Remote Sensing of Environment 265 (2021) 112667

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

The QAA-RGB: A universal three-band absorption and backscattering retrieval algorithm for high resolution satellite sensors. Development and implementation in ACOLITE

Jaime Pitarch [a,*], Quinten Vanhellemont [b]

[a] Consiglio Nazionale delle Ricerche (CNR), Istituto di Scienze Marine (ISMAR), Via del Fosso del Cavaliere 100, 00133 Rome, Italy
[b] Royal Belgian Institute of Natural Sciences, Operational Directorate Natural Environments, Vautierstraat 29, 1000 Brussels, Belgium

ARTICLE INFO

Editor: Menghua Wang

Keywords:
Semianalytical algorithm
Coastal
Inland
Optical
Landsat
Sentinel
PACE
Absorption
Backscattering
Diffuse attenuation
Secchi

ABSTRACT

We present the "QAA-RGB", a semianalytical algorithm that retrieves total absorption, particle backscattering, as well as the diffuse attenuation coefficient and the Secchi disk depth, from eighteen metre and decametre satellite sensors, including present and heritage Landsat's data, Sentinel 2 at 10 m and an array of commercial satellites such as PlanetScope, Pléiades and Worldview. The QAA-RGB is a minimal version of the Quasi-Analytical Algorithm (QAA), and therefore keeps its robustness and general validity across different water types. It ingests remote-sensing reflectance at only three bands, centred on red, green and blue wavelengths. Retrievals of non-water absorption at 555 nm ($a_{nw}(555)$) are satisfactory, with a standard deviation of $\sigma \sim 50\%$ up to $a_{nw}(555) < 2$ m^{-1}, based on in situ datasets. Careful calibration results in unbiased retrievals across all sensors. Validation to independent in situ data is also provided. The QAA-RGB has been implemented in the ACOLITE generic processor, and is openly accessible to the community. Consistency of near-coincident observations from different satellite sensors is demonstrated.

1. Introduction

Since the launch of the Thematic Mapper (TM) on board Landsat 4 in 1982 and Landsat 5 in 1984, geophysical parameter retrieval from satellite imagery is possible at 30 m spatial resolution with a 16 day track repeat cycle. The band configuration and signal to noise ratio are adequate to allow derivation of the remote-sensing reflectance (R_{rs}; sr^{-1}) and hence aquatic parameter retrieval, if the water is neither too clear nor dominated by dissolved substances (Pahlevan et al., 2018). Successors of these early sensors were built ensuring overlap for intercalibration and absence of temporal gaps, with the Enhanced Thematic Mapper (ETM) and Operational Land Imager (OLI) on board of Landsat 7 (1999) and Landsat 8 (2013), presently in operation. Three decades after the launch of TM, the European Space Agency launched a set of Landsat-like sensors within the Copernicus Programme, i.e. the MultiSpectral Imager (MSI) onboard Sentinel 2 A and B (S2A and S2B), leading to a 5 day global revisit period with the two units in opposing phase. MSI has a spatial resolution of 10 m for three bands in the visible domain and one in the Near Infrared (NIR) and other bands at 20 and 60 m. Both Earth Observation programs will continue for the next decades, and hence interoperability of past, current, and future sensors is crucial (Pahlevan et al., 2019; Xu et al., 2020). In parallel, several commercial operators launched satellites with even finer spatial resolution, allowing for metre-scale acquisitions. Although these latter satellites were developed mostly for cartographic, rapid response, and military purposes, it was recently shown that there is potential for using these sensors for aquatic applications, e.g. mapping of near shore R_{rs} and turbidity (Vanhellemont, 2019b, 2020; Vanhellemont and Ruddick, 2018) and bottom depth (Gabr et al., 2020; Poursanidis et al., 2019).

Algorithm developers for this class of high resolution sensors (HRS) have often made redundant efforts by proposing empirical algorithms tuned with local data, with unknown applicability outside the calibration conditions, leading to algorithm proliferation and confusion among non-expert users. Clearly, there is a need for algorithms with known uncertainties that have a clearly known and ideally wide range of applicability, in order to monitor many, potentially disconnected, water

* Corresponding author.
E-mail address: jaime.pitarchportero@artov.ismar.cnr.it (J. Pitarch).

https://doi.org/10.1016/j.rse.2021.112667
Received 15 April 2021; Received in revised form 18 August 2021; Accepted 19 August 2021
Available online 26 August 2021

Table 1
Summary of relative spectral responses (RSR) of the RGB bands of the different sensors. For every sensor and band combination, the band number, band averaged wavelength calculated and full width at half maximum (between brackets) is given.

	B	G	R	Spatial res. (m)
Landsat 4	1–486(66)	2–571(80)	3–660(69)	30
Landsat 5	1–486(66)	2–570(81)	3–660(67)	30
Landsat 7	1–479(73)	2–561(82)	3–661(61)	30
Landsat 8	2–483(60)	3–561(57)	3–665(37)	30
Sentinel 2A	2–492(64)	3–560(35)	4–665(31)	10
Sentinel 2B	2–492(65)	3–559(35)	4–665(30)	10
Pléiades 1A	1–501(80)	2–561(83)	3–650(80)	2*
Pléiades 1B	1–505(78)	2–558(83)	3–663(80)	2*
PlanetScope 0c	1–493(70)	2–542(100)	3–621(100)	3*
PlanetScope 0d05	1–493(70)	2–542(100)	3–621(100)	3*
PlanetScope 0d06	1–493(70)	2–542(100)	3–621(100)	3*
PlanetScope 0e	1–517(70)	2–552(100)	3–663(110)	3*
PlanetScope 0f	1–506(70)	2–546(100)	3–625(100)	3*
PlanetScope 22	1–492(53)	2–566(40)	3–666(33)	3*
RapidEye	1–477(70)	2–556(70)	3–658(55)	5*
Worldview2	2–479(61)	3–548(70)	5–659(60)	1.85**
WorldViev3	2–482(59)	3–547(69)	5–660(60)	1.24**
VENμS_VSSC	3–492(38)	4–555(38)	7–666(28)	5.35**

(* as resampled by the satellite operator, ** at nadir).

bodies with different optical characteristics that may lack in situ calibration data. In the marine optics community, efforts have been directed in this direction, but unfortunately, knowledge transfer to the Landsat and inland water communities has been poor. This deficient connectivity has likely been amplified by the limited spectral information provided by sensors that are restricted to three bands in the visible range for heritage Landsat satellites and the high cost of Landsat imagery before the opening of the archive (Wulder et al., 2012).

The current paper aims at bridging this knowledge gap and at providing the water quality monitoring community with tools to more fully exploit high resolution satellite imagery. Good candidates for universal application are the so-called semianalytical algorithms, which are partly based on relationships that derive from radiative transfer theory. The Quasi-Analytical Algorithm (QAA) was developed for moderate resolution imagery (multispectral data) such as SeaWiFS, MODIS and MERIS (Lee et al., 2002). Its last update is version 6 (Lee, 2014). Due to its transparent and modular structure, it has become the most popular and consolidated semianalytical algorithm within the marine optics community. It has been validated by independent researchers (Brewin et al., 2015; Pitarch et al., 2020; Zheng et al., 2014) and is implemented in operational processing chains (Colella et al., 2021; Jackson, 2020).

Although moderate resolution imagery provides high-quality data over oceans, near-coastal waters and large lakes, HRS are preferred for many inland and near-shore waters. Moreover, they are often the only option due to the small spatial scales to be resolved. Adaptation of the QAA to HRS is hence desirable, as it was made by Lee et al. (2016) for Landsat 8. A similar development for the rest of the Landsat series would be beneficial for long-term monitoring, and the consideration of sensors from other agencies and companies would increase the temporal frequency and spatial resolution.

In the present paper we make a generic adaptation of the QAA that ingests data at three bands available in the visible domain, covering approximately red, green and blue wavelengths (RGB), a requirement that is fulfilled by all sensors of interest. Our "QAA-RGB" is calibrated to a common standard, namely the regular QAA that is commonly applied to moderate resolution imagery. This procedure is made for every HRS individually as the empirical coefficients in the QAA depend on the relative spectral response function (RSR) of the sensor.

This article summarizes the development of the QAA-RGB and reports the specific calibration for eighteen HRS, including Landsat 4, 5, 7 and 8, Sentinel 2 at 10 m (RGB bands), as well as for a number of commercial sensors (Table 1). It will be shown that, despite the limitations of using only three bands, the QAA-RGB is able to retrieve the bulk inherent optical properties (IOPs) absorption (a; m^{-1}) and backscattering (b_b; m^{-1}). The phytoplankton and non-algal absorption fractions are not retrieved due to the lack of a band at 412 nm or similar. However, the bulk IOPs are already useful: the particle backscattering (b_{bp}; m^{-1}) is directly related to turbidity, the spectral slope of b_{bp}, assuming a power-law dependence (η; unitless) can be used as a proxy for particle size distribution, and a can be used in combination with other variables external to ocean colour like temperature or salinity to derive further information. a and b_b are combined to estimate the diffuse attenuation coefficient (K_d; m^{-1}) (Lee et al., 2013). K_d represents the attenuation of sunlight in the water and is relevant for phytoplankton bloom modeling and thermal heating in aquatic models, and can also be readily linked to other parameters like euphotic depth or visibility, that have many applications in coastal and inland waters. In addition, K_d relates to the Secchi disk depth (z_{SD}; m), a water quality parameter that is widely measured, has a long historic archive in coastal and inland waters, and has recently gained attention from a remote sensing perspective (Lee et al., 2015). For these reasons, a seamless z_{SD} retrieval across all HRS, derived from the QAA-RGB, is presented.

2. Methods

2.1. HRS overview

The QAA-RGB is calibrated for the eighteen HRS listed in Table 1. The full Landsat series since Landsat 4 and both Sentinel 2 sensors are included. Their RSR were downloaded from the USGS and ESA websites, respectively. In addition, to extend the application to metre scale sensors, a number of commercial and demonstration satellite sensors with RGB bands are also included, namely the Pléiades, PlanetScope, RapidEye, WorldView and VENμS sensors. Their RSR were downloaded from the ACOLITE GitHub page at https://github.com/acolite.

The RGB bands on the different HRS are described in Table 1 in terms of band number, center wavelength and full width at half maximum (FWHM). There are clear differences in the position and width of the RGB bands across the sensors, which will impact algorithm coefficients and performance.

2.2. The QAA-RGB

The QAA has been widely analyzed and recalibrated in many articles and an exhaustive description is not needed. Briefly, the QAA is a sequential flow of steps, some of which are empirical relationships between R_{rs} and some IOPs, whereas others follow from radiative transfer theory. A key initial step is the linkage of a predictor (χ; unitless), derived from R_{rs}, to the non-water absorption coefficient at 555 nm, $a_{nw}(555)$. The band 555 nm is chosen because it is relatively insensitive to pigment variability. This step is similar to the derivation of chlorophyll from band ratios, as done in the OCx algorithms (O'Reilly and Werdell, 2019).

After $a_{nw}(555)$ retrieval, $b_b(555)$ is derived analytically by forcing R_{rs} closure. Then, b_b is extended to the whole spectral range by assuming a power law spectral shape, whose exponent (η) is derived empirically. Finally, absorption is derived at all bands by forcing closure again. This workflow is also used in the QAA-RGB and the required adaptations are described in the following subsections.

The updated QAA (Lee, 2014) switches to 670 nm as a reference wavelength to start the retrieval for turbid waters when $R_{rs}(670) > 0.0015\ sr^{-1}$. This switch is not implemented in the QAA-RGB because we found that a single relationship, described next, is sufficient to build a robust link between R_{rs} and the IOPs, based on the empirical evidence we could collect.

Table 2
List of all datasets used in this article, detailing dataset name or PI and references, number of samples (N), date range, included variables, some basic statistics in the form of percentiles for absorption, covered geographic areas and the usage in the present study.[a, b]

Dataset and references	N	Date range	Variables	$a_{nw}(555)$ percentiles: p_{10}-p_{50}-p_{90} (m^{-1})	Area(s)	Usage
Lee	869	Undisclosed	R_{rs}[a]	Not applicable	Undisclosed	- Obtain relationships between SeaWiFS/Satlantic R_{rs} and high resolution sensor R_{rs} - Harmonize the B/G ratio to the SeaWiFS reference - Fine tune the sensor-specific z_{SD} to the SeaWiFS reference
PACE Schaeffer (Casey et al., 2020) a.r.t.	398	Jan 2010 – Nov. 2011	R_{rs}[a], $a_{nw}(555)$	0.1374, 0.2712, 0.9105	Northern Gulf of Mexico estuaries, USA	- Calibrate the $a_{nw}(555)$ vs. χ relationship
PACE Mouw (Casey et al., 2020) a.r.t.	71	May 2013 – May 2016	R_{rs}[a], $a_{nw}(555)$	0.0101, 0.0294, 0.0805	Lake Superior, USA	- Calibrate the $a_{nw}(555)$ vs. χ relationship
PACE Craig (Casey et al., 2020) a.r.t.	43	Feb. 2009 – Mar. 2010	R_{rs}[a], $a_{nw}(555)$	0.0362, 0.0683, 0.1166	Coastal northwestern Atlantic ocean	- Calibrate the $a_{nw}(555)$ vs. χ relationship
Valente19 (Valente et al., 2019)	317	Mar. 1999 – Oct. 2006	R_{rs}[b], $a_{nw}(555)$	0.0033, 0.0285, 0.2045	Multiple areas, mostly coastal	- Calibrate the $a_{nw}(555)$ vs. χ relationship
PACE BIOSOPE (Casey et al., 2020) a.r.t.	15	Nov. 2004 – Dec. 2004	R_{rs}[a], $a_{nw}(555)$	0.0011, 0.0029, 0.0270	Southeastern Pacific ocean	- Compare to the calibrated $a_{nw}(555)$ vs. χ relationship
CNR-ISMAR (Pitarch et al., 2016)	15	Mar. 2015 – Apr. 2015	R_{rs}[b], $a_{nw}(555)$	0.0072, 0.0155, 0.0639	Western Adriatic sea	- Compare to the calibrated $a_{nw}(555)$ vs. χ relationship
CNR-IREA (Bresciani et al., 2018; Bresciani et al., 2013; Manzo et al., 2015)	35	Sep. 2014 – Jul. 2017	R_{rs}[a], $a_{nw}(555)$	0.0477, 0.1701, 0.4076	Mantua and Iseo lakes, northern Italy	- Compare to the calibrated $a_{nw}(555)$ vs. χ relationship

a.r.t.: and references therein.
[a] Hyperspectral.
[b] At SeaWiFS/Satlantic bands.

2.2.1. The χ to $a_{nw}(555)$ relationship

The estimation of $a_{nw}(555)$ is made by reducing several bands to a certain absorption predictor χ. The multispectral QAA algorithm (Lee, 2014) uses two blue bands, one green band and one red band (443 nm, 490 nm, 555 nm, 670 nm) and proposes $\chi = log_{10}\left(\frac{r_{rs}(443)+r_{rs}(490)}{r_{rs}(555)+\frac{5r_{rs}(670)^2}{r_{rs}(490)}}\right)$, where r_{rs} (sr^{-1}) is the below-surface remote-sensing reflectance.

That relationship has to be modified due to the limited spectral information that is available (Table 1). We propose the following analog expression for the predictor χ as a function of the RGB bands:

$$\chi = log\left(\frac{2B}{G+\frac{5R^2}{B}}\right) \quad (1)$$

In Eq. (1), R, G and B represent above-surface R_{rs} at the respective red, green and blue bands. This expression is solely based on band ratios, which effectively capture the absorption influence while being quite immune to backscattering (Lee et al., 1998). For blue waters, Eq. (1) reduces to a B/G ratio like in the OCx algorithms (O'Reilly and Werdell, 2019), and for the high range of absorption, it reduces to a B/R ratio. By combining in a predictive functional form several band ratios that capture the variability at different parts of the range, the range of applicability is enlarged (Lee et al., 1998).

The predictor χ is related to $\log_{10}[a_{nw}(555)]$ through a polynomial of χ, i.e., $P(\chi) = \sum_{k=0}^{N}(p_k\chi^k)$, with specific coefficients for each sensor:

$$log_{10}[a_{nw}(555)\,] = P(\chi) \quad (2)$$

The coefficients of $P(\chi)$ in Eq. (2) are found by ordinary least squares regression, using matched reflectance and absorption data, as shown in Section 3.1.

2.2.2. Harmonization of the blue-green ratio

Raman scattering correction of R_{rs} prior to ingestion by the QAA was recommended (Lee et al., 2013). The effect of this step is negligible for turbid waters but significant in clear waters, to the point that QAA backscattering retrievals from satellite R_{rs} showed better agreement to in situ backscattering if Raman scattering was corrected for (Pitarch et al., 2020). Operationally, Lee et al. (2013) provided formulas that depend on the 443-to-555 ratio of R_{rs}, but for the HRS, the bluest band is positioned near 490 nm. Given this mismatch and the differences in the spectral responses of HRS, the B/G ratio is intrinsically different across them and hence, a harmonization to a common standard is needed in order to standardize expressions that involve the B/G ratio across sensors. The same harmonization is needed for the retrieval of the particle backscattering slope η, estimated in the QAA with the saturating exponential model $\eta = 2\left\{1 - 1.2exp\left[-0.9\frac{r_{rs}(443)}{r_{rs}(555)}\right]\right\}$.

The harmonization is performed by replacing the B/G ratio in the algorithm with a polynomial of itself $Q\left(\frac{B}{G}\right)$ that is sensor-specific. The polynomial coefficients are found by comparing simulated 443-to-555 ratio to HRS B/G ratio, as shown in Section 3.2. Once $Q\left(\frac{B}{G}\right)$ has been determined, the corresponding QAA formulas can be used without changing the coefficients. For example, η is retrieved in the QAA-RGB with the formula:

$$\eta = 2\left\{1 - 1.2exp\left[-0.9Q\left(\frac{B}{G}\right)\right]\right\} \quad (3)$$

2.2.3. Diffuse attenuation and Secchi disk depth retrieval

The QAA-RGB proceeds further by estimating spectral b_b, followed by spectral a, that is solved analytically from b_b and R_{rs}. Here, the values of the pure water absorption (a_w) and backscattering (b_{bw}) are needed for the bands of every HRS and they are provided in Appendix A. Strictly speaking, the QAA only covers the retrieval of absorption and backscattering, but given the high relevance of the following quantities, they are considered here as well.

From a and b_b, K_d is estimated at all HRS bands:

$$K_d = a + m_1\left(1 - \gamma\frac{b_{bw}}{b_b}\right)(1 - m_2e^{-m_3a})b_b \quad (4)$$

where m_1 (unitless), m_2 (unitless) m_3 (m) and γ (unitless) are known

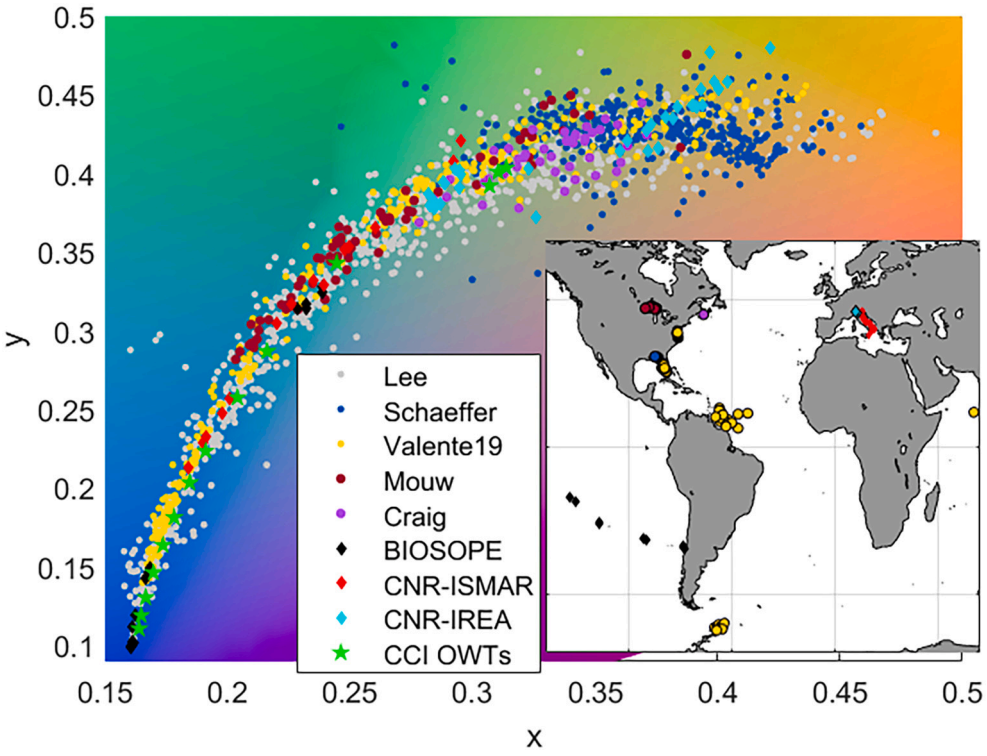

Fig. 1. A comprehensive view of all remote sensing reflectance datasets used in this article, projected onto the CIE1931 (x,y) colour space. See Pitarch et al. (2019) for a precise explanation on the generation and interpretation of the spectral projection onto the CIE1931 scale. The optical water types used for classification within the OC-CCI project (Jackson et al., 2017) have been included to highlight the range covered by each dataset. Outliers evidence uncommon spectral shapes, likely caused by measurement errors. A station map is included.

constants (Lee et al., 2013).

Now, we specifically focus on the Secchi disk depth z_{SD} (Pitarch, 2020), a very popular parameter in coastal and inland water monitoring that in some places has been recorded for multiple decades (Capuzzo et al., 2015). z_{SD} is derived from K_d using the following equation (Lee et al., 2015):

$$z_{SD,b} = \frac{log\left(\frac{|0.14 - R_{rs}(\lambda_m)|}{0.013}\right)}{2.5 K_d(\lambda_m)} \quad (5)$$

where λ_m is the band where the spectral minimum of K_d is located. This derivation is analytical and unifies the z_{SD} derivation for HRS, thereby avoiding empirical formulas driven by regional data.

The subscript "b" in Eq. (5) has been intentionally included to indicate that the retrieval still includes a small bias caused by the RSR differences of each HRS. The harmonization described above, i.e. the sensor-specific χ to $a_{nw}(555)$ calibration (Eq. (1)) and the blue to green ratio adjustment to a common standard (Eq. (3)), gives a seamless IOP and z_{SD} retrieval regardless of the sensor. Still, the retrieved z_{SD} may slightly differ between sensors. Firstly, there can be residual differences caused by the association of the estimated $a_{nw}(555)$ to the green band, which may be not exactly centered at 555 nm and generally covers a rather broad spectral range. Secondly, a coarse spectral sampling can lead to an overestimation of the minimum K_d and subsequent underestimation of z_{SD}. Therefore, in order to remove potentially remaining biases, a polynomial of the biased estimate, $S(z_{SD,\,b})$, is proposed to obtain the unbiased estimate:

$$z_{SD} = S\left(z_{SD,b}\right) \quad (6)$$

The coefficients of $S(z_{SD,\,b})$ are found by comparing QAA-RGB and QAA z_{SD} retrievals and regressing the former to the latter (Section 3.3).

2.3. In situ data

In the previous section, the procedure to retrieve IOPs and derived quantities from RGB data has been described. It has been shown that, despite the highly analytical approach, calibration is needed in some steps. This section describes the data that were gathered for that purpose. An overview is given in Table 2. The reader will notice that the word "simulated" is profusely used throughout this document, but this term refers to the generation of satellite-comparable R_{rs} from in situ R_{rs} by convolution to the RSR. All R_{rs} used in this paper were measured in situ.

Simulated HRS data were generated by convolving hyperspectral R_{rs} with the RSR of each sensor. Indeed, this procedure requires R_{rs} to be hyperspectral because the convolution involves continuous spectral integrals. Additionally, matched $a_{nw}(555)$ data are needed in order to calibrate the χ predictor in Eq. (2).

Insufficient data sharing in the community has led to a lack of open field data for algorithm development and testing. The lack is more pronounced for datasets containing matched radiometry and IOPs. However, NASA's open access approach has led to the recent release of the PACE datasets (Casey et al., 2020). These include hyperspectral R_{rs} and other optical variables in a similar file format. R_{rs} is however not resampled to the same bands, the composition of the absorption budget varies among datasets and moreover, R_{rs} and IOPs are not matched in a single file. This makes the pre-processing time consuming, and has to be repeated separately for every dataset. Therefore, we prioritized the data mining to datasets with a decent amount of $R_{rs} - a_{nw}(555)$ matchups. $a_{nw}(555)$ was further restricted to samples where all absorption components were measured from discrete water samples. After screening and matching up, the retained datasets were Schaeffer ($N = 398$), Mouw ($N = 71$) and Craig ($N = 43$). The BIOSOPE dataset ($N = 15$) was used as well, though only for validation.

All R_{rs} in these datasets are comprehensively plotted in Fig. 1, alongside with a station map. Table 2 displays some basic metrics. It is derived that PACE data are mostly limited to green-brown waters. More data that cover the blue-green range are needed, but unfortunately, no more open access coincident hyperspectral R_{rs} and $a_{nw}(555)$ could be found. For this reason the Valente19 dataset (Valente et al., 2019) was considered. This dataset, which is the result of a large collective effort, contains in situ R_{rs} resampled to some typical multispectral sensor

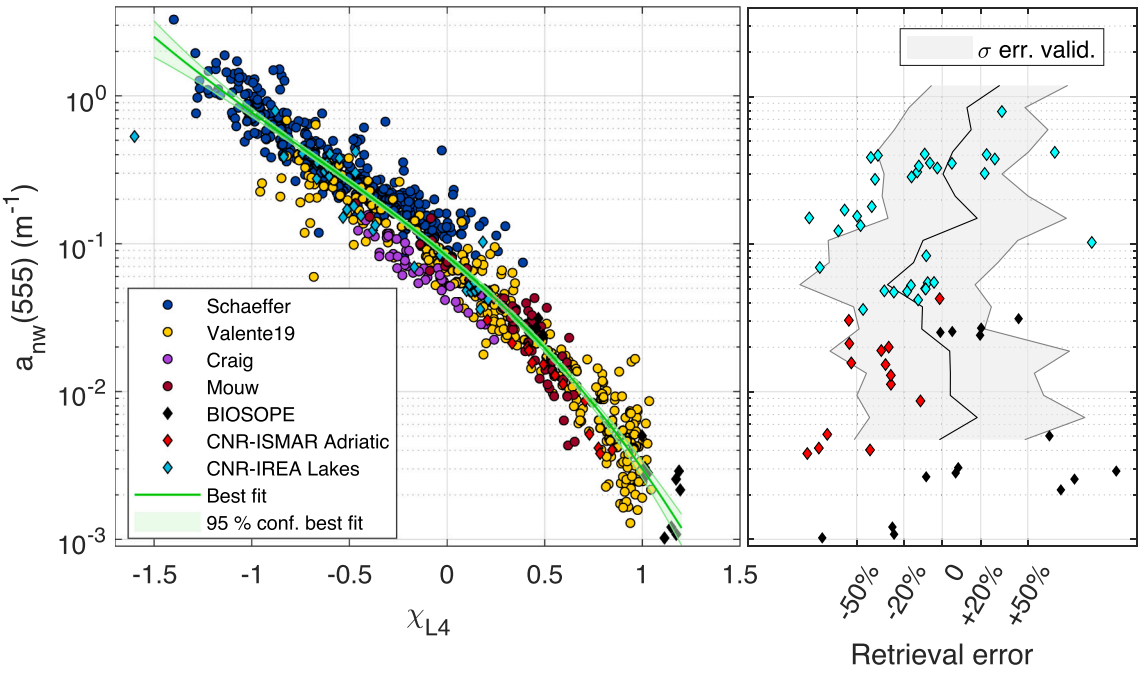

Fig. 2. Left panel: χ parameter derived from simulated Landsat 4 RGB data (Eq. (1)) vs. in situ non-water absorption at 555 nm, $a_{nw}(555)$. Dots were used to calibrate the average polynomial curve (green), shown with its 95% confidence interval. The diamond-shaped markers indicate data that were not used for calibration. Right panel: Retrieval errors and one standard deviation (σ) with respect to the validation dataset (grey lines and shaded area) with the independent dataset (diamonds), plotted on top. Low data density and likely high measurement errors for $a_{nw}(555) < 5{\cdot}10^{-3}$ m^{-1} prevents from extending the shaded area beyond this limit with confidence. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

bands. Valente19 is conveniently organized by combining all data in a spreadsheet, which allows easy selection of data verifying given criteria. We chose IOPs and AOPs from their "±6 nm"version, which aggregated spectral data within this interval. The requirements of having R_{rs} data at the six SeaWiFS bands (412/443/490/510/555/670 nm) and absorption from discrete samples that could be added up to calculate $a_{nw}(555)$ led to a dataset of $N = 317$ matchups.

Since Valente19 provides multispectral R_{rs}, their data cannot be convolved to the RSR in order to produce simulated RGB data. Instead, RGB data was generated from the Valente19 R_{rs} using neural networks (NNs) that ingest R_{rs} at the SeaWiFS bands. These NNs were trained with an independent and large hyperspectral R_{rs} dataset ($N = 869$), encompassing the highest dynamic range (Fig. 1), kindly provided by Zhongping Lee. The NNs were found to produce simulated HRS RGB data that best agreed to the RGB data directly generated from convolution of hyperspectral R_{rs}, compared to other methods (see Appendix B). The procedure is rather complex and since it is only used to generate the calibration dataset, it can be skipped by readers who are only interested in the QAA-RGB explanation.

A few independent matchup datasets were used to independently evaluate the $a_{nw}(555)$ retrieval. Firstly, the PACE BIOSOPE cruise data across ultra-oligotrophic Pacific waters, present in fragments within the "Bricaud" and "Lewis" datasets, was unified. Despite its small size ($N = 15$), it was regarded interesting to evaluate the detection ability at the low concentration end.

An internal "CNR-ISMAR" dataset (N = 15), collected in the western Adriatic Sea during a single campaign was used to provide further validation in coastal waters. Multispectral radiometry was obtained using Satlantic OCR-507 radiometers (Pitarch et al., 2016). It was converted to the HRS RGB bands by using similar NNs to those produced for the Valente19 dataset, but adapted to the Satlantic bands. $a_{nw}(555)$ data, on the other hand, are used for the first time in this article.

Given the great potential of this algorithm for inland water monitoring, an evaluation using lake data would provide further confidence on the universal applicability. Our CNR-IREA colleagues Claudia Giardino and Mariano Bresciani kindly provided $N = 35$ matchups for northern Italy lakes Mantua (Bresciani et al., 2013; Manzo et al., 2015) and Iseo (Bresciani et al., 2018). Radiometry in this dataset was measured with well-established hyperspectral radiometers (ASD and TriOs).

2.4. Validation statistics

Agreement of retrievals by different sensors (x and y) is assessed with the following statistical metrics:

Median absolute signed difference:

$$\delta_a = p_{50}(y - x) \tag{7}$$

Median percent signed difference:

$$\delta_r = 200{\cdot}p_{50}\left(\frac{y - x}{y + x}\right) \tag{8}$$

Median absolute unsigned difference:

$$\Delta_a = p_{50}(|y - x|) \tag{9}$$

Median percent unsigned difference:

$$\Delta_r = 200{\cdot}p_{50}\left(\left|\frac{y - x}{y + x}\right|\right) \tag{10}$$

here, p_{50} is the 50-th percentile or median of a given sample.

3. Results

Because the same procedure is performed for eighteen HRS (Table 1), the detailed account for all of them would be unnecessarily repetitive. The procedure is fully described here for the Thematic Mapper on board Landsat 4 only, and the coefficients for all HRS are reported in the Appendix A.

3.1. The $a_{nw}(555)$ vs. χ relationship

From R_{rs} in the calibration datasets of Table 2, simulated Landsat 4 R_{rs} were generated and the χ predictor was computed (Eq. (1)), which was associated to $\log_{10}[a_{nw}(555)]$ (Fig. 2, left), $N = 892$. A tight relationship was found and a functional predictor in the form of a third degree polynomial was fitted (Table A1) with $r^2 = 0.94$. The errors around this fitting curve followed a near-Gaussian distribution with a standard deviation (σ) in logarithmic scale equivalent to 50% in linear absorption units.

The polynomial fit was evaluated using a Monte Carlo procedure. For each of 1000 iterations, random 50%–50% calibration-validation subsets were used to compute and evaluate the polynomial fit, respectively. The shaded band around the central curve in the left panel of Fig. 2 represents the 95% confidence interval of the regression, calculated as twice the standard deviation for the 1000 fits. The error envelope in the right panel of Fig. 2 represents one standard deviation of the relative errors between each iteration's curve and the respective validation subset. This latter envelope indicates that the bulk 50% spread found for the regression is fairly constant across the dynamic range.

The BIOSOPE and CNR matchups were not used in the calibration, and are plotted on each panel for independent comparison, showing a general agreement with the results obtained using the other datasets.

Apart from the differences caused by measurement errors and different methodologies, the dot spread in Fig. 2 is caused by optical diversity in the form of a variety of specific phytoplankton absorption spectra, spectral slopes, different composition of the absorption budget, etc. Still, the data cloud and the fit show that the link between the χ predictor and $a_{nw}(555)$ is satisfactory from the clearest waters until at least $a_{nw}(555) = 2\ m^{-1}$. In the lower absorption range, it is well known that the B/G ratio that dominates χ is very sensitive to absorption variations (Lee et al., 1998), so the dot spread is mostly caused by the difficulties in measuring very small $a_{nw}(555)$ with accuracy. In the higher absorption range, no signs of loss of sensitivity are visible, but it is a well-known fact that R_{rs} in the visible range loses sensitivity to turbidity for extremely turbid waters (Luo et al., 2018). If the data range reached much higher absorption values, it is very likely that at some point the data cloud would spread and "bend upwards", therefore losing predictability from χ. This could probably be avoided with the inclusion of a NIR band in Eq. (1), or using blended approaches (Shi and Wang, 2017, 2019). This would require datasets from very highly absorbing waters with hyperspectral R_{rs} that covered the HRS NIR band. Unfortunately, such data are lacking even in the datasets used here. For example, R_{rs} in the Schaeffer dataset, that covers the turbid range, are only provided up to 735 nm. Therefore, for the time being we choose to stick to the empirical evidence we could collect and simply advise to raise a flag if $a_{nw}(555) > 2\ m^{-1}$. This value is not directly convertible to other parameters because many different water constituents can be combined to sum up to it. In the Coastcolour simulated dataset (Nechad et al., 2015a; Nechad et al., 2015b), $a_{nw}(555) = 2\ m^{-1}$ results in a total suspended matter concentration (TSM) of about 200 g m^{-3} for sediment-dominated waters. If coloured dissolved organic matter (CDOM) is dominant, then we obtain $a_{CDOM}(443)$ of about 3 m^{-1}. If absorption is dominated by phytoplankton, then the chlorophyll concentration (CHL) is in the range of 200 mg m^{-3}. Nevertheless, these values are only roughly indicative as there is significant specific absorption variability worldwide and only the limitation $a_{nw}(555) < 2\ m^{-1}$ applies.

Derived coefficients for all eighteen HRS are shown in Table A1, highlighting the impact of the differences between the RSR. Note the high similarities between Landsat 4 and 5 and between Sentinel 2A and 2B, due to the fact that they carry instruments with the same design, though slight differences in the measured RSR led to differences in these coefficients.

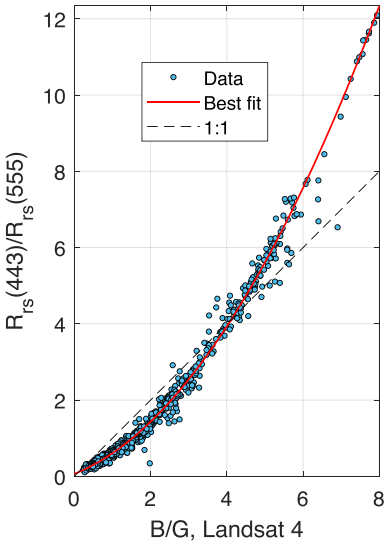

Fig. 3. Blue to green ratio of simulated Landsat 4 data vs. 443-to-555 ratio, using Lee's dataset. A second degree polynomial was fitted (Table A2) to correct for this deviation (Eq. (3)). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

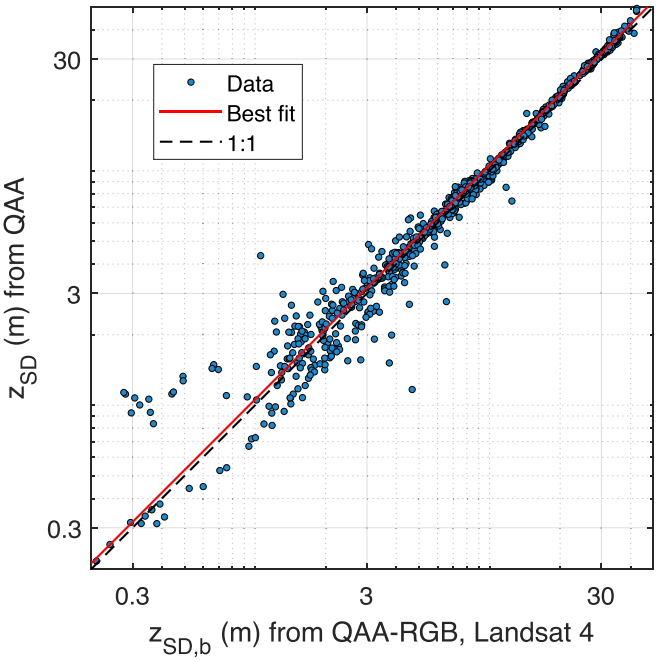

Fig. 4. Biased Secchi disk depth ($z_{SD,b}$) from simulated Landsat 4 data vs. z_{SD} from simulated SeaWiFS data, using Lee's dataset. A linear fit, forcing zero intercept, was fitted (Table A3) to correct for this deviation (Eq. (6)).

3.2. Harmonization of the blue to green ratio

To establish the relationship between the B/G ratio of HRS used in the QAA-RGB to the 443-to-555 ratio used in the QAA, Lee's R_{rs} dataset was convolved to the SeaWiFS bands and to the RSR of each HRS. The B/G ratios were computed using the 443 nm and 555 nm SeaWiFS bands, and the HRS B and G bands, respectively.

Fig. 3 shows that the HRS B/G ratio is indeed different from the 443-to-555 ratio across the whole range, but that a second degree polynomial fits well across the encountered range and allows for a numerical prediction of the deviation. For most HRS, a second degree curve sufficed.

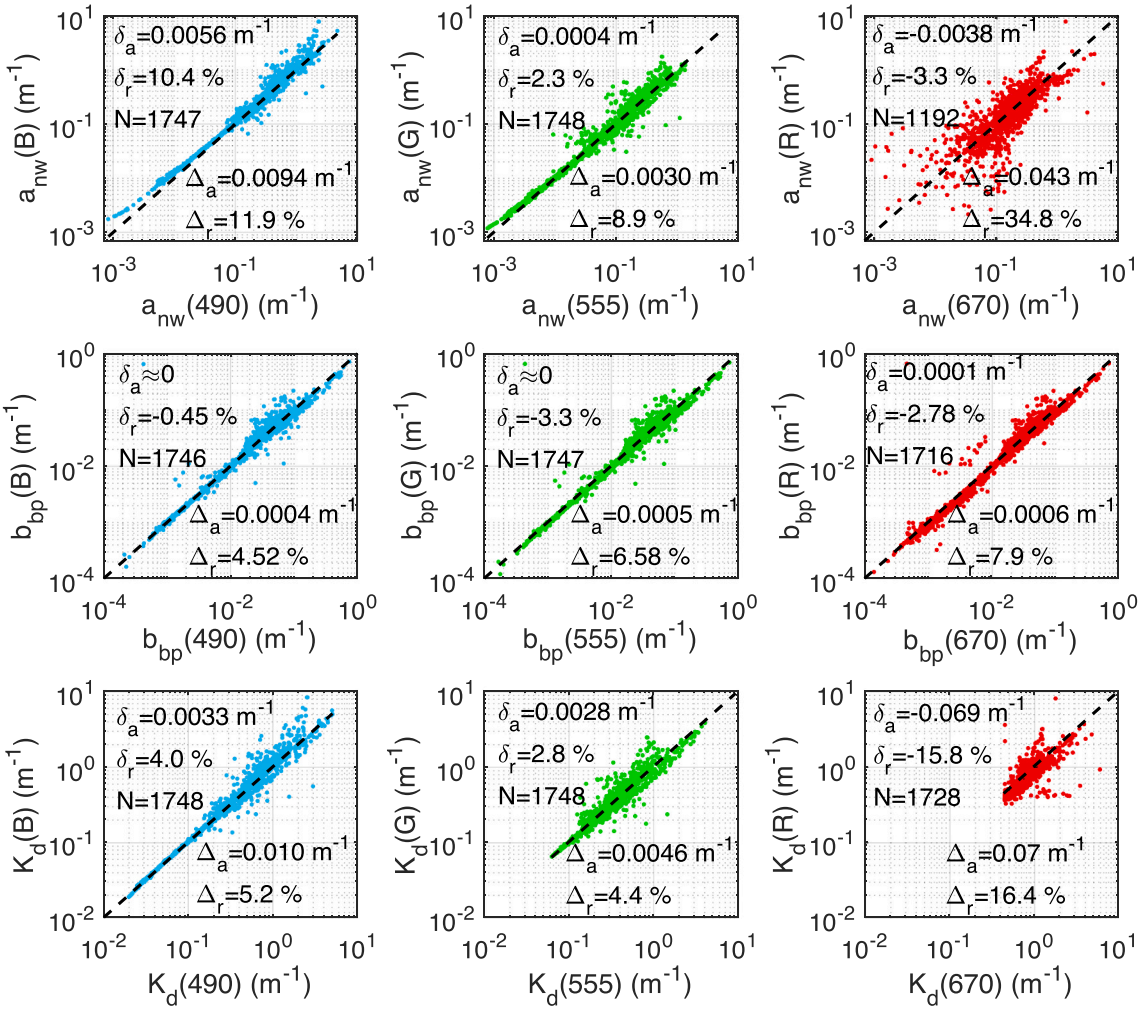

Fig. 5. Comparison of non-water absorption (a_{nw}), particle backscattering (b_{bp}) and diffuse attenuation coefficient (K_d) between the QAA (horizontal axes) and the QAA-RGB for Landsat 8 (vertical axes). Source data are the same hyperspectral R_{rs} for both, resampled to the respective sensor bands. Comparison is made between closest bands. Eqs. (7–10) define the statistical metrics that are embedded in the plots.

Coefficients for all of them are given in Table A2.

3.3. Secchi disk depth fine tuning

To establish a reference z_{SD}, Lee's R_{rs} dataset was resampled to the SeaWiFS bands, the QAA was applied and z_{SD} was calculated (Eqs. (4–5)). For full consistency, the $a_{nw}(555)$ vs. χ relationship in the QAA was recalibrated using the data in Section 2.3. This led to a replacement of the original QAA coefficients of the second-degree polynomial (−0.469,-1.366,-1.146) with new ones (−0.189,-1.252,-1.191). Lee's R_{rs} dataset was also resampled to the RSR of the different HRS, the QAA-RGB was applied and the "biased" quantity $z_{SD,b}$ was retrieved (Eq. (5)).

The agreement between QAA's z_{SD} and QAA-RGB's $z_{SD,b}$ is very high, within 5–10% for most HRS (see Fig. 4 for Landsat 4), which validates the calibration and harmonization procedures described in the previous sections. This agreement is remarkable given the very coarse sampling of the spectrum provided by the three RGB bands (Table 1). Moreover, the relationship between both quantities can be well modelled with a linear or a quadratic polynomial, except Pléiades and PlanetScope, which require higher degree functions. This modeling was applied as a bias correction to $z_{SD,b}$ (Eq. (6)). For Landsat 4, a linear model with zero intercept and a slope of ~1.059 was fitted to the data (Fig. 4). Coefficients for all HRS are given in Table A3.

This bias correction is no longer valid for high z_{SD} values, as the spectral K_d minimum shifts towards wavelengths that are shorter than Landsat 4's blue band at ~490 nm. In these cases, the z_{SD} estimation shows a pattern of increasing underestimation and a decreased sensitivity. In the absence of atmospheric correction errors, this occurs at about $z_{SD} = 40$ m, which can be set as an upper bound for the z_{SD} retrieval, except for the PlanetScope 0e, whose suboptimal band characteristics reduce this limit to about 30 m. Nevertheless, these upper bounds are irrelevant in practice, as the real limitation for HRS in very clear waters is the signal to noise ratio.

4. Validation

In an effort to increase the confidence in the QAA-RGB performance, this section provides two validation tests: (1) a retrieval consistency check respect to the QAA, and (2) a comparison to independent in situ IOP data.

4.1. Retrieval consistency respect to the QAA

This section shows evidence that the QAA and the QAA-RGB retrievals are consistent. A large R_{rs} dataset was pooled together by merging the Lee, Schaeffer, Mouw, Craig, Valente19, BIOSOPE and CNR-IREA datasets. It was resampled to both SeaWiFS and two chosen HRS, namely Landsat 8 and Sentinel 2A, after which the QAA and the

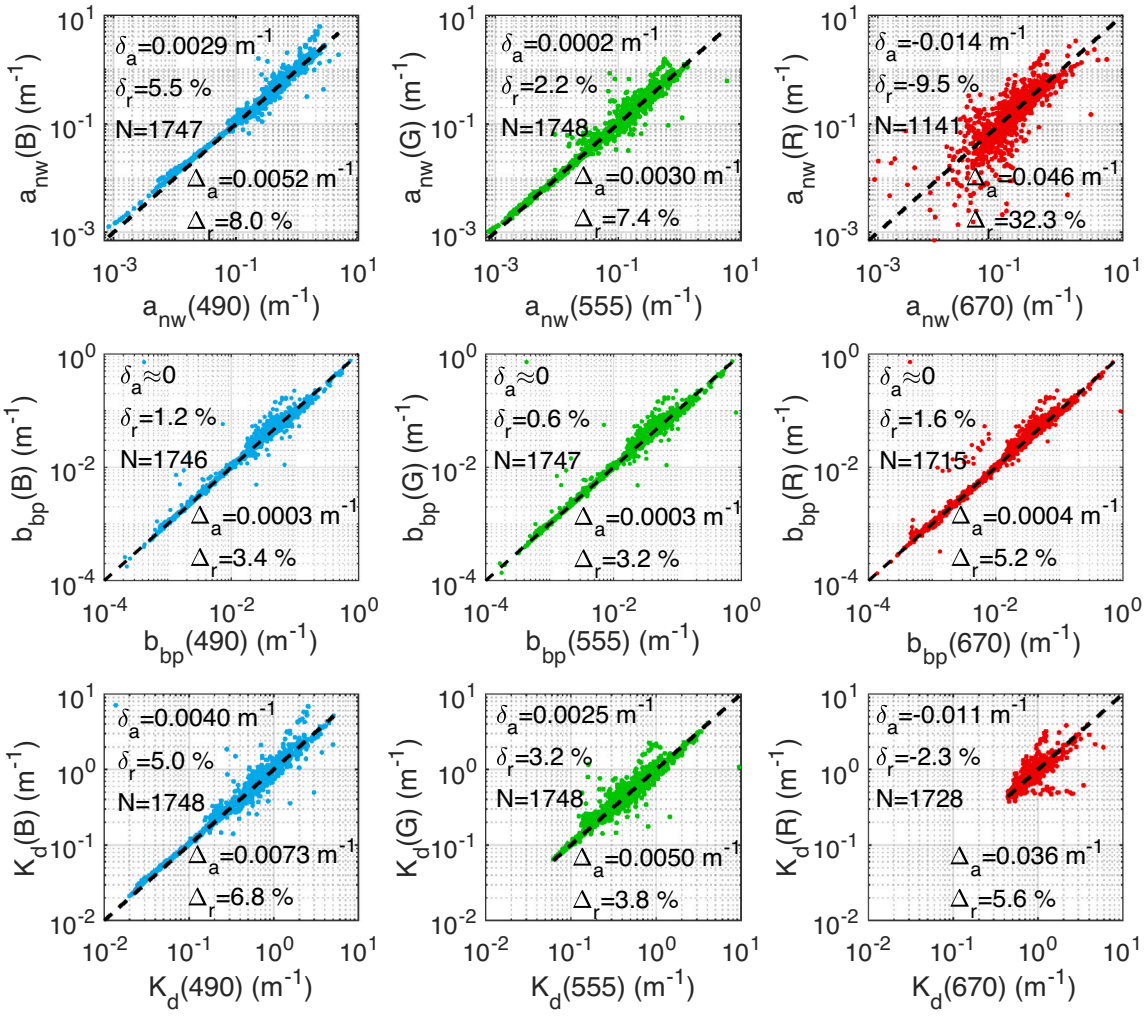

Fig. 6. Same as Fig. 5, but for Sentinel 2A.

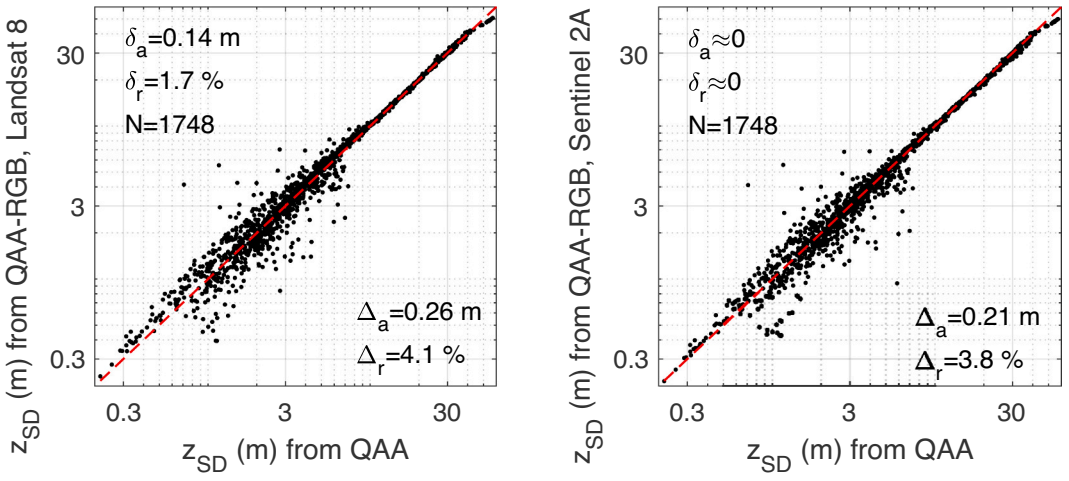

Fig. 7. Left panel: comparison of the Secchi disk depth (z_{SD}) between the QAA (horizontal axes) and the QAA-RGB for Landsat 8 after bias correction (vertical axes). Source data are the same hyperspectral R_{rs} for both, resampled to the respective sensor bands. Eqs. (7–10) define the statistical metrics that are embedded in the plots. Right panel: same as left panel but for Sentinel 2A.

QAA-RGB were respectively applied.

Retrievals at the HRS bands (B,G,R) are compared to the respective retrievals at the SeaWiFS closest bands (490/555/670 nm). Figs. 5 and 6 show the comparison of a_{nw}, b_{bp} and K_d for Landsat 8 and Sentinel 2A, respectively. These show good agreement, with lowest differences for the green and largest for the red bands. Results at the blue and red bands

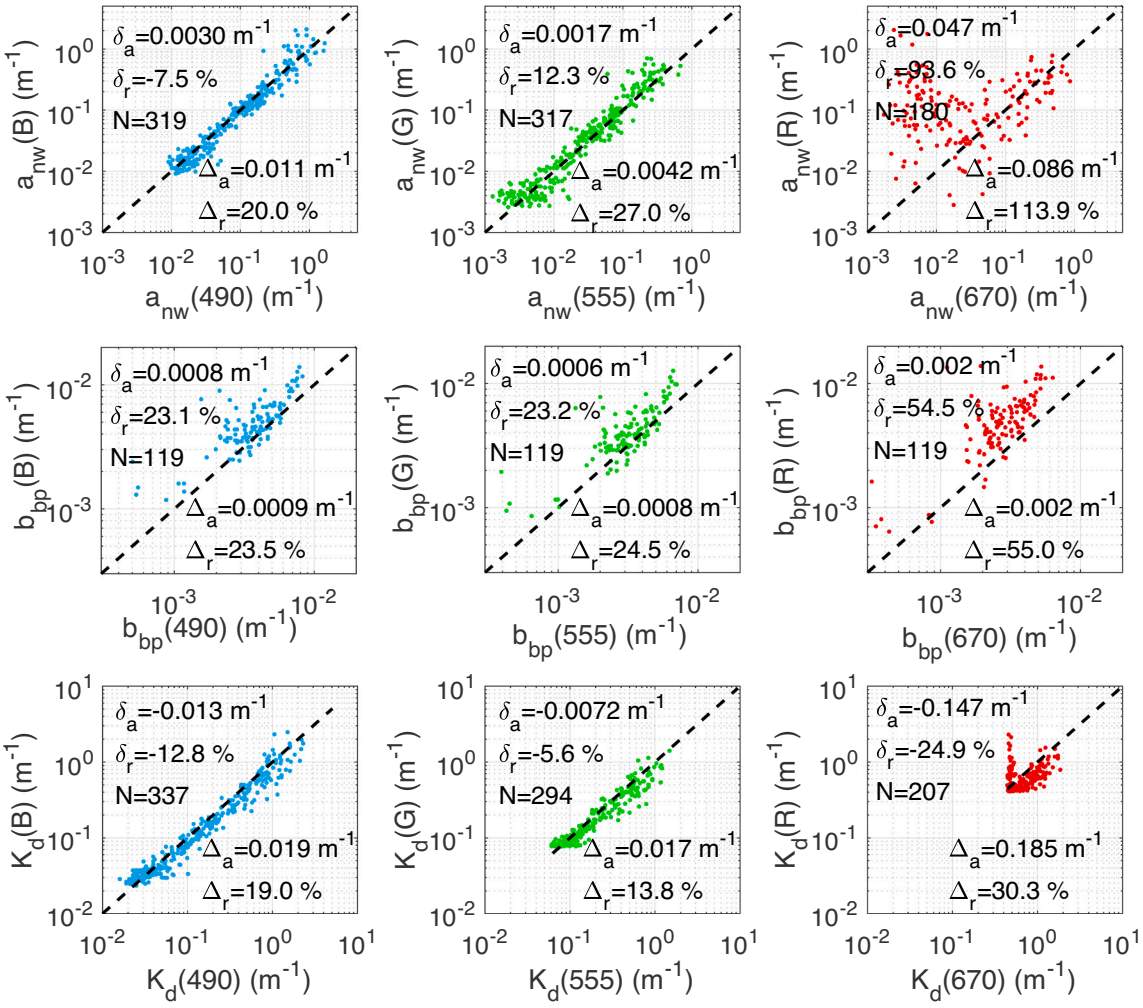

Fig. 8. Comparison of non-water absorption (a_{nw}), particle backscattering (b_{bp}) and diffuse attenuation coefficient (K_d) between in situ data from the Valente19 dataset and QAA-RGB retrievals from matched R_{rs} (vertical axes). Comparison is made for closest bands. Eqs. (7–10) define the statistical metrics that are embedded in the plots.

are influenced by spectral mismatch between the HRS and SeaWiFS, which becomes more relevant at low turbidities. Agreement is slightly better for Sentinel 2A than for Landsat 8 due to the former's narrower spectral responses.

For z_{SD} (Fig. 7 for Landsat 8 and Sentinel 2A), median unsigned differences between the QAA and the QAA-RGB are very low (<0.3 m, ~4%), while median signed differences are negligible, thanks to the residual bias correction. The agreement degrades for the most transparent waters $z_{SD} > 40$ m as anticipated.

4.2. Assessment with independent in situ data

In the preceding section, the QAA-RGB has been demonstrated to be in agreement with the QAA. In this section, a comparison of QAA-RGB outputs from simulated Landsat 4 R_{rs} to independent in situ IOP data from the Valente19 dataset is made. This time, a comparison to real data carries along increased mismatches, impacted by the low levels of standardization and the lack of uncertainty characterization of in situ IOP measurements.

As in the previous section, retrievals at the HRS bands (B,G,R) are compared to the respective retrievals at the SeaWiFS closest bands (490/555/670 nm). Results in Fig. 8 shows good overall agreement. For a_{nw}, the lowest differences are found at the blue and green bands. In the red band, the retrieval is consistent for $a_{nw} > \sim 0.02$ m^{-1}. Here, retrieval is hampered by the steep increase in the absorption coefficient, especially for HRS whose red band extends beyond 700 nm. It should be noted that $a_{nw}(555)$ was used for calibration, but neither $a_{nw}(490)$ nor $a_{nw}(670)$ were. Similar relative differences with general overestimation are observed for b_{bp}. K_d retrievals compare very well to in situ data except for very low values at the red band. In this latter case, in situ K_d measurements are problematic because downwelling irradiance is subject to the opposing effects of a very high attenuation and Raman scattering, which leads to strong depth-dependence of K_d and hence variability depending on the depths that are used for its estimation.

5. Implementation in ACOLITE

The Dark Spectrum Fitting algorithm (DSF, Vanhellemont and Ruddick, 2018) as implemented in ACOLITE (Vanhellemont, 2019a) is a commonly used atmospheric correction algorithm for aquatic applications of high resolution sensors, such as turbidity mapping and bathymetry retrieval. The QAA-RGB has been implemented in Python and integrated in ACOLITE. Hence, ACOLITE now allows for the retrieval of bulk absorption and backscattering properties, K_d and Secchi disk depth from the HRS it can process. With the open availability of ACOLITE and the QAA-RGB, a community-based evaluation is expected to follow in the coming years.

This section presents a demonstration of QAA-RGB retrievals over

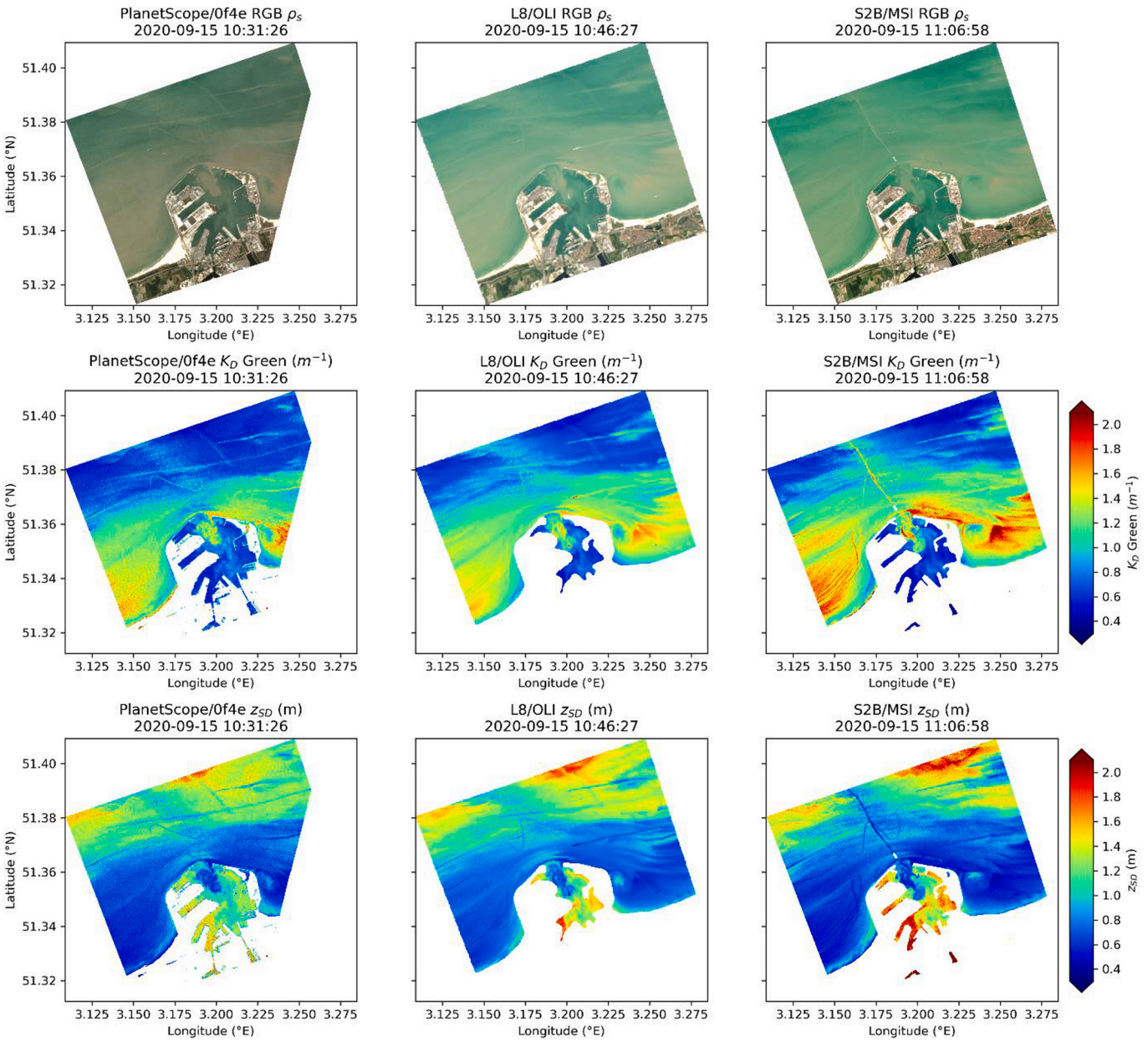

Fig. 9. RGB composites (top), and QAA-RGB retrievals of K_d for the green band (middle) and z_{SD} (bottom) for the turbid waters around Zeebrugge, Belgium on 2020-09-15 from PlanetScope (left, 10:31 UTC) Landsat 8 (middle, 10:46 UTC) and Sentinel 2B (right, 11:06 UTC). The images were taken during flood tide, and some horizontal advection and resuspension of sediments has occurred in the time between images. Stable sediment plumes are caused by monitoring poles installed in the seabed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

selected areas. Over the Belgian port of Zeebrugge, we used three near-coincident images from Landsat 8, Sentinel 2B and PlanetScope, taken on 2020-09-15 (Fig. 9). RGB composites and the retrieval of K_d and z_{SD} are provided. These results show a good correspondence between the three sensors.

An intercomparison among these satellite products was performed. Output data were reprojected to the 30 m Landsat 8 grid and compared to Landsat 8 outputs for 1000 random sampling points. It results that K_d and z_{SD} are in good agreement for these near-coincident images: 6%-15% for the signed differences and 13%-20% for the unsigned differences. These differences include those caused by strong tidal forces, ship movement, horizontal advection and top-of-atmosphere calibration.

Application to older Landsat sensors is demonstrated in Fig. 10, showing absorption and backscattering from a Landsat 5 image on 1984-05-10 over lake Erie near Toledo, and from an image over the Mississippi river delta on 1999-11-27. Similar ranges of absorption are retrieved, while the backscatter in Lake Erie is significantly higher. Notable are the spatial features of the surface sediments, that impact the retrieved IOPs, and are well resolved using these high resolution sensors.

6. Conclusions

We have presented the QAA-RGB, a three band (Red/Green/Blue) adaptation of the QAA to high resolution optical satellite sensors. The QAA-RGB performs semianalytical retrieval of absorption and backscattering, and derived properties, such as K_d and z_{SD}, from historic and current high resolution optical satellite imagery. Small lakes and rivers can now be monitored with a consolidated algorithm specifically adapted to metre and decametre sensors, after the application of an appropriate atmospheric correction algorithm. Calibration of the QAA-

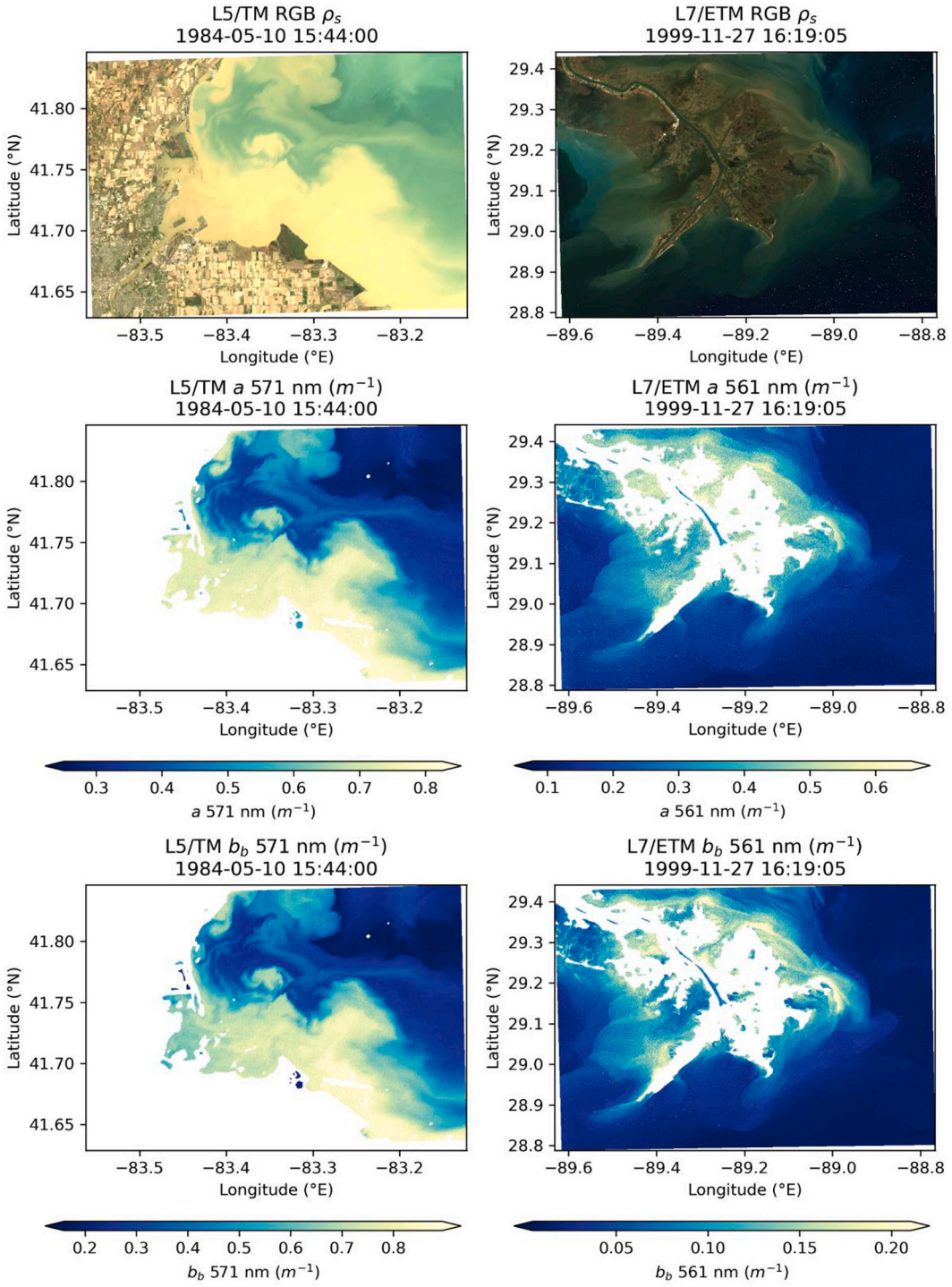

Fig. 10. Application of the QAA-RGB to imagery from Landsat 5 over Lake Erie near Toledo (top row, 1984-05-10) and Landsat 7 over the Mississippi river mouth (bottom row, 1999-11-27). From left to right: RGB composite, QAA-RGB absorption and QAA-RGB backscatter in the green band. Note the differences in the colour bars. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

RGB using in situ measurements showed a satisfactory performance across a wide dynamic range and high consistency with the classic QAA and independent in situ measurements. The calibration presented here may be updated in the future if additional open access datasets are made available to the community. This contribution aims at aligning developments in the ocean optics community and the needs of the Landsat and inland water communities. The QAA-RGB is not intended as a correction to or a replacement of any existing algorithm, but as a clear approach for IOP retrievals from spectrally limited sensors. A major conclusion is that, even for semianalytical algorithms, adaptation to a new sensor leads to new coefficients even if its bands and those of the reference sensor are close.

The algorithm is integrated in ACOLITE and is hence openly available to the community, and can applied to the free Landsat and Sentinel 2 archives to provide data from the mid-1980s to the present. The algorithm is simple, and hence can be easily implemented in any programming language or platform, including cloud computing environments.

Code availability

The QAA-RGB has been coded in Python and integrated into the ACOLITE generic processor, retrievable at https://github.com/acolite/. MATLAB scripts are available at doi:https://doi.org/10.5281/zenodo.4923880.

Funding

Jaime Pitarch was funded by the European Centre for Medium-Range Weather Forecasts (ECMWF), contract nr. C3S_511: Copernicus Climate Change Service, Quality Assessment of ECV Products.

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

Journal editor-in-chief Menghua Wang and three anonymous reviewers are thanked for the time they spent suggesting changes and finding errors in the former versions of this article. We thank Claudia Giardino, Mariano Bresciani, and Annalisa di Cicco for providing validation data to this study. Gianluca Volpe and Simone Colella helped in the data collection and processing, and provided feedback to an early version of the manuscript. ZhongPing Lee is thanked for the large reflectance dataset that was of great aid. Data contributors to the PACE and Valente19 datasets are thanked for their great service to the community. Satellite data was kindly provided free of charge by their respective agencies. Jaime Pitarch acknowledges support by the Coordinadora de Microorganismos Partisanos.

Appendix A. QAA-RGB coefficients

This Appendix provides all model coefficients for every HRS in tables.

Table A1
Coefficients of the "P" polynomial in Eq. (2) that links the χ parameter to $\log_{10}[a_{nw}(555)]$.

Sensor	p_3	p_2	p_1	p_0
Landsat 4	−0.10419	−0.23184	−1.10221	−1.08595
Landsat 5	−0.10249	−0.23418	−1.10967	−1.08523
Landsat 7	−0.08258	−0.26497	−1.15697	−1.14697
Landsat 8	−0.06989	−0.24566	−1.17869	−1.15467
Sentinel 2A	−0.08409	−0.35707	−1.33678	−1.09651
Sentinel 2B	−0.0699	−0.34549	−1.34071	−1.09689
PHR 1A	−0.2255	−0.43238	−1.30193	−1.13823
PHR 1B	−0.29946	−0.59577	−1.36584	−1.06297
PlanetScope 0c	−0.50593	−0.74629	−1.5122	−1.30412
PlanetScope 0d05	−0.52847	−0.77675	−1.51752	−1.30208
PlanetScope 0d06	−0.51825	−0.75991	−1.51414	−1.30413
PlanetScope 0e	−0.7746	−0.80376	−1.46749	−1.27228
PlanetScope 0f	−0.67964	−0.74304	−1.45322	−1.28734
PlanetScope 22	−0.06047	−0.26792	−1.2441	−1.08911
RapidEye	−0.07515	−0.27579	−1.19543	−1.18533
WorldView 2	−0.12796	−0.3797	−1.25527	−1.17702
WorldView 3	−0.1223	−0.3755	−1.25754	−1.13707
VENμS VSSC	−0.11458	−0.39764	−1.32554	−1.08491

Table A2
Coefficients of the "Q" polynomial of the blue-to-green ratio to ensure compatibility with the formulas involving the 443-to-555 ratio.

Sensor	q_4	q_3	q_2	q_1	q_0
Landsat 4	0	0	0.14173	0.400392	0.064038
Landsat 5	0	0	0.149681	0.390775	0.068354
Landsat 7	0	−0.00564	0.202401	0.47588	0.037058
Landsat 8	0	0	0.167207	0.548575	0.022365
Sentinel 2A	0	0.010022	0.226931	0.540187	−0.02085
Sentinel 2B	0	0.009593	0.238763	0.539832	−0.02551
PHR 1A	0	0.020095	0.611234	−0.05321	0.126072
PHR 1B	0	0.020095	0.611234	−0.05321	0.126072
PlanetScope 0c	0.622116	−2.05399	3.708664	−2.25303	0.658542
PlanetScope 0d05	0.623035	−2.05764	3.714245	−2.25686	0.659462
PlanetScope 0d06	0.623035	−2.05764	3.714245	−2.25686	0.659462
PlanetScope 0e	1.078401	−3.59734	6.227313	−4.06846	1.072136
PlanetScope 0f	0.873197	−2.83881	5.026056	−3.22297	0.868836
PlanetScope 22	0	0.00954	0.125348	0.565848	0.011577
RapidEye	0	−0.00564	0.216527	0.578923	−0.00475
WorldView 2	0	−0.00255	0.322795	0.54172	−0.03806
WorldView 3	0	0.004014	0.345076	0.508684	−0.04121
VENμS VSSC	0	0.035071	0.151045	0.644829	−0.06704

Table A3
Coefficients of the "S" polynomial for the residual bias removal of the Secchi disk depth retrievals between the QAA-RGB and the QAA.

Sensor	s_3	s_2	s_1	s_0
Landsat 4	0	0	1.058675	0
Landsat 5	0	0	1.049492	0
Landsat 7	0	0	1.061503	0
Landsat 8	0	0	1.047961	0
Sentinel 2A	0	0.002532	1.023179	0
Sentinel 2B	0	0.002628	1.025141	0
PHR 1A	0	0.008802	0.977379	0
PHR 1B	0	0.012123	0.920493	0
PlanetScope 0c	0.00169	−0.01957	0.900316	0
PlanetScope 0d05	0.001695	−0.02016	0.905782	0
PlanetScope 0d06	0.00169	−0.01974	0.901322	0
PlanetScope 0e	0.00303	−0.03536	1.026617	0
PlanetScope 0f	0.00208	−0.02245	0.925919	0
PlanetScope 22	0	0	1.029139	0
RapidEye	0	0	1.076275	0
WorldView 2	0	0.004488	0.961	0
WorldView 3	0	0.004075	0.975182	0
VENμS VSSC	0	0.003201	0.992756	0

Table A4
Absorption coefficient of pure water. Directly sampled from hyperspectral tables by selecting data at the center wavelength of each sensor. Source until 550 nm is from Mason et al. (2016). From 550 nm: Pope and Fry (1997).

Sensor	B	G	R
Landsat 4	0.01336	0.07104	0.41
Landsat 5	0.01336	0.0695	0.41
Landsat 7	0.011955	0.06236	0.4138
Landsat 8	0.01274	0.06236	0.371
Sentinel 2A	0.01545	0.0619	0.429
Sentinel 2B	0.01545	0.06144	0.429
PHR 1A	0.021575	0.06236	0.34
PHR 1B	0.02546	0.06098	0.4214
PlanetScope 0c	0.015965	0.04882	0.27708
PlanetScope 0d05	0.015965	0.04882	0.27708
PlanetScope 0d06	0.015965	0.04882	0.27708
PlanetScope 0e	0.038495	0.057614	0.29736
PlanetScope 0f	0.02668	0.05224	0.2834
PlanetScope 22	0.01545	0.06526	0.431
RapidEye	0.011575	0.06006	0.3944
WorldView 2	0.011955	0.05425	0.4022
WorldView 3	0.01254	0.053245	0.41
VENμS VSSC	0.01545	0.0596	0.431

Table A5
Backscattering coefficient of pure water. Directly sampled from hyperspectral tables by selecting data at the center wavelength of each sensor. Source is from Zhang et al. (2009).

Sensor	B	G	R
Landsat 4	0.001482	0.000753	0.000412
Landsat 5	0.001482	0.000759	0.000412
Landsat 7	0.001576	0.000811	0.000409
Landsat 8	0.001522	0.000811	0.000425
Sentinel 2A	0.001407	0.000817	0.000399
Sentinel 2B	0.001407	0.000823	0.000399
PHR 1A	0.001304	0.000811	0.000439
PHR 1B	0.001261	0.000829	0.000404
PlanetScope 0c	0.001396	0.000937	0.000531
PlanetScope 0d05	0.001396	0.000937	0.000531
PlanetScope 0d06	0.001396	0.000937	0.000531
PlanetScope 0e	0.001142	0.000868	0.00049
PlanetScope 0f	0.00125	0.000908	0.000517
PlanetScope 22	0.001407	0.000781	0.000397
RapidEye	0.001604	0.000842	0.000417
WorldView 2	0.001576	0.000895	0.000415
WorldView 3	0.001535	0.000901	0.000412
VENμS VSSC	0.001407	0.000848	0.000397

Table A6
"α" parameter for the Raman scattering correction. Linearly interpolated from Lee et al. (Lee et al., 2013) to the center wavelength of each sensor.

Sensor	B	G	R
Landsat 4	0.010065	0.016775	0.0178
Landsat 5	0.010107	0.016757	0.017803
Landsat 7	0.009116	0.016442	0.01783
Landsat 8	0.009687	0.016699	0.017853
Sentinel 2A	0.010879	0.016856	0.017908
Sentinel 2B	0.010839	0.016818	0.017914
PHR 1A	0.010668	0.016106	0.01754
PHR 1B	0.011112	0.015975	0.017568
PlanetScope 0c	0.0105	0.01471	0.016955
PlanetScope 0d05	0.010501	0.01471	0.016955
PlanetScope 0d06	0.010501	0.01471	0.016955
PlanetScope 0e	0.011534	0.014905	0.017409
PlanetScope 0f	0.011288	0.014899	0.017392
PlanetScope 22	0.010785	0.017008	0.017912
RapidEye	0.008888	0.016312	0.01783
WorldView 2	0.009215	0.015899	0.017825
WorldView 3	0.009584	0.015861	0.017829
VENμS VSSC	0.010998	0.016612	0.017928

Table A7
"β_1" parameter for the Raman scattering correction. Linearly interpolated from Lee et al. (Lee et al., 2013) to the center wavelength of each sensor.

Sensor	B	G	R
Landsat 4	0.011143	0.01	0.01
Landsat 5	0.01111	0.01	0.01
Landsat 7	0.011645	0.01	0.01
Landsat 8	0.011243	0.01	0.01
Sentinel 2A	0.010752	0.01	0.01
Sentinel 2B	0.010772	0.01	0.01
PHR 1A	0.011072	0.01	0.010063
PHR 1B	0.010832	0.01	0.010071
PlanetScope 0c	0.011065	0.010194	0.010067
PlanetScope 0d05	0.011065	0.010194	0.010067
PlanetScope 0d06	0.011065	0.010194	0.010067
PlanetScope 0e	0.010879	0.010162	0.010026
PlanetScope 0f	0.011003	0.010182	0.010032
PlanetScope 22	0.010664	0.01	0.01
RapidEye	0.011742	0.010002	0.01
WorldView 2	0.011509	0.01	0.01
WorldView 3	0.011294	0.01	0.01
VENμS VSSC	0.010416	0.01	0.01

Table A8
"β_2" parameter for the Raman scattering correction. Linearly interpolated from Lee et al. (Lee et al., 2013) to the center wavelength of each sensor.

Sensor	B	G	R
Landsat 4	−0.04764	−0.07816	−0.0808
Landsat 5	−0.0478	−0.07809	−0.0808
Landsat 7	−0.04369	−0.07674	−0.08083
Landsat 8	−0.04596	−0.07812	−0.08085
Sentinel 2A	−0.05106	−0.07903	−0.08091
Sentinel 2B	−0.05089	−0.07886	−0.08091
PHR 1A	−0.05043	−0.0752	−0.07995
PHR 1B	−0.05227	−0.07456	−0.07983
PlanetScope 0c	−0.04963	−0.06864	−0.07769
PlanetScope 0d05	−0.04964	−0.06864	−0.07769
PlanetScope 0d06	−0.04964	−0.06864	−0.07769
PlanetScope 0e	−0.05421	−0.06944	−0.07971
PlanetScope 0f	−0.05319	−0.06944	−0.07961
PlanetScope 22	−0.05052	−0.07963	−0.08091
RapidEye	−0.04272	−0.07625	−0.08083
WorldView 2	−0.04403	−0.07431	−0.08083
WorldView 3	−0.04554	−0.07413	−0.08083
VENμS VSSC	−0.05128	−0.07789	−0.08093

Appendix B. Neural network description and performance

Neural networks (NNs) were trained to generate high resolution sensors (HRS) R_{rs} data from multispectral R_{rs} (SeaWiFS-like). For this method to be practical, NNs should return HRS-equivalent R_{rs} that are more exact than R_{rs} obtained by simpler means, such as linearly interpolating the multispectral R_{rs} and then convoluting these spectra with the RSR, or by simply selecting the bands 490/555/670 nm of a multispectral sensor (bands 3, 5 and 6 for SeaWiFS), which are the closest to the HRS RGB bands, and assuming them equivalent.

The MATLAB® Neural Network fitting tool (MATLAB, 2021) was used to create two-layer feed-forward NNs with sigmoid hidden neurons and linear output neurons. NNs were trained with Lee's hyperspectral R_{rs} dataset resampled to SeaWiFS and HRS bands as respectively the input and target datasets (Fig. B1). Ten hidden neurons were chosen and the NN parameters were fitted with the Levenberg-Marquardt method.

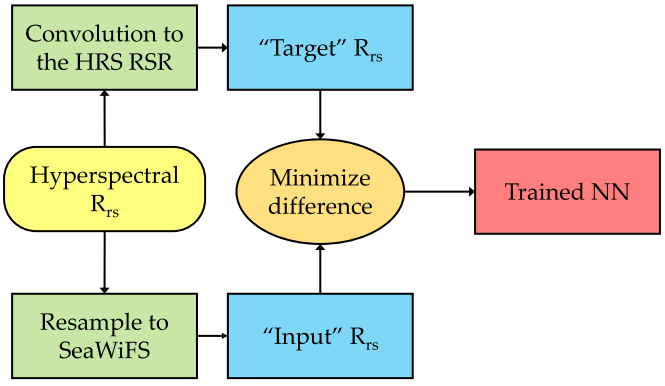

Fig. B1. Flow chart describing the neural network calibration.

To assess the performance of the NN and other resampling methods, the hyperspectral datasets other than Lee's, namely Schaeffer, Craig, Mouw, BIOSOPE and CNR-IREA, were resampled to SeaWiFS and to the HRS bands. The HRS resampled data was then used as reference for comparing three different methods: (1) NN outputs from SeaWiFS inputs, (2) linearly interpolated SeaWiFS data, and then resampled to the HRS RSR, and (3) the closest SeaWiFS bands (3, 5 and 6). We present a graphical explanation with a flow chart in Fig. B2.

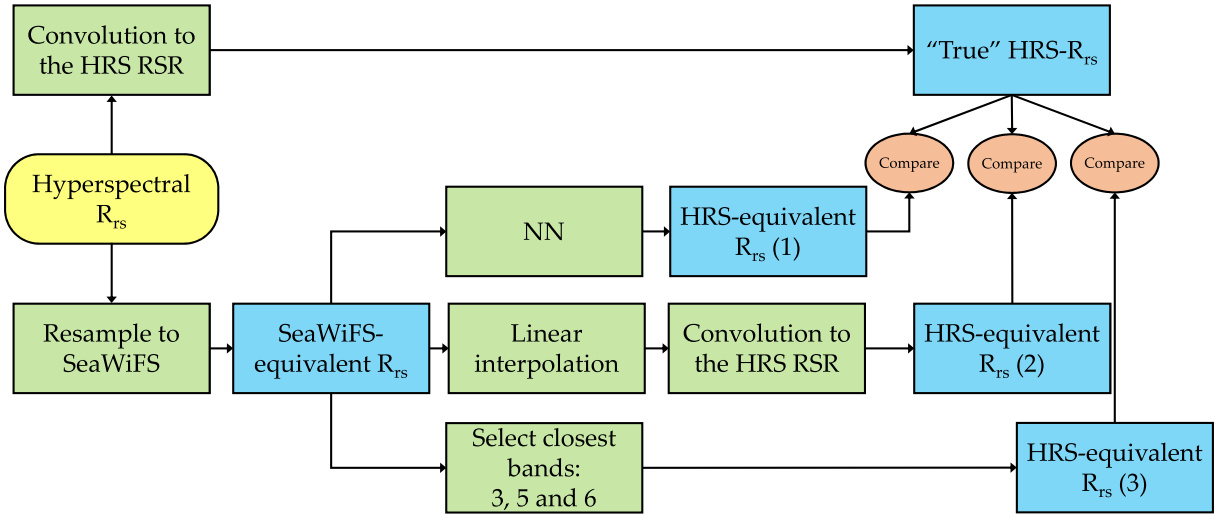

Fig. B2. Flow chart of three different methods to generate high resolution-sensor data from SeaWiFS-equivalent data, and the comparison to the "true" data.

Performance of the three methods is presented in Fig. B3 for Landsat 4. In terms of R_{rs}, NN provide significantly better results in the green and red bands. For the blue band, the linear interpolation method provides slightly lower noise, likely due to the dense coverage by SeaWiFS in the blue spectral region. In terms of $a_{nw}(555)$ after QAA-RGB application, the NN introduce an RMS error of 10.4%, or 0.039 m^{-1} in absolute units, significantly lower than the other two methods.

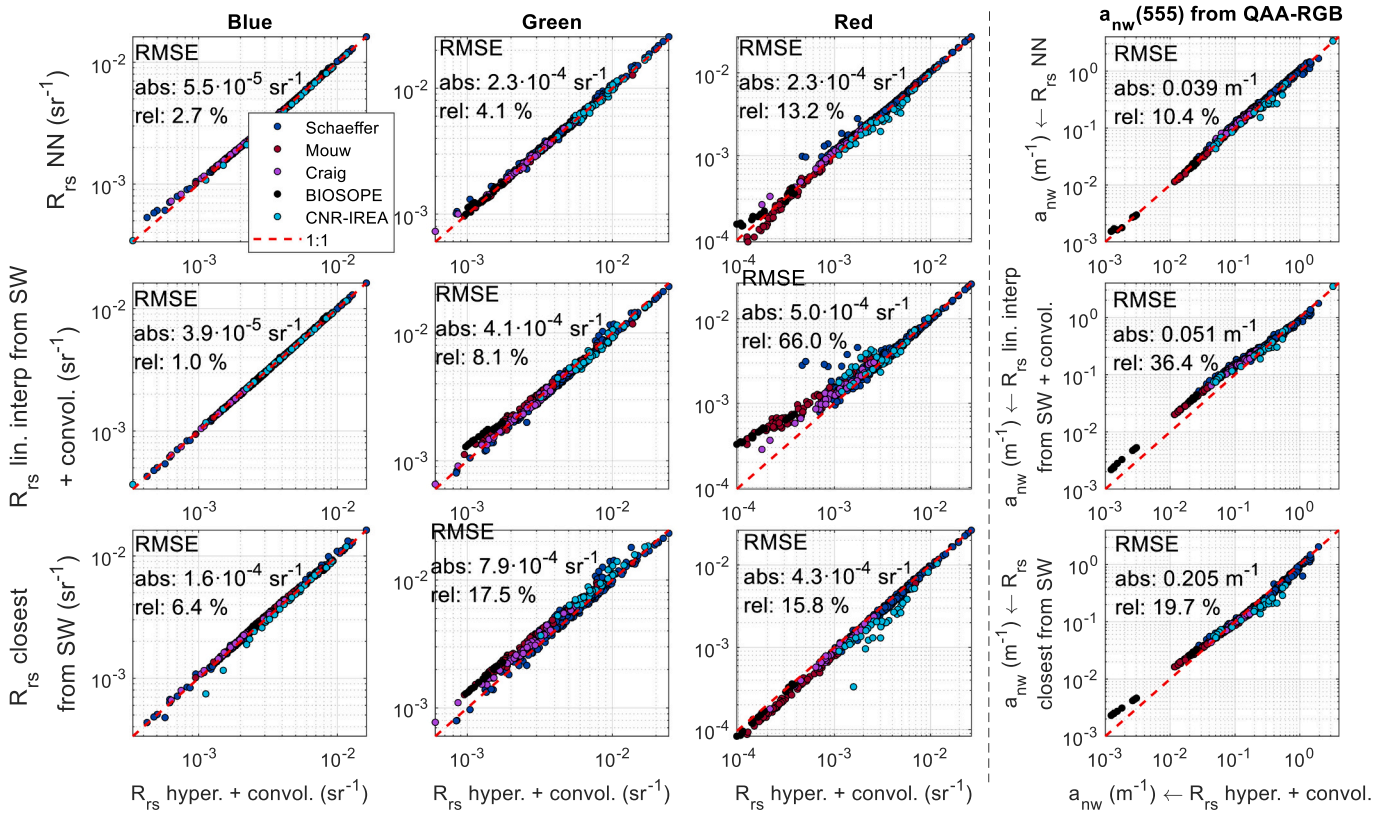

Fig. B3. First three columns: comparison between the "true" Landsat 4 R_{rs} generated by resampling hyperspectral R_{rs} to Landsat 4 (horizontal axis) and three different methods to convert SeaWiFS-equivalent R_{rs} to Landsat 4 data. Last column: same comparison in absorption units, after application of Eq. (1) of the QAA-RGB (main manuscript body). Vertical axis in upper row: using the neural network developed with Lee's dataset. Vertical axis in mid row: filling the spectral gaps by linearly interpolating the SeaWiFS- equivalent R_{rs} between 412 and 670 nm and then convolving to Landsat 4. Vertical axis in lower row: selecting the closest SeaWiFS bands to the Landsat 4 bands, namely bands 3, 5 and 6. RMS errors in absolute and relative units for the pooled datasets and each band are included.

References

Bresciani, M., Rossini, M., Morabito, G., Matta, E., Pinardi, M., Cogliati, S., Julitta, T., Colombo, R., Braga, F., Giardino, C., 2013. Analysis of within- and between-day chlorophyll-a dynamics in Mantua superior Lake, with a continuous spectroradiometric measurement. Mar. Freshw. Res. 64, 303–316.

Bresciani, M., Cazzaniga, I., Austoni, M., Sforzi, T., Buzzi, F., Morabito, G., Giardino, C., 2018. Mapping phytoplankton blooms in deep subalpine lakes from sentinel-2A and Landsat-8. Hydrobiologia 824, 197–214.

Brewin, R.J.W., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P.-Y., Devred, E., Doerffer, R., Fomferra, N., Franz, B., Grant, M., Groom, S., Horseman, A., Hu, C., Krasemann, H., Lee, Z., Maritorena, S., Mélin, F., Peters, M., Platt, T., Regner, P., Smyth, T., Steinmetz, F., Swinton, J., Werdell, J., White, G.N., 2015. The ocean colour climate change initiative: III. A round-robin comparison on in-water bio-optical algorithms. Remote Sens. Environ. 162, 271–294.

Capuzzo, E., Stephens, D., Silva, T., Barry, J., Forster, R.M., 2015. Decrease in water clarity of the southern and Central North Sea during the 20th century. Glob. Chang. Biol. 21, 2206–2214.

Casey, K.A., Rousseaux, C.S., Gregg, W.W., Boss, E., Chase, A.P., Craig, S.E., Mouw, C.B., Reynolds, R.A., Stramski, D., Ackleson, S.G., Bricaud, A., Schaeffer, B., Lewis, M.R., Maritorena, S., 2020. A global compilation of in situ aquatic high spectral resolution inherent and apparent optical property data for remote sensing applications. Earth Syst. Sci. Data 12, 1123–1139.

Colella, S., Brando, V.E., Cicco, A.d., D'Alimonte, D., Forneris, V., Bracaglia, M., 2021. Ocean Colour Production Centre, Ocean Colour Mediterranean and Black Sea Observation Product. Copernicus Marine Environment Monitoring Centre. Quality Information Document. https://resources.marine.copernicus.eu/documents/QUID/CMEMS-OC-QUID-009-038to045-071-073-078-079-095-096.pdf.

Gabr, B., Ahmed, M., Marmoush, Y., 2020. PlanetScope and Landsat 8 imageries for bathymetry mapping. J. Mar. Sci. Eng. 8, 143.

Jackson, T., 2020. ESA Ocean colour climate change initiative – phase 3. In: Product User Guide for v5.0 Dataset. https://esa-oceancolour-cci.org/sites/esa-oceancolour-cci.org/alfresco.php?file=a68aa514-3668-4935-9235-fca10f7e8bee&name=OC-CCI-PUG-v4.1-v1.pdf.

Jackson, T., Sathyendranath, S., Mélin, F., 2017. An improved optical classification scheme for the ocean colour essential climate variable and its applications. Remote Sens. Environ. 203, 152–161.

Lee, Z., Shang, S., Qi, L., Yan, J., Lin, G., 2016. A semi-analytical scheme to estimate Secchi-disk depth from Landsat-8 measurements. Remote Sens. Environ. 177, 101–106.

Lee, Z.P., 2014. Update of the Quasi-Analytical Algorithm (QAA_v6). http://www.ioccg.org/groups/Software_OCA/QAA_v6_2014209.pdf.

Lee, Z.P., Carder, K.L., Steward, R.G., Peacock, T.G., Davis, C.O., Patch, J.S., 1998. An empirical algorithm for light absorption by ocean water based on color. J. Geophys. Res. Oceans 103, 27967–27978.

Lee, Z.P., 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.P., 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.

Lee, Z.P., Shang, S., Hu, C., Du, K., Weidemann, A., Hou, W., Lin, J., Lin, G., 2015. Secchi disk depth: a new theory and mechanistic model for underwater visibility. Remote Sens. Environ. 169, 139–149.

Luo, Y., Doxaran, D., Ruddick, K., Shen, F., Gentili, B., Yan, L., Huang, H., 2018. Saturation of water reflectance in extremely turbid media based on field measurements, satellite data and bio-optical modelling. Opt. Express 26, 10435–10451.

Manzo, C., Bresciani, M., Giardino, C., Braga, F., Bassani, C., 2015. Sensitivity analysis of a bio-optical model for Italian lakes focused on Landsat-8, Sentinel-2 and Sentinel-3. Eur. J. Remote Sens. 48, 17–32.

Mason, J.D., Cone, M.T., Fry, E.S., 2016. Ultraviolet (250–550 nm) absorption spectrum of pure water. Appl. Opt. 55, 7163–7172.

MATLAB, 2021. Neural Network Fitting Tool. https://es.mathworks.com/help/deeplearning/gs/fit-data-with-a-neural-network.html.

Nechad, B., Ruddick, K., Schroeder, T., Blondeau-Patissier, D., Cherukuru, N., Brando, V. E., Dekker, A.G., Clementson, L., Banks, A., Maritorena, S., Werdell, P.J., Sá, C., Brotas, V., Caballero de Frutos, I., Ahn, Y.-H., Salama, S., Tilstone, G., Martinez-Vicente, V., Foley, D., McKibben, M., Nahorniak, J., Peterson, T.D., Siliò-Calzada, A., Röttgers, R., Lee, Z., Peters, M., Brockmann, C., 2015a. CoastColour round Robin datasets, version 1. In, in supplement to: Nechad, Bouchra; Ruddick, Kevin; Schroeder, Thomas; Oubelkheir, Kadija; Blondeau-Patissier, David; Cherukuru, Nagur; Brando, Vittorio E; Dekker, Arnold G; Clementson, Lesley; Banks, Andrew; Maritorena, Stéphane; Werdell, P Jeremy; Sá, Carolina; Brotas, Vanda; Caballero de Frutos, Isabel; Ahn, Yu-Hwan; Salama, Suhyb; Tilstone, Gavin; Martinez-Vicente,

victor; Foley, David; McKibben, Morgaine; Nahorniak, jasmine; Peterson, Tawnya D; Siliò-Calzada, Ana; Röttgers, Rüdiger; Lee, Zhongping; Peters, Marco (2015): CoastColour round Robin data sets: a database to evaluate the performance of algorithms for the retrieval of water quality parameters in coastal waters. Earth Syst. Sci. Data 7 (2), 319–348. https://doi.org/10.5194/essd-7-319-2015. PANGAEA.

Nechad, B., Ruddick, K., Schroeder, T., Oubelkheir, K., Blondeau-Patissier, D., Cherukuru, N., Brando, V., Dekker, A., Clementson, L., Banks, A.C., Maritorena, S., Werdell, P.J., Sá, C., Brotas, V., Caballero de Frutos, I., Ahn, Y.H., Salama, S., Tilstone, G., Martinez-Vicente, V., Foley, D., McKibben, M., Nahorniak, J., Peterson, T., Siliò-Calzada, A., Röttgers, R., Lee, Z., Peters, M., Brockmann, C., 2015b. CoastColour round Robin data sets: a database to evaluate the performance of algorithms for the retrieval of water quality parameters in coastal waters. Earth Syst. Sci. Data 7, 319–348.

O'Reilly, J.E., Werdell, P.J., 2019. Chlorophyll algorithms for ocean color sensors - OC4, OC5 & OC6. Remote Sens. Environ. 229, 32–47.

Pahlevan, N., Balasubramanian, S.V., Sarkar, S., Franz, B.A., 2018. Toward long-term aquatic science products from heritage Landsat missions. Remote Sens. 10, 1337.

Pahlevan, N., Chittimalli, S.K., Balasubramanian, S.V., Vellucci, V., 2019. Sentinel-2/Landsat-8 product consistency and implications for monitoring aquatic systems. Remote Sens. Environ. 220, 19–29.

Pitarch, J., 2020. A review of Secchi's contribution to marine optics and the foundation of Secchi disk science. Oceanography 33 (3).

Pitarch, J., Volpe, G., Colella, S., Santoleri, R., Brando, V., 2016. Absorption correction and phase function shape effects on the closure of apparent optical properties. Appl. Opt. 55, 8618–8636.

Pitarch, J., van der Woerd, H.J., Brewin, R.J.W., Zielinski, O., 2019. Optical properties of Forel-Ule water types deduced from 15 years of global satellite ocean color observations. Remote Sens. Environ. 231, 111249.

Pitarch, J., Bellacicco, M., Organelli, E., Volpe, G., Colella, S., Vellucci, V., Marullo, S., 2020. Retrieval of particulate backscattering using field and satellite radiometry: assessment of the QAA algorithm. Remote Sens. 12, 77.

Pope, R.M., Fry, E.S., 1997. Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements. Appl. Opt. 36, 8710–8723.

Poursanidis, D., Traganos, D., Chrysoulakis, N., Reinartz, P., 2019. Cubesats allow high spatiotemporal estimates of satellite-derived bathymetry. Remote Sens. 11, 1299.

Shi, W., Wang, M., 2017. Characterization of particle backscattering of global highly turbid waters from VIIRS Ocean color observations. J. Geophys. Res. Oceans 122, 9255–9275.

Shi, W., Wang, M., 2019. A blended inherent optical property algorithm for global satellite ocean color observations. Limnol. Oceanogr. Methods 17, 377–394.

Valente, A., Sathyendranath, S., Brotas, V., Groom, S., Grant, M., Taberner, M., Antoine, D., Arnone, R., Balch, W.M., Barker, K., Barlow, R., Bélanger, S., Berthon, J. F., Beşiktepe, Ş., Borsheim, Y., Bracher, A., Brando, V., Canuti, E., Chavez, F., Cianca, A., Claustre, H., Clementson, L., Crout, R., Frouin, R., García-Soto, C., Gibb, S.W., Gould, R., Hooker, S.B., Kahru, M., Kampel, M., Klein, H., Kratzer, S., Kudela, R., Ledesma, J., Loisel, H., Matrai, P., McKee, D., Mitchell, B.G., Moisan, T., Muller-Karger, F., O'Dowd, L., Ondrusek, M., Platt, T., Poulton, A.J., Repecaud, M., Schroeder, T., Smyth, T., Smythe-Wright, D., Sosik, H.M., Twardowski, M., Vellucci, V., Voss, K., Werdell, J., Wernand, M., Wright, S., Zibordi, G., 2019. A compilation of global bio-optical in situ data for ocean-colour satellite applications – version two. Earth Syst. Sci. Data 11, 1037–1068.

Vanhellemont, Q., 2019a. Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives. Remote Sens. Environ. 225, 175–192.

Vanhellemont, Q., 2019b. Daily metre-scale mapping of water turbidity using CubeSat imagery. Opt. Express 27, A1372–A1399.

Vanhellemont, Q., 2020. Sensitivity analysis of the dark spectrum fitting atmospheric correction for metre- and decametre-scale satellite imagery using autonomous hyperspectral radiometry. Opt. Express 28, 29948–29965.

Vanhellemont, Q., Ruddick, K., 2018. Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications. Remote Sens. Environ. 216, 586–597.

Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: how free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ. 122, 2–10.

Xu, Y., Feng, L., Zhao, D., Lu, J., 2020. Assessment of Landsat atmospheric correction methods for water color applications using global AERONET-OC data. Int. J. Appl. Earth Obs. Geoinf. 93, 102192.

Zhang, X., Hu, L., He, M.-X., 2009. Scattering by pure seawater: effect of salinity. Opt. Express 17, 5698–5710.

Zheng, G., Stramski, D., Reynolds, R.A., 2014. Evaluation of the quasi-analytical algorithm for estimating the inherent optical properties of seawater from ocean color: comparison of Arctic and lower-latitude waters. Remote Sens. Environ. 155, 194–209.