002; Chang and Dickey, 2004; Jiao et al., 2010; Hansell and Carlson, 2014). Oceanic sources of DOM include microbial degradation of phytoplankton (Rochelle-Newall and Fisher, 2002; Nelson et al., 2004) and phytoplankton secretions for photo-protection (Castillo et al., 2010), e.g., mycosporine-like amino acids (MAAs) as described by Morrison and Nelson (2004) and Jessup et al. (2009). Weathering of terrestrial material injects DOM into aquatic ecosystems, and rivers transport DOM from inland to estuarine and coastal waters (Del Castillo et al., 2000). The optically active component of DOM, termed colored dissolved organic matter (CDOM), has been used as a DOM tracer (Spencer et al., 2012; Vantrepotte et al., 2015) and enables remote sensing (space or airborne) of ecosystem processes, e.g., for analysis of large run-off events (Kudela and Chavez, 2004) or for measurement of oceanic circulation rates (Nelson et al., 2010).

* Corresponding author.
E-mail addresses: hhouskee@ucsc.edu (H.F. Houskeeper), stanford.b.hooker@nasa.gov (S.B. Hooker), kudela@ucsc.edu (R.M. Kudela).

https://doi.org/10.1016/j.rse.2020.112155
Received 11 February 2020; Received in revised form 12 October 2020; Accepted 20 October 2020

Decadal analysis of satellite imagery has revealed global fluctuations in ocean color observations related to basin-scale oscillations in oceanic CDOM content (Nelson and Siegel, 2013). Improving remote estimation of CDOM has been a key objective for aquatic remote sensing because the accuracy of remote estimation of chlorophyll *a* concentration (Chl *a*), which is a proxy for phytoplankton biomass, degrades in waters with elevated CDOM loads (Carder et al., 1989; Gregg and Casey, 2004; Siegel et al., 2005; Brown et al., 2008; Sauer et al., 2012; Siegel et al., 2013). Traditional Chl *a* remote-sensing algorithms, i.e., blue-green band ratios (O'Reilly et al., 1998), target high Chl *a*-specific absorption of blue light, a spectral characteristic that also describes CDOM absorption, in which spectral dependencies are modeled (Bricaud et al., 1981) as follows:

$$a_{\mathrm{CDOM}}(\lambda) \;=\; a_{\mathrm{CDOM}}(440)\; e^{-S(\lambda-440)}\,, \tag{1}$$

where λ is wavelength, $a_{\mathrm{CDOM}}(440)$ is the absorption coefficient of CDOM at 440 nm, and S is the spectral slope of CDOM absorption. Natural variability in S corresponds to differences in lability (humic versus fulvic) or source (marine versus riverine) of DOM (Carder et al., 1989). S values ranging from 0.015–0.03 nm^{-1} are common among oceanic waters (Nelson and Siegel, 2013), although discrepancies exist in the literature due, in part, to differences in wavelength intervals and fitting (Twardowski et al., 2004). An expanded range in S of 0.0095–0.0410 nm^{-1} for oceanic, coastal, and inland waters was reported in Hooker et al. (2020).

Remote estimation of $a_{\mathrm{CDM}}(440)$, or the combined absorption of CDOM and detritus at 440 nm, has been most successful in the optically simple oligotrophic ocean using inversion or semi-analytical methods, e.g., the Garver-Siegel-Maritorena (GSM) algorithm (Maritorena et al., 2002) or the Quasi-Analytical Algorithm (QAA) by Lee et al. (2002). Smaller-scale optimizations of existing semi- or quasi-analytical algorithms have improved remote sensing of regional coastal ecosystems (Aurin and Dierssen, 2012; Joshi and D'Sa, 2018), and improvements to the GSM and QAA by Matsuoka et al. (2013), Dong et al. (2013), Wang et al., 2017, and other researchers have enabled the separation of the dissolved and detrital components, i.e., $a_{\mathrm{CDOM}}(440)$ from $a_{\mathrm{CDM}}(440)$. As a computationally simpler alternative to semi- or quasi-analytical algorithms, band-ratio approaches using neighboring wavelengths (e.g., 412 and 443 nm or 443 and 555 nm) have produced similar results as algorithms with greater spectral and computational requirements in oceanic environments (Morel and Gentili, 2009), and band-ratio approaches have also been applied to coastal (e.g., Kahru and Mitchell, 2001; Mannino et al., 2014) and lacustrine (e.g., Kutser et al., 2005; Brezonik et al., 2015) environments.

The performance of band-ratio algorithms is anticipated to improve by using more spectrally separated wavebands because the dynamic range of ocean color measurements as a function of $a_{\mathrm{CDOM}}(440)$ increases within the ultraviolet (UV) and near-infrared (NIR) domains, i.e., the spectral end members, as shown by Hooker et al. (2013). The algorithmic approach based on greater spectral separation within waveband ratios, hereafter referred to as end-member analysis (EMA), also has the advantage of minimizing the effects of photosynthetic pigments, which have strong spectral dependencies (including nonlinearities with Chl *a*) across the visible (VIS) domain (Bricaud et al., 1995). Using VIS data products, Mannino et al. (2014) also found that increasing the spectral separation of waveband ratios improved the performance of $a_{\mathrm{CDOM}}(440)$ band-ratio satellite algorithms, and Zhu et al. (2014) used in situ above-water instruments to show that including the longest available red wavelength in $a_{\mathrm{CDOM}}(440)$ algorithms improved estimation of $a_{\mathrm{CDOM}}(440)$ within freshwater environments.

For in situ measurements across a global range of water masses (e.g., oceanic, estuarine, lacustrine, and riverine) and using a broad spectral range (UV-NIR) of in-water data products, Hooker et al. (2020) found that EMA based on diffuse attenuation coefficients for spectral downward irradiance, $K_d(\lambda)$, enabled estimation of $a_{\mathrm{CDOM}}(440)$ within a 2% root mean square difference (RMSD), with consistent $\log_{10}$-scale accuracy for $a_{\mathrm{CDOM}}(440)$ values spanning over three decades in range. The efficacy of a similarly expansive (UV–NIR) EMA algorithmic approach derived from above-water (as opposed to in-water) data products, i.e., the normalized water-leaving radiance, $[L_W(\lambda)]_{\mathrm{N}}$, is evaluated herein using bio-optical theory, three independent in situ datasets, and airborne matchups to in situ water samples.

2. Materials and methods

2.1. Derivation of synthetic case-1 bio-optical datasets

The radiance distribution and the optically relevant constituents of optically simple waters may be estimated to first order based solely on Chl *a* content because the optically active constituents (including algal and non-algal particles, CDOM, etc.) covary; a condition defined in optical oceanography as case-1 (Morel and Prieur, 1977), with all other conditions described herein as case-2. Under this perspective, biogeochemical relationships between $a_{\mathrm{CDOM}}(440)$ and Chl *a*, for example as provided by Prieur and Sathyendranath (1981), enable derivation of a case-1 model relating $a_{\mathrm{CDOM}}(440)$ to Chl *a* and, thus, to the bio-optical constituents and the theoretical light field of a water mass. A synthetic, case-1 $a_{\mathrm{CDOM}}(440)$ dataset was derived using the parameterizations of Morel (2009) as follows:

$$a_{\mathrm{CDOM}}(440) \;=\; 0.032\,[\mathrm{Chl}\ a]^{0.63}\,. \tag{2}$$

Although the case-1 scenario is not upheld in optically complex waters, algorithmic agreement with case-1 relationships may be considered as a baseline for evaluating whether an algorithm is internally consistent with the predefined bio-optical relationships of case-1 waters. The parameterizations used in this study provided a general framework for considering the basis of an EMA algorithm within the parameterized ranges in Chl *a* and were not intended as a thorough review or recommendation of specific bio-optical models.

Seawater absorption coefficients were obtained from Smith and Baker (1981) for red and NIR wavelengths and from Morel et al. (2007) for blue and UV wavelengths. Backscattering contributions of seawater were obtained following Twardowski et al. (2007). The absorption by particles (including algal and non-algal) were parameterized with spectral dependencies according to Bricaud et al. (1998), with coefficients selected from the HydroLight Mid-range UV option of the *New Case-1 Model* (Mobley and Sundman, 2008). The scattering contributions of particles were parameterized according to Huot et al. (2008). In keeping with these parameterizations, the derivations for two $K_d(\lambda)$ and one $[L_W(\lambda)]_{\mathrm{N}}$ synthetic case-1 datasets are briefly described below.

2.2. Biogeochemical parameterizations of $K_d(\lambda)$

$K_d(\lambda)$ may be treated as an inherent optical property (IOP), wherein it is dependent on the constituents of a water mass and not on the ambient light field, if $K_d(\lambda)$ is normalized by the downwelling distribution function, $D_o(\lambda)$, which relates the subsurface scalar downwelling irradiance to the subsurface downwelling irradiance (Gordon, 1989). The correction allows for the partitioning of $K_d(\lambda)$ among optically relevant constituents (e.g., water, particles, and CDOM). Least-squares fitting relates $K_d(\lambda)$ to absorption, $a(\lambda)$, and back-scattering, $b_b(\lambda)$, coefficients (Gordon, 1989), such that:

$$\frac{K_d(\lambda)}{D_o(\lambda)} \;=\; 1.0395\left[\sum_{i=1}^{m} a(\lambda,i) \;+\; \sum_{i=1}^{m} b_b\left(\lambda,i\right)\right], \tag{3}$$

where $a(\lambda)$ and $b_b(\lambda)$ are partitioned into i of m components (e.g., water, particles, and CDOM). By approximating the spectral fraction of downward irradiance contained within the solar beam at relevant wavelengths following Pan and Zimmerman (2010), the uncertainty relevant to an EMA algorithm due to spectral dependencies in D_o, i.e.,

variability in the ratios $D_o(670)/D_o(412)$ and $D_o(780)/D_o(320)$, was less than other sources of uncertainty in the case-1 parameterizations, e.g., natural variability in S. The parameterizations of $a(\lambda,i)$ and $b_b(\lambda,i)$ described in Sect. 2.1, therefore, enabled a case-1 approximation based on Eq. (3) relating $K_d(\lambda)$ band ratios to $a_{\mathrm{CDOM}}(440)$. This synthetic dataset is hereafter referred to as GSyn.

$K_d(\lambda)$ has also been parameterized as a power-law function of Chl a within the VIS domain for waters spanning Chl a of 0.035–5.45 mg m^{-3} (Morel and Maritorena, 2001) and of 0.02–30 mg m^{-3} (Morel, 1988), as:

$$K_d(\lambda) = K_w(\lambda) + \chi(\lambda)[\mathrm{Chl}\ a]^{e(\lambda)}, \quad (4)$$

where $K_w(\lambda)$ is the diffuse attenuation coefficient of pure water, and $\chi(\lambda)$ and $e(\lambda)$ are fitting coefficients. Despite a sparsity of available data, Morel and Antoine (1994) were able to extend the $\chi(\lambda)$ and $e(\lambda)$ coefficients to NIR wavelengths by assuming no biological contributions above 750 nm, i.e., $\chi(\lambda)$ decreases linearly between 700 and 750 nm, and to UV wavelengths by assuming a chlorophyll-specific $K_d(\lambda)$ that resembles the specific absorption of non-algal particles (in clear waters) or of a pure algal suspension (in eutrophic waters). A second synthetic dataset of $K_d(\lambda)$ ratios was derived from Eq. (4) following Morel and Maritorena (2001) and Morel and Antoine (1994). This synthetic dataset is hereafter referred to as MSyn.

2.3. Biogeochemical parameterizations of $[L_W(\lambda)]_N$

The principal above-water data product used herein is $[L_W(\lambda)]_N$, which was derived by normalizing the water-leaving radiance, $L_W(\lambda)$, for influences from the atmosphere and Sun, including geometry (e.g., the Earth-Sun distance). The case-1 model was used to derive a synthetic $[L_W(\lambda)]_N$ dataset based on the approach of Gordon et al. (1988), with the refractive index of water obtained from Bashkatov and Genina (2003) and with the other relevant terms and coefficients consistent with those used in Sect. 2.2. The case-1 $[L_W(\lambda)]_N$ model was compared with a HydroLight model (Liu et al., 1999) using the relative percent difference (RPD), with HydroLight as the reference in the difference calculations. The mean RPD value as a function of the spectral domains and total Chl a spanning 0.02–20 mg m^{-3} was 1.1%, and the overall absolute percent difference (APD) was 13.9%. The comparison of spectral shape was significantly better, with values of the Pearson correlation coefficient, ρ, for eutrophic, mesotrophic, and oligotrophic waters of 0.991, 0.966, and 0.987, respectively. The synthetic $[L_W(\lambda)]_N$ dataset was derived using the Processing of Radiometric Observations of Seawater using Information Technologies (PROSIT) software (Hooker et al., 2018b), and is hereafter referred to as PSyn.

2.4. Description of above- and in-water data products

Ratios of $[L_W(\lambda_1)]_N/[L_W(\lambda_2)]_N$, hereafter $\Lambda^{\lambda_1}_{\lambda_2}$, were selected for comparison based on the following: a) recent advances (Hooker et al., 2018a, 2018b, 2018c) in remote and in situ capabilities (Λ^{320}_{780}); b) legacy remote and in situ capabilities (Λ^{412}_{670}); c) proposed band-ratio approaches (Λ^{443}_{555}); and d) legacy capabilities for sensors with configurations not consistent with the common ocean color wavebands (Λ^{465}_{625}). Accurate remote sensing of data products corresponding to all $\Lambda^{\lambda_1}_{\lambda_2}$ ratios evaluated herein was possible using existing capabilities for low-altitude airborne remote sensing, including for the UV and NIR domains, recently documented in Hooker et al. (2018a), Kudela et al. (2019), and Guild et al. (2020).

The primary objective of this study is to evaluate algorithms for airborne applications, although additional waveband pairs corresponding to configurations of various ocean color platforms are also shown in Table A.3 of Appendix A. The applicable $\Lambda^{\lambda_1}_{\lambda_2}$ pairs listed above were obtained from the case-1 synthetic datasets as well as from three in situ datasets containing coincident observations of $a_{\mathrm{CDOM}}(440)$ and $[L_W(\lambda)]_N$, with the latter spanning up to 320–780 nm, which are described below.

An in situ, oceanic $[L_W(\lambda)]_N$ dataset was created by reprocessing data obtained with legacy instruments, hereafter referred to as the OCEAN dataset, wherein the number of observations, N, was 113. The reprocessing was performed for 16 wavebands spanning 320–780 nm with 10 nm central bandwidths. The OCEAN dataset was obtained from both rocket-shaped profiler designs as well as the prototype Compact-Optical Profiling System (C-OPS), which used a novel kite-shaped backplane (Morrow et al., 2010). Consequently, vertical sampling resolution (VSR) was usually not less than 1 cm, and the first upwelling radiance observations, $L_u(z,\lambda)$, were typically obtained deeper than 0.3 m (the nominal first measurement depth of the C-OPS downward-pointing radiometer). The geographical areas sampled ranged from the Southern Ocean to tropical waters, plus northern mid-latitudes. The majority of the measurements were made in deep oceanic waters (63%) with the remainder in shallower coastal waters (37%); no inland waters were sampled, and the depths of the profiles were to the 1% light level or more. Additional details regarding the sampling and processing parameters of the OCEAN dataset are provided in Appendix B.

A separate, in situ, global $[L_W(\lambda)]_N$ dataset of oceanic, coastal, and inland waters was obtained using an advanced C-OPS design with an average VSR of 6.0 mm (0.9 mm within very shallow or attenuating waters) and with the first $L_u(z,\lambda)$ observations routinely obtained at approximately 0.3 m (Hooker et al., 2020). The C-OPS spectral configuration contained 19 wavebands with 10 nm bandwidths, and was obtained with the Compact-Propulsion Option for Profiling Systems (C-PrOPS) digital thrusters accessory (Hooker et al., 2018a), which enabled sampling of shallow, non-navigable waters, and which mitigated sampling difficulties related to adjacency, shading, and physical perturbations. For the in-water data presented herein, data collection was only initiated when the downward-pointing radiance aperture defined the principal solar plane with no intervening obstacles, i.e., was closest to the Sun. Deployment of the in-water profiler in the appropriate geometry with respect to the solar illumination was simplified by the use of the C-PrOPS thruster accessory, which allows the operator to steer the profiler into the principal plane with the downward-pointing radiance aperture always being closest to the Sun. The global dataset was designed with the intention of enabling algorithm development and validation in optically complex water bodies, and sampling was uniformly spread across oceanic (31.2%), coastal (36.6%), and inland water (32.2%) ecosystems. The uniform distribution achieved is important for unbiased curve fitting of global waters. The diversity of water bodies represented within this dataset is also demonstrated by the range in $a_{\mathrm{CDOM}}(440)$ of 0.001–2.305 m^{-1}, as well as the aforementioned expansive range in S of 0.0095–0.0410 nm^{-1} (Hooker et al., 2020).

Factual observations regarding essential water mass attributes were recorded prior to collection of the global dataset to establish a subjective classification scheme (Hooker et al., 2020). The rationale for prior subjective classification was not to exclude optically complex water bodies, but rather to identify waters with nonconservative perturbations (i.e., modifications not consistent with a linear mixing of parent water masses), which would not be expected to adhere to a global algorithmic approach. Examples of relevant categorical information include the persistence of a severe drought, which cuts off inflow to lakes and rivers and can increase resuspension due to shallowing of the water level, or flooding of lakes and rivers beyond nominal fill levels, which can introduce scoured terrigenous materials. The subjective classification scheme does not signify that these conditions are ignored in the evaluation of algorithm efficacy. Rather, the classifications provide extra information to enable testing of whether particular environments (e.g., a drought-stricken lake) require an alternate or regionalized algorithmic approach. Measurements from the global dataset with no subjective classification assignments (conservative water bodies relevant to a global algorithmic perspective) are hereafter referred to as the GLOBC dataset (N = 612), while observations with nonconservative

classifications are referred to as the GLOBN dataset ($N = 524$) and are considered separately in Section 3.4.

An additional in situ dataset containing 308 $[L_W(\lambda)]_N$ EMA pairs for legacy (VIS) wavelengths (Λ_{670}^{412}, Λ_{555}^{443}, and Λ_{625}^{465}) and using both above- and in-water instruments was obtained from the NASA bio-Optical Marine Algorithm Dataset (NOMAD), using version 2a (https://seabass.gsfc.nasa.gov/wiki/NOMAD), which is part of a long-standing repository of bio-optical measurements produced for ocean optics algorithm modeling and validation (Hooker et al., 1994; Werdell and Bailey, 2002; Bailey, 2005). In-water data products from the NOMAD repository were previously applied to evaluate EMA based on $K_d(\lambda)$ in Hooker et al. (2020), which found that the legacy products confirmed the in-water algorithm after quality control of the dataset, for example by removing coefficients clearer than pure water. The in situ dataset consisting of NOMAD $[L_W(\lambda)]_N$ pairs is hereafter referred to as the NOMAD dataset. A synthetic $K_d(\lambda)$ dataset was also produced from NOMAD $[L_W(\lambda)]_N$ products using Jamet et al. (2012), which estimates in-water products from above-water observations using a neural network approach, and is hereafter referred to as the NNet dataset. The synthetic NNet dataset is discussed in Sect 4.1 to investigate whether remote-sensing methods to predict $a_{CDOM}(440)$ would benefit from working through $K_d(\lambda)$ intermediates rather than directly estimating $a_{CDOM}(440)$ from above-water measurements.

2.5. Quality control and partitioning of datasets

Initial quality control for the GLOBC and GLOBN datasets was based on prior subjective classification (Sect. 2.4). For the OCEAN dataset, in which prior subjective classification information was not available, an objective approach for partitioning $K_d(\lambda)$ spectra based on fuzzy c-means analysis (Hooker et al., 2020) was applied following Moore et al. (2001). Median spectra that satisfied criteria corresponding with conservative water bodies in Hooker et al. (2020) and did not contain $K_d(\lambda)$ below that of natural water, were retained.

Additional quality control was performed for the OCEAN, GLOBC, and GLOBN datasets due to the inherent difficulty of measuring $L_u(z,\lambda)$ from an in-water profiler within extremely attenuating waters, i.e., $K_d(\lambda) > 10\ m^{-1}$. Although the VSR was optimized in the GLOBC and GLOBN datasets by technological improvements designed to increase surface loitering and stabilize aperture planar geometries (Hooker et al., 2018a), the shallowest possible $L_u(z,\lambda)$ measurements were limited by the length of the downward-pointing radiometer. Application of a $K_d(\lambda) < 10\ m^{-1}$ filter was applied to account for increased uncertainty in deriving $[L_W(\lambda)]_N$ data products, e.g., compared to $K_d(\lambda)$, in highly attenuating water masses, which removed roughly 1% of the OCEAN dataset, 10% of the GLOBC dataset, and 30% of the GLOBN dataset. More observations were removed from the GLOBN dataset because the nonconservative water bodies contained a greater range in turbidity. For GLOBC validation comparisons herein (Sect. 3.3), the $K_d(\lambda)$ threshold was relaxed in order to consider the complete dataset range in $a_{CDOM}(440)$. The subset of water bodies that satisfied $K_d(\lambda) < 10\ m^{-1}$ was previously considered in Houskeeper (2020).

2.6. Description of airborne datasets

The Compact-Airborne Environmental Radiometers for Oceanography (C-AERO) instrument suite uses above-water (airborne) radiometers, which typically have 19 wavelengths (nominally with 10 nm bandwidths), spanning the UV–NIR with 16 wavelengths plus the shortwave infrared (SWIR) with three wavelengths (Hooker et al., 2018a). The instruments are used to derive normalized data products by simultaneously measuring the total radiance $L_T(\lambda)$ from the water surface at a specified angle with respect to nadir (typically 40°), the indirect (sky) radiance $L_i(\lambda)$ measured in the same plane and at a complementary zenith angle as $L_T(\lambda)$, and the global solar irradiance $E_s(\lambda)$. The radiance radiometers have a narrow field of view (2.5° full view angle) and are fitted with a shroud to reduce long-wavelength scattering at the aperture. A spectrally-dependent synthetic or predictive dark current method—based on an operational range of environmental and instrument parameters characterized in the laboratory—improved radiometric agreement between C-AERO and in situ measurements (Guild et al., 2020). The $L_W(\lambda)$ term was derived from the $L_T(\lambda)$ observations by filtering out sun glint in the $L_T(\lambda)$ data (the C-AERO instrument suite enabled rapid sampling of 15 or 30 Hz depending on the acquisition date for superior glint discretization and rejection), removing the sky reflection based on a spectral reflectance model depending on the viewing geometry (i.e., pointing angle of the radiometers) and wind speed, and computing normalized forms following published protocols (Hooker et al., 2002, 2004).

The C-AERO instrument suite was flown aboard a Twin Otter aircraft operated by the Center for Interdisciplinary Remotely Piloted Aircraft Studies (CIRPAS) at the lowest safe altitude (LSA), which was as low as 100 ft (30.5 m) in order to exclude atmospheric effects, and measurements were screened for improper pointing relative to the solar azimuth. C-AERO observations were obtained in the following major water bodies: a) San Francisco Bay Delta (SFBD), a tidally mixed estuary; b) Lake Tahoe (LT), an oligotrophic alpine lake; and c) Monterey Bay (MB), a coastal marine sanctuary partially sheltered from upwelling favorable winds by the coastline geometry. All flights coincided with water sampling for $a_{CDOM}(440)$ in which samples were filtered through a Whatman GD/X 0.2 μm syringe filter and then measured with a Cary Varian 50 spectrophotometer with a 10 cm path-length (Hooker et al., 2020). The locations of the C-AERO data products and water sampling sites are shown in Fig. 1.

Signal-to-noise ratios for the C-AERO measurements (including for the UV and NIR domains) were shown to exceed those of other relevant remote platforms within SFBD and LT (Kudela et al., 2019), and coincident in-water C-OPS data products demonstrated radiometric $[L_W(\lambda)]_N$ agreement of $\rho = 0.989$ and 0.999 at LT ($N = 4$) and SFBD ($N = 3$), respectively. Geospatial interpolations (two-dimensional with 50 m resolution) were produced from the C-AERO measurements (approximately 3.4 m surface spot size, i.e., the sensor's footprint at the water surface, at LSA for native resolution) using the natural neighbor method (Sibson, 1981) without extrapolation beyond the boundary of the airborne measurements and with the closest nearshore retrievals masked.

2.7. Derivation of algorithm fits for $[L_W(\lambda)]_N$ EMA

The $\Lambda_{\lambda_2}^{\lambda_1}$ pairs indicated log-linear association with coincident $a_{CDOM}(440)$ observations, and a power-law relationship was applied to model $a_{CDOM}(440)$ from $[L_W(\lambda)]_N$ as follows:

$$a_{CDOM}(440) = A\left[\Lambda_{\lambda_2}^{\lambda_1}\right]^B, \tag{5}$$

where the coefficients A and B were derived by minimizing an absolute deviation cost function, and their associated uncertainties were estimated by nonparametric bootstrapping. The distribution of $a_{CDOM}(440)$ observations within the OCEAN dataset emphasized a narrow range in $a_{CDOM}(440)$, and an iterative random thinning of the oversampled range was performed to obtain a median curve fit that was representative of the full dataset range in $a_{CDOM}(440)$. Additional $a_{CDOM}(440)$ products were derived based on the GSM (Maritorena et al., 2002) and an updated version (v5) of the QAA (Lee et al., 2009), with the contributions of the dissolved and detrital components partitioned following Matsuoka et al. (2013) and Dong et al. (2013), respectively. For our implementation of the GSM following Matsuoka et al. (2013), the spectral slopes for backscattering of particles and for absorption of the dissolved and detrital organic pools were derived from individual spectra and without applying regional (i.e., arctic) tunings for phytoplankton absorption. The GSM and QAA algorithms required wavelengths spanning 412 to

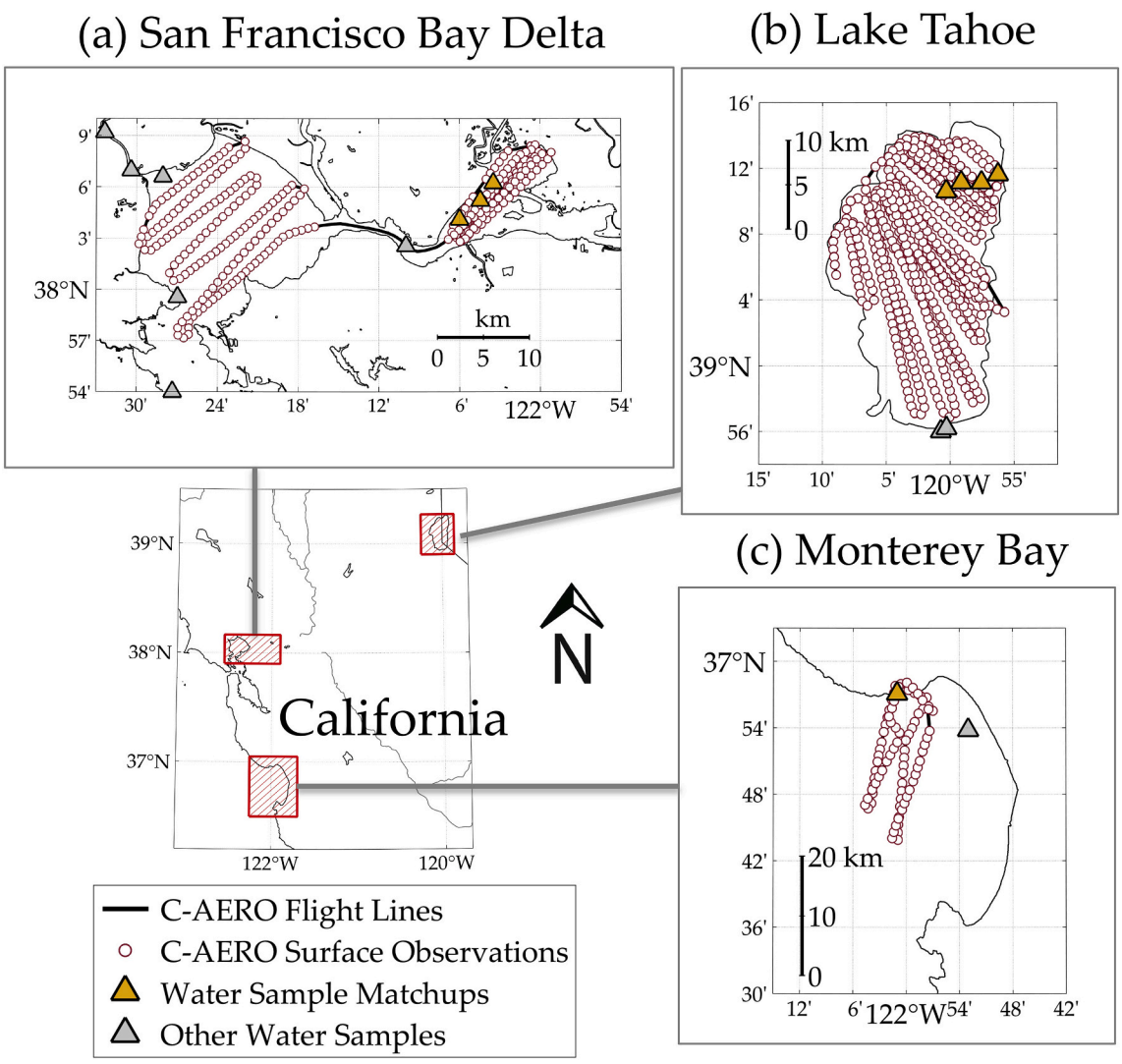

Fig. 1. Locations of C-AERO surface observations obtained as a sequence of flight lines aligned with respect to the solar geometry: (a) SFBD in California; (b) LT in California and Nevada; and (c) MB in California. Twin Otter flight lines are indicated in solid black, with the location of C-AERO data products (generated approximately every 15 s of flight time) shown as red circles, and water sampling match-up sites indicated as orange triangles. Gray triangles indicate water sampling sites that were not considered for match-up statistics due to spatial or temporal separation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

555 nm and 412 to 670 nm, respectively, including the central VIS wavebands provided by legacy ocean color sensors. Differences in optical properties between 440 nm and 443 nm were considered negligible for the purposes of this work.

2.8. Summary of algorithm validation statistics

Algorithm performance was assessed using RMSD statistics, as follows:

$$\mathrm{RMSD} = \left[\frac{\sum_{i=1}^{N}(X_i - Y_i)^2}{N}\right]^{1/2}, \quad (6)$$

where X_i and Y_i correspond to the algorithm and in situ values of $a_{\mathrm{CDOM}}(440)$, respectively. When reported as a percentage, RMSD was normalized by the range in the in situ values. For comparison of match-up statistics and evaluation of subjective subcategories, algorithm uncertainty was also assessed with the root mean square log-difference (RMSLD), which was used for evaluating the success of the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) mission, e.g., by Gregg and Casey (2004), and was derived as follows:

$$\mathrm{RMSLD} = \left[\frac{\sum_{i=1}^{N}(log_{10}(X_i) - log_{10}(Y_i))^2}{N}\right]^{1/2}. \quad (7)$$

Mean absolute deviation (MAD), mean bias (MBIAS), and percent wins (PWINS) were assessed following Seegers et al. (2018), with PWINS defined as the percentage of observations in which an algorithm was the best performing of those evaluated. MAD and MBIAS were derived as follows:

$$\mathrm{MAD} = 10^{\left[\frac{\sum_{i=1}^{N}|log_{10}(X_i) - log_{10}(Y_i)|}{N}\right]}, \quad (8)$$

and

$$\mathrm{MBIAS} = 10^{\left[\frac{\sum_{i=1}^{N}(log_{10}(X_i) - log_{10}(Y_i))}{N}\right]}. \quad (9)$$

Airborne matchups, which had a lower number of observations, were assessed using unbiased percent difference (UPD), derived as follows:

$$\mathrm{UPD} = \frac{200}{N} \sum_{i=1}^{N} \frac{|X_i - Y_i|}{(X_i + Y_i)} \,. \tag{10}$$

Uncertainty requirements to within 35% were described in Hooker and McClain (2000) for operational ocean color products (i.e., Chl *a*). Allotting half of the total uncertainty budget to in situ activities (which includes radiometry and algorithm development), and assuming quadrature summing of uncertainties, the target uncertainty for a near-surface activity (C-AERO operated at LSA) would be to within 25% in order to satisfy calibration and validation requirements for legacy oceanic measurements (Hooker et al., 2007). For future NASA missions, e.g., the Phytoplankton, Aerosol, Cloud, ocean Ecosystem mission (PACE), calibration and validation activities within open ocean waters may aim to achieve uncertainties within 15% (Hooker et al., 2007). In smaller and more challenging inland water bodies, which were not targeted by legacy ocean color missions, next-generation activities for remote sensing may aim to satisfy the 25% criteria.

The relative performance of the different EMA algorithms was also compared by propagating radiometric and model uncertainties for the algorithms based on mean properties observed at LT, MB, and SFBD, which represent a gradient of dark to bright water targets. Briefly, the uncertainty in the EMA product was estimated as the quadrature sum of the radiometric and model contributions, each derived through partial differentiation, and consistent with the presentation of McKinna et al. (2019). Spectrally dependent radiometric uncertainties were estimated in Hooker et al. (2004, 2018b), and an expanded range of spectrally dependent radiometric uncertainties were considered herein as sensitivity testing in order to compare various end-member combinations and water targets. Coefficient uncertainties for the EMA algorithms were estimated by nonparametric bootstrapping.

3. Results

3.1. Case-1 evaluation of a $K_d(\lambda)$ end-member approach

A straightforward algorithmic approach for estimating $a_{\mathrm{CDOM}}(440)$ was evaluated in Hooker et al. (2020), in which ratios of spectrally separated $K_d(\lambda)$ pairs were predictors of $a_{\mathrm{CDOM}}(440)$. The case-1 relationships between $K_d(\lambda)$ band ratios and $a_{\mathrm{CDOM}}(440)$ were evaluated herein by comparing the GSyn and MSyn datasets with the $K_d(\lambda)$ EMA algorithms derived using the Hooker et al. (2020) dataset. EMA algorithms based on linear (Hooker et al., 2020) and power-law fits were considered for $K_d(320)/K_d(780)$ and $K_d(412)/K_d(670)$, respectively, which produced normally distributed residuals between algorithm and data pairs, shown in Fig. 2. GSyn and MSyn datasets were not extended beyond the Chl *a* value of 25 mg m^{-3}, consistent with the range considered in Bricaud et al. (1995, 1998), and were compared to the EMA algorithms using UPD statistics. The GSyn a_{CDOM} slope coefficients were varied between 0.01 and 0.03 nm^{-1} for sensitivity testing, shown in the Fig. 2 panel inlays.

For the UV-NIR $K_d(320)/K_d(780)$ pair, the Hooker et al. (2020) EMA relationship was to within 25% UPD of MSyn for $a_{\mathrm{CDOM}}(440)$ above 0.006 m^{-1} (Chl $a > 0.07$ mg m^{-3}) and also to within 25% of GSyn for $a_{\mathrm{CDOM}}(440)$ between 0.009 and 0.098 m^{-1} ($0.13 <$ Chl $a < 5.91$ mg m^{-3}). For the VIS $K_d(412)/K_d(670)$ pair, the power-law EMA relationship was to within 25% UPD of MSyn for $a_{\mathrm{CDOM}}(440)$ between 0.062 and 0.201 m^{-1} ($2.88 <$ Chl $a < 18.52$ mg m^{-3}), and to within 25% of GSyn for $a_{\mathrm{CDOM}}(440)$ above 0.076 m^{-1} (Chl $a > 3.96$ mg m^{-3}).

The $K_d(412)/K_d(670)$ EMA power-law algorithm produced lower $a_{\mathrm{CDOM}}(440)$ values than the synthetic case-1 datasets, although both remained within the spread of the data due to the increased variability of the $K_d(412)/K_d(670)$ relationship compared to $K_d(320)/K_d(780)$. For GSyn, an expanded range in *S* produced greater variability in $K_d(320)/K_d(780)$ than was apparent in the in situ data products, while the effect of *S* on $K_d(412)/K_d(670)$ was less than the variability in the in situ data products for low values of $K_d(412)/K_d(670)$.

3.2. EMA algorithms derived for $[L_W(\lambda)]_N$ pairs

EMA algorithms for UV-NIR (Λ_{780}^{320}) and VIS (Λ_{670}^{412}, Λ_{555}^{443}, and Λ_{625}^{465}) $[L_W(\lambda)]_N$ pairs were derived from the OCEAN, GLOBC, and NOMAD datasets (Table 1). For all datasets, the dynamic range in $\Lambda_{\lambda_2}^{\lambda_1}$ increased with expanded spectral separation between wavelength pairs, with the greatest ranges observed in Λ_{780}^{320} for the OCEAN and GLOBC datasets and in Λ_{670}^{412} for the NOMAD dataset, which lacked the relevant UV and NIR wavelengths. The $\Lambda_{\lambda_2}^{\lambda_1}$ EMA algorithm fits are overlaid in solid black on the OCEAN and GLOBC datasets in Figs. 3 and 4, respectively, with the PSyn synthetic case-1 dataset overlaid in solid gray for relevant legacy wavelengths.

Compared to the coefficients derived from the GLOBC dataset, the OCEAN dataset normalizing coefficients (*A*) were 8.9% greater (Λ_{780}^{320}) and 0.3% less (Λ_{670}^{412}), and the exponential coefficients (*B*) were 2.9% (Λ_{780}^{320}) and 18.1% (Λ_{670}^{412}) greater (Table 1). NOMAD *A* and *B* coefficients for the Λ_{670}^{412} algorithm were elevated by 17.7% and 33.7%, respectively,

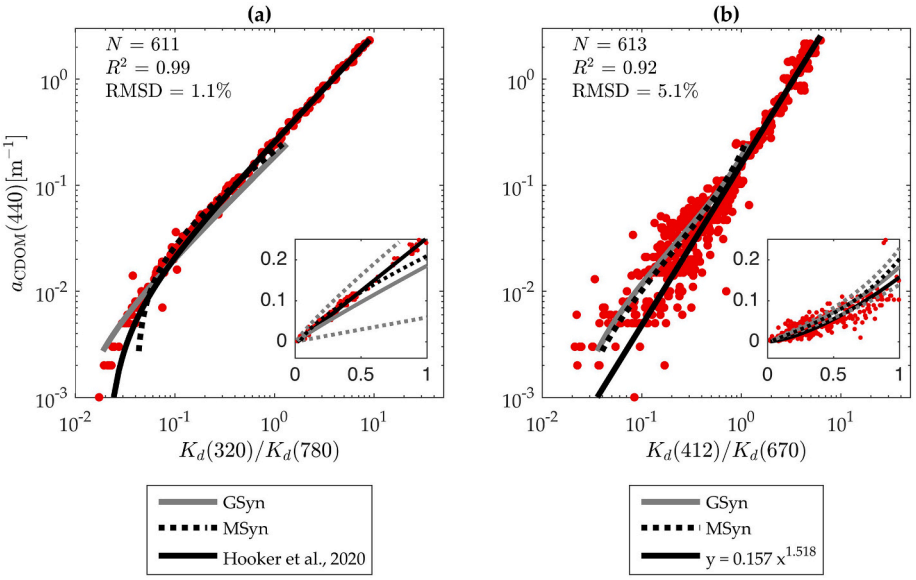

Fig. 2. $K_d(\lambda_1)/K_d(\lambda_2)$ relationships to $a_{\mathrm{CDOM}}(440)$, with the $K_d(\lambda)$ dataset that was presented in Hooker et al. (2020) shown in red, and the coefficient of determination, R^2, reported for $\log_{10}$-transformed values. The data products are overlaid, respectively, as follows: (a) GSyn with $S = 0.018$ nm^{-1} in solid gray, MSyn in dashed black, and EMA from Hooker et al. (2020) in solid black; and (b) GSyn with $S = 0.018$ nm^{-1} in solid gray, MSyn in dashed black, and a least absolute deviation power-law fit in solid black. Panel inlays show relationships in linear space for a reduced range, with sensitivity testing for GSyn included in dashed gray, wherein $S = 0.01$ and 0.03 nm^{-1}. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

compared with the algorithm fit to the GLOBC dataset. The coefficient differences between datasets for the VIS $\Lambda_{\lambda_2}^{\lambda_1}$ pairs were in some instances greater than the standard errors of the coefficients, indicating that the uncertainties associated with the VIS algorithms were underestimated, likely due to differences in the technology, distribution, or range of the datasets. Differences in the range of a dataset would alter the coefficients derived for a power-law fit, for example, if a power-law model was not appropriate. This scenario is supported for the VIS (but not the UV-NIR) algorithms, as follows: First, the data products shown in Fig. 4 indicate log-space nonlinearities in the clustering of points between the VIS $\Lambda_{\lambda_2}^{\lambda_1}$ pairs and $a_{CDOM}(440)$ when viewed across an expanded $a_{CDOM}(440)$ range. Second, if the GLOBC dataset were confined to the $a_{CDOM}(440)$ range in the OCEAN dataset (0.004 – 0.613 m^{-1}), B would increase for the Λ_{670}^{412} algorithm by about 11%, and the coefficient uncertainties for the GLOBC and OCEAN algorithms would overlap. Polynomial algorithms were not pursued herein given the size and variability of the in situ datasets, and because within the more spectrally expansive GLOBC dataset, power-law fitting was reasonable for the Λ_{780}^{320} relationship. The log-linearity of the Λ_{780}^{320} algorithmic relationship was also robust to whether or not the $K_d(\lambda) > 10$ m^{-1} data products were included during model fitting.

The GLOBC dataset contains the greatest range in $a_{CDOM}(440)$ with the most uniform distribution, and was obtained with extrapolation intervals nearer to the surface and with finer VSR to improve measurement at end-member wavelengths. Algorithms fit to the GLOBC dataset indicate that RMSD increased with decreasing spectral separation of the wavebands, and the dynamic range in $\Lambda_{\lambda_2}^{\lambda_1}$ decreased by about two orders of magnitude between Λ_{780}^{320} and Λ_{555}^{443}, in agreement with the perspective of Hooker et al. (2013). Due to decreasing sensitivity for the Λ_{555}^{443} and Λ_{625}^{465} algorithms within the global waters dataset, the subsequent analyses herein focus on validation for the Λ_{780}^{320} and Λ_{670}^{412} algorithms. However, high R^2 for the Λ_{625}^{465} relationship indicates that sensors with configurations that are not consistent with ocean color instruments may have potential oceanographic applications, which have not been widely exploited.

Table 1
The power-law coefficients derived from in situ datasets. R^2 is derived from $\log_{10}$-transformed values. See Appendix A for additional waveband combinations.

λ Pair	Fit	N	R^2	$u_m(A)$ [m^{-1}]	$u_m(B)$	Dataset
$[L_W(320)]_N/[L_W(780)]_N$ ($x = \Lambda_{780}^{320}$)	$y = 0.281x^{-0.542}$[a]	112	0.89	0.041	0.066	OCEAN
	$y = 0.259x^{-0.558}$[b]	550	0.87	0.009	0.018	GLOBC
$[L_W(412)]_N/[L_W(670)]_N$ ($x = \Lambda_{670}^{412}$)	$y = 0.242x^{-0.787}$	112	0.89	0.023	0.058[c]	OCEAN
	$y = 0.242x^{-0.961}$	606	0.92	0.010	0.036[c]	GLOBC
	$y = 0.285x^{-0.638}$	497	0.89	0.010[c]	0.039[c]	NOMAD
$[L_W(443)]_N/[L_W(555)]_N$ ($x = \Lambda_{555}^{443}$)	$y = 0.066x^{-1.523}$	112	0.89	0.006	0.192[c]	OCEAN
	$y = 0.063x^{-1.764}$	609	0.87	0.008	0.129[c]	GLOBC
	$y = 0.065x^{-1.399}$	864	0.66	0.003	0.096[c]	NOMAD
$[L_W(465)]_N/[L_W(625)]_N$ ($x = \Lambda_{625}^{465}$)	$y = 0.349x^{-0.996}$	112	0.87	0.041[c]	0.069[c]	OCEAN
	$y = 0.430x^{-1.320}$	609	0.80	0.023[c]	0.080[c]	GLOBC
	$y = 0.128x^{-0.564}$	133	0.34	0.043[c]	0.100[c]	NOMAD

[a] Algorithm applied to GLOBC validation (Sect. 3.3).
[b] Algorithm applied to GLOBN analysis (Sect. 3.4) and to airborne imagery (Sect 3.5).
[c] The differences between coefficients are greater in many cases than the uncertainties obtained by bootstrapping (discussed in Sect. 3.2, 3.5, and 4.1). This is true in all instances for B within the VIS (but not UV-NIR) algorithms, perhaps indicating that a power-law model is not appropriate for the Λ_{670}^{412}, Λ_{555}^{443}, and Λ_{625}^{465} VIS algorithms.

3.3. Validation of $a_{CDOM}(440)$ algorithms

The performance of the EMA (Λ_{780}^{320} and Λ_{670}^{412}), GSM, and QAA algorithms was evaluated using the GLOBC dataset, which included oceanic, coastal, and inland water bodies and spanned more than three decades of dynamic range (0.001–2.305 m^{-1}) in $a_{CDOM}(440)$. The EMA algorithms considered were those fit to the OCEAN dataset (Table 1) in order to maintain independence from the GLOBC validation dataset. Results for each algorithm are shown in Fig. 5 using the GLOBC observations in which all of the relevant wavebands (spanning 320 to 780 nm) were available.

The EMA $a_{CDOM}(440)$ algorithms produced higher PWINS plus lower RMSD, RMSLD, and MAD values than the GSM and QAA algorithms across the full GLOBC range in $a_{CDOM}(440)$, shown in Fig. 5. Algorithm performance varied greatly for the GSM and QAA algorithms as a function of $a_{CDOM}(440)$ range, and additional results are presented below for data within low ($\leq$0.1 m^{-1}) and high (>0.1 m^{-1}) ranges of $a_{CDOM}(440)$ in situ observations, which approximately partition the GLOBC dataset into oceanic versus coastal/inland water bodies. Within the high $a_{CDOM}(440)$ partition, the GSM algorithm did not converge to a solution for 10 (5.1%) of the $[L_W(\lambda)]_N$ spectra, while in the low $a_{CDOM}(440)$ partition, the GSM algorithm produced 3 (0.7%) negative retrievals. These problematic retrievals were not considered for the GSM statistics presented, but were included when presenting the performance of the EMA and QAA algorithms.

RMSD within the high (>0.1 m^{-1}) $a_{CDOM}(440)$ range was 0.277, 0.379, 0.497, and 0.540 m^{-1} for the Λ_{780}^{320}, Λ_{670}^{412}, GSM, and QAA algorithms, respectively, which corresponded to 12.6%, 17.2%, 22.6%, and 24.5% of the range in the high $a_{CDOM}(440)$ partition. Within this upper range, RMSD for the QAA was degraded by over-prediction of $a_{CDOM}(440)$, with positive bias of 1.368, although the method indicated high R^2 (0.78 using $\log_{10}$-transformed values), perhaps indicating its potential for regional optimization. RMSLD values also indicated better performance of the EMA algorithms, with values of 0.133, 0.184, 0.278, and 0.236 for the Λ_{780}^{320}, Λ_{670}^{412}, GSM, and QAA algorithms, respectively.

Within the low $a_{CDOM}(440)$ partition ($\leq$0.1 m^{-1}), the relative performance of the algorithms was mixed across the statistical metrics reported. For example, RMSD values were lowest for the Λ_{780}^{320} and QAA algorithms, with 0.025, 0.032, 0.035, and 0.027 m^{-1} for the Λ_{780}^{320}, Λ_{670}^{412}, GSM, and QAA algorithms, respectively, which corresponded to 25.0%, 32.3%, 36.2%, and 27.9% of the range in the low $a_{CDOM}(440)$ partition. In contrast, RMSLD values were slightly lower for the Λ_{670}^{412} and GSM algorithms, with 0.243, 0.226, 0.234, and 0.316 for the Λ_{780}^{320}, Λ_{670}^{412}, GSM, and QAA algorithms, respectively. Performance of the Λ_{780}^{320} EMA algorithm was degraded relative to the Λ_{670}^{412} algorithm within the lowest $a_{CDOM}(440)$ waters in part by the difficulty of retrieving $[L_W(780)]_N$ from an in-water profiler within clear waters (because the first L_u observations are nominally at a depth of 0.3 m). New technologies that would mitigate this difficulty are described in Sect. 4.2.

The results above are presented for the Matsuoka et al. (2013) and Dong et al. (2013) implementations of the GSM and QAA algorithms, respectively, which separate the detrital and dissolved contributions, or partition $a_{CDOM}(440)$ from $a_{CDM}(440)$. However, when the GSM and QAA algorithms were implemented without accounting for the detrital signal, i.e., equating $a_{CDOM}(440)$ with $a_{CDM}(440)$—an approximation for open ocean waters—the performance improved in the clearer-water partition, but worsened in the higher $a_{CDOM}(440)$ waters and led to an over-prediction bias (i.e., elevated MBIAS).

The performance metrics for the EMA algorithms were most improved relative to the GSM and QAA algorithms (for all

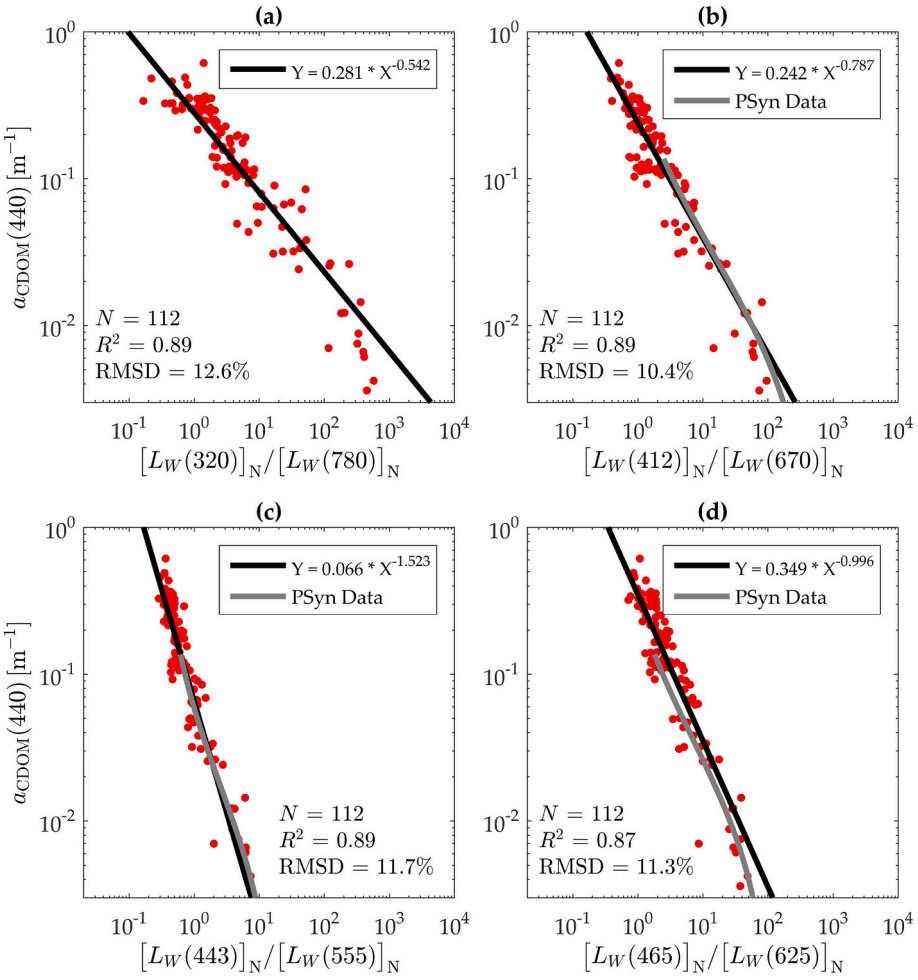

Fig. 3. OCEAN dataset (red dots), EMA algorithm (solid black line), and PSyn case-1 dataset (solid gray line) for: (a) Λ_{780}^{320}; (b) Λ_{670}^{412}; (c) Λ_{555}^{443}; and (d) Λ_{625}^{465}. The R^2 values are reported for $\log_{10}$-transformed values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

implementaions evaluated herein) in the higher $a_{CDOM}(440)$ water bodies. Although the $a_{CDOM}(440) > 0.1$ m^{-1} partition represents a small fraction of global waters by area, it constitutes a large portion of the total dynamic range in $a_{CDOM}(440)$ relevant to a global algorithmic perspective. For example, over 40% or 95% (in $\log_{10}$- or linear-space, respectively) of the $a_{CDOM}(440)$ range for conservative waters exceeds an $a_{CDOM}(440)$ value of 0.1 m^{-1} based on the complete in situ range in the GLOBC dataset. Decreased performance of the GSM and QAA implementations using validation on a global dataset was anticipated given that both were tuned for oceanic or coastal ecosystems, and roughly one third of the GLOBC dataset corresponded to inland water bodies, which are in general more optically variable than oceanic or coastal waters and less likely to match the internal bio-optical parameterizations of oceanic algorithms. The GLOBC dataset contained roughly uniform density of $a_{CDOM}(440)$ across the full range in $\log_{10}$-space, which was not in itself an objective of the sampling, but which resulted from the effort to represent a wide diversity of water bodies (e.g., lakes, rivers, bays, and marshes), and which was made possible by the prior subjective classification information.

3.4. *Efficacy of $a_{CDOM}(440)$ retrievals in nonconservative water bodies*

Observations of nonconservative waters from the GLOBC dataset were assessed within classifications defined in Hooker et al. (2020). The classifications do not necessarily indicate that optical or ecological conditions invalidate a global algorithmic approach, but they apply available prior information in order to quantify algorithm performance in conditions where algorithms might otherwise be expected to degrade. The classifications are briefly summarized as follows: a) resuspension (shallow regions with strong mixing, e.g., from tides or winds, which resuspends bottom material); b) refilled or flooded (lakes or rivers with an above average mean high-water line or which have recently been replenished); c) drought-stricken (a water body with significantly below normal elevation or water line); d) harbor (a sheltered area for docking vessels); e) harmful algal bloom or HAB (elevated concentrations of phytoplankton, which may produce toxic compounds); f) wetland or marsh (shallow, tidally influenced estuarine areas, often with brackish properties); g) polluted (waters containing anthropogenic sources that alter the natural water properties); h) alkaline lake (lacustrine ecosystems containing significantly elevated loads of salt compounds and generally high pH); i) river mouth (a region where a riverine ecosystems mixes with a larger water body, including lakes and bays); and j) invasive species (an ecosystem perturbed by anthropogenic introduction of noxious macroalgae). The classifications are consistent with the presentation of Hooker et al. (2020), in which a more detailed discussion is provided on the optically relevant water mass modifications anticipated for each classification. The EMA performance results for the nonconservative (GLOBN) dataset subcategories were assessed using EMA algorithms fit to the conservative (GLOBC) dataset, and are shown in Table 2. The statistics shown indicate whether particular environments are expected to be consistent with the EMA algorithms derived herein, or whether a regionalized approach would potentially be supported (e.g., if

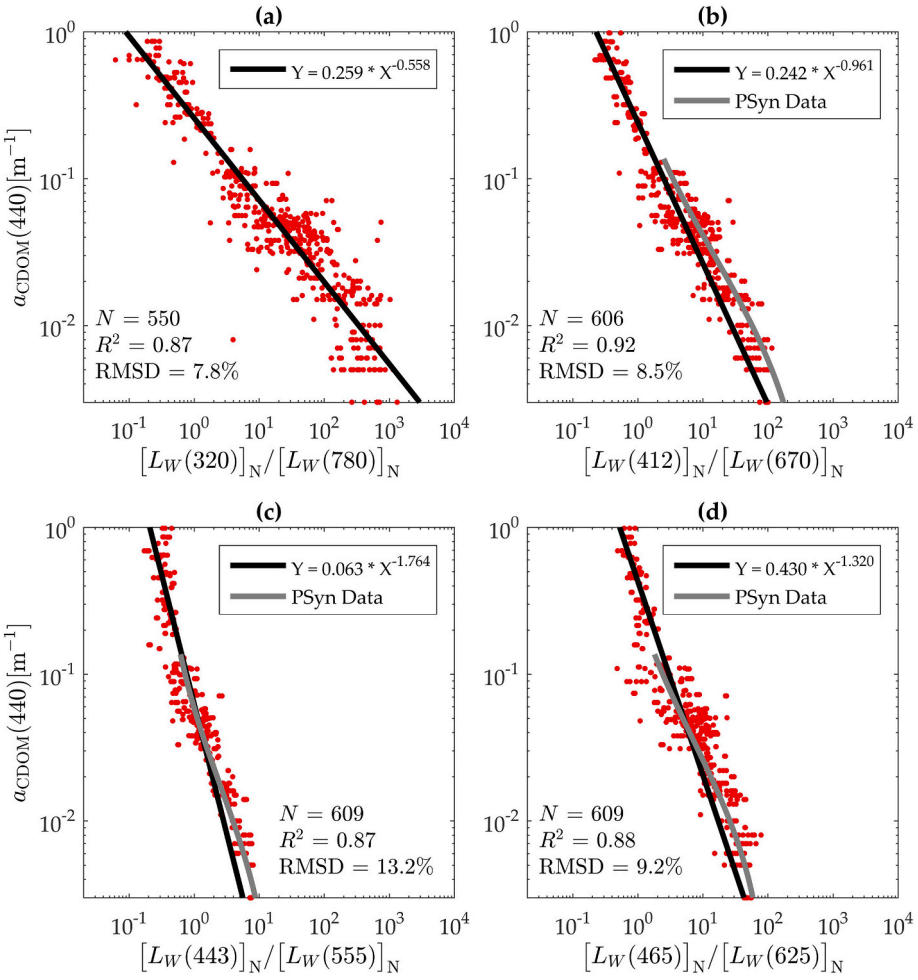

Fig. 4. GLOBC dataset (red dots), EMA algorithm (solid black line), and PSyn case-1 dataset (solid gray line) for: (a) Λ_{780}^{320}; (b) Λ_{670}^{412}; (c) Λ_{555}^{443}; and (d) Λ_{625}^{465}. The R^2 values are reported for $\log_{10}$-transformed values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

MAD approximates MBIAS, R^2 is significant, and N is large).

Despite overall decreased algorithm efficacy for the nonconservative waters, some subjective categories indicated RMSLD within that derived in the GLOBC validation comparison (Sect. 3.3), including (for the Λ_{780}^{320} algorithm) harbors, river mouths, and refilled or flooded lakes and rivers. Other water bodies, such as HABs and drought-stricken lakes, indicated the potential for the development of a regionalized algorithmic approach based on high correlation (R^2) between $a_{CDOM}(440)$ and Λ_{780}^{320}, RMSLD near the GLOBC validation results, similar magnitude in MAD and MBIAS, and the relatively large number of observations considered ($N = 60$, 11) for an inland waters, ecosystem-specific dataset.

For drought-stricken lakes, harbors, and river mouths, water mass classification is possible from a remote perspective. For example, maps of harbors and river mouths are accessible, and lake levels can be assessed through satellite imagery or by applying external knowledge of drought conditions. For other environments, classification is more difficult. For example, tidal estuaries are strongly affected by tidal cycles, which may alternate source water masses or lead to changes in depth and resuspension processes. An operator in a boat can better avoid mud flats or vegetation, although the validation dataset primarily sampled high tide conditions in these environments because access by boat was sometimes not possible during low tide.

3.5. Analysis of airborne imagery

The C-AERO EMA data products were evaluated with respect to contemporaneous water sampling, with the Λ_{780}^{320} EMA algorithm indicating UPD less than 25% for all sites and consistently below that obtained by applying the Λ_{670}^{412} algorithm, although water sample matchups were sparse (N = 4, 1, and 3 unique replicates at LT, MB, and SFBD, respectively). The sites spanned over an order of magnitude in $a_{CDOM}(440)$, with mean in situ values of 0.023 m^{-1}, 0.103 m^{-1}, and 0.560 m^{-1} at LT, MB, and SFBD, respectively. Matchups between airborne observations and in situ water samples indicated UPD of 23.3%, 24.5%, and 11.9% for the Λ_{780}^{320} algorithm and 25.8%, 34.2%, and 25.7% for the Λ_{670}^{412} algorithm, at LT, MB, and SFBD, respectively. Nearest neighbor spatial interpolations based on the Λ_{780}^{320} EMA algorithm using the C-AERO flight data are shown in Fig. 6.

Coincident water samples for the SFBD C-AERO flights were obtained in Grizzly Bay and not in San Pablo Bay, although the observed structure in $a_{CDOM}(440)$ within San Pablo Bay was consistent with sampling in a prior year, as well as information on tidal cycles and bathymetry. In particular, a transition between roughly 0.3 m^{-1} (green) and 0.5 m^{-1} (yellow) $a_{CDOM}(440)$ was indicated in the San Pablo C-AERO observations near the 2 m bathymetric contour (the bottom depth was greater than the 0.1% light level for the Λ_{780}^{320} wavebands within the regions shown). A similar $a_{CDOM}(440)$ gradient was observed in situ during prior sampling, with $a_{CDOM}(440)$ of 0.554 m^{-1} observed in the northern edge of San Pablo Bay, compared to 0.337 m^{-1} observed roughly 13 km due south. Sampling at four sites farther south into San Francisco Bay recorded a mean value of 0.207 m^{-1}. Within the Carquinez Strait that joins San Pablo Bay and Grizzly Bay, a mean $a_{CDOM}(440)$ value of 0.553

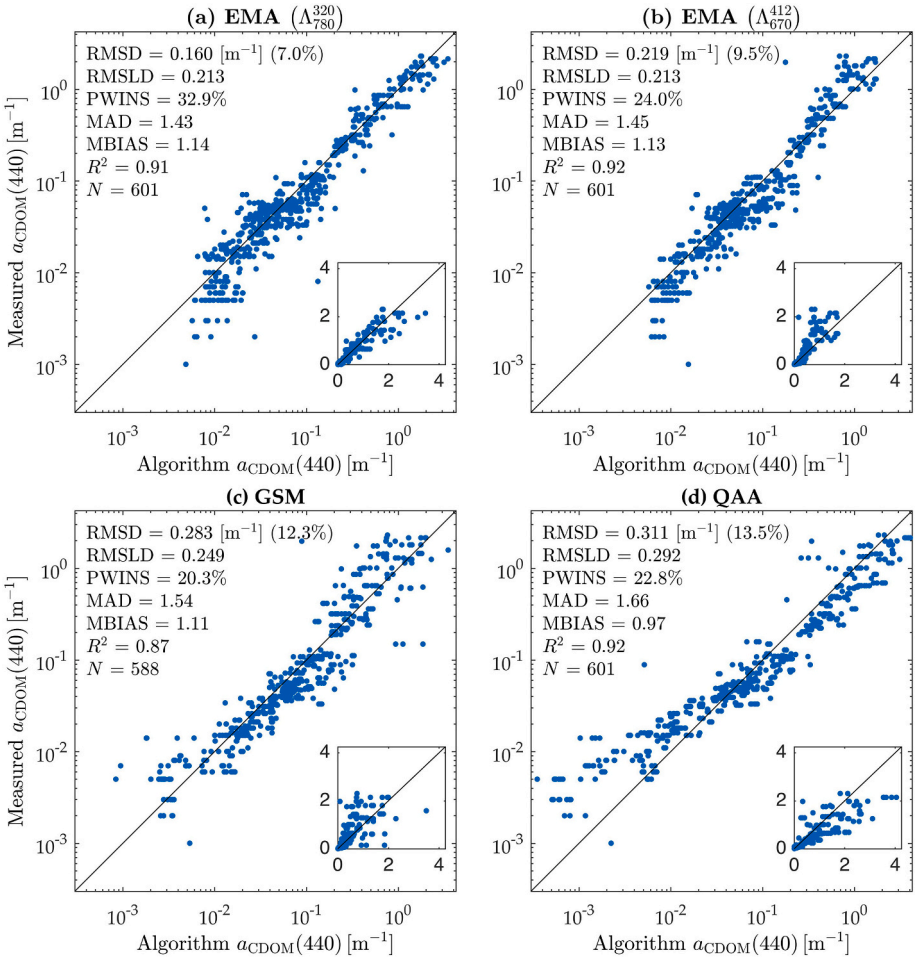

Fig. 5. Algorithm versus in situ $a_{CDOM}(440)$ from the GLOBC dataset for: (a) Λ_{780}^{320}; (b) Λ_{670}^{412}; (c) GSM; and (d) QAA. Linear axes are included as inlays, with a one-to-one line shown in solid black for all panels. R^2 is derived from $\log_{10}$-transformed values.

Table 2
Performance of $[L_W(\lambda)]_N$ end-member algorithms within various water mass classifications.

Algorithm	Watermass classification	*N*	R^2	RMSD [m^{-1}]	RMSLD	MAD	MBIAS
Λ_{780}^{320}	Resuspension	66	0.09	0.031	0.308	1.771	0.767
	Refilled or Flooded	15	**0.69**	0.219	0.183	1.386	0.790
	Drought-Stricken	60	**0.71**	0.194	0.269	1.698	1.573
	Harbor	48	**0.66**	0.120	0.162	1.358	1.050
	HAB	11	**0.70**	0.150	0.250	1.515	0.721
	Wetland or Marsh	28	0.16	0.146	0.399	2.089	1.859
	Polluted	38	0.14	0.142	0.407	1.895	1.596
	Alkaline Lake	34	**0.38**	0.165	0.374	2.176	2.095
	River Mouth	16	**0.94**	0.113	0.103	1.202	0.949
	Invasive Species	10	**0.71**	0.165	0.504	3.023	3.023
Λ_{670}^{412}	Resuspension	66	**0.11**	0.053	0.364	2.060	0.698
	Refilled or Flooded	15	0.22	0.281	0.274	1.645	0.745
	Drought-Stricken	60	**0.72**	0.169	0.268	1.654	1.515
	Harbor	48	**0.66**	0.095	0.177	1.383	1.222
	HAB	11	**0.94**	0.095	0.137	1.303	0.814
	Wetland or Marsh	28	0.01	0.213	0.487	2.510	2.182
	Polluted	38	**0.53**	0.065	0.263	1.597	1.257
	Alkaline Lake	34	**0.27**	0.174	0.384	2.186	2.062
	River Mouth	16	**0.94**	0.100	0.130	1.277	0.845
	Invasive Species	10	**0.98**	0.105	0.289	1.936	1.936

The R^2 values were derived from $\log_{10}$-transformed data, with bold indicating satisfaction of $P \leq 0.01$.

m^{-1} and decreased salinity (9 ppt) relative to the San Francisco Bay water (28–30 ppt) was recorded in situ on the same day as the C-AERO flights, but high spatial heterogeneity within the strait combined with an increased spatial offset from the airborne observations precluded match-up analysis for this point. The C-AERO observations of SFBD were obtained on a flood tide with an eastward-flowing current at the in-water

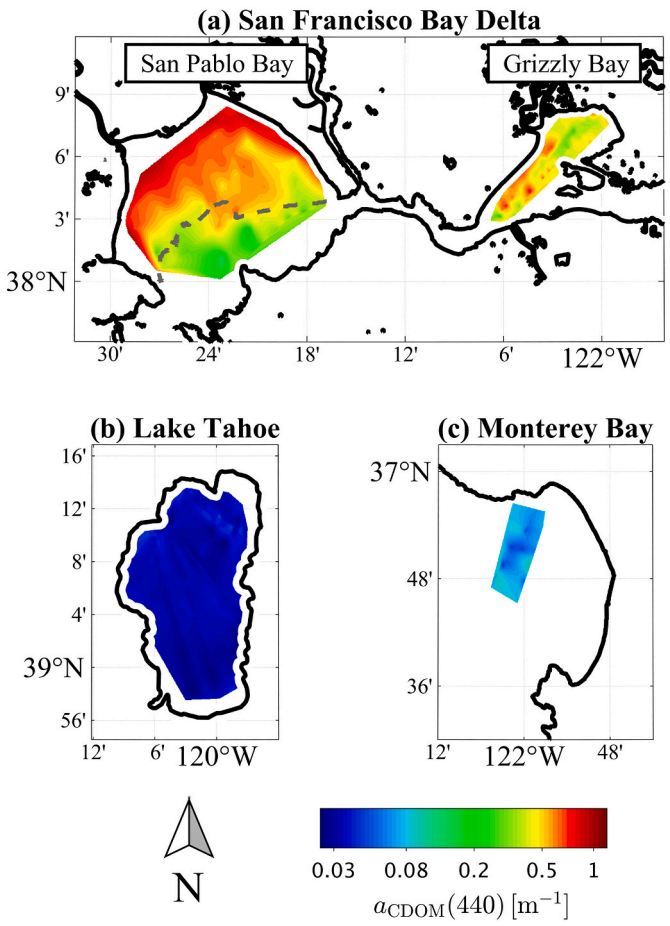

Fig. 6. The Λ_{780}^{320} EMA algorithm applied to C-AERO measurements obtained at LSA and under clear-sky conditions at the following sites: (a) SFBD with larger San Pablo Bay on the left and smaller Grizzly Bay on the right (the 2 m bathymetric contour for San Pablo Bay is shown as a dashed gray line); (b) LT; and (c) MB.

sampling locations. The Grizzly Bay SFBD match-up site was in a tidally-mixed system, potentially denoting the nonconservative Resuspension water mass classification.

Uncertainty propagation for the EMA algorithms based on the GLOBC coefficient uncertainties presented in Table 1 and using a range of radiometric uncertainties exceeding those outlined in Hooker et al. (2018b) and Hooker et al. (2004) suggested that the Λ_{780}^{320} and Λ_{670}^{412} algorithms would satisfy 15% uncertainty for targets resembling LT, MB, and SFBD. Target brightness was a primary factor determining the relative performance of Λ_{780}^{320} compared to Λ_{670}^{412}, due in part to the low (although nonzero) signal of $[L_W(780)]_N$ in clear (i.e., resembling LT) waters. For example, the estimated uncertainties for the Λ_{780}^{320} algorithm were greater than those derived for the Λ_{670}^{412} algorithm in dark water bodies resembling LT, but were lower in bright water bodies resembling SFBD. For an intermediate target resembling MB, uncertainties were less for the Λ_{780}^{320} algorithm as long as radiometric uncertainties assigned to the end-member wavelengths were not severely degraded relative to those assigned to the visible wavelengths.

The uncertainty propagation resulted in lower uncertainties for the more spectrally expansive Λ_{780}^{320} algorithm compared with the Λ_{670}^{412} algorithm for non-oligotrophic targets in part because of the flattened exponential slope (i.e., a less negative B coefficient) for the Λ_{780}^{320} algorithm, which resulted from the increased dynamic range of the more spectrally separated end-member ratio. If the uncertainty propagation were performed for the algorithms fit to the OCEAN dataset, in which the Λ_{670}^{412} algorithm B coefficients were less negative than for those fit to the GLOBC dataset (possibly because of the narrower $a_{CDOM}(440)$ range in the OCEAN dataset), the uncertainty for the Λ_{670}^{412} algorithm would decrease. The model uncertainties obtained for the Λ_{670}^{412} algorithm, however, were likely underestimated because the differences in the Λ_{670}^{412} algorithm B coefficients between the datasets were greater than the B coefficient uncertainties determined by bootstrapping. As described in Sect. 3.2, this discrepancy is perhaps due to log-space nonlinearities in the VIS algorithmic relationships (Fig. 4) combined with different ranges in $a_{CDOM}(440)$ between the in situ datasets.

4. Discussion

4.1. Consistency of EMA with the case-1 parameterizations

A band-ratio algorithmic approach in which spectrally separated $K_d(\lambda)$ or $[L_W(\lambda)]_N$ end-member pairs are related to $a_{CDOM}(440)$ was evaluated using synthetic bio-optical datasets derived from case-1 parameterizations. EMA algorithms based on ratios of UV-NIR and VIS data products indicated that case-1 parameterizations produced similar algorithmic relationships as those derived using a global in situ dataset of in-water measurements within a Chl a range relevant to case-1 parameterizations. $K_d(\lambda)$ EMA algorithm residuals were within the variability due to a global range in S for UV-NIR—but not VIS—band ratios, and the UV-NIR $K_d(\lambda)$ relationship was fit using a linear model while the VIS $K_d(\lambda)$ relationship was nonlinear.

Considering band ratios of Eq. (3) helps to illustrate the importance of spectral range by separating the effects of various optically relevant constituents within each of the UV, VIS, and NIR domains. For example, if a case-1 condition is considered in which $K_d(\lambda)$ is approximately proportional to the total absorption coefficient (e.g., sun at zenith with $a \gg b_b$), the greatest spectral dependency in a UV-NIR $K_d(\lambda)$ band ratio is due to CDOM, i.e., $a_{CDOM}(780) \ll a_{CDOM}(320)$, and an approximately linear relationship emerges between the Eq. (3) ratio of $K_d(\lambda)$ end-members and $a_{CDOM}(440)$. If the same logic is applied to a VIS $K_d(\lambda)$ band ratio, the absorption contribution by phytoplankton increases (relative to that of CDOM), contributing nonlinear spectral dependencies (e.g., Bricaud et al., 1995). This logic is supported by the linear and nonlinear (power-law) algorithms shown in Fig. 2. A similar result is indicated for the $[L_W(\lambda)]_N$ algorithms by the similarity of the B coefficients between the in situ datasets for the Λ_{780}^{320} algorithms, compared with the B coefficients that differ by more than the fitting uncertainties for the Λ_{670}^{412} algorithms. Although not evaluated herein, the differences in B coefficients for the Λ_{670}^{412} algorithms (Table 1) may be due in part to log-space nonlinearities in the relationship between the VIS $[L_W(\lambda)]_N$ ratios and $a_{CDOM}(440)$.

Compared to the $K_d(\lambda)$ algorithms, the $[L_W(\lambda)]_N$ algorithms are more sensitive to differences in brightness, as well as to technological limitations of in-water measurements used herein (i.e., no L_u observations within the upper 0.3 m of the water column). Variability in target brightness (e.g., due to particle loading) increases uncertainty of the $[L_W(\lambda)]_N$ EMA relative to the $K_d(\lambda)$ EMA, but an $[L_W(\lambda)]_N$ algorithmic approach is useful because $[L_W(\lambda)]_N$ is derived directly from remote measurements, whereas $K_d(\lambda)$ derivations are spectrally incomplete (Lee et al., 2013), despite ongoing improvements, e.g., Cao et al. (2014). An alternate approach of estimating $K_d(\lambda)$ from above-water measurements and then deriving $a_{CDOM}(440)$ was also considered, wherein a neural network dataset of synthetic $K_d(\lambda)$ products (NNet) was derived from NOMAD above-water measurements, based on Jamet et al. (2012). Comparison of the EMA algorithm using the NNet $K_d(\lambda)$ data products with the EMA algorithm using the original $[L_W(\lambda)]_N$ data products is shown in Fig. 7. The neural network solution indicated close agreement between NNet and MSyn, but RMSD increased and R^2 decreased for the EMA algorithm that used the NNet $K_d(\lambda)$ data products compared with the EMA algorithm that used the original, above-water data products.

4.2. Performance of EMA algorithms

Three sets of $[L_W(\lambda)]_N$ EMA algorithms were derived using the OCEAN, GLOBC, and NOMAD datasets, which contained coincident $a_{CDOM}(440)$ and $[L_W(\lambda)]_N$ in situ observations. For all datasets, the dynamic range in $\Lambda_{\lambda_2}^{\lambda_1}$ increased with increasing spectral separation between wavelength pairs, which increased the sensitivity and robustness of the algorithmic approach. Scalar coefficients (A) were in general similar between the OCEAN and GLOBC datasets for the Λ_{780}^{320}, Λ_{670}^{412}, and Λ_{555}^{443} algorithm pairs. Exponential coefficients (B) obtained for VIS wavelength pairs were variable between datasets, with the GLOBC dataset consistently deriving more negative (steeper in log-log space) B values than the OCEAN dataset, which in turn produced more negative B values than the NOMAD dataset. The differences in coefficients were in some instances greater than the coefficient uncertainties derived by bootstrapping, indicating that model uncertainties were underestimated. The in-water datasets may have contained measurement or processing differences, or differences in $a_{CDOM}(440)$ range and distribution may have caused differences in curve fitting. It is possible that the $[L_W(\lambda)]_N$ EMA relationships for the VIS wavelength pairs would be more accurately modeled with a polynomial function, which was not evaluated herein, but which is similar to the $K_d(\lambda)$ EMA perspective, in which the VIS algorithm required a higher-order relationship (power-law) than the UV-NIR algorithm (linear), shown in Fig. 2. Nonuniform sampling within the OCEAN dataset across the full range in $a_{CDOM}(440)$ also modified the curve fitting compared to the GLOBC dataset, but was corrected by thinning the OCEAN dataset to produce a consistent log-scale distribution of observations.

The $[L_W(\lambda)]_N$ EMA algorithms fit to the OCEAN dataset were compared with existing $a_{CDOM}(440)$ algorithms using the GLOBC dataset, which contained an approximately equal number of observations from inland, coastal, and oceanic waters. Across the full GLOBC range in $a_{CDOM}(440)$ values (0.001–2.305 m^{-1}), the Λ_{780}^{320} EMA algorithm produced the highest PWINS and the lowest RMSD, with RMSD corresponding to 7.0%, 9.5%, 12.3%, and 13.5% of the range in $a_{CDOM}(440)$ for the Λ_{780}^{320}, Λ_{670}^{412}, GSM and QAA algorithms, respectively. The RMSD values obtained by validation on the GLOBC dataset but without the more attenuating observations, i.e., excluding observations wherein $K_d(\lambda) > 10\ m^{-1}$, were shown in Houskeeper, 2020, which indicated that the result for the VIS Λ_{670}^{412} algorithm was consistent with that reported by Mannino et al. (2014) within coastal waters of the eastern United States.

Within the clear-water fraction, $a_{CDOM}(440) \leq 0.1\ m^{-1}$, the performance metrics for all algorithms were similar, and their rankings varied by the statistical tests considered. For example, RMSD was lower but RMSLD was higher for the Λ_{780}^{320} algorithm in the clear-water fraction. Uncertainty may have been elevated for the Λ_{780}^{320} algorithm within the clearest waters by the difficulty of deriving $[L_W(780)]_N$, e.g., compared to $[L_W(670)]_N$ or $K_d(780)$, from an in-water profiler in clear waters due to the strong attenuation scale for long wavelengths plus $L_u(z,\lambda)$ measurement limitations. For each of the in situ datasets used, the shallowest $L_u(z,\lambda)$ observations were obtained at least 0.3 m from the surface, which limited the signal available in highly attenuating water bodies. The recent development of a Compact-Hybridspectral Radiometer (C-HyR), which combines microradiometer instruments with a spectrograph aperture separated from the top of a C-OPS backplane by a rigid, radiance collector assembly (Hooker et al., 2018a), enables

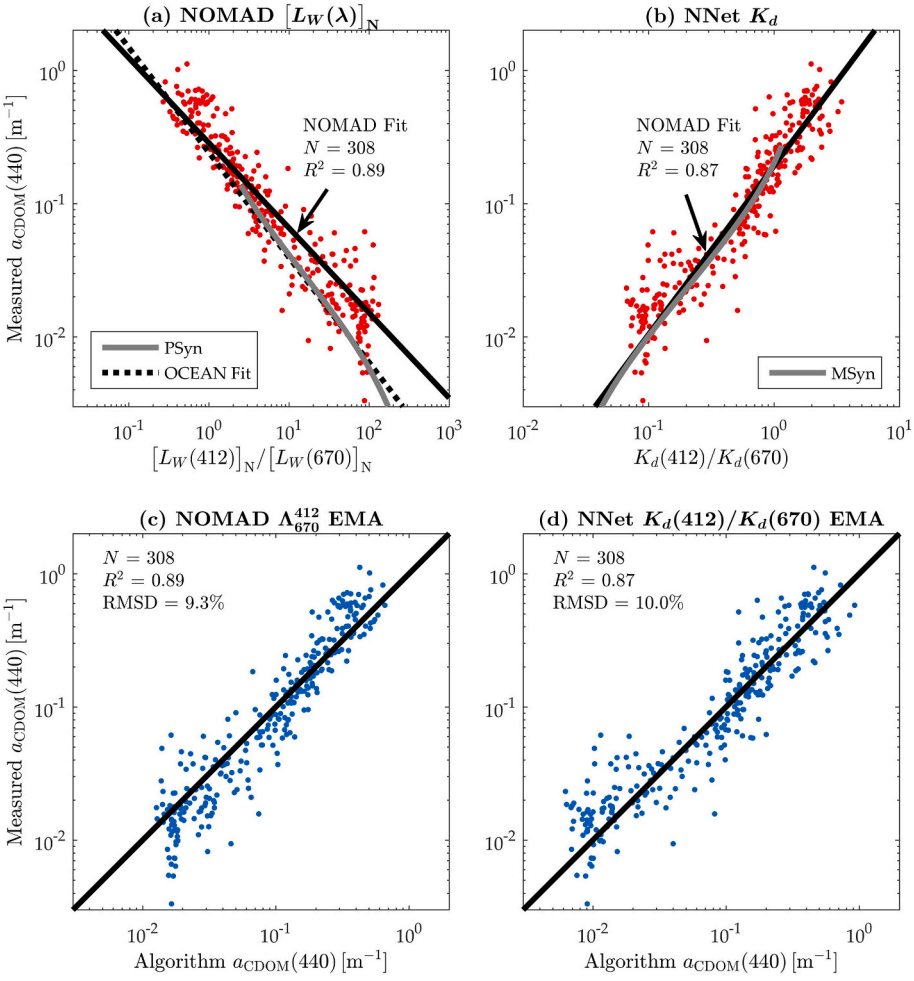

Fig. 7. Algorithm intercomparisons as follows: (a) NOMAD dataset shown in red for the Λ_{670}^{412} products, with the PSyn synthetic case-1 dataset overlaid in solid gray, the Λ_{670}^{412} EMA algorithm fit to the OCEAN dataset overlaid in dashed black, and a least absolute deviation power-law fit of the NOMAD data in solid black; (b) NNet dataset shown in red for synthetic $K_d(412)/K_d(670)$ products, with the MSyn synthetic case-1 dataset overlaid in solid gray and a least absolute deviation power-law fit of the NNet data in solid black; (c) Validation results from the NOMAD Λ_{670}^{412} EMA with a one-to-one line indicated in solid black; and (d) Validation results from the NNet $K_d(412)/K_d(670)$ EMA with a one-to-one line indicated in solid black. R^2 is derived for all panels using $\log_{10}$-transformed values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

measurement of $L_u(z,\lambda)$ starting just below the sea surface. In addition, new technology to increase the sensitivity of radiometers measuring AOPs has been developed wherein so-called hybridnamic micro-radiometers provide 14 decades of dynamic range (Hooker et al., 2018a). Both of these new technologies expand the vertical sampling interval within attenuating water bodies, thereby improving measurement of $L_u(z,\lambda)$.

For higher $a_{\mathrm{CDOM}}(440)$ waters (>0.1 m^{-1}), RMSD values for the Λ_{780}^{320} and Λ_{670}^{412} EMA algorithms were more improved relative to GSM and QAA, with RMSD of 12.6% and 17.2% derived for the Λ_{780}^{320} and Λ_{670}^{412} algorithms, respectively, and 22.6% and 24.5% for the GSM and QAA algorithms, respectively. The statistics obtained for GSM in the high $a_{\mathrm{CDOM}}(440)$ portion also benefited from the removal of 10 (5.1%) divergent solutions. Overall, the EMA approach was most useful relative to the GSM and QAA algorithms in the coastal and inland waters portion of the dataset or when the dynamic range of water bodies was expanded.

Validation results understate the usefulness of GSM and QAA, which derive other relevant bio-optical parameters, such as phytoplankton absorption and particle backscattering, for which EMA has not yet been applied. For deriving $a_{\mathrm{CDOM}}(440)$ across a global range of water bodies, the GSM and QAA methods generated large positive residuals (in addition to the GSM producing failed or nonphysical retrievals), indicating

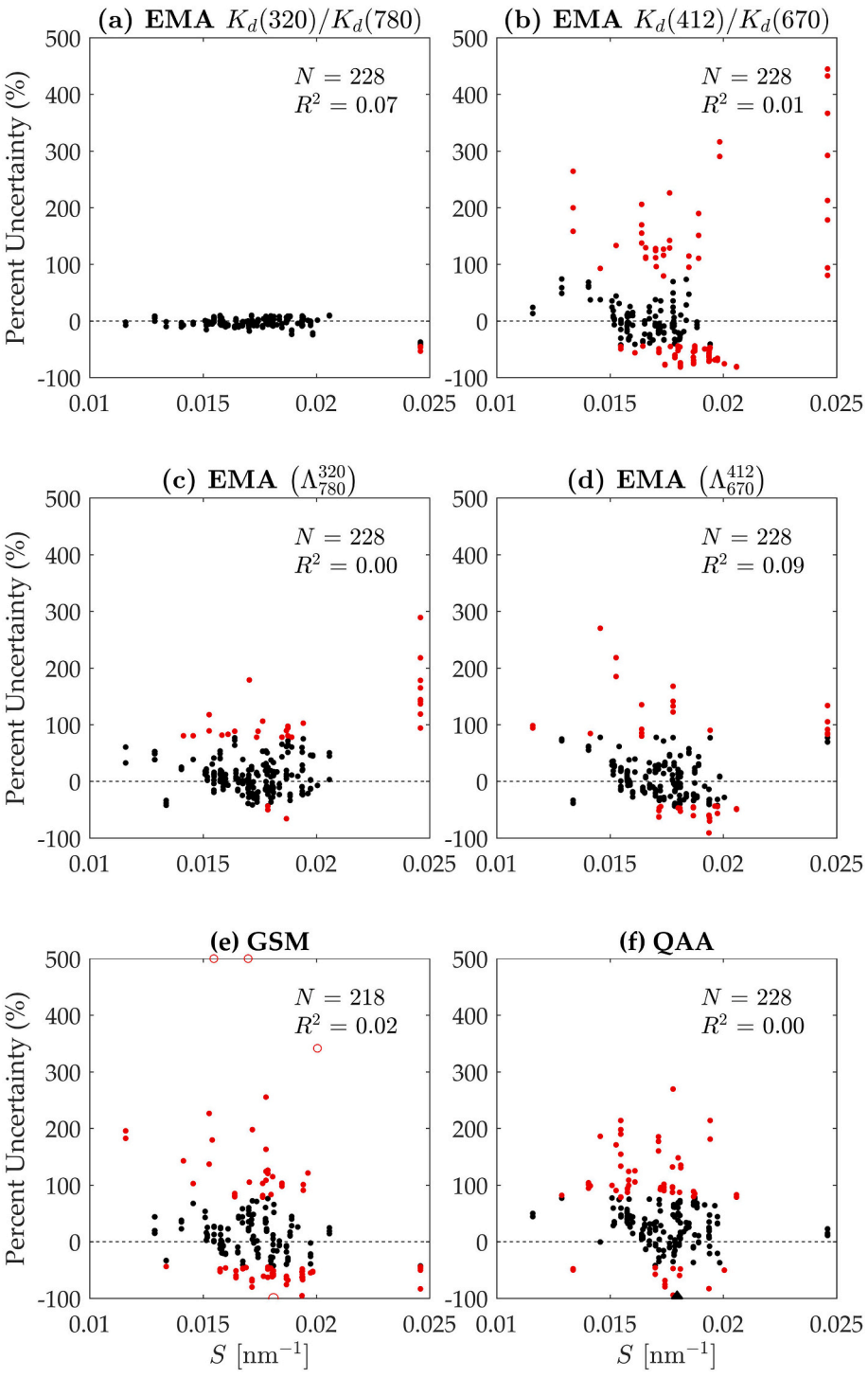

Fig. 8. The relative percent uncertainty from the western U.S. portion of the GLOBC dataset (positive indicates overestimation) as a function of a_{CDOM} spectral slope as follows: (a) $K_d(320)/K_d(780)$ EMA using a linear fit described in Hooker et al. (2020); (b) $K_d(412)/K_d(670)$ EMA using a least absolute deviation power-law fit; (c) Λ_{780}^{320} EMA fit to the OCEAN dataset; (d) Λ_{670}^{412} EMA fit to the OCEAN dataset; (e) GSM with positive failed retrievals shown as open circles at the top of the *y*-axis; and (f) QAA with the median derived a_{CDOM} slope shown as a black diamond on the *x*-axis. Data points with RMSLD below 0.25 are shown in black, and all others are shown in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that internal algorithm parameterizations, which were selected for oceanic or coastal waters, were not consistent across a global range in $a_{CDOM}(440)$ and generally led to overestimation of $a_{CDOM}(440)$. The updates to the GSM and QAA models based on Matsuoka et al. (2013) and Dong et al. (2013), respectively, which allowed for the separation of detrital and dissolved organic matter contributions, improved algorithm performance in nonoceanic waters. For example, if the detrital contributions were not accounted for in oceanic waters, the RMSLD for the GSM and QAA algorithms would increase within the high $a_{CDOM}(440)$ fraction (>0.1 m^{-1}) of waters from 0.278 to 0.376 and from 0.236 to 0.379, respectively.

Additional regionalized improvements to the GSM and QAA algorithms have been proposed, but testing each of these approaches was not the focus of this study, which evaluated a global, rather than regional, algorithmic perspective. A globally consistent set of bio-optical parameters that could be used to extend GSM and QAA across the data range shown herein is unlikely to exist due to variability in the spectral slopes of dissolved and detrital signals, the specific absorption spectra of phytoplankton, and other natural sources of variability across global water bodies, e.g., Bricaud et al. (2012). Methods to estimate parameters including the spectral slope of $a_{CDM}(440)$ or $a_{CDOM}(440)$, however, are presented in Matsuoka et al. (2013) and Dong et al. (2013), respectively, which may improve the applicability of the GSM and QAA algorithms across a greater range of water bodies. The empirical relationship between absorption by non-algal particles and backscattering by particles implemented herein for the GSM algorithms was derived from Arctic waters (Matsuoka et al., 2007), but it did not result in strong bias in the GSM estimation of $a_{CDOM}(440)$ across global water types. The validation results presented herein also suggest that the quality of a_{CDOM} is not a primary source of algorithm uncertainty, because no correlation was found between the spectral slope S of a_{CDOM} and the algorithm residuals, as shown in Fig. 8 using the western U.S. portion of the GLOBC dataset, which was processed by a single laboratory and encompassed the full range in $a_{CDOM}(440)$ and S.

Results from nonconservative waters suggested that the global Λ_{780}^{320} EMA algorithmic approach is applicable to harbors, river mouths, and refilled or flooded lakes and rivers, based on RMSLD within that reported for the validation with the GLOBC dataset. Other water bodies including HABs and drought-stricken lakes are potential candidates for developing regionally tuned Λ_{780}^{320} algorithmic approaches. Waters with high resuspension showed decreased correlation between $a_{CDOM}(440)$ and Λ_{780}^{320}, consistent with expectations based on the sensitivity of above-water methods to differences in target brightness. In keeping with the case-1 versus case-2 perspective for oceanic ecosystems, the conservative versus nonconservative classifications provided a logical partition for algorithm development within inland waters, in which potentially problematic water bodies could be individually assessed for agreement with a global approach or else evaluated for regional tuning.

The Λ_{780}^{320} and Λ_{670}^{412} algorithms were applied to airborne $[L_W(\lambda)]_N$ observations of three environments, which spanned over an order of magnitude in $a_{CDOM}(440)$. Unbiased percent differences at each site were 23.3%, 24.5%, and 11.9% for the Λ_{780}^{320} algorithm and 25.8%, 34.2%, and 25.7% for the Λ_{670}^{412} algorithm at LT, MB, and SFBD, respectively. Increased uncertainty for the Λ_{780}^{320} algorithm at LT was anticipated because of low signal at $[L_W(780)]_N$ and because RMSLD for the Λ_{780}^{320} algorithm was greatest in the clearest water fraction of the GLOBC validation comparison (Sect. 3.3). Results from MB may have been impacted by the difficulty of matching up in situ observations from a wharf (MB) compared to a small vessel (LT and SFBD). Caution is warranted for interpreting the airborne validation results shown herein due to the limited number of available matchups. Comparing the matchup results with the uncertainty propagation derived from spectrally dependent radiometric uncertainties, however, may provide verification in that the results agree as follows: performance was increased for the more spectrally separated (Λ_{780}^{320}) pair within the brighter (SFBD) target (due to the expanded dynamic range of the end-member ratio), while performance was similar between the Λ_{780}^{320} and Λ_{670}^{412} pairs within the darker target (LT), likely due to greater signal in $[L_W(412)]_N$ and $[L_W(670)]_N$ compared with $[L_W(320)]_N$ and $[L_W(780)]_N$.

Previous work subsequent to Hooker et al. (2013) has indicated that the inclusion of longer wavelengths (i.e., above 600 nm) in $a_{CDOM}(440)$ algorithms improves $a_{CDOM}(440)$ observations in complex inland or freshwater environments by helping to separate the effects of detrital particles (Zhu et al., 2014), which is in agreement with the findings herein, as well as those of Mannino et al. (2014). Other studies have suggested that $a_{CDOM}(440)$ algorithm performance may decrease with the addition of information from short wavelengths (i.e., 412 and 443 nm) due to variability in phytoplankton absorption, decreasing radiometric sensitivity, and increasing difficulty of atmospheric correction for data products at shorter wavelengths (e.g., Kutser et al., 2005; Menken et al., 2006; Brezonik et al., 2015). Results discussed herein are not in opposition with these findings because the extension to UV wavelengths is anticipated to mitigate the relative importance of phytoplankton pigmentation, and because satellite radiometric capabilities and accuracy of atmospheric correction were not evaluated herein. However, this work indicates that provided accurate radiometry of end-member wavelengths, increasing the spectral separation within $a_{CDOM}(440)$ algorithms improves the robustness of remote $a_{CDOM}(440)$ observations across a global diversity of water types. Presently, existing airborne capabilities enable high signal-to-noise retrievals at relevant end-member wavelengths (Kudela et al., 2019) to enable remote sensing of $a_{CDOM}(440)$ from airborne platforms based on the ratios of UV and NIR spectral end-members evaluated herein.

5. Conclusion

A band-ratio EMA algorithmic approach—in which spectrally separated end-member pairs, either $[L_W(\lambda)]_N$ or $K_d(\lambda)$, are related to $a_{CDOM}(440)$—was shown to agree with case-1 relationships as well as to predict $a_{CDOM}(440)$ with comparable log-scale statistics across more than three decades of $a_{CDOM}(440)$ dynamic range (0.001–2.305 m^{-1}), including within oceanic, coastal, and inland conservative water bodies. Greater spectral separation of the $\Lambda_{\lambda_2}^{\lambda_1}$ waveband pairs increased the dynamic range expressed as a function of $a_{CDOM}(440)$, which improved sensitivity and flattened the exponential slope of the Λ_{780}^{320} algorithm relative to the Λ_{670}^{412} algorithm, therefore lowering the model uncertainty contribution of the Λ_{780}^{320} algorithm in the uncertainty propagation analysis. The performance of the Λ_{780}^{320} and Λ_{670}^{412} EMA algorithms was greatest relative to the GSM and QAA in the higher $a_{CDOM}(440)$ partition (>0.1 m^{-1}) of the global conservative waters dataset, a range generally associated with case-2 or inland waters. For the nonconservative GLOBN dataset, the Λ_{780}^{320} EMA algorithm produced statistics for harbors, refilled or flooded lakes and rivers, and river mouths that were similar to those produced for the conservative GLOBC dataset.

Airborne remote sensing matchups indicated that the Λ_{780}^{320} algorithm satisfied 25% uncertainty (UPD) in coastal and inland waters spanning an order of magnitude in $a_{CDOM}(440)$. From an airborne remote-sensing perspective, the EMA approach evaluated here also confers multiple practical advantages for inland waters. First, inland waters are overall more optically diverse than oceanic waters, and EMA does not require internal bio-optical parameterizations to specific, conservative water bodies. Second, inland waters contain a greater range in $a_{CDOM}(440)$ than oceanic waters (LT to SFBD spans over an order of magnitude), and EMA maintained stable log-scale uncertainty across a global range in $a_{CDOM}(440)$ for the conservative waters considered in this analysis. Third, many inland waters require increased spatial resolution of remote sensors because of their smaller areas and greater adjacency effects. Sensors with high spatial resolution often sacrifice spectral resolution, which is needed for more complex inversion methods but not for two-band algorithms such as EMA. Fourth, inland waters are frequently shallow, and bottom effects are problematic for central (VIS)

wavelengths. Λ_{780}^{320} is relatively insensitive to bottom effects (as is Λ_{670}^{412} within moderately turbid waters) due to the high attenuation of light by water at UV and NIR wavelengths.

Finally, the robustness of EMA across more than three decades of dynamic range in $a_{\mathrm{CDOM}}(440)$ suggests that algorithms based on widely separated wavelengths, or end members, are less sensitive to variations in optical complexity than algorithms which utilize the central, VIS wavelengths. Future work to improve an understanding of EMA may influence the technological goals for remote sensing of optically complex inland water systems, for example by emphasizing the quality of retrievals at end-member wavebands. Advancing the airborne perspective presented here to an ocean color capability (Appendix A) would require parallel efforts to improve detector technology as well as atmospheric correction, without preconceived brightness approximations (e.g., no black-pixel assumptions) in order to preserve the relevant UV and NIR information.

Funding

This work was supported by the NASA C-HARRIER campaign (grant number NNX17AK89G), as well as by the California State Water Resources Control Board (grant number A17-0553).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful for contributions to this work from (in alphabetical order) Jim Brown (University of Miami), Liane Guild (NASA Ames), Randall Lind (Biospherical Instruments Inc.), Atsushi Matsuoka (University of New Hampshire), John Morrow (Biospherical Instruments Inc.), Kendra Negrey (University of California Santa Cruz), Koji Suzuki (Hokkaido University) and Youhei Yamashita (Hokkaido University). Useful feedback on this manuscript was provided by Meredith McPherson (University of California Santa Cruz), and John Ryan (Monterey Bay Aquarium Research Institute). Finally, this manuscript was greatly improved by feedback provided from three reviewers.

Appendix A. Additional power-law coefficients derived from in situ datasets

Table A.3
The power-law coefficients derived from in situ datasets.

λ Pair	Fit	N	R^2	$u_m(A)$ [m^{-1}]	$u_m(B)$	Dataset
$[L_W(340)]_N/[L_W(780)]_N$	$y = 0.432x^{-0.586}$	112	0.80	0.086	0.067	OCEAN
($x = \Lambda_{780}^{340}$)	$y = 0.394x^{-0.589}$	562	0.87	0.019	0.028	GLOBC
$[L_W(395)]_N/[L_W(710)]_N$	$y = 0.237x^{-0.689}$	112	0.89	0.022	0.049	OCEAN
($x = \Lambda_{710}^{395}$)	$y = 0.244x^{-0.679}$	605	0.92	0.010	0.025	GLOBC
$[L_W(412)]_N/[L_W(710)]_N$	$y = 0.343x^{-0.717}$	112	0.87	0.038	0.049	OCEAN
($x = \Lambda_{710}^{412}$)	$y = 0.359x^{-0.719}$	606	0.90	0.016	0.026	GLOBC

Appendix B. Sampling and processing parameters for the OCEAN dataset

The original legacy datasets (Table B.4) were reprocessed using PROSIT and an EMA perspective, wherein all wavelengths spanning 320 to 780 nm produced data products within the upper 1 m of the water column while respecting the requirements for planar geometry and high vertical sampling resolution, i.e., low descent velocity. These requirements resulted in an approximately 15% reduction in the number of stations with EMA data products. The general sampling requirement was to obtain three optical casts in quick succession at each station. As a result, the removed stations equate to approximately 44% of the original optical casts. To keep self-shading perturbations to a correctable minimum, the C-OPS profiler equipped with the C-PrOPS accessory was deployed with the requisite solar geometry by positioning the L_u aperture towards the Sun and into the principal plane by using differential thrust. For all other profilers, the L_u aperture was correctly positioned using a combination of ship and manual maneuvers with subsequent dissipation time included prior to profiling for any induced turbulence.

Table B.4
Sampling and processing parameters for the OCEAN dataset, as follows: the number of stations for the original (N_O^S) and reprocessed (N_R^S) datasets; the number of optical casts for the original (N_O^C) and reprocessed (N_R^C) datasets; the average bottom depth (z_2) for the extrapolation interval; plus the average vertical tilt (ϕ) and average descent velocity (v) of the profiler in the extrapolation interval. If $\phi \leq 5^\circ$ and $v \leq 0.15$ m s^{-1}, the VSR is 1 cm or less, which ensures the efficacy of deriving data products using an EMA perspective.

Principal Sampling Region	N_O^S	N_R^S	N_O^C	N_R^C	z_2 [m]	ϕ [°]	v [m s^{-1}]
Tropical Pacific Ocean	27	19	39	32	0.80	3.1	0.14
Southern (Indian) Ocean	11	10	37	14	0.76	2.9	0.12
Mid-Atlantic Bight (MAB)	53	43	222	131	0.72	1.4	0.11
Gulf of Maine and MAB	96	87	309	163	0.67	1.0	0.09
Combined	187	159	606	340	0.71	1.4	0.10

Appendix C. Glossary

$a_{CDOM}(440)$	Absorption coefficient of CDOM at 440 nm
$a_{CDM}(440)$	Absorption coefficient of CDOM and detritus at 440 nm
$a(\lambda)$	Total spectral absorption coefficient
$b_b(\lambda)$	Total spectral backscattering coefficient
C-AERO	Compact-Airborne Environmental Radiometers for Oceanography
C-OPS	Compact-Optical Profiling System
C-PrOPS	Compact-Propulsion Option for Profiling Systems
CDOM	Colored dissolved organic matter
Chl *a*	Chlorophyll *a*
CIRPAS	Center for Interdisciplinary Remotely Piloted Aircraft Studies
$D_o(\lambda)$	Downwelling Distribution Function
$E_s(\lambda)$	Global solar irradiance
EMA	End-member analysis
GSM	Garver-Siegel-Maritorena algorithm (Maritorena et al., 2002)
IOP	Inherent Optical Property
$K_d(\lambda)$	Diffuse attenuation coefficient for spectral downward irradiance
$L_i(\lambda)$	Indirect (sky) radiance
$L_T(\lambda)$	Total radiance from the water surface
$L_u(z,\lambda)$	Upwelling radiance
$L_W(\lambda)$	Water-leaving radiance
$[L_W(\lambda)]_N$	Normalized water-leaving radiance
$\Lambda_{\lambda_2}^{\lambda_1}$	The ratio $[L_W(\lambda_1)]_N/[L_W(\lambda_2)]_N$
LSA	Lowest Safe Altitude
LT	Lake Tahoe, California, USA
MB	Monterey Bay, California, USA
NIR	Near-infrared
NOMAD	NASA bio-Optical Marine Algorithm Dataset
S	Spectral slope of CDOM
SFBD	San Francisco Bay Delta, California, USA
SWIR	Short-wave infrared
QAA	Quasi-Analytical-Algorithm (Lee et al., 2002)
UV	Ultra-violet
VIS	Visible
VSR	Vertical Sampling Resolution

References

Aurin, D.A., Dierssen, H.M., 2012. Advantages and limitations of ocean color remote sensing in cdom-dominated, mineral-rich coastal and estuarine waters. Remote Sens. Environ. 125, 181–197.

Bailey, S., 2005. An improved bio-optical data set for ocean color algorithm development and satellite data product variation. Remote Sens. Environ. 98, 122–140.

Bashkatov, A.N., Genina, E.A., 2003. Water refractive index in dependence on temperature and wavelength: a simple approximation. In: Saratov Fall Meeting 2002: Optical Technologies in Biophysics and Medicine IV. International Society for Optics and Photonics, pp. 393–395.

Brezonik, P.L., Olmanson, L.G., Finlay, J.C., Bauer, M.E., 2015. Factors affecting the measurement of cdom by remote sensing of optically complex inland waters. Remote Sens. Environ. 157, 199–215.

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, 43–53.

Bricaud, A., Babin, M., Morel, A., Claustre, H., 1995. Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: analysis and parameterization. J. Geophys. Res. Oceans 100, 13321–13332.

Bricaud, A., Morel, A., Babin, M., Allali, K., Claustre, H., 1998. Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: analysis and implications for bio-optical models. J. Geophys. Res. Oceans 103, 31033–31044.

Bricaud, A., Ciotti, A., Gentili, B., 2012. Spatial-temporal variations in phytoplankton size and colored detrital matter absorption at global and regional scales, as derived from twelve years of seawifs data (1998–2009). Glob. Biogeochem. Cycles 26.

Brown, C.A., Huot, Y., Werdell, P.J., Gentili, B., Claustre, H., 2008. The origin and global distribution of second order variability in satellite ocean color and its potential applications to algorithm development. Remote Sens. Environ. 112, 4186–4203.

Cao, F., Fichot, C.G., Hooker, S.B., Miller, W.L., 2014. Improved algorithms for accurate retrieval of uv/visible diffuse attenuation coefficients in optically complex, inshore waters. Remote Sens. Environ. 144, 11–27.

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, 68–81.

Castillo, C.R., Sarmento, H., Alvarez-Salgado, X.A., Gasol, J.M., Marraséa, C., 2010. Production of chromophoric dissolved organic matter by marine phytoplankton. Limnol. Oceanogr. 55, 446–454.

Chang, G.C., Dickey, T.D., 2004. Coastal Ocean optical influences on solar transmission and radiant heating rate. J. Geophys. Res. Oceans 109.

Del Castillo, C.E., Gilbes, F., Coble, P.G., Müller-Karger, F.E., 2000. On the dispersal of riverine colored dissolved organic matter over the West Florida shelf. Limnol. Oceanogr. 45, 1425–1432.

Dong, Q., Shang, S., Lee, Z., 2013. An algorithm to retrieve absorption coefficient of chromophoric dissolved organic matter from ocean color. Remote Sens. Environ. 128, 259–267.

Gordon, H.R., 1989. Can the lambert-beer law be applied to the diffuse attenuation coefficient of ocean water? Limnol. Oceanogr. 34, 1389–1409.

Gordon, H.R., Brown, O.B., Evans, R.H., Brown, J.W., Smith, R.C., Baker, K.S., Clark, D. K., 1988. A semianalytic radiance model of ocean color. J. Geophys. Res.-Atmos. 93, 10909–10924.

Gregg, W.W., Casey, N.W., 2004. Global and regional evaluation of the seawifs chlorophyll data set. Remote Sens. Environ. 93, 463–479.

Guild, Liane S., Kudela, Raphael M., Hooker, Stanford B., Palacios, Sherry L., Houskeeper, Henry F., 2020. Airborne radiometry for calibration, validation, and research in oceanic, coastal, and inland waters. Front. Environ. Sci. 8, 200. https://doi.org/10.3389/fenvs.2020.585529. https://www.frontiersin.org/article/10.3389/fenvs.2020.585529, 2296-665X.

Hansell, D.A., Carlson, C.A., 2014. Biogeochemistry of Marine Dissolved Organic Matter. Academic Press.

Hooker, S., McClain, C., 2000. The calibration and validation of seawifs data. Prog. Oceanogr. 45, 427–465.

Hooker, S., McClain, C., Firestone, J., Westphal, T., Yeh, E.N., Ge, Y., Firestone, E., 1994. The SeaWiFS bio-optical archive and storage system (SeaBASS), part 1. In: Technical Report TM-104566. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland.

Hooker, S., Morrow, J., Matsuoka, A., 2013. Apparent optical properties of the Canadian Beaufort Sea–part 2: the 1% and 1 cm perspective in deriving and validating aop data products. Biogeosciences 10, 4511–4527.

Hooker, S.B., Lazin, G., Zibordi, G., McLean, S., 2002. An evaluation of above-and in-water methods for determining water-leaving radiances. J. Atmos. Ocean. Technol. 19, 486–515.

Hooker, S.B., Zibordi, G., Berthon, J.F., Brown, J.W., 2004. Above-water radiometry in shallow coastal waters. Appl. Opt. 43, 4254–4268.

Hooker, S.B., McClain, C.R., Mannino, A., Center, G., 2007. NASA strategic planning document: a comprehensive plan for the long-term calibration and validation of oceanic biogeochemical satellite data. In: Technical Report SP-2007-214152. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland.

Hooker, S.B., Lind, R.N., Morrow, J.H., Brown, J.W., Suzuki, K., Houskeeper, H.F., Hirawake, T., Maúre, E.D.R., 2018a. Advances in above- and in-water radiometry, volume 1: enhanced legacy and state-of-the-art instrument suites. In: Technical

Report TP-2018-219033. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland.
Hooker, S.B., Lind, R.N., Morrow, J.H., Brown, J.W., Kudela, R.M., Houskeeper, H.F., Suzuki, K., 2018b. Advances in above- and in-water radiometry, volume 2: autonomous atmospheric and oceanic observing systems. In: Technical Report TP-2018-219033. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland.
Hooker, S.B., Lind, R.N., Morrow, J.H., Brown, J.W., Kudela, R.M., Houskeeper, H.F., Suzuki, K., 2018c. Advances in above-and in-water radiometry, volume 3: hybridspectral next-generation optical instruments. In: Technical Report TP-2018-219033. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland.
Hooker, S.B., Matsuoka, A., Kudela, R.M., Yamashita, Y., Suzuki, K., Houskeeper, H.F., 2020. A global end-member approach to derive $a_{CDOM}(440)$ from near-surface optical measurements. Biogeosciences 17, 475–497. URL: https://www.biogeosciences.net/17/475/2020/ https://doi.org/10.5194/bg-17-475-2020.
Houskeeper, H.F., 2020. Advances in Bio-Optics for Observing Aquatic Ecosystems. Ph.D. thesis. UC Santa Cruz, Santa Cruz, CA, USA. https://escholarship.org/uc/item/17n392pq.
Huot, Y., Morel, A., Twardowski, M., Stramski, D., Reynolds, R., 2008. Particle optical backscattering along a chlorophyll gradient in the upper layer of the eastern south pacific ocean. Biogeosciences 5, 495–507.
Jamet, C., Loisel, H., Dessailly, D., 2012. Retrieval of the spectral diffuse attenuation coefficient kd (λ) in open and coastal ocean waters using a neural network inversion. J. Geophys. Res. Oceans 117.
Jessup, D.A., Miller, M.A., Ryan, J.P., Nevins, H.M., Kerkering, H.A., Mekebri, A., Crane, D.B., Johnson, T.A., Kudela, R.M., 2009. Mass stranding of marine birds caused by a surfactant-producing red tide. PLoS One 4.
Jiao, N., Herndl, G.J., Hansell, D.A., Benner, R., Kattner, G., Wilhelm, S.W., Kirchman, D. L., Weinbauer, M.G., Luo, T., Chen, F., et al., 2010. Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. Nat. Rev. Microbiol. 8, 593.
Joshi, I.D., D'Sa, E.J., 2018. An estuarine-tuned quasi-analytical algorithm (qaa-v): assessment and application to satellite estimates of spm in Galveston bay following hurricane Harvey. Biogeosciences 15, 4065–4086.
Kahru, M., Mitchell, B.G., 2001. Seasonal and nonseasonal variability of satellite-derived chlorophyll and colored dissolved organic matter concentration in the California current. J. Geophys. Res. Oceans 106, 2517–2529.
Kudela, R.M., Chavez, F.P., 2004. The impact of coastal runoff on ocean color during an el nino year in Central California. Deep-Sea Res. II Top. Stud. Oceanogr. 51, 1173–1185.
Kudela, R.M., Hooker, S.B., Houskeeper, H.F., McPherson, M., 2019. The influence of signal to noise ratio of legacy airborne and satellite sensors for simulating next-generation coastal and inland water products. Remote Sens. 11, 2071.
Kutser, T., Pierson, D.C., Kallio, K.Y., Reinart, A., Sobek, S., 2005. Mapping lake cdom by satellite remote sensing. Remote Sens. Environ. 94, 535–540.
Lee, Z., Carder, K.L., Arnone, R.A., 2002. Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. Appl. Opt. 41, 5755–5772.
Lee, Z., Lubac, B., Werdell, J., Arnone, R., 2009. An update of the quasi-analytical algorithm (qaa_v5). In: International Ocean Color Group Software Report, pp. 1–9.
Lee, Z., Hu, C., Shang, S., Du, K., Lewis, M., Arnone, R., Brewin, R., 2013. Penetration of uv-visible solar radiation in the global oceans: insights from ocean color remote sensing. J. Geophys. Res. Oceans 118, 4241–4255.
Liu, C.C., Woods, J.D., Mobley, C.D., 1999. Optical model for use in oceanic ecosystem models. Appl. Opt. 38, 4475–4485.
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 us coast. Remote Sens. Environ. 152, 576–602.
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.
Matsuoka, A., Huot, Y., Shimada, K., Saitoh, S.I., Babin, M., 2007. Bio-optical characteristics of the western arctic ocean: implications for ocean color algorithms. Can. J. Remote. Sens. 33, 503–518.
Matsuoka, A., Hooker, S., Bricaud, A., Gentili, B., Babin, M., 2013. Estimating absorption coefficients of colored dissolved organic matter (cdom) using a semi-analytical algorithm for southern Beaufort Sea waters: application to deriving concentrations of dissolved organic carbon from space. Biogeosciences 10, 917–927.
McKinna, L., Cetinic, I., Chase, A., Werdell, J., 2019. Approach for propagating radiometric data uncertainties through nasa ocean color algorithms. Front. Earth Sci. 7, 176.
Menken, K., Brezonik, P., Bauer, M., 2006. Influence of chlorophyll and humic color on reflectance spectra of lakes: implications for measurement of lake-water properties by remote sensing. Lake Reserv. Manag. 22, 179–190.
Mobley, C., Sundman, L., 2008. Hydrolight 5 Ecolight 5 users' Guide. Sequoia Scientific Inc., US.
Moore, T.S., Campbell, J.W., Feng, H., 2001. A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. IEEE Trans. Geosci. Remote Sens. 39, 1764–1776.
Morel, A., 1988. Optical modeling of the upper ocean in relation to its biogenous matter content (case i waters). J. Geophys. Res. Oceans 93, 10749–10768.
Morel, A., 2009. Are the empirical relationships describing the bio-optical properties of case 1 waters consistent and internally compatible? J. Geophys. Res. Oceans 114.
Morel, A., Antoine, D., 1994. Heating rate within the upper ocean in relation to its bio–optical state. J. Phys. Oceanogr. 24, 1652–1665.
Morel, A., Gentili, B., 2009. A simple band ratio technique to quantify the colored dissolved and detrital organic material from ocean color remotely sensed data. Remote Sens. Environ. 113, 998–1011.
Morel, A., Maritorena, S., 2001. Bio-optical properties of oceanic waters: a reappraisal. J. Geophys. Res. Oceans 106, 7163–7180.
Morel, A., Prieur, L., 1977. Analysis of variations in ocean color. Limnol. Oceanogr. 22, 709–722.
Morel, A., Gentili, B., Claustre, H., Babin, M., Bricaud, A., Ras, J., Tieche, F., 2007. Optical properties of the "clearest" natural waters. Limnol. Oceanogr. 52, 217–229.
Morrison, J.R., Nelson, N.B., 2004. Seasonal cycle of phytoplankton uv absorption at the Bermuda Atlantic time-series study (bats) site. Limnol. Oceanogr. 49, 215–224.
Morrow, J., Booth, C., Lind, R., Hooker, S., 2010. The compact-optical profiling system (c-ops). Advances in measuring the apparent optical properties (AOPs) of optically complex waters, NASA tech. Memo 215856, 42–50.
Nelson, N.B., Siegel, D.A., 2013. The global distribution and dynamics of chromophoric dissolved organic matter. Annu. Rev. Mar. Sci. 5, 447–476.
Nelson, N.B., Carlson, C.A., Steinberg, D.K., 2004. Production of chromophoric dissolved organic matter by sargasso sea microbes. Mar. Chem. 89, 273–287.
Nelson, N.B., Siegel, D.A., Carlson, C.A., Swan, C.M., 2010. Tracing global biogeochemical cycles and meridional overturning circulation using chromophoric dissolved organic matter. Geophys. Res. Lett. 37.
O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., McClain, C., 1998. Ocean color chlorophyll algorithms for seawifs. J. Geophys. Res. Oceans 103, 24937–24953.
Pan, X., Zimmerman, R.C., 2010. Modeling the vertical distributions of downwelling plane irradiance and diffuse attenuation coefficient in optically deep waters. J. Geophys. Res. Oceans 115.
Pegau, W.S., 2002. Inherent optical properties of the central arctic surface waters. J. Geophys. Res. Oceans 107 (SHE-16).
Prieur, L., Sathyendranath, S., 1981. An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter, and other particulate materials 1. Limnol. Oceanogr. 26, 671–689.
Rochelle-Newall, E., Fisher, T., 2002. Production of chromophoric dissolved organic matter fluorescence in marine and estuarine environments: an investigation into the role of phytoplankton. Mar. Chem. 77, 7–21.
Sauer, M.J., Roesler, C., Werdell, P., Barnard, A., 2012. Under the hood of satellite empirical chlorophyll a algorithms: revealing the dependencies of maximum band ratio algorithms on inherent optical properties. Opt. Express 20, 20920–20933.
Seegers, B.N., Stumpf, R.P., Schaeffer, B.A., Loftin, K.A., Werdell, P.J., 2018. Performance metrics for the assessment of satellite data products: an ocean color case study. Opt. Express 26, 7404–7422.
Sibson, R., 1981. Interpreting multivariate data. John Wiley & Sons.
Siegel, D., Maritorena, S., Nelson, N., Behrenfeld, M., McClain, C., 2005. Colored dissolved organic matter and its influence on the satellite-based characterization of the ocean biosphere. Geophys. Res. Lett. 32.
Siegel, D.A., Behrenfeld, M.J., Maritorena, S., McClain, C.R., Antoine, D., Bailey, S.W., Bontempi, P.S., Boss, E.S., Dierssen, H.M., Doney, S.C., et al., 2013. Regional to global assessments of phytoplankton dynamics from the seawifs mission. Remote Sens. Environ. 135, 77–91.
Smith, R.C., Baker, K.S., 1981. Optical properties of the clearest natural waters (200–800 nm). Appl. Opt. 20, 177–184.
Spencer, R.G., Butler, K.D., Aiken, G.R., 2012. Dissolved organic carbon and chromophoric dissolved organic matter properties of rivers in the usa. J. Geophys. Res. Biogeosci. 117.
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, 69–88.
Twardowski, M.S., Claustre, H., Freeman, S.A., Stramski, D., Huot, Y., 2007. Optical backscattering properties of the "clearest" natural waters. Biogeosci. Discuss. 4, 2441–2491. URL: https://hal.archives-ouvertes.fr/hal-00330254.
Vantrepotte, V., Danhiez, F.P., Loisel, H., Ouillon, S., Mériaux, X., Cauvin, A., Dessailly, D., 2015. Cdom-doc relationship in contrasted coastal waters: implication for doc retrieval from ocean color remote sensing observation. Opt. Express 23, 33–54.
Wang, Y., Shen, F., Sokoletsky, L., Sun, X., 2017. Validation and calibration of qaa algorithm for cdom absorption retrieval in the changjiang (yangtze) estuarine and coastal waters. Remote Sens. 9, 1192.
Werdell, P., Bailey, S., 2002. The SeaWiFS bio-optical archive and storage system (SeaBASS): current architecture and implementation. In: Technical Report TM-2002-211617. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland.
Zhu, W., Yu, Q., Tian, Y.Q., Becker, B.L., Zheng, T., Carrick, H.J., 2014. An assessment of remote sensing algorithms for colored dissolved organic matter in complex freshwater environments. Remote Sens. Environ. 140, 766–778.