Article

Validation of Remote-Sensing Algorithms for Diffuse Attenuation of Downward Irradiance Using BGC-Argo Floats

Charlotte Begouen Demeaux * and **Emmanuel Boss**

School of Marine Sciences, University of Maine, Orono, ME 04469, USA
* Correspondence: charlotte.begouen@maine.edu

Abstract: Estimates of the diffuse attenuation coefficient (K_d) at two different wavelengths and band-integrated (PAR) were obtained using different published algorithms developed for open ocean waters spanning in type from explicit-empirical, semi-analytical and implicit-empirical and applied to data from spectral radiometers on board six different satellites (MODIS-Aqua, MODIS-Terra, VIIRS–SNPP, VIIRS-JPSS, OLCI-Sentinel 3A and OLCI-Sentinel 3B). The resultant K_ds were compared to those inferred from measurements of radiometry from sensors on board autonomous profiling floats (BGC-Argo). Advantages of BGC-Argo measurements compared to ship-based ones include: 1. uniform sampling in time throughout the year, 2. large spatial coverage, and 3. lack of shading by platform. Over 5000 quality-controlled matchups between K_ds derived from float and from satellite sensors were found with values ranging from 0.01 to 0.67 m^{-1}. Our results show that although all three algorithm types provided similarly ranging values of K_d to those of the floats, for most sensors, a given algorithm produced statistically different K_d distributions from the two others. Algorithm results diverged the most for low K_d (clearest waters). Algorithm biases were traced to the limitations of the datasets the algorithms were developed and trained with, as well as the neglect of sun angle in some algorithms. This study highlights: 1. the importance of using comprehensive field-based datasets (such as BGC-Argo) for algorithm development, 2. the limitation of using radiative-transfer model simulations only for algorithm development, and 3. the potential for improvement if sun angle is taken into account explicitly to improve empirical K_d algorithms. Recent augmentation of profiling floats with hyper-spectral radiometers should be encouraged as they will provide additional constraints to develop algorithms for upcoming missions such as NASA's PACE and SBG and ESA's CHIME, all of which will include a hyper-spectral radiometer.

Keywords: radiometry; diffuse attenuation coefficient; algorithm validation; ocean optics; BGC-Argo

1. Introduction

The spectral diffuse attenuation coefficient of downward irradiance ($K_d(\lambda)$, see Table 1 for notation) describes how radiation is attenuated near the ocean's surface and hence is important in regulating physical and biogeochemical processes in the upper ocean such as heating, photosynthesis and photo-chemistry. As an apparent optical property (AOP), $K_d(\lambda)$ primarily varies with the water's inherent optical properties (IOPs), such as the absorption and the volume scattering function coefficients, and is affected, to a much smaller extent by sun angle, clouds, surface waves and other environmental conditions, therefore characterizing the optical properties of a water body [1].

Retrieving $K_d(\lambda)$ from remote sensing provides constraints on the above processes in assimilative global biogeochemical models. Both NASA and ESA provide the diffuse attenuation at 490 nm, $K_d(490)$, as a Level-2 product, using empirical algorithms that are a function of the ratio of blue and green wavelengths of the remotely sensed reflectance (R_{rs}) computed for a given sensor. The primary purpose of our study is to perform a comprehensive evaluation of the $K_d(\lambda)$ products on a global scale.

Profiling BGC-Argo floats provide a novel and extensive dataset to assess algorithms' performance with as they are unbiased temporally (e.g., sample equally in winters and summers) and span the globe. Additionally, radiometers are mounted in a way such as not to be shadowed by the floats, unlike the inevitable shading (of, at a minimum, some skylight) present when radiometers are deployed from research vessels.

In this paper, we expand on the pioneering analysis of [2] who similarly used radiometry data on profiling floats to test $K_d(490)$ and $K_d(PAR)$ algorithms by: 1. including a significantly larger dataset (40,738 profiles from 225 different floats), 2. comparing in situ K_ds with six different satellite missions, 3. using Level 2 satellite data, 4. evaluating the algorithm for $K_d(412)$, and 5. including a published algorithm based on a neural network [3] for $K_d(\lambda)$ as well as an additional empirical algorithm [4] to compute $K_d(PAR)$.

Table 1. Symbols and definitions.

Symbol	Definition
APD	Absolute Percentage Difference
$a(\lambda)$	Absorption
$b_b(\lambda)$	Backscattering
Chla	Chlorophyll *a* concentration
$E_d(\lambda)$	Downwelling irradiance
$E_d(\lambda, 0^-)$	Downwelling irradiance below the surface
$E_d(\lambda)^*$	Simulated downwelling irradiance with spectral response function
ηw	Relative contribution of molecular scattering to total scattering
iPAR	Instantaneous Downwelling Photosynthetically Available Radiation
$K_{bio}(490)$	Diffuse attenuation coefficient at 490 nm due to material co-varying with chlorophyll
K_d^{ESA}	ESA L2 operational algorithm based on Morel
$(K_d^{float}, K_d^{R_{rs}})$	Centroid pair , average of $K_d^{R_{rs}}$ and K_d^{float}
$K_d^{float}(PAR)$	E_d-retrieved $K_d(PAR)$
$K_d(\lambda)$	The layer averaged diffuse attenuation coefficient
$K_d(\lambda)_{z_{pd}}$	The layer averaged diffuse attenuation coefficient from the surface to the penetration depth
K_d^{Lee}	Lee's semi-analytical algorithm
$K_d^{Lee05}(PAR)$	Lee's semi-analytical algorithm retrieving PAR
$K_d^{Morel07}(PAR)$	Morel's empirical algorithm retrieving PAR
K_d^{NASA}	NASA L2 operational algorithm based on Austin& Werdell
K_d^{NN}	Jamet's neural network algorithm
$K_d^{R_{rs}}(\lambda)$	R_{rs}-retrieved $K_d(\lambda)$
$K_d^{R_{rs}}(PAR)$	R_{rs}-retrieved $K_d(PAR)$
$K_w(490)$	Diffuse attenuation coefficient at 490 nm of pure water
L_w^*	Simulated water leaving radiance corrected for the spectral response function
L_w	Water leaving radiance
PAR	Downwelling Photosynthetically Available Radiation
$R(0^-)$	Irradiance reflectance below the surface
R_{rs}	Remote Sensing Reflectance
R_{rs}^T	Satellite Remote sensing reflectance
$\sigma_{490,560}$	Water leaving reflectance ratio of 490/560
θ	Sun zenith angle over the surface of the ocean
χ_N, η_N, α_N	Constants for computation of $K_d(PAR)^{Lee05}$
z_{pd}	Penetration depth

2. Materials and Methods

2.1. Bio-Argo Data

A portion of the fleet of BGC-Argo floats were equipped with a multispectral downwelling irradiance radiometer (OCR-504, Sea-Bird Scientific, Bellevue, WA, USA) that measures E_d at 3 different channels (380 (or 443), 412, 490 nm [W m^{-2}]) and *PAR* (the integrated irradiance from 400 to 700 nm, [μmol quanta m^{-2} s^{-1}], [5]). Most floats analyzed

in this study acquired a vertical profile of radiance from 0–250 m and surfaced once a day around solar noon.

2.1.1. Retrieval of BGC-Argo Data

A large number of Argo float profiles (40,738, Table A1) with downwelling irradiance ($E_d(\lambda)$), or downwelling Photosynthetically Available Radiation (*PAR*), were retrieved in April 2022 from the Argo Global Assembly Center (GDACs) of Ifremer, where profiles are available $\approx$ 24 h after acquisition (ftp://ftp.ifremer.fr/ifremer/argo/etc/argo-synthetic-profile, accessed on 3 April 2022).

2.1.2. Processing of BGC-Argo Data

After retrieval, all profiles were quality controlled (QC) following [6]. This QC removes a constant dark signal along the $E_d(\lambda)$ or *PAR* by identifying the depth below which all subsequent measurements were normally distributed. The mean $E_d(\lambda)$ measured at that depth is then considered the dark value and subsequently removed from all $E_d(\lambda)$ values of a given profile. These dark values are on the order of 0.001% of surface values, and thus this procedure has a negligible effect on results here. Profiles where the polynomial fit had an $R^2 > 0.995$ were kept. The mean of the residuals from the fit was computed for each profile that was kept, and data points with departures from the fit $> 2 \times \sigma$ (standard deviation) from the mean of the residuals were identified as the product of clouds/spikes and removed from the profile, per the procedure described in [2,6]. A second order polynomial was then performed, and profiles with an $R^2 > 0.998$ passed. Within each profile that passed this second QC step, data points with a residual $> \sigma$ (one standard deviation) from the mean of the residuals were deemed affected by waves focusing and smaller clouds (not identified by the first polynomial fitting) and were removed. Only profiles flagged as "1", i.e., considered as good and not requiring any further modification were used in this study. A total of 29,004 profiles of E_d and/or *PAR* (with data for at least one channel) passed this quality control (QC) (Tables 2 and A1, Figure 1).

Table 2. Number of BGC-Argo float profiles passing the QC tests for each wavelength.

Wavelengths (λ)	N ($K_d(\lambda)$)
380 nm	22,167
412 nm	19,813
490 nm	14,876
PAR	12,552
Total unique float profiles	29,004

2.1.3. Computation of the Diffuse Attenuation Coefficients

The surface layer averaged diffuse attenuation coefficient $K_d(\lambda)$ and $K_d(PAR)$, from right below the surface (0^-) to a depth z are defined as:

$$K_d(\lambda)_z = \frac{1}{z} \times \ln\left(\frac{E_d(0^-)}{E_d(z)}\right) \tag{1}$$

$$K_d(PAR)_z = \frac{1}{z} \times \ln\left(\frac{iPAR(0^-)}{iPAR(z)}\right) \tag{2}$$

with *iPAR* the instantaneous Photosynthetically Available Radiation (Table 1) measured by BGC-Argo floats. In many studies, $K_d(\lambda)$ is defined as the layer averaged $K_d(\lambda)$ from surface to the penetration depth (z_{pd}, Table 1); z_{pd} is the depth above which 90% of the irradiance originates [7] and is the depth where downwelling irradiance reaches $1/e$ of the surface's value ($E_d(\lambda, z_{pd}) = E_d(\lambda, 0^-) \times e^{-1}$) [7]. The z_{pd} is spectrally dependant and must therefore be computed for each wavelength. Pressure was converted to depth using the MATLAB function pts2z.m from USCD's Matlab Oceanography toolbox, which computes pressure from depth, temperature and salinity. To derive $K_d(\lambda)_{z_{pd}}$ from $E_d(\lambda)$, a

non-linear, least-squares exponential fit to measure $E_d(\lambda)$ was computed adapting a Matlab function from our lab (http://misclab.umeoce.maine.edu/software.php, last accessed on 21 July 2022). This method (hereon lsq-method) assumes that $K_d(\lambda)_z$ is constant from 0 to z_{pd}. To test the robustness of this method retrieval of $K_d(\lambda)$, two other methods were also evaluated. Since BGC-Argo floats do not acquire radiometry at $z = 0$, $E_d(\lambda, 0^-)$ was computed for each profile at each available wavelength by either extrapolating to the surface using a second degree polynomial fit of $\ln(E_d(\lambda)_z)$ as a function of z in the upper 10 m of the water column (polynomial method) or by extrapolating a linear regression of $\ln(E_d(\lambda, z))$ following [2] (linear method). Inputting $z_{pd}(\lambda)$ in Equation (1) results in the following expression of $K_d(\lambda)_{z_{pd}}$:

$$K_d(\lambda)_{z_{pd}} = \frac{1}{z_{pd}}. \tag{3}$$

Those three methods were applied only if there were more than 5 data points in the upper 10 m of the water column, in order to get a robust extrapolation [2] and to make sure that the least square fit is not driven by outliers.

After comparing the results of $K_d(\lambda)_{z_{pd}}$ retrieval by the three different methods (Figure A3), and assessing that there was no major difference/bias associated with one given method in $K_d(\lambda)_{z_{pd}}$ retrieval, we decided to use the lsq-method that did not involve extrapolating to the surface. Profiles that resulted in $K_d(\lambda)_{z_{pd}} < K_d(\lambda)$ of pure water ($K_w(490) = 0.016$ m^{-1}, [8]) were removed as they were deemed outliers (160 profiles out of the 29,004 retrieved profiles passing QC (Table 2)).
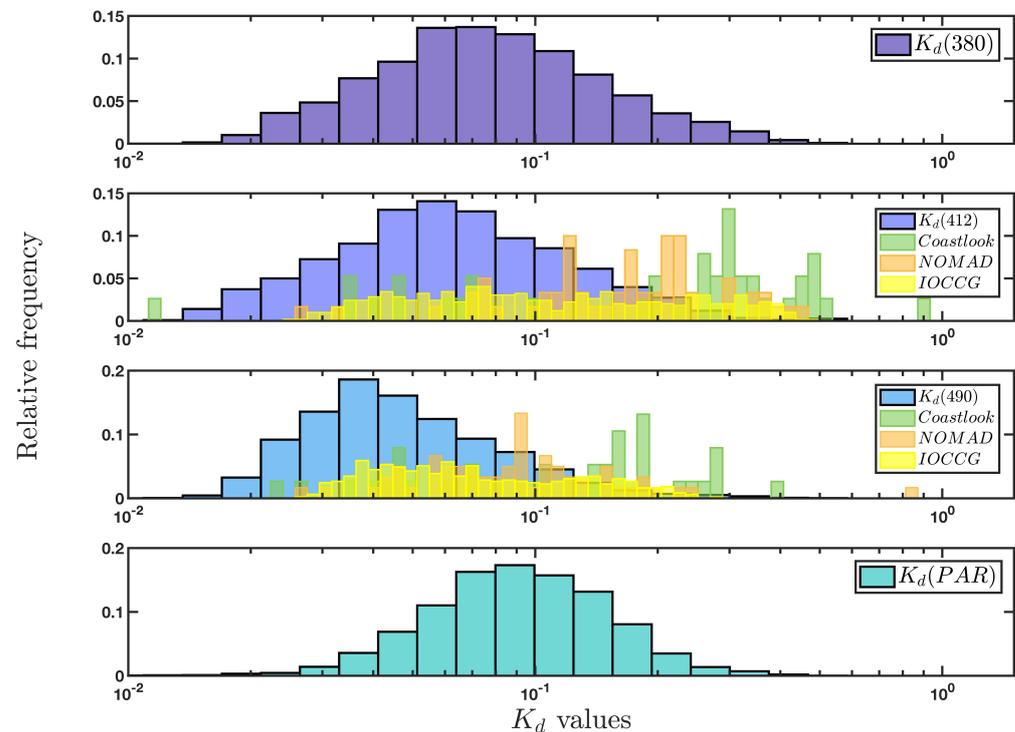
Figure 1. Histogram of the frequency distribution of $K_d^{float}(\lambda)$ values for the BGC-argo floats. The vertical axis represents the probability of the occurrence of $K_d^{float}(\lambda)$ within a specific value bin relative to the total number of profiles (N) for each specific wavelength and database. For 412 nm and 490 nm, the relative frequency of $K_d(\lambda)$ was added for the Case 1 waters ($\frac{R_{rs}(490)}{R_{rs}(555)} > 0.85$) present in the NOMAD, COASTLOOC and the IOCCG (simulated) datasets.

From here on, $K_d(\lambda)_{z_{pd}}$ derived from float's $E_d(\lambda)$ profiles using the lsq-method will be referred to as $K_d^{float}(\lambda)$. The $iPAR(0^-)$ was calculated following [2] by regressing a second degree polynomial of $\ln(iPAR(z))$ with depth (z) in the upper 10 m and extrapolating to the surface. $K_d(PAR)$ was then calculated over the z_{pd} of PAR:

$$K_d(PAR)^{float}_{z_{pd}} = \frac{1}{z_{pd}(PAR)} \times \ln\left(\frac{iPAR(0^-)}{iPAR(z_{pd})}\right). \tag{4}$$

We recognize that even in a layer where IOPs are constant, $K_d(PAR)$ varies significantly with depth, much more so than $K_d(\lambda)$ [9].

In total, $K_d(\lambda)$ and/or $K_d(PAR)$ were computed for 25,090 profiles out of the 29,004 passing the QC (Table 2). Specifics on how many profiles were filtered at each QC step can be found in the Appendix A (Table A1). $K_d(380)$ was computed from $E_d(380)$ data acquired from BGC-Argo floats, but no sensor had compatible $R_{rs}(\lambda)$ wavelengths, so its retrieval from Ocean Color Satellite is not evaluated in this study.

2.2. Additional Data

In situ databases used to calibrate and validate empirical algorithms of K_d were also used to compare to the extent of the float's coverage and values, as well as to assess the correct implementation of published algorithms.

The NOMAD (NASA bio-Optical Marine Algorithm Data set (https://seabass.gsfc.nasa.gov/wiki/NOMAD, last accessed on 14 February 2022) is the in situ bio-optical dataset used in the development of the operational $K_d(490)^{R_{rs}}$ algorithm of NASA and consists of approximately 3400 stations with water-leaving radiances, surface irradiance, E_d measurements and the associated auxiliary metadata, all collected before 2002 [10]. In this study, only the stations that had both $E_d(\lambda)$ and $R_{rs}(\lambda)$ were used, resulting in nearly 1000 $K_d(490)$ data points all located in the North Atlantic (Figure 2).

The COASTLOOC dataset consists of 338 stations from the 6 COASTLOOC campaigns which took place in European waters (Adriatic Sea, Baltic Sea, English Channel, Mediterranean Sea and North Sea) and the Atlantic coast in 1997 and 1998, at which were measured several IOPs over a large variety of waters, including $K_d(\lambda)$ values and irradiance reflectance below the surface $R(0^-)$ at several wavelengths over a large variety of water types.

A synthetic dataset, from here on called the IOCCG dataset, was developed by a committee of the International Ocean Color Coordinating Group (IOCCG). This dataset contains both IOPs and AOPs (derived from IOPs using the Hydrolight radiative transfer code) and covering a wide range of variability encountered in natural waters with the aim of validating and evaluating algorithms' retrieval of IOPs from R_{rs} [11]. 1000 different R_{rs} spectra (and associated $K_d(\lambda)$) were retrieved with a solar zenith angle of either 30° or 60° from https://ioccg.org/what-we-do/ioccg-publications/ioccg-reports/synthesized-dataset-from-ioccg-report-5/, last accessed on 18 February 2022.

2.3. Regional Analysis

The global open-ocean biomes in which each float was localized were identified. These biomes represent regions of similar biogeochemical characteristics based mainly on Mixed Layer Depth (MLD), Chlorophyll a (Chl *a*), Sea Surface Temperature (SST) and ice coverage over 1998–2010 [12]. Out of the 17 biomes characterized in [12], float profiles were present in 11 biomes at all of the measured wavelengths. The Mediterranean Sea was not represented in the original publication and was therefore added. Biome 18 ('W MED') corresponds to the Western Mediterranean and biome 19 ('E MED') to the Eastern Mediterranean + the Black Sea region, with the division between the two basins being the Sicily Strait (Figure 2). Since floats were not homogeneously distributed spatially, an analysis within biomes in addition to the global analysis allowed us to determine whether biases associated with specific algorithms were local or global in nature. Float profiles that were not in any biomes

($N = 297$) were taken into account in the global analysis but excluded from the regional analysis. The profiles not located in any biome were in regions that did not fit in the abovementioned criteria for any biome, usually because of land influence or being part of a marginal sea (e.g., the Red Sea, the Caribbean Sea, the Gulf of Mexico) [12].
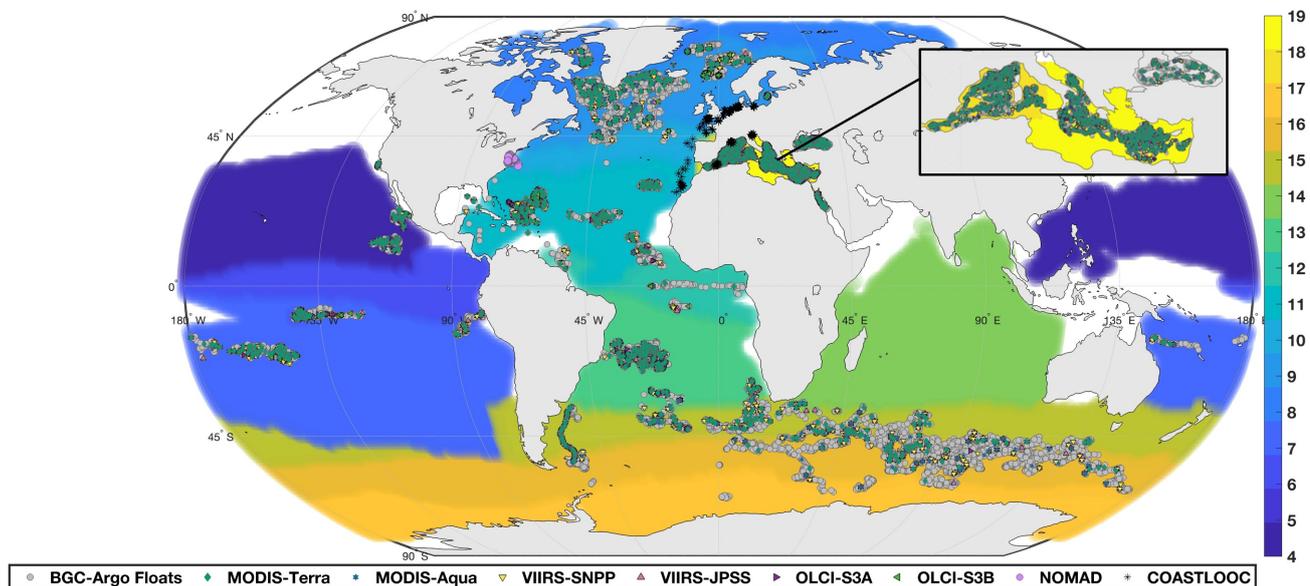
Figure 2. Map of Bio-Argo float observations and in situ observations used for algorithm development colored by matchups with individual satellite sensors or dataset. Background represents oceanic biomes 1–17 of [12]. Insert represent the Mediterranean Sea and the two biomes we added. The names associated with each biome number and color are listed in Table A2.

2.4. Satellite Data Products

Level 2 Remote Sensing Reflectance (R_{rs}, sr^{-1}) and $K_d(490)$ products were retrieved from three sensors onboard six satellites: MODIS (MODerate resolution Imaging Spectroradiometer, pixel size = 1×1 km^2) onboard the Aqua and Terra satellites, VIIRS (Visible Infrared Imaging Radiometer Suite, pixel = 750×750 m^2) onboard Suomi-NPP and JPSS and OLCI (Ocean and Land Color Instrument, pixel = 300×300 m^2) on board Sentinel-3A and Sentinel-3B. Scenes were downloaded using the getOC utility (https://github.com/OceanOptics/getOC, accessed on 16 March 2021). Using this utility, we acquired all the satellite images that for any given day coincided with the location of a BGC-Argo Float profile within a radius of less than 1 nautical mile.

To matchup satellite and float data, we used published criteria [13,14] to assess the suitability of the satellite data. The satellite image must have been taken within 3 h from float surfacing (floats acquire measurements on their upward profile) and the solar zenith at the time of the satellite overpass must be less than 75°. A pixel box around the float location of 25 km^2 is then selected (5×5 for MODIS, 7×7 for VIIRS, 15×15 for OLCI) in which half of the pixel must be unflagged. The quality flags selected are the standard Level 2 Ocean color for MODIS and VIIRS (https://oceancolor.gsfc.nasa.gov/atbd/ocl2flags/, accessed on 18 May 2021) and the recommended Level 2 flags from the EUMETSAT product for OLCI (https://www-cdn.eumetsat.int/files/2020-04/pdf_s3_pn_olci_l2m_001.pdf, accessed on 11 November 2021). Additionally, the coefficient of variation for the R_{rs} bands between 412 nm and 555 nm and for the aerosol optical thickness (aot) at 865 nm for the selected pixels must be inferior to 15%. In total, we found 1802 matchups for MODIS-Terra, 2144 matchups for MODIS-Aqua, 3290 matchups for VIIRS- SNPP, 2445 matchups for VIIRS-JPSS, 651 matchups for OLCI on Sentinel 3A and 382 on Sentinel 3B.

When a BGC-Argo Float profile matched-up with several different images from the same satellite, only the matchup with the image that was taken closest in time to the float surfacing was retained to minimize the matchups bias of location. Since some of the K_d algorithms were developed specifically for optically clear water (Case 1), only the R_{rs} spectra that emulated Case 1 waters were used forward. Case 1 waters were defined as waters where $\frac{R_{rs}(490)}{R_{rs}(555)} > 0.85$ [15].

Quality Control of the Satellite Derived R_{rs} Spectra

A further quality assurance step was conducted on the R_{rs} spectra of the matchups following [16]. Each R_{rs} spectrum was evaluated based on the harmonic mean across a range of previously identified wavelengths, effectively resulting in one "summary color". When compared to the range of the spectra's averages observed in global optically deep waters, spectra with a harmonic mean outside the expected range can be identified and removed. We removed all spectra that had an a score whose absolute value was larger than 0.2 [16]. Less than 3.2% of matchup spectra were removed during this quality control step (Table A1).

2.5. K_d Algorithms Tested

2.5.1. Explicit Empirical Algorithms

The algorithms tested include NASA's $K_d(\lambda)$ operational product (based on the algorithm of Austin and Petzold [17]) and available as a L2 product for the MODIS and VIIRS sensors https://oceancolor.gsfc.nasa.gov/atbd/kd_490/, accessed on 18 February 2022. It is an empirical relationship derived from in situ measurements of $K_d(490)$ based on the NOMAD dataset using blue-to-green band ratios of R_{rs}. In situ radiometer data were processed similarly to the float profiles (e.g., for NOMAD). NASA products for each satellite are based on derived empirical relationships between log-transformed $K_d(490)$ and the log-transformed R_{rs} band-ratio (Equations (5) and (6), [4,8,10]). This algorithm was developed solely for $K_d(490)$ and will be referred to at as K_d^{NASA}.

$$\log_{10}(K_{bio}(490)) = a_0 + \Sigma_{i=1}^{4} A_i \left(\log_{10} \left(\frac{R_{rs}(\lambda_{blue})}{R_{rs}(\lambda_{green})} \right) \right)^i \tag{5}$$

$$K_d^{NASA} = K_{bio}(490) + K_w(490), \tag{6}$$

where $K_w(490) = 0.0166 \, \text{m}^{-1}$ denotes $K_d(490)$ due to seawater only (pure water). The OLCI product also utilizes a similar empirical algorithm to compute $K_d(490)$ [4] based on the blue-to-green ratio of normalized water leaving reflectance ($\sigma_{490,560}$), and the coefficients A_i tuned to OLCI. The A_i coefficients were also derived using the NOMAD dataset, where measured R_{rs} data were input into the algorithm, and the retrieved $K_d(490)^{R_{rs}}$ was compared to the measured $K_d(490)$ from the database. The coefficients that resulted in the best fit were identified and kept. This OLCI product (https://sentinel.esa.int/documents/247904/0/OLCI_L2_ATBD_Ocean_Colour_Products_Case-1_Waters.pdf/4e1c1cd4-697e-4491-b574-777a791b5141, accessed on 20 May 2021) is referred to as K_d^{ESA}:

$$K_d(490)^{ESA} = K_w(490) + 10^{\Sigma_{i=0}^{4} A_i (\log_{10}(\sigma_{490,560}^i))} \tag{7}$$

2.5.2. Lee's Semi-Analytical Algorithm

The second type of algorithm evaluated is the semi-analytical algorithm described in [18]. This algorithm relies on the assumption that the variability of the diffuse attenuation coefficient comes from the variability in the IOPs and the solar zenith angle (θ) and is derived from radiative-transfer computations using the Hydrolight model. The absorption $a(\lambda)$ and backscattering coefficients $b_b(\lambda)$ at any measured wavelength were retrieved from $R_{rs}(\lambda)$ using the Quasi-analytical algorithm (QAA) version 6 (https://www.ioccg.org/groups/Software_OCA/QAA_v6_2014209.pdf, accessed on 17 November 2021), [19])

and are used as inputs to compute $K_d(\lambda)$ using the following equation, revised in 2013 [20]. This algorithm will be referred to as K_d^{Lee05}:

$$K_d^{Lee05}(\lambda) = (1 + 0.005\theta) \times a(\lambda) + 4.259 \times (1 - 0.265\eta_w(\lambda)) \times (1 - 0.52 \times e^{-10.8a(\lambda)}) \times b_b(\lambda), \tag{8}$$

with $\eta_w(\lambda) = \frac{b_{bw}(\lambda)}{b_b(\lambda)}$ the relative contribution of molecular scattering to the total backscattering coefficient. Model parameters were derived from Hydrolight simulations. Before inverting with QAA for IOPs, the effects of Raman scattering had to be accounted for, as the bias caused by Raman scattering on R_{rs} can be as high as 20% at visible wavelengths [21]. Satellite R_{rs} data were corrected for Raman scattering following [20].

2.5.3. Jamet's Neural Network Algorithm

The third type of algorithm evaluated is empirically derived using a neural network with $R_{rs}(\lambda)$ as an input parameter. This 2-layer algorithm was originally designed with the visible bands of the SeaWiFs sensor [3]; however, a newer version, used here, was developed for MODIS and OLCI [22]. It is denoted as K_d^{NN}. This new version distinguishes between Case 1 and Case 2 water using the ratio of $R_{rs}(490)/R_{rs}(555)$ [15]. Only the Case 1 version is analyzed here. No version has been published for the VIIRS sensor; thus this algorithm was not used with VIIRS data in this study. The algorithm was trained on the IOCCG, NOMAD and BOUM datasets, and its performance was validated using the COASTLOOC dataset [3].

2.6. PAR Algorithms

Two published algorithms to estimate $K_d(PAR)$ from remote sensing data were evaluated here. The first, denoted as "Morel07", is based on an empirical relationship between $K_d(490)$ and $K_d(PAR)$ that takes into account variation of the sun zenith angle (angle between the sun and the horizon) and was derived from an in situ dataset (the LOV dataset, containing hyperspectral diffuse attenuation coefficient and irradiance reflectance derived from planar irradiance measurements [4]). The following equation was derived for a layer of depth $z = z_{pd}$ and with $K_d^{Lee}(490)$ used for the $K_d(490)$ value as per [23]:

$$K_d^{Morel07}(PAR) = 0.0864 + 0.884 \times K_d^{Lee}(490) - \frac{0.00137}{K_d^{Lee}(490)} \tag{9}$$

The second algorithm (denoted as "Lee05") is a semi-analytical algorithm based on the retrieval of IOPs using QAA [19] (as above). The layer average $K_d^{Lee05}(PAR, z)$ can be retrieved between the surface and any depth (z) but here is computed for the $0 - z_{pd}$ layer:

$$K_d^{Lee05}(PAR) = K_1 + \frac{K_2}{\sqrt{1 + z_{pd}}} \tag{10}$$

K_1 and K_2 are computed using $a(490)$, $b_b(490)$ and θ as inputs, and the constants χ_{0-2}, ζ_{0-2} and α_{0-2} were obtained based on Hydrolight simulations [24]:

$$K_1 = \left[\chi_0 + \chi_1 \times a(490)^{0.5} + \chi_2 \times b_b(490)\right](1 + \alpha_0 \times sin(\theta)) \tag{11}$$

$$K_2 = [\zeta_0 + \zeta_1 \times a(490) + \zeta_2 \times b_b(490)](\alpha_1 + \alpha_2 \times cos(\theta)). \tag{12}$$

2.7. Statistical Performance Metrics

The following statistical metrics were used to evaluate the performance of each of the studied algorithms. These specific metrics are widely used to validate $K_d^{R_{rs}}$ performance across the literature (e.g., [3,14,18,25]).

1. Absolute Percentage Difference (APD) of the log-transformed $K_d^{R_{rs}}$ and K_d^{float} [18]. The APD gives an equal weight to over and under estimations of K_d^{float} while covering a range of value larger than one order of magnitude.

$$APD = 100 \times exp\left(mean \left| ln \frac{K_d^{R_{rs}}}{K_d^{float}} \right| \right) - 1. \tag{13}$$

2. Root Mean Squared Difference (RMSD) is calculated as follows with N the total number of matchups:

$$RMSD = \sqrt{\frac{\Sigma_{i=1}^{N}(K_d^{float} - K_d^{R_{rs}})^2}{N}}. \tag{14}$$

3. The bias is defined as the median of the ratio of float and remote observations:

$$Bias = median\left(\frac{K_d^{float}}{K_d^{R_{rs}}} \right). \tag{15}$$

4. The r is Pearson's correlation coefficient.

Several regressions methods were tested, with limited difference in both the slope and the correlation coefficient (r). Because there are uncertainties in the retrieval of both K_d^{float} and $K_d^{R_{rs}}$, a type-2 least-squares fit regression was applied by computing the geometric mean of the type-1 regression of K_d^{float} vs. $K_d^{R_{rs}}$ and the type-1 regression of $K_d^{R_{rs}}$ vs. K_d^{float}, effectively resulting in a slope that passes by the centroid $\left(\overline{K_d^{float}}, \overline{K_d^{R_{rs}}} \right)$.

The significance of the differences between the distributions of different $K_d^{R_{rs}}$ (either between sensors or algorithms) was evaluated using a two-sample non-parametric Kolmogorov–Smirnov (K-S) test, with a chosen statistically significant level of 5%.

To test that we implemented the algorithms correctly (and that our method of computing K_d with in situ data were consistent), we evaluated them with the datasets used by the authors (Figure A1), and the results retrieved were in accordance with the original publications [3,18].

3. Results

3.1. $K_d(\lambda)$: Global Scale Match-Ups

$K_d(490)^{R_{rs}}$ retrieved from each of the algorithms generally followed the 1:1 line (Figure 3). The operational products ($K_d(490)^{NASA}$ for MODIS and VIIRS, $K_d(490)^{ESA}$ for OLCI) had the best statistical results for the VIIRS and OLCI sensors, with the lowest Bias, APD and RMSE for each sensor (Table 3), and $K_d(490)^{Lee05}$ had the best results for the MODIS sensors. $K_d(490)^{NASA}$ also retrieved the slope closest to one for all four sensors. $K_d(490)^{Lee05}$ systematically overestimated $K_d(490)$ at low K_d values (<0.025 m^{-1}) and had a few outliers for the MODIS sensors (not plotted on Figure 3 but used in statistics) that impacted its r-score. $K_d(490)^{NN}$ had a slope furthest from one for the MODIS sensors and also showed a systematic overestimation at very low values ($K_d(\lambda) < 0.025$). The slopes were below one for all the sensors with a significant non-zero intercept. The $K_d(490)^{R_{rs}}/K_d(490)^{float}$ ratio (Figure 3b) for low K_d values is larger for K_d values with a small zenith angle $\approx10°$), but for a given K_d value, a higher solar zenith angle resulted in a larger $K_d(490)^{R_{rs}}/K_d(490)^{float}$ ratio.

The $K_d(412)^{float}$ range of values is 0.0126–0.7 m^{-1} (Figures 1 and 4). The operational products are not computed at 412 nm and we therefore only compare K_d^{Lee05} and $K_d(490)^{NN}$ to $K_d(490)^{float}$. $K_d(412)^{NN}$ performed significantly worse than $K_d(412)^{Lee05}$ for the MODIS-Aqua, MODIS-Terra, and OLCI-S3B sensors, with a lower r-score and a higher bias, APD, and RMSD for all sensors (Table 4). For OLCI-S3A, $K_d(412)^{NN}$ performed better. The slopes are closer to one for each of the sensors than at $\lambda = 490$ nm while still showing

a systematic overestimation for small $K_d(412) < 0.026$ values (Figure 3b), along with a significant non-zero intercept.

Both $K_d(412)^{Lee05}$ and $K_d(412)^{NN}$ had a lower APD at 412 nm than at 490 nm for all three sensors; however, $K_d(412)^{NN}$ exhibited a larger RMSD. The slope is closer to one (all are >0.91). The ratio K_d^{Rrs}/K_d^{float} was closer to one at 412 nm than 490 nm (Figure 4).
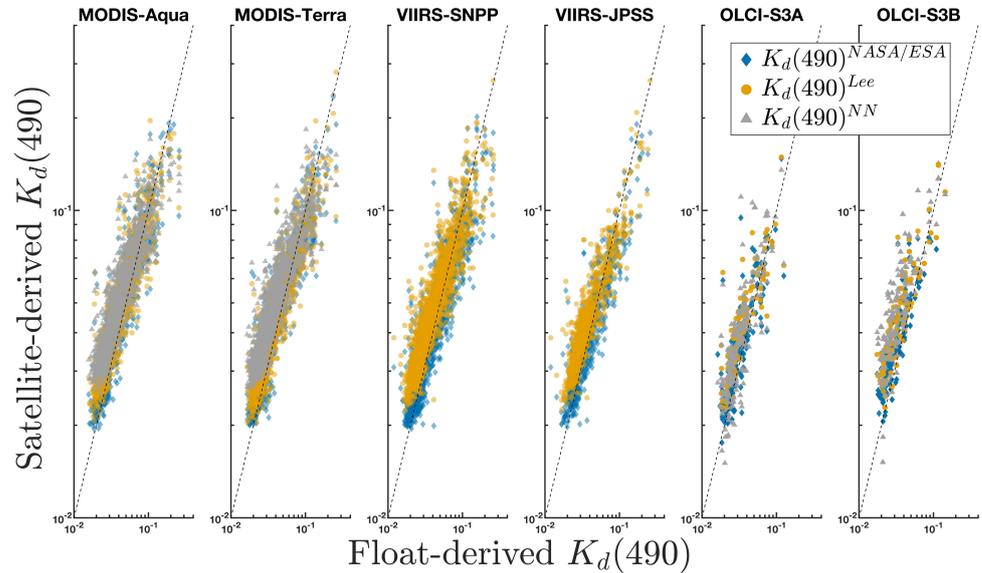
(**a**)
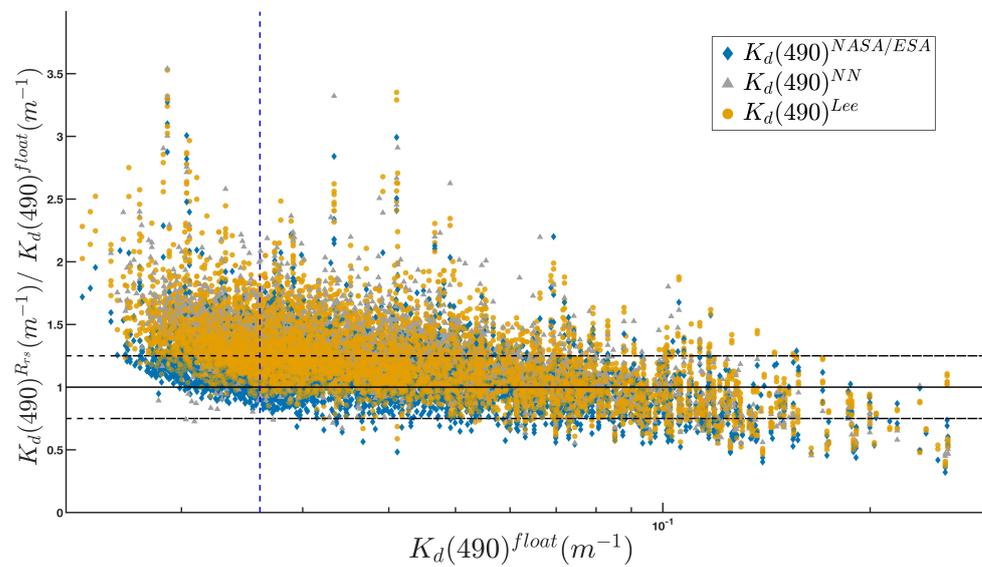
(**b**)

Figure 3. Comparison of satellite-derived and float-derived $K_d(490)$ for the MODIS-Aqua, MODIS-Terra, VIIRS-JPSS, VIIRS-SNPP, OLCI-S3A and OLCI-S3B sensors: (**a**) $K_d^{Rrs}(490)$ computed using the 3 different algorithms compared to $K_d^{float}(490)$; the black dashed line is the 1:1 line; (**b**) $K_d^{Rrs}(490)/K_d^{float}(490)$ for each of the 3 evaluated algorithms (color coded) for all sensors; the solid black line is a ratio of 1, and the dashed black lines are the 0.75 (Bottom) and 1.25 (Top) ratio. The vertical dashed blue line indicates the minimum value of $K_d(490)$ present in the NOMAD dataset (0.026).

Table 3. Comparison of performance statistics at the global scale of the K_d(490 nm) for the MODIS, VIIRS, and OLCI sensors at the global scale and K_d(490 nm) algorithms. See Methods section for the definitions of the metrics. All distributions within a given sensor are statistically different. N represents the number of matchups with data at 490 nm.

Sensor & Algo	BIAS	APD (%)	RMSD (m⁻¹)	r	Slope	Intercept	N
MODIS-Terra: K_d^{Lee05}	1.08	18.62	0.01	0.90	0.78	0.010	
MODIS-Terra: K_d^{NN}	1.31	33.72	0.02	0.87	0.76	0.018	2144
MODIS-Terra: K_d^{NASA}	1.13	20.37	0.01	0.90	0.84	0.010	
MODIS-Aqua: K_d^{Lee05}	1.11	19.58	0.01	0.89	0.82	0.011	
MODIS-Aqua: K_d^{NN}	1.27	31.19	0.02	0.86	0.79	0.017	1802
MODIS-Aqua: K_d^{NASA}	1.12	19.67	0.01	0.89	0.89	0.009	
VIIRS-SNPP: K_d^{Lee05}	1.16	22.46	0.02	0.88	0.77	0.013	3290
VIIRS-SNPP: K_d^{NASA}	1.06	17.36	0.02	0.88	0.78	0.010	
VIIRS-SNPP: K_d^{Lee05}	1.16	22.46	0.02	0.88	0.77	0.013	2445
VIIRS-SNPP: K_d^{NASA}	1.06	17.36	0.02	0.88	0.78	0.010	
OLCI-S3A: K_d^{Lee05}	1.16	21.46	0.01	0.84	0.79	0.012	651
OLCI-S3A: K_d^{NN}	1.19	26.62	0.01	0.77	0.91	0.008	
OLCI-S3A: K_d^{ESA}	1.08	17.85	0.01	0.83	0.82	0.008	
OLCI-S3B: K_d^{Lee05}	1.25	27.73	0.01	0.91	0.68	0.018	
OLCI-S3B: K_d^{NN}	1.42	43.24	0.02	0.85	0.84	0.019	382
OLCI-S3B: K_d^{ESA}	1.18	20.88	0.01	0.92	0.71	0.013	

Table 4. Comparison of performance statistics at the global scale of the K_d(412 nm) for the MODIS and OLCI sensors at the global scale and K_d(412 nm) algorithms. See Methods section for the definitions of the metrics.

Sensor & Algo	BIAS	APD (%)	RMSD	r	Slope	Intercept	N
MODIS-Terra: K_d^{Lee05}	1.13	11.60	0.06	0.68	0.91	0.010	1633
MODIS-Terra: K_d^{NN}	1.19	19.49	0.03	0.87	1.01	0.018	
MODIS-Aqua: K_d^{Lee05}	1.10	8.91	0.02	0.86	0.88	0.011	1384
MODIS-Aqua: K_d^{NN}	1.15	16.01	0.03	0.88	1.09	0.017	
OLCI-S3A: K_d^{Lee05}	1.19	20.68	0.02	0.93	0.98	0.012	269
OLCI-S3A: K_d^{NN}	1.15	13.67	0.02	0.88	0.92	0.008	
OLCI-S3B: K_d^{Lee05}	1.23	23.23	0.02	0.92	0.85	0.018	326
OLCI -S3B: K_d^{NN}	1.34	35.02	0.02	0.87	1.20	0.019	

3.2. $K_d(PAR)$: Global Scales Matchups

The number of matchups between sensors-derived and floats-derived $K_d(PAR)$ was 832 for MODIS-Aqua, 944 for MODIS-Terra, 1402 for VIIRS-SNPP, 613 for VIIRS-JPSS, 155 for OLCI-S3A and 227 for OLCI-S3B (Figure 5) resulting in a total of 4173 matchups between float and satellite. For all sensors, there was an underestimation for small values: $K_d(PAR)^{float} < 0.038$ m⁻¹ for $K_d(PAR)^{Morel}$ and $K_d(PAR)^{float} < 0.048$ m⁻¹ for $K_d(PAR)^{Lee05}$ representing 11% and 20% of the full dataset, respectively. For those values, independent of the sensors, $K_d(PAR)^{float} < K_d(PAR)^{Rrs}$, with the ratio increasing as $K_d(PAR)^{float}$ decreased (Figure 5b). The regression slopes are <1 for both algorithms, and there was a significant intercept for both of them (Table 5).

$K_d(PAR)^{Morel}$ had a lower bias, lower APD, lower RMSD and a higher r than $K_d(PAR)^{Lee05}$ (Table 5). It also had a slope closer to one. For high values, $K_d(PAR)^{Lee05} > K_d(PAR)^{Morel}$, whereas for low values, $K_d(PAR)^{Morel} > K_d(PAR)^{Lee05}$ (Figure 5). The biggest discrepancy between the two algorithms occurs when the Solar zenith angle is low (<20°), but for a given $K_d(PAR)^{float}$ value, the higher the solar zenith angle, the bigger the difference.

(**a**)

(**b**)
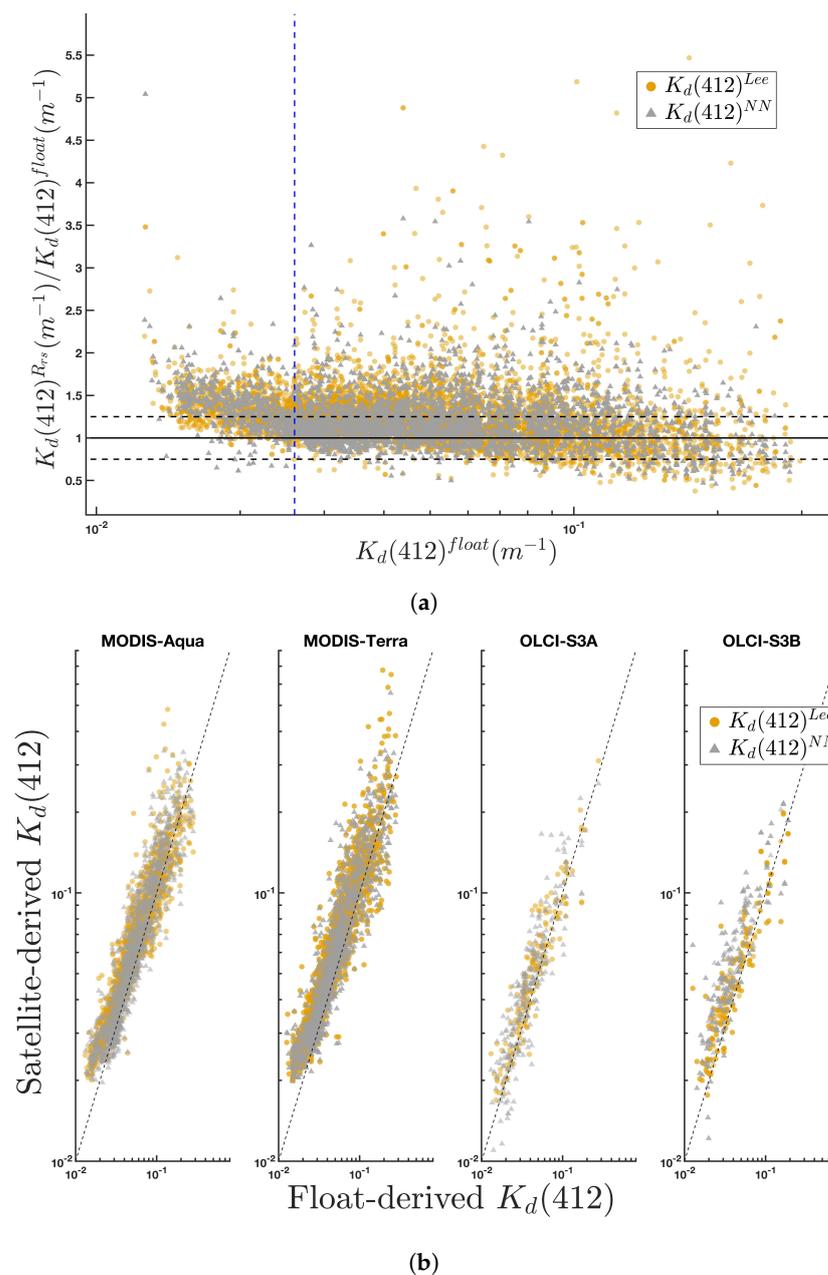
Figure 4. Comparison of satellite-derived and float-derived $K_d(412)$ for the two MODIS and the two OLCI sensors: (**a**) $K_d^{Rrs}(412)$ computed using the 2 different algorithms compared to $K_d^{float}(412)$; the black dashed line is the 1:1 line. (**b**) K_d^{Rrs}/K_d^{float} for the matchups; the solid black line is a ratio of 1, and the dashed black lines denote ratios of 0.75 (Bottom) and 1.25 (Top) ratio; the dashed blue line indicates the minimum value of $K_d(411)$ present in the NOMAD dataset (0.026).

3.3. Variability in Performance between Satellite Sensors

We performed a Kolmogorov–Smirnov (K-S) test to assess whether the distributions of $K_d^{Rrs}(\lambda)$ retrieved by a given sensor using different algorithms were different. The K-S test indicates whether the K_d values retrieved by different algorithms have a different distribution within a given confidence interval (here 5%). The distributions of $K_d(490)^{NN}$ vs. $K_d(490)^{Lee05}$ retrieved by the OLCI-S3A sensor were not statistically different from each other (Table 3), whereas, they were different for the other sensors. The distribution of $K_d(490)^{Lee05}$ vs. $K_d(490)^{NASA/ESA}$ was statistically different for all sensors as was the case for the distribution of $K_d(490)^{NN}$.

(**a**)

(**b**)
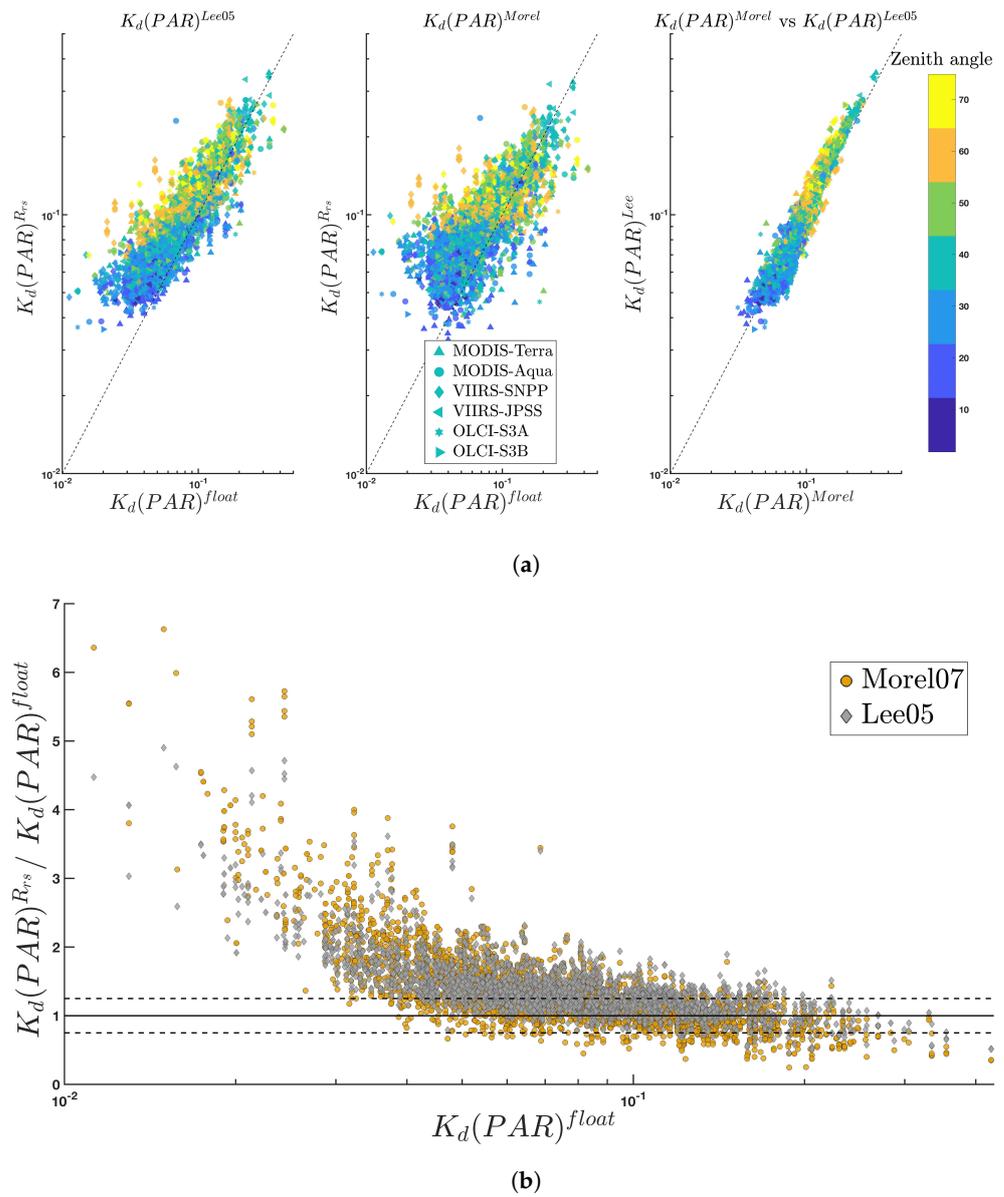
Figure 5. Results of the comparison between the satellite-derived $K_d(PAR)^{Rrs}$ and the float-retrieved $K_d(PAR)^{float}$, for two different PAR algorithms: (**a**) $K_d(PAR)^{Rrs}$ vs. $K_d(PAR)^{float}$ colored by solar zenith angle with each marker shape indicating a different sensor; the dashed line indicates the 1:1 line; $K_d(PAR)^{Lee05}$ vs. $K_d(PAR)^{float}$ (left), $K_d(PAR)^{Morel}$ vs. $K_d(PAR)^{float}$ (center) $K_d(PAR)^{Lee05}$ vs. $K_d(PAR)^{Morel}$ (right); (**b**) ratio for each of the two algorithms against $K_d(PAR)^{float}$; the solid line is a ratio of 1 and the dashed black lines denote ratios of 0.75 and 1.25.

Table 5. Summary statistics for all satellite sensors at the global scale for both PAR algorithms. See Methods section for definitions of statistical metrics.

	MODIS-Terra		MODIS-Aqua		VIIRS-SNPP		VIIRS-JPSS		OLCI-S3A		OLCI-S3B	
	Lee05	Morel07	Lee05	Morel07	Lee05	Morel07	Lee05	Morel07	Lee05	Morel07	Lee05	Morel07
Bias	1.24	1.20	1.28	1.23	1.28	1.26	1.28	1.24	1.21	1.17	1.25	1.28
ADP	23.88	21.05	28.72	26.34	30.89	29.81	28.74	28.07	20.77	15.97	33.73	33.31
RMSD	0.027	0.032	0.031	0.035	0.031	0.037	0.029	0.035	0.028	0.034	0.026	0.028
r	0.87	0.75	0.86	0.76	0.86	0.75	0.88	0.77	0.81	0.65	0.92	0.85
Slope	0.83	0.61	0.90	0.68	0.82	0.57	0.77	0.53	0.85	0.53	0.94	0.69
Intercept	0.029	0.044	0.028	0.044	0.035	0.053	0.036	0.053	0.024	0.045	0.025	0.042

The difference in the distribution of $K_d^{Rrs}(\lambda)$ values across the six different sensors for a given algorithm was also assessed by Kolmogorov–Smirnov test. A difference between sensors would mean that there are either differences in the specifications of each sensor across different satellites or that there are different distributions of floats-satellite matchups (potentially due to the timing of the overpass of the sensor or to the different launch dates between satellites, resulting in datasets covering varying time periods).

For all three algorithms tested, the only ones that had a similar distribution were for the $K_d(490)^{NASA/ESA}$ algorithm of the VIIRS-SNPP/OLCI-S3B sensor pair, the $K_d(490)^{Lee05}$ algorithm between the OLCI-S3B/VIIRS-SNPP and for the OLCI-S3B/VIIRS-JPSS pairs. All other sensor pairs had statistically different distributions.

3.4. Regional Analysis

Matchups between floats and sensors were not distributed homogeneously throughout the biomes (Figure 2, Table A2). Biomes 18 (Western Mediterranean) and 19 (Eastern Mediterranean) had the most matchups (representing 33% and 28% of the overall matchups, respectively). The next biome with the highest proportion of floats-sensors matchups was biome 13 (South Atlantic Subtropical Permanently Stratified, Table A2) with 8%. Between the different sensors, a two-sample Kolmogorov–Smirnov showed a different distribution of the matchups amongst all the biomes between OLCI-S3A and all the other sensors except for OLCI-S3B, as well as a different distribution between OLCI-S3B and all the other sensors, except for OLCI-S3A, at the 5% significance level. All other sensor pairs did not show a significant difference in distribution between biomes.

A statistical analysis was performed when removing the influence of the Mediterranean Sea (biomes 18 and 19) to assess the impact of not including the Mediterranean data on global statistics (Table A5); this analysis did not result in a significant improvement of algorithm performance.

No obvious biome-based bias in algorithm performance was observed (Table 6), with biome 9 having the lowest bias and RMSD across all three algorithms but having slopes significantly different from one. Similarly, biome 18 had slopes close to one for $K_d(490)^{Lee05}$ and $K_d(490)^{NN}$ but lower *r* than other biomes, such as biome 19. Note that some biomes (e.g., 6, 8, 10, 12, and 14) had a low number of matchups and limited dynamic range, resulting in (non-significant) negative slopes and *r*.

Table 6. Summary statistics at $\lambda = 490$ nm for each of the biomes defined in Table A2, with all sensors grouped together. The *NASA* empirical algorithm (see Methods section) was applied for the MODIS and the VIIRS sensors, whereas the *ESA* empirical algorithm was applied on the OLCI sensors. As they are both empirical algorithms, they were grouped together for the overall statistical analysis. In parentheses are the numbers of matchups between $K_d(490)^{float}$ and $K_d(490)^{Rrs}$ in each of the biomes for each algorithm. For the definition of the bias, APD, RMSD and r (Pearson's correlation coefficient) see the Methods section.

	Biome 4 (N = 113)			Biome 6 (N = 35)			Biome 7 (N = 239)			Biome 8 (N = 76)			Biome 9 (N = 435)		
	Lee	Jamet	Austin	Lee	Jamet	Austin	Lee	Jamet	Austin	Lee	Jamet	Austin	Lee	Jamet	Austin
BIAS	1.13	1.25	1.08	1.16	1.35	1.20	1.18	1.31	1.11	0.95	1.04	0.88	0.97	1.01	0.96
ADP	17.57	29.58	13.65	18.38	35.04	21.08	19.92	35.38	14.38	25.38	30.85	25.90	21.03	25.30	20.72
RMSD	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.02	0.03	0.02	0.03	0.03	0.03
r	0.55	0.52	0.75	0.54	0.90	0.66	0.90	0.87	0.89	−0.05	−0.32	0.00	0.85	0.77	0.84
Slope	0.51	0.34	0.57	0.41	0.87	0.43	0.77	0.74	0.86	−0.31	−0.50	−0.18	0.60	0.45	0.57
Intercept	0.02	0.03	0.02	0.03	0.02	0.03	0.01	0.02	0.01	0.09	0.11	0.08	0.03	0.05	0.03

Table 6. *Cont.*

	Biome 10 (N = 6)			Biome 11 (N = 225)			Biome 12 (N = 21)			Biome 13 (N = 308)			Biome 14 (N = 10)		
	Lee	Jamet	Austin	Lee	Jamet	Austin	Lee	Jamet	Austin	Lee	Jamet	Austin	Lee	Jamet	Austin
BIAS	1.14	1.31	1.11	1.13	1.27	1.03	1.07	1.12	1.05	1.17	1.42	1.07	1.06	1.25	0.99
ADP	22.21	43.08	15.02	19.21	32.40	14.55	15.77	17.32	11.97	19.35	40.94	13.39	7.71	23.60	11.31
RMSD	0.02	0.03	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.01
r	0.61	−0.65	0.82	0.53	0.50	0.60	0.63	0.59	0.69	0.83	0.70	0.84	0.84	0.12	0.64
Slope	0.43	−0.47	0.65	0.43	0.46	0.49	0.52	0.56	0.62	0.64	0.59	0.70	1.11	0.24	1.29
Intercept	0.04	0.12	0.03	0.02	0.02	0.01	0.02	0.02	0.01	0.01	0.02	0.01	0.00	0.04	−0.01

	Biome 15 (N = 246)			Biome 16 (N = 184)			Biome 18 (N = 1554)			Biome 19 (N = 1986)		
	Lee	Jamet	Austin	Lee	Jamet	Austin	Lee	Jamet	Austin	Lee	Jamet	Austin
BIAS	1.07	1.17	1.08	1.11	1.28	1.07	1.05	1.17	1.04	1.10	1.27	1.05
ADP	23.50	30.17	22.35	19.52	34.44	17.00	17.59	24.43	16.80	17.58	29.20	15.42
RMSD	0.03	0.03	0.03	0.01	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.01
r	0.57	0.52	0.57	0.85	0.88	0.85	0.85	0.82	0.85	0.92	0.89	0.90
Slope	0.49	0.49	0.52	0.69	0.87	0.73	0.73	0.73	0.78	0.74	0.80	0.73
Intercept	0.03	0.04	0.03	0.02	0.02	0.02	0.01	0.02	0.01	0.01	0.02	0.01

4. Discussion

4.1. Observed Biases in K_d

All four algorithms had a slope < 1 at λ = 490 (Table 3) because $K_d(490)^{Rrs} > K_d(490)^{float}$ at small K_d values. This overestimation of $K_d(490)$ effectively leads to an underestimation of the depth to which light penetrates in the water column, potentially resulting in an underestimation of heat transfer to depth and other depth-derived products from $K_d(490)$. On the other hand, for $K_d(490) > 0.1$, $K_d(490)^{Rrs} < K_d(490)^{float}$ (Figure 3). This underestimation of K_d will result in the overestimation of $K_d(490)$-derived products. The overestimation at small values is also found at λ = 412, with a systematic overestimation for $K_d(412)^{NN}$ and $K_d(412)^{Lee05}$ at values < 0.025 (Figure 4). However, there is no persistent underestimation for larger $K_d(412)$ values (Table 4).

It is also relevant to note that there is a strong relationship between $K_d(490)^{float}$ and Rrs-retrieved Chl a (r-score of 0.84 over the full matchup dataset), which is only slightly lower than the correlation score between the full $K_d(490)^{float}$ and $K_d(490)^{Lee05}$ (0.89 over the full dataset) which asks the question about redundancy between the offered Satellite L2 products.

4.2. Limitation of Datasets Used to Train Empirical Algorithm

$K_d^{float}(490)$ values in this study ranged within 0.012–0.42 m^{-1} (Figure 1), spanning from oligotrophic open ocean waters to near-coastal more eutrophic waters. The lowest values measured were actually lower than the literature values used for $K_d(490)$ of pure seawater (0.016, [4]). For $K_d(PAR)$, the lowest values obtained from floats measurements were not below the values of $K_d(PAR)$ of pure seawater (0.0185) computed with Morel07 (See Methods section).

The low end of $K_d(490)$ values is lower than those included in the databases used to train empirical algorithms. The range of the NOMAD and the IOCCG synthetic $K_d(490)$ datasets was [0.026 : 4.26], and its distribution was nearly uniform as opposed to nearly lognormal as seen with the floats (Figure 1). K_d^{NN} was not trained with values <0.026 m^{-1} [3] and therefore does not provide an output below 0.026 m^{-1} for any of the MODIS satellites sensors (Figure 3), resulting in an overestimation of the K_d^{NN} in clear oligotrophic waters. For the OLCI sensors, K_d^{NN} did retrieve values below 0.026 m^{-1} but nonetheless consistently

overestimated the small K_d values measured by the floats. Similarly, $K_d^{Lee05}(490)$'s training datasets had a range spanning $[0.04:4]$ m^{-1} [18], consistent with the biggest difference between K_d^{Lee05} and K_d^{float} occurring at low values (Figure 3b).

We found the slopes of the regression between algorithm and float to typically be significantly less than one. If the regression intercept is forced to zero, the slope of $K_d(490)^{Rrs}$ vs. $K_d(490)^{float}$ is closer to one, regardless of algorithm or sensors (Table A3). It ranges from $[1.03-1.17]$ for $K_d(490)^{Lee05}$, $[1.15-1.30]$ for $K_d(490)^{NN}$, and $[0.98-1.14]$ for $K_d(490)^{NASA/ESA}$. It is apparent that the small values that are not sufficiently represented in the original range drive the slope offset we found.

Overall, for $K_d(490)^{Lee05}$, 74% of the values were within $\pm25\%$ of $K_d(PAR)$float, 49% for $K_d(490)^{NN}$, and 80% for $K_d(490)^{NASA/ESA}$ (Figure 3b). Therefore, the performances of the algorithms were significantly lower than on the original datasets they were based on, indicating that they could be improved. At 412 nm, 64% of $K_d(412)^{Lee05}$ were within $\pm25\%$ of $K_d(412)^{float}$ versus 65% for $K_d(412)^{NN}$.

From the discussion above, it follows that an important source of bias in empirical algorithms originates from the limitations of the range of the datasets they were tuned with. It has long been known that the operational algorithms (K_d^{Austin} and K_d^{Morel}) were designed for and are only applicable to Case-1 waters, with $K_d > 0.25$ resulting in poor K_d retrieval [3,26]. This study shows that within Case-1 water, with either semi-analytical or empirical algorithms there is still significant bias that could be corrected.

Similarly, $K_d(PAR)^{Morel}$ and $K_d(PAR)^{Lee05}$ were designed using available in situ databases (NOMAD, among others mentioned above) or the IOCCG dataset, resulting in similar biases to those observed for $K_d(412)$ and $K_d(490)$. Clear-water biases are more important than for $K_d(\lambda)$: 53% of $K_d(PAR)^{Lee05}$ values were within $\pm25\%$ of Kd$(PAR)^{float}$ and 20% of were consistently overestimated ($K_d(PAR)^{float} < 0.048$). Some 53% of the $K_d(PAR)^{Morel}$ were within $\pm25\%$ of $K_d(PAR)^{float}$, and 11% were systematically overestimated ($K_d(PAR)^{float} < 0.039$). Note that the overestimation of $K_d(PAR)$ was first pointed out by the authors of [2].

This overestimation of $K_d(PAR)$ for the very oligotrophic waters is different from previous findings [23] which expected the presence of a DCM (Deep Chlorophyll Maximum) below the z_{pd} to cause an underestimation of $K_d(PAR)^{Rrs}$, as it would not be taken into account by the satellite-based algorithms.

4.3. Influence of the Solar Zenith Angle

As an AOP, $K_d(\lambda)$ will vary with solar zenith angle, with instantaneous values of $K_d(490)$ within a single day ranging from 25% to 250% of the mean diurnal averaged $K_d(490)$ [27] (over the daytime). Arguments based on radiative transfer theory suggest that a higher zenith angle (sun closer to the horizon) will result in an increased path length for photons, effectively increasing $K_d(\lambda)$ for the same IOPs.

The solar zenith angle for satellite algorithm was limited to 75° in our methodology following published criteria [10]. The ±3 h window criterion between the solar zenith when the float surfaced and the one of the sensor overpass meant that the solar zenith angle for the float profile varied from 0.53–84.7°. As found in a previous study [14], the difference in solar zenith angle (of up to 40° in our dataset) between the float surfacing and the satellite overpass did not correlate with the difference $K_d^{Rrs}-K_d^{float}$ (Figure A2), suggesting that the matchup quality was not affected by the difference in time and/or change in solar zenith angle between the satellite overpass and the float surfacing.

The solar zenith angle at the time of satellite overpass does appear to have an impact on the magnitude of the residuals when coupled with the amplitude of $K_d(490)$. The variability in the residuals of $K_d(490)^{Rrs}-K_d(490)^{float}$ increases as the solar zenith angle increases for all three studied algorithms (Figure 3), but not relative to the value of $K_d(PAR)$. Thus, the bigger the solar zenith angle, the bigger the potential for an error in $K_d(490)^{Rrs}$, especially for large values of $K_d(490)^{float}$ (Figure A2). However, the mean relative residual of the retrieval of $K_d(PAR)$ or $K_d(\lambda)$ is closer to 0 for larger zenith angle and further from 0

from zenith angle ranging from 0–10° (Figure 6), likely due to the small number of data points in the original datasets with small solar angle.
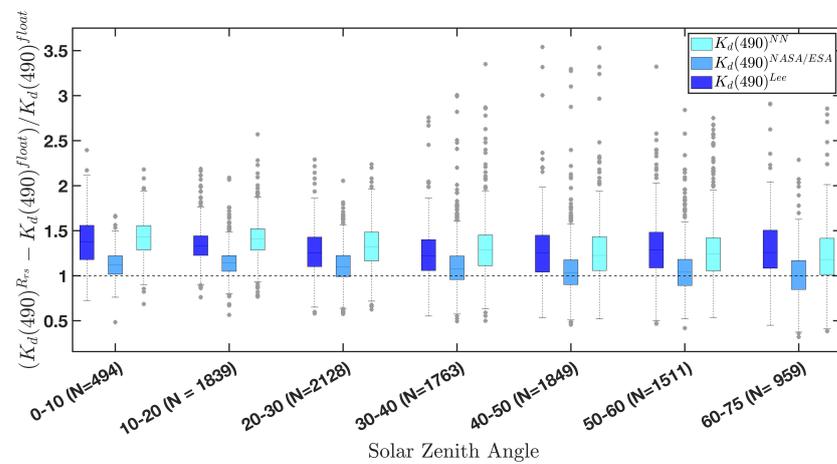
(**a**)
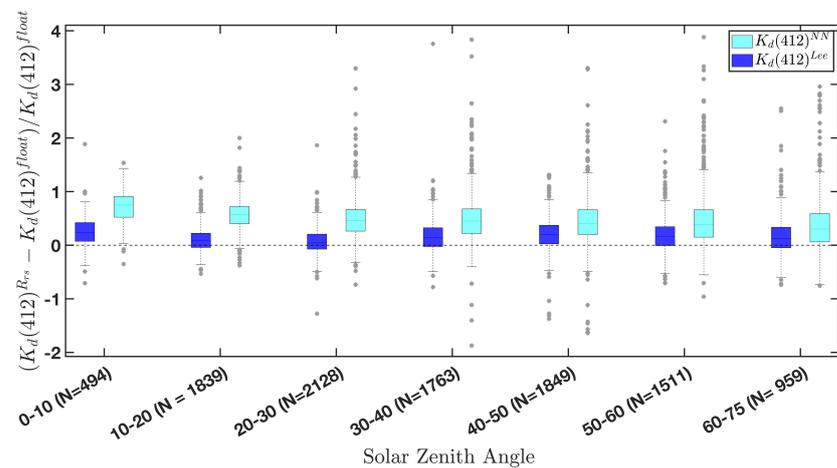
(**b**)
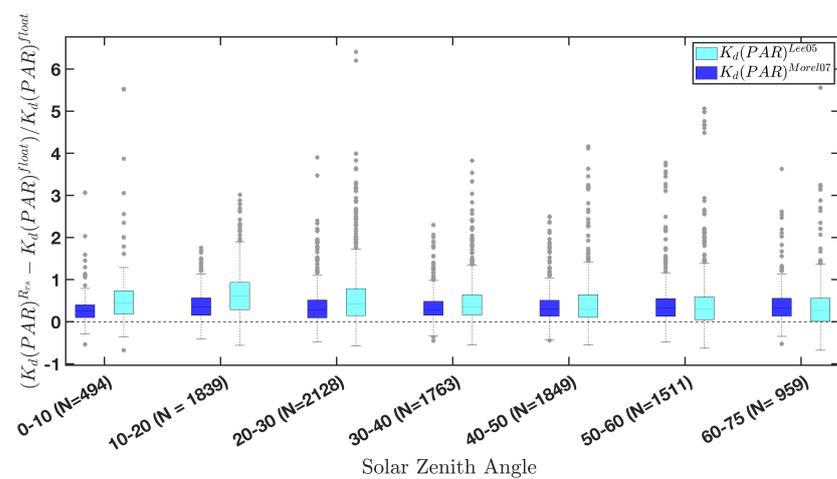
(**c**)

Figure 6. Boxplot of the relative residuals between satellite-derived and float-derived $K_d(\lambda)$ and its dependence on the solar zenith angle for: (**a**) 490 nm; (**b**) 412 nm; and (**c**) PAR. Color denotes the algorithm used. Datapoints for 412 nm with residuals value higher than 0.8 were considered to be outliers and were not plotted here for clarity.

We thus conclude that including the value of solar zenith angle in empirical algorithms is likely to improve them. However, the effect is likely to change spectrally [28], as the change in the mean cosine of the light near the surface depends on, besides the sun angle, competition between scattering (which tends to diffuse light) and absorption which tends to collimate light vertically and thus on the single scattering albedo which is smaller in red compared to blue wavelengths.

5. Conclusions

In this study, we examined published algorithms for the diffuse attenuation coefficient derived from remotely sensed ocean color using a novel dataset comprised of irradiance profiles measured with sensors on profiling floats. We found significant bias in these algorithms, particularly at low values of K_d, which we attribute to the limitations of previous datasets used in the design of the algorithms.

We recommend that a revision of these algorithms be implemented to remove bias, and that empirical algorithms that do not take into account the solar zenith angle at the time of satellite pass do so as their performance is likely to improve. Finally, we recommend, given the hyperspectral nature of the upcoming PACE, SBG and CHIME missions, that hyperspectral radiometers be deployed on floats (e.g., [29]) as this will increase the potential for algorithm validation.

Author Contributions: Conceptualization, E.B.; methodology, E.B. and C.B.D.; software, E.B. and C.B.D.; validation, C.B.D.; formal analysis, C.B.D.; data curation, C.B.D.; writing—original draft preparation, C.B.D.; writing—review and editing, C.B.D. and E.B.; visualization, C.B.D.; supervision, E.B.; project administration, E.B.; funding acquisition, E.B. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Floats Quality Control and Distribution

A total of 40,738 raw Argo floats profiles containing radiometric measurements were downloaded on the 18th of March 2022. Out of those, 29,004 passed the QC of [6] (71.2%). A minimum of five measurements in the upper 10 m are required to extrapolate to $E_d(\lambda, 0^-)$ or $iPAR(0^-)$ [2]. A total of 25,090 profiles (86.5% of QC-ed profiles, 61.59% of the original profiles) were therefore used to compute $K_d^{float}(\lambda)$ at a minimum of one wavelength and/or $K_d^{float}(PAR)$.

Table A1. Number of downloaded profiles and images.

BGC-Argo	Profiles downloaded	40,738
	QC from Organelli et al.	29,004
	5 $E_d(\lambda)$ values within the upper 10 m	25,090
Satellite	Images downloaded	364,813
	Images passing matchups criteria	10,873
	Images passing QWIP [16]	10,805

Appendix B. Biomes Separation

Table A2. List of the biomes according to Fay and McKinley, 2014, with the associated number. Biomes 18 and 19 were added for Mediterranean Sea with the East/West division at the Sicily Strait. Percentage of all profiles for all of the float-sensors matchups grouped together, regardless of wavelength. Oceanic coverage proportion is area relative to the area covered by all the biomes grouped together.

Biome Number	Biome Acronym	Biome Name	Float Proportion (in %)	Oceanic Coverage Proportion (in%)
1	'NP ICE'	North Pacific Ice	0	1.37
2	'NP SPSS'	North Pacific Subpolar Seasonally Stratified	0	3.85
3	'NP STSS'	North Pacific Subtropical Seasonally Stratified	0	2.04
4	'NP STPS'	North Pacific Subtropical Permanently Stratified	1.21	12.29
5	'PEQU-W'	West pacific Equatorial	0	3.50
6	'PEQU-E'	East Pacific Equatorial	1.36	4.46
7	'SP STPS'	South Pacific Subtropical Permanently Stratified	4.65	15.79
8	'NA ICE'	North Atlantic Ice	2.67	1.64
9	'NA SPSS'	North Atlantic Subpolar Seasonally Stratified	6.53	3.01
10	'NA STSS'	North Atlantic Subtropical Seasonally Stratified	0.51	1.79
11	'NA STPS'	North Atlantic Subtropical Permanently Stratified	4.69	5.23
12	'AEQU'	Atlantic Equatorial	0.18	2.22
13	'SA STPS'	South Atlantic Subtropical Permanently Stratified	8.88	5.41
14	'IND STPS'	Indian Ocean Subtropical Permanently Stratified	0.17	10.76
15	'SO STSS'	Southern Ocean Subtropical Seasonally Stratified	4.62	8.89
16	'SO SPSS'	Southern Ocean Subpolar Seasonally Stratified	3.49	11.87
17	'SO ICE'	Southern Ocean Ice	0.036	5.59
18	'W MED'	Western Mediterranean	32.75	0.22
19	'E MED'	Eastern Mediterranean	28.25	0.56

Appendix C. COASTLOOC and NOMAD Comparison

To ensure that the algorithms perform the same way they were designed, they were run on the same datasets for which they were originally designed, i.e., NOMAD and COAST-LOOC. The NOMAD dataset was accessed from NASA (https://seabass.gsfc.nasa.gov/wiki/NOMAD, accessed on 14 February 2022), and the COASTLOOC dataset was provided by Marcel Babin ([30]. $R(0^-)$ and was converted to R_{rs} using $R_{rs} = 0.133 \times R(0^-)$ [15]. Wavelengths differing from the MODIS or the Sea-WIFs wavelengths used as inputs in the NN were interpolated to the input wavelengths using a spline interpolation, and any $K_d(490) < 0.016$ ($K_d(490)$ of pure water [8]) were removed.

(**a**)

(**b**)

(**c**)

(**d**)
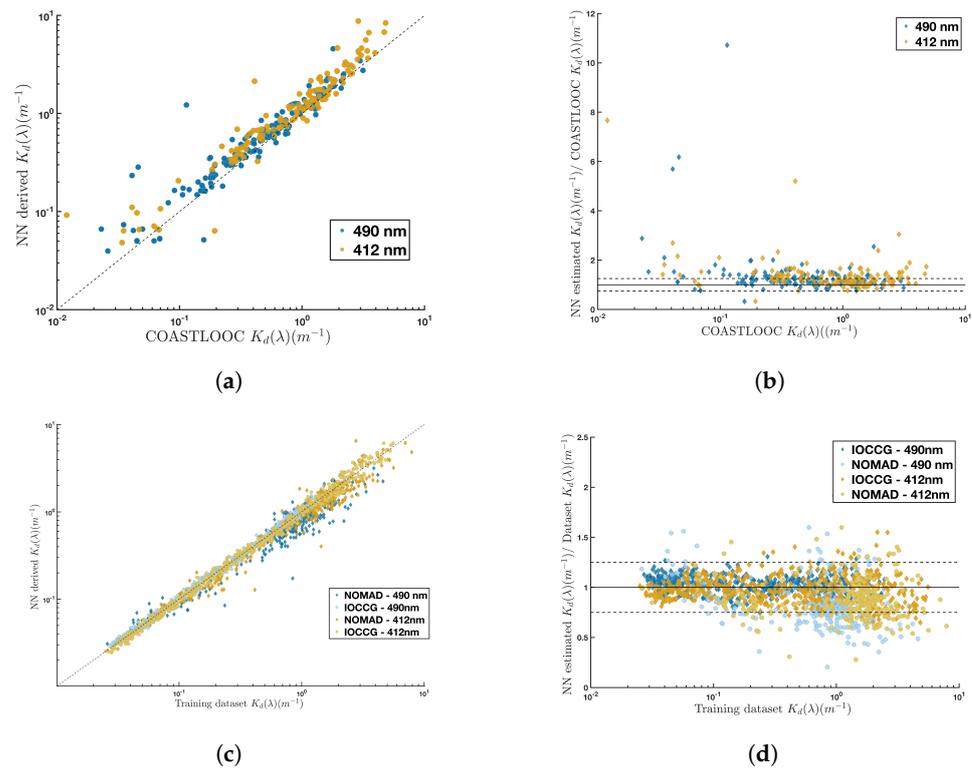
Figure A1. Results retrieved from the COASTLOOC the NOMAD & IOCCG synthetic dataset at 412 nm and 490 nm: (**a**) scatterplot of estimated versus desired $K_d(\lambda)$ values for the COASTLOOC dataset; the dashed line is the 1:1 line; (**b**) ratio of the estimated versus desired $K_d(\lambda)$; the horizontal line is a ratio of 1; (**c**) scatterplot of estimated versus desired $K_d(\lambda)$ values; the dashed line is the 1:1 line; (**d**) ratio of the estimated versus desired $K_d(\lambda)$; the horizontal line is a ratio of 1.

Table A3. Summary statistics for the retrieval of $K_d(490)$ and $K_d(412)$ from Jamet's neural network algorithm using NOMAD & IOCCG initial dataset and the validation dataset of COASTLOOC. Root Mean Square Difference (RMSD) is the square root of the mean square error between the retrieved K_d and the measured/simulated K_d; r is Pearson's correlation coefficient.

	COASTLOOC		NOMAD		IOCCG	
	490	412	490	412	490	412
Slope	1.10	1.43	0.68	0.82	1.00	1.02
Intercept	0.060	−0.132	0.103	0.095	−0.002	−0.019
r	0.92	0.92	0.89	0.90	0.99	0.99
RMSD	0.319	0.855	0.373	0.530	0.062	0.206

There were no significant biases at low K_d values visible from Subdataset 2 (Figure A1), and COASTLOOC did not have enough point at low K_d ($< 10^{-1}$) to conclude. Generally speaking, for all datasets the algorithm performed better retrieving $K_d(490)$ than $K_d(412)$ (Table A3).

Appendix D. Satellite-Floats Retrieval Slopes Forced to 0

Table A4. Results at 490 nm when the intercept is forced to 0. See Methods section for definitions of statistical performance metrics.

Sensor & Algorithm	BIAS	APD	RMSD	r	Slope
MODIS-Terra: K_d^{Lee05}	1.19	27.14	0.018	0.87	1.11
MODIS-Terra: K_d^{NN}	1.24	30.76	0.020	0.84	1.16
MODIS-Terra: $K_d^{NASA/ESA}$	1.08	19.97	0.018	0.86	1.04
MODIS-Aqua: K_d^{Lee05}	1.229	29.25	0.026	0.705	1.144
MODIS-Aqua: K_d^{NN}	1.219	29.88	0.021	0.804	1.132
MODIS-Aqua: $K_d^{NASA/ESA}$	1.081	20.08	0.019	0.830	1.049
VIIRS-SNPP: K_d^{Lee05}	1.289	34.65	0.022	0.828	1.190
VIIRS-SNPP: K_d^{Austin}	1.017	19.46	0.021	0.833	0.979
VIIRS-JPSS: K_d^{Lee05}	1.272	32.91	0.021	0.836	1.174
VIIRS-JPSS: $K_d^{NASA/ESA}$	1.007	18.89	0.020	0.840	0.969
OLCI-S3A: K_d^{Lee05}	1.280	30.73	0.019	0.708	1.205
OLCI-S3A: K_d^{NN}	1.154	25.91	0.020	0.648	1.073
OLCI-S3A: $K_d^{NASA/ESA}$	1.050	18.34	0.018	0.715	1.013
OLCI-S3B: K_d^{Lee05}	1.384	39.70	0.017	0.895	1.314
OLCI-S3B: K_d^{NN}	1.360	40.97	0.020	0.818	1.304
OLCI-S3A: $K_d^{NASA/ESA}$	1.116	19.44	0.013	0.905	1.061

Appendix E. Variation in Solar Zenith Angle
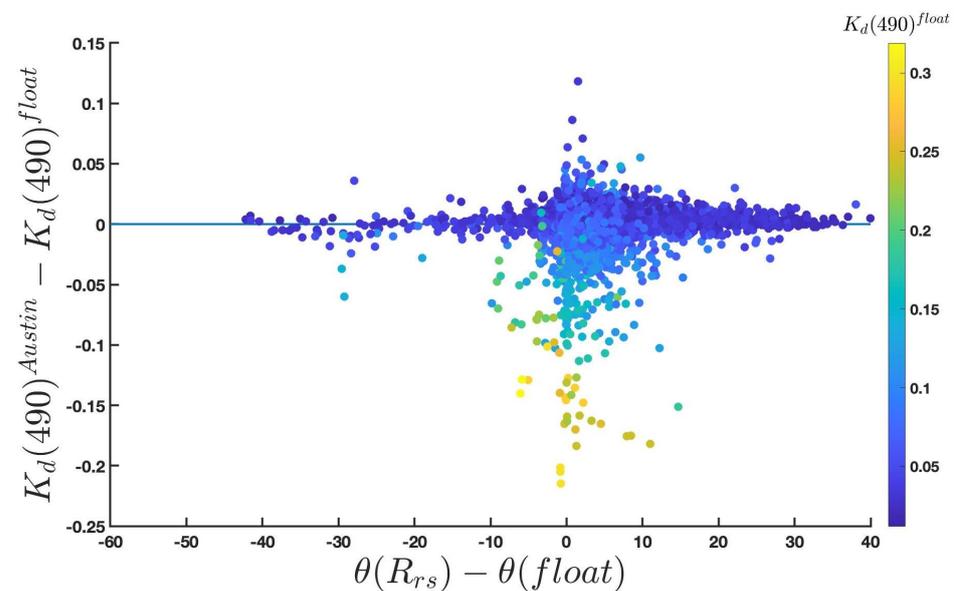
Figure A2. Residuals between for the operational algorithm ($K_d(490)^{NASA/ESA}$) and $K_d(490)^{float}$ as a function of the difference between the solar zenith angle at the time of the sensor overpass ($\theta(R_{rs})$) and the solar zenith angle when the BGC-Argo float surfaced ($\theta(float)$). Scatter points are colored by the value of $K_d(490)^{NASA/ESA}$.

Table A5. Summary statistics at 490 nm when biomes 18 and 19 are removed from the full dataset—5425 profiles removed, corresponding to the ones located in the Mediterranean and the Black Sea.

Sensor & Algorithm	BIAS	APD	RMSD	r	Slope	Intercept
MODIS-Terra: K_d^{Lee05}	1.21	29.49	0.023	0.85	0.88	0.012
MODIS-Terra: K_d^{NN}	1.27	34.11	0.025	0.82	0.88	0.014
MODIS-Terra: $K_d^{NASA/ESA}$	1.07	19.83	0.023	0.85	0.91	0.006
MODIS-Aqua: K_d^{Lee05}	1.24	32.05	0.038	0.62	0.86	0.016
MODIS-Aqua: K_d^{NN}	1.24	34.04	0.028	0.79	0.75	0.020
MODIS-Aqua: $K_d^{NASA/ESA}$	1.08	21.98	0.025	0.82	0.89	0.009
VIIRS-SNPP: K_d^{Lee05}	1.30	35.65	0.027	0.84	0.84	0.018
VIIRS-SNPP: $K_d^{NASA/ESA}$	1.02	19.76	0.026	0.85	0.83	0.008
VIIRS-JPS: K_d^{Lee05}	1.28	33.79	0.026	0.85	0.92	0.013
VIIRS-JPSS: $K_d^{NASA/ESA}$	1.01	18.98	0.025	0.86	0.85	0.007
OLCI-S3A: K_d^{Lee05}	1.26	28.99	0.021	0.71	0.93	0.009
OLCI-S3A: K_d^{NN}	1.15	25.41	0.022	0.66	0.88	0.007
OLCI-S3A: $K_d NASA/ESA$	1.02	18.42	0.021	0.71	0.74	0.008
OLCI-S3B: K_d^{Lee05}	1.32	34.68	0.016	0.93	0.61	0.022
OLCI-S3B: K_d^{NN}	1.25	32.64	0.018	0.90	0.61	0.020
OLCI-S3B: $K_d NASA/ESA$	1.08	17.45	0.015	0.93	0.63	0.014

Appendix F. Retrieval Method for Float $K_d(\lambda)$

To ensure the accurate retrieval of $K_d(\lambda)$ from $E_d(\lambda)$ by BGC-Argo floats, three different methods to compute $K_d(\lambda)$ were tested. The first one uses only the $E_d(\lambda)$ measurements between the surface and $z_{pd}(lambda)$ and performs an iterative least square fit on the exponential value of the $E_d(\lambda)$ profile. From here on, it is called the "lsq no extrapolation" method. This method assumes that $K_d(\lambda)$ is constant within Z_{pd}. The second method is the one used in [2] and extrapolates the $E_d(\lambda)$ profile to the surface in order to acquire $E_d(0^-)$ by performing a linear fit ("Linear fit"). This method also assumes that $K_d(\lambda)$ is constant within the z_{pd}. Lastly, $E_d(\lambda)$ was again extrapolated to the surface to acquire $E_d(\lambda, 0^-)$ but using a second-order polynomial, which implies that there are some variations with depth of $K_d(\lambda)$ within z_{pd}.
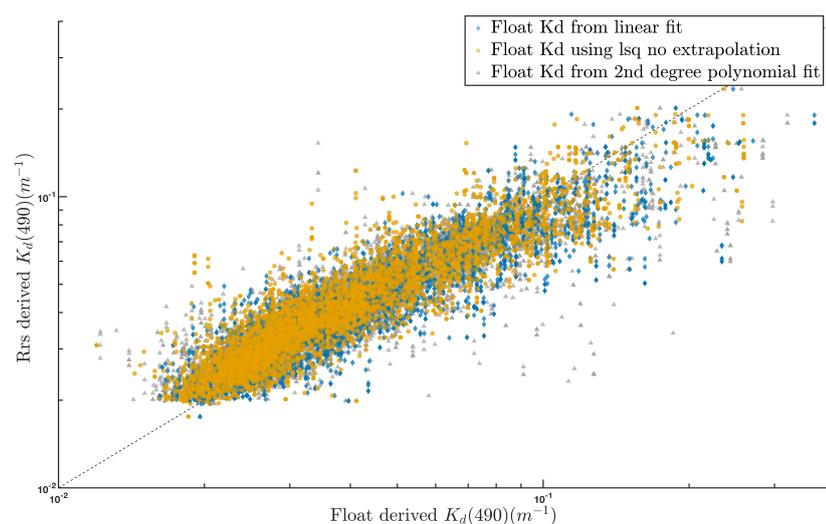
Figure A3. Test of the three potential methods for $K_d(\lambda)$ retrieval from $E_d(\lambda)$ measured by floats at 490 nm, i.e., a linear fit to extrapolate to $E_d(0^-)$ (blue), a second degree polynomial to extrapolate to $E_d(0^-)$ (grey) and an iterative least-square fit on existing $E_d(490)$ measurements (yellow).

Appendix G. Biomes Correlation and Analysis

Table A6. Results for each of the biomes for each sensor. Individual results for each of the satellite sensors within each of the biomes, with the red color indicating poor performance for a given satellite sensor and the green good statistical performance for a given sensor within a biome. The *NASA* empirical algorithm (see Methods section) was applied for the MODIS and the VIIRS sensors, whereas the *ESA* empirical algorithm was applied to the OLCI sensors. As they are both empirical algorithm, they were grouped together for the overall statistical analysis.

MODIS-TERRA

	Biome 4			Biome 6			Biome 7			Biome 8			Biome 9			Biome 11			Biome 13			Biome 15			Biome 16			Biome 18			Biome 19		
	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA
BIAS	1.14	1.27	1.07	1.27	1.26	1.25	1.45	1.50	1.13	0.99	1.13	0.91	1.02	1.00	0.95	1.27	1.31	1.00	1.40	1.49	1.07	1.16	1.15	1.12	1.32	1.35	1.14	1.11	1.12	1.05	1.21	1.25	1.07
ADP	22.15	31.17	15.84	22.61	28.31	30.18	36.78	41.05	18.28	51.63	52.92	45.05	25.64	29.51	24.12	28.41	32.04	19.49	38.76	46.87	15.80	27.76	27.33	22.25	28.97	33.40	18.51	26.61	28.05	21.95	25.18	29.51	18.59
RMSD	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.04	0.04	0.03	0.04	0.05	0.04	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01
r	0.64	0.52	0.73	0.02	0.32	0.26	0.74	0.80	0.85	−0.65	−0.70	−0.70	0.79	0.73	0.78	0.55	0.45	0.47	0.34	0.32	0.30	0.58	0.59	0.59	0.84	0.83	0.84	0.85	0.82	0.85	0.91	0.90	0.90
Slope	0.48	0.49	0.51	1.12	1.05	1.04	0.82	0.78	1.15	−0.55	−0.63	−0.22	0.60	0.49	0.55	0.56	0.47	0.54	0.49	0.48	0.47	0.44	0.41	0.47	1.35	1.14	1.19	0.72	0.72	0.76	0.80	0.83	0.67
Intercept	0.02	0.02	0.02	0.00	0.01	0.01	0.01	0.01	0.00	0.12	0.13	0.09	0.03	0.04	0.03	0.02	0.02	0.01	0.02	0.02	0.01	0.04	0.04	0.03	−0.01	0.01	0.00	0.02	0.02	0.01	0.01	0.01	0.01

MODIS-AQUA

	Biome 4			Biome 6			Biome 7			Biome 8			Biome 9			Biome 11			Biome 13			Biome 15			Biome 16			Biome 18			Biome 19		
	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA
BIAS	1.44	1.39	1.21	1.27	1.26	1.25	1.32	1.37	1.00	1.16	1.19	1.07	1.07	1.03	1.00	1.33	1.39	1.00	1.44	1.52	1.08	1.13	1.11	1.02	1.26	1.34	1.13	1.09	1.08	1.00	1.22	1.24	1.03
ADP	40.01	37.00	24.63	22.61	28.31	30.18	38.86	38.41	15.05	34.26	38.25	26.67	25.92	28.71	23.96	35.42	35.33	19.79	34.62	41.40	13.50	31.23	31.09	24.66	33.05	38.08	26.42	27.56	26.79	20.16	27.95	28.51	18.09
RMSD	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.03	0.03	0.02	0.04	0.04	0.04	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01
r	−0.56	−0.46	−0.29	0.02	0.32	0.26	0.92	0.90	0.94	−0.10	−0.15	−0.05	0.75	0.70	0.74	0.64	0.68	0.69	0.84	0.81	0.87	0.51	0.47	0.53	0.80	0.79	0.76	0.81	0.79	0.82	0.87	0.87	0.84
Slope	−0.72	−0.75	−0.45	1.12	1.05	1.04	0.89	0.72	1.12	−0.93	−1.00	−0.77	0.56	0.45	0.58	0.68	0.64	0.75	0.77	0.76	0.87	0.47	0.43	0.51	1.14	1.02	1.04	0.84	0.78	0.88	0.85	0.79	0.73
Intercept	0.06	0.06	0.05	0.00	0.01	0.01	0.01	0.02	0.00	0.15	0.16	0.13	0.04	0.04	0.04	0.02	0.02	0.02	0.01	0.02	0.00	0.04	0.04	0.03	0.01	0.01	0.00	0.02	0.02	0.01	0.01	0.02	0.01

VIIRS SNPP

	Biome 4			Biome 6			Biome 7			Biome 8			Biome 9			Biome 11			Biome 13			Biome 15			Biome 16			Biome 18			Biome 19		
	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA
BIAS	1.34	NaN	1.07	1.33	NaN	1.11	1.37	NaN	0.97	1.12	NaN	0.94	1.05	NaN	0.91	1.41	NaN	1.01	1.47		1.01	1.31	NaN	1.02	1.31	NaN	0.95	1.36	NaN	1.03	1.36	NaN	1.03
ADP	36.85	NaN	11.91	38.28	NaN	22.57	45.05	NaN	14.01	30.78	NaN	33.20	28.43	NaN	25.23	41.28	NaN	14.28	42.20		14.48	33.26	NaN	20.21	40.02	NaN	22.51	31.46	NaN	20.01	36.50	NaN	18.68
RMSD	0.01	NaN	0.00	0.01	NaN	0.01	0.01	NaN	0.00	0.03	NaN	0.03	0.04	NaN	0.04	0.01	NaN	0.00	0.01		0.01	0.02	NaN	0.02	0.02	NaN	0.02	0.02	NaN	0.02	0.02	NaN	0.02
r	0.63	NaN	0.73	0.64	NaN	0.49	0.84	NaN	0.86	−0.08	NaN	−0.12	0.73	NaN	0.74	0.67	NaN	0.69	0.78		0.80	0.58	NaN	0.59	0.57	NaN	0.53	0.77	NaN	0.77	0.81	NaN	0.80
Slope	0.81	NaN	0.79	0.29	NaN	0.39	0.92	NaN	0.96	−0.48	NaN	−0.40	0.61	NaN	0.58	0.68	NaN	0.60	0.87		0.87	0.53	NaN	0.52	0.74	NaN	0.65	0.74	NaN	0.73	0.66	NaN	0.56
Intercept	0.02	NaN	0.01	0.04	NaN	0.03	0.01	NaN	0.00	0.11	NaN	0.09	0.04	NaN	0.03	0.02	NaN	0.01	0.01		0.00	0.04	NaN	0.03	0.02	NaN	0.02	0.02	NaN	0.01	0.02	NaN	0.01

VIIRS JPSS

	Biome 4			Biome 6			Biome 7			Biome 8			Biome 9			Biome 11			Biome 13			Biome 15			Biome 16			Biome 18			Biome 19		
	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA
BIAS	1.36	NaN	1.06	1.34	NaN	1.12	1.28	NaN	0.94	0.99	NaN	0.85	1.07	NaN	0.90	1.38	NaN	0.98	1.39	NaN	0.96	1.23	NaN	1.02	1.23	NaN	0.94	1.14	NaN	0.93	1.33	NaN	1.02
ADP	32.62	NaN	10.48	36.67	NaN	16.84	38.38	NaN	11.95	22.87	NaN	29.39	26.94	NaN	24.07	40.59	NaN	15.13	39.60	NaN	14.95	33.88	NaN	22.69	36.31	NaN	19.07	30.09	NaN	19.25	34.61	NaN	18.16
RMSD	0.01	NaN	0.00	0.01	NaN	0.01	0.01	NaN	0.00	0.02	NaN	0.03	0.04	NaN	0.04	0.01	NaN	0.01	0.01	NaN	0.01	0.03	NaN	0.03	0.02	NaN	0.01	0.02	NaN	0.02	0.02	NaN	0.01
r	0.53	NaN	0.66	0.59	NaN	0.57	0.83	NaN	0.85	−0.05	NaN	−0.07	0.75	NaN	0.76	0.23	NaN	0.23	0.80	NaN	0.83	0.51	NaN	0.52	0.88	NaN	0.84	0.79	NaN	0.78	0.80	NaN	0.80
Slope	0.68	NaN	0.65	0.86	NaN	1.04	0.89	NaN	0.94	−0.56	NaN	−0.47	0.67	NaN	0.61	0.57	NaN	0.47	0.76	NaN	0.75	0.47	NaN	0.47	0.98	NaN	0.86	0.77	NaN	0.76	0.67	NaN	0.56
Intercept	0.02	NaN	0.01	0.02	NaN	0.00	0.01	NaN	0.00	0.12	NaN	0.10	0.03	NaN	0.04	0.02	NaN	0.01	0.01	NaN	0.01	0.01	NaN	0.01	0.04	NaN	0.03	0.01	NaN	0.01	0.02	NaN	0.01

OLCI-S3A

	Biome 4			Biome 6			Biome 7			Biome 8			Biome 9			Biome 11			Biome 13			Biome 15			Biome 16			Biome 18			Biome 19		
	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA
BIAS	1.31	1.17	1.05	1.25	1.22	1.04	1.21	1.07	0.96	0.83	0.78	0.74	1.01	1.00	0.89	1.34	1.15	1.04	1.40	1.26	1.08	1.23	1.24	1.02	1.23	1.24	0.94	1.04	0.92	0.85	1.23	1.05	0.99
ADP	26.90	22.85	17.54	43.62	44.44	25.28	24.28	23.40	17.45	52.60	78.07	74.93	12.40	14.11	15.95	29.48	21.40	14.75	32.93	26.03	12.64	33.88	28.51	22.69	36.31	28.51	19.07	32.75	32.76	20.51	32.88	25.26	17.77
RMSD	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.04	0.05	0.05	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.03	0.01	0.03	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.02	0.01
r	0.52	0.53	0.56	0.94	0.97	0.96	0.67	0.58	0.64	NAN	NaN	NaN	0.93	0.90	0.93	0.65	0.61	0.78	0.19	−0.03	0.19	0.51	0.87	0.52	0.88	0.87	0.84	0.85	0.68	0.89	0.65	0.58	0.64
Slope	0.64	0.74	0.57	1.41	2.36	1.41	0.95	0.97	0.81	NAN	NaN	NaN	1.06	0.92	1.09	1.02	1.08	0.84	1.56	−1.57	0.99	0.47	0.79	0.47	0.98	0.79	0.86	0.62	0.83	0.63	0.71	0.84	0.67
Intercept	0.02	0.01	0.01	0.00	−0.03	−0.01	0.01	0.00	0.00	NAN	NaN	NaN	0.00	0.01	−0.01	0.01	0.00	0.00	0.01	0.06	0.00	−0.01	0.06	0.00	0.04	0.02	0.03	0.01	0.02	0.01	0.02	0.01	0.01

OLCI-S3B

	Biome 4			Biome 6			Biome 7			Biome 8			Biome 9			Biome 11			Biome 13			Biome 15			Biome 16			Biome 18			Biome 19		
	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA	Lee	NN	NASA/ESA
BIAS	1.19	1.24	1.08	1.23	1.22	1.08	1.29	0.48	1.02	1.27	0.49	1.01	1.28	1.15	1.05	1.38	1.36	1.12	1.44	1.28	1.12	1.31	1.35	1.08	1.45	1.34	1.15	0.79	0.79	0.69	1.31	1.41	1.25
ADP	27.14	30.76	19.97	29.13	30.00	20.19	34.71	158.21	19.46	32.89	140.55	18.76	30.73	25.91	18.34	39.70	40.97	19.44	38.86	33.61	17.74	30.59	36.75	14.38	32.17	27.25	14.13	58.61	56.02	71.42	12.90	23.58	11.17
RMSD	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.04	0.04	0.05	0.01	0.02	0.01
r	0.87	0.84	0.86	0.83	0.81	0.83	0.82	0.73	0.82	NaN	NaN	NaN	0.71	0.65	0.71	0.90	0.82	0.91	0.70	0.46	0.87	1.00	1.00	1.00	0.97	0.97	0.97	0.79	1.00	1.00	0.96	0.81	0.96
Slope	0.79	0.78	0.80	0.83	0.75	0.86	0.79	0.91	0.76	NaN	NaN	NaN	0.92	0.83	0.72	0.87	0.67	0.63	0.55	0.51	0.52	0.10	1.40	0.00	0.82	1.08	0.79	0.08	0.07	0.07	0.94	0.78	1.01
Intercept	0.01	0.02	0.01	0.02	0.02	0.01	0.02	−0.01	0.01	NaN	NaN	NaN	0.01	0.01	0.01	0.02	0.02	0.02	0.03	0.02	0.02	0.04	0.00	0.04	0.02	0.01	0.01	0.06	0.06	0.05	0.01	0.03	0.00

References

1. Mobley, C. *The Oceanic Optics Book*; International Ocean Colour Coordinating Group (IOCCG): Monterey, CA, USA, 2022. [CrossRef]
2. Xing, X.; Boss, E.; Zhang, J.; Chai, F. Evaluation of Ocean Color Remote Sensing Algorithms for Diffuse Attenuation Coefficients and Optical Depths with Data Collected on BGC-Argo Floats. *Remote Sens.* **2020**, *12*, 2367. [CrossRef]
3. Jamet, C.; Loisel, H.; Dessailly, D. Retrieval of the spectral diffuse attenuation coefficient $K_d(\lambda)$ in open and coastal ocean waters using a neural network inversion: Retrieval of Diffuse Attenuation. *J. Geophys. Res. Ocean.* **2012**, *117*, C10023. [CrossRef]
4. Morel, A.; Huot, Y.; Gentili, B.; Werdell, P.J.; Hooker, S.B.; Franz, B.A. Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. *Remote Sens. Environ.* **2007**, *111*, 69–88. [CrossRef]
5. Incorporated, S. *OCR 504 User Manual*; Technical Report. Available online: https://www.seabird.com/asset-get.download.jsa?id=54627868876 (accessed on 21 July 2022)
6. Organelli, E.; Claustre, H.; Bricaud, A.; Barbieux, M.; Uitz, J.; D'Ortenzio, F.; Dall'Olmo, G. Bio-optical anomalies in the world's oceans: An investigation on the diffuse attenuation coefficients for downward irradiance derived from Biogeochemical Argo float measurements: WORLD'S OCEAN BIO-OPTICAL ANOMALIES. *J. Geophys. Res. Ocean.* **2017**, *122*, 3543–3564. [CrossRef]
7. Gordon, H.R.; McCluney, W.R. Estimation of the Depth of Sunlight Penetration in the Sea for Remote Sensing. *Appl. Opt.* **1975**, *14*, 413. [CrossRef] [PubMed]
8. Mueller, J. *SeaWiFS Algorithm for the Diffuse Attenuation Coefficient, K(490), Using Water-Leaving Radiances at 490 and 555 nm*; Technical Report Part 3; NASA Goddard Space Flight Center: Greenbelt, MD, USA, 2000.
9. Morel, A. Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). *J. Geophys. Res. Ocean.* **1988**, *93*, 10749–10768. [CrossRef]
10. Werdell, P.J.; Bailey, S.W. An improved in-situ bio-optical dataset for ocean color algorithm development and satellite data product validation. *Remote Sens. Environ.* **2005**, *98*, 122–140. [CrossRef]
11. IOCCG. *Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications*; Technical Report 5; IOCCG: Dartmouth, NS, Canada, 2006.
12. Fay, A.R.; McKinley, G.A. Global open-ocean biomes: Mean and temporal variability. *Earth Syst. Sci. Data* **2014**, *6*, 273–284. [CrossRef]
13. Bailey, S.W.; Werdell, P.J. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sens. Environ.* **2006**, *102*, 12–23. [CrossRef]
14. Bisson, K.M.; Boss, E.; Westberry, T.K.; Behrenfeld, M.J. Evaluating satellite estimates of particulate backscatter in the global open ocean using autonomous profiling floats. *Opt. Express* **2019**, *27*, 30191. [CrossRef]
15. Zhang, T.; Fell, F. An empirical algorithm for determining the diffuse attenuation coefficient K_d in clear and turbid waters from spectral remote sensing reflectance: K_d in clear and turbid waters. *Limnol. Oceanogr. Methods* **2007**, *5*, 457–462. [CrossRef]
16. Dierssen, H.M.; Vandermeulen, R.A.; Barnes, B.B.; Castagna, A.; Knaeps, E.; Vanhellemont, Q. QWIP: A Quantitative Metric for Quality Control of Aquatic Reflectance Spectral Shape Using the Apparent Visible Wavelength. *Front. Remote. Sens.* **2022**, *3*, 869611. [CrossRef]
17. Austin, R.W.; Petzold, T.J. The Determination of the Diffuse Attenuation Coefficient of Sea Water Using the Coastal Zone Color Scanner. In *Oceanography from Space*; Gower, J.F.R., Ed.; Springer: Boston, MA, USA, 1981; pp. 239–256. [CrossRef]
18. Lee, Z.P. Diffuse attenuation coefficient of downwelling irradiance: An evaluation of remote sensing methods. *J. Geophys. Res.* **2005**, *110*, C02017. [CrossRef]
19. Lee, Z.; Carder, K.L.; Arnone, R.A. Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Appl. Opt.* **2002**, *41*, 5755. [CrossRef]
20. Lee, Z.; Hu, C.; Shang, S.; Du, K.; Lewis, M.; Arnone, R.; Brewin, R. Penetration of UV-visible solar radiation in the global oceans: Insights from ocean color remote sensing: Penetration of UV-Visible Solar Light. *J. Geophys. Res. Ocean.* **2013**, *118*, 4241–4255. [CrossRef]
21. Westberry, T.K.; Boss, E.; Lee, Z. Influence of Raman scattering on ocean color inversion models. *Appl. Opt.* **2013**, *52*, 5552. [CrossRef]
22. Loisel, H.; Stramski, D.; Dessailly, D.; Jamet, C.; Li, L.; Reynolds, R.A. An Inverse Model for Estimating the Optical Absorption and Backscattering Coefficients of Seawater From Remote-Sensing Reflectance Over a Broad Range of Oceanic and Coastal Marine Environments: Inversion of Seawater IOPS. *J. Geophys. Res. Ocean.* **2018**, *123*, 2141–2171. [CrossRef]
23. Xing, X.; Boss, E. Chlorophyll-Based Model to Estimate Underwater Photosynthetically Available Radiation for Modeling, *In-Situ*, and Remote-Sensing Applications. *Geophys. Res. Lett.* **2021**, *48*, e2020GL092189. [CrossRef]
24. Lee, Z. Penetration of solar radiation in the upper ocean: A numerical model for oceanic and coastal waters. *J. Geophys. Res.* **2005**, *110*, C09019. [CrossRef]
25. Brewin, R.J.; Sathyendranath, S.; Müller, D.; Brockmann, C.; Deschamps, P.Y.; Devred, E.; Doerffer, R.; Fomferra, N.; Franz, B.; Grant, a.; et al. The Ocean Colour Climate Change Initiative: III. A round-robin comparison on in-water bio-optical algorithms. *Remote Sens. Environ.* **2015**, *162*, 271–294. [CrossRef]
26. Yu, X.; Salama, M.S.; Shen, F.; Verhoef, W. Retrieval of the diffuse attenuation coefficient from GOCI images using the 2SeaColor model: A case study in the Yangtze Estuary. *Remote Sens. Environ.* **2016**, *175*, 109–119. [CrossRef]

27. Zhang, Y.; Xu, Z.; Yang, Y.; Wang, G.; Zhou, W.; Cao, W.; Li, Y.; Zheng, W.; Deng, L.; Zeng, K.; et al. Diurnal Variation of the Diffuse Attenuation Coefficient for Downwelling Irradiance at 490 nm in Coastal East China Sea. *Remote Sens.* **2021**, *13*, 1676. [CrossRef]

28. Zheng, X.; Dickey, T.; Chang, G. Variability of the downwelling diffuse attenuation coefficient with consideration of inelastic scattering. *Appl. Opt.* **2002**, *41*, 6477. [CrossRef] [PubMed]

29. Organelli, E.; Leymarie, E.; Zielinski, O.; Uitz, J.; D'Ortenzio, F.; Hervé, C. Hyperspectral Radiometry on Biogeochemical-Argo Floats: A Bright Perspective for Phytoplankton Diversity. *Oceanography* **2021**, *34*, 90–91. [CrossRef]

30. Babin, M. Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. *J. Geophys. Res.* **2003**, *108*, 3211. [CrossRef]