Remote Sensing of Environment 258 (2021) 112366

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters

Nima Pahlevan [a,b,*], Antoine Mangin [c], Sundarabalan V. Balasubramanian [d], Brandon Smith [a,b], Krista Alikas [e], Kohei Arai [f], Claudio Barbosa [g], Simon Bélanger [h], Caren Binding [i], Mariano Bresciani [j], Claudia Giardino [j], Daniela Gurlin [k], Yongzhen Fan [l], Tristan Harmel [m], Peter Hunter [n], Joji Ishikaza [o], Susanne Kratzer [p], Moritz K. Lehmann [q], Martin Ligi [e], Ronghua Ma [r], François-Régis Martin-Lauzer [c], Leif Olmanson [s], Natascha Oppelt [t], Yanqun Pan [h,u], Steef Peters [v], Nathalie Reynaud [w], Lino A. Sander de Carvalho [x], Stefan Simis [y], Evangelos Spyrakos [n], François Steinmetz [z], Kerstin Stelzer [aa], Sindy Sterckx [ab], Thierry Tormos [ac], Andrew Tyler [n], Quinten Vanhellemont [ad], Mark Warren [y]

[a] NASA Goddard Space Flight Center, Greenbelt, MD, USA
[b] Science Systems and Applications, Inc. (SSAI), Lanham, MD, USA
[c] ACRI-ST, Sophia Antipolis, France
[d] Geosensing and Imaging Solution Consultancy, Trivandrum, Kerala, India
[e] Tartu Observatory, University of Tartu, Tartumaa, Estonia
[f] Department of Information Science, Saga University, Saga, Japan
[g] Instrumentation Lab for Aquatic Systems (LabISA), National Institute for Space Research (INPE), São José dos Campos, Brazil
[h] Université du Québec à Rimouski, Département de Biologie, Chimie et Géographie, Group BOREAS and Québec-Océan, Rimouski, QC, Canada
[i] Environment and Climate Change Canada, Burlington, ON, Canada
[j] National Research Council of Italy, Institute for Electromagnetic Sensing of the Environment, CNR-IREA, Italy
[k] Wisconsin Department of Natural Resources, Madison, WI, USA
[l] Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, NJ, USA
[m] Géosciences Environment Toulouse (GET), Toulouse, France
[n] Earth and Planetary Observation Sciences (EPOS), Department of Biological and Environmental Sciences, University of Stirling, Stirling, UK
[o] Institute for Space-Earth Environmental Research (ISEE), Nagoya University, Nagoya, Japan
[p] Department of Ecology, Environment and Plant Sciences (DEEP), Stockholm University, Stockholm, Sweden
[q] Xerra Earth Observation Institute and the University of Waikato, New Zealand
[r] Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Science, Nanjing, China
[s] Department of Forest Resources, University of Minnesota, St. Paul, MN, USA
[t] Earth Observation and Modelling, Kiel University, Department of Geography, Kiel, Germany
[u] ARCTUS, Rimouski, QC, Canada
[v] Water Insight, Wageningen, the Netherlands
[w] UR RECOVER - Pôle ECLA, INRAE, Aix-en-Provence, France
[x] Meteorology Department, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil
[y] Plymouth Marine Laboratory, Plymouth, UK
[z] Hygeos, Euratechnologies, Lille, France
[aa] Brockmann Consult GmbH, Hamburg, Germany
[ab] Flemish Institute for Technological Research (VITO), Remote Sensing Unit, Mol, Belgium
[ac] Unité ECosystèmes LAcustres – Pôle ECLA, Office Français pour le Biodiversité (OFB), Aix-en-Provence, France
[ad] Royal Belgian Institute of Natural Sciences (RBINS), Operational Directorate Natural Environments, Brussels, Belgium

A B S T R A C T

Atmospheric correction over inland and coastal waters is one of the major remaining challenges in aquatic remote sensing, often hindering the quantitative retrieval of biogeochemical variables and analysis of their spatial and temporal variability within aquatic environments. The Atmospheric Correction Intercomparison Exercise (ACIX-Aqua), a joint NASA – ESA activity, was initiated to enable a thorough evaluation of eight state-of-the-art atmospheric correction (AC) processors available for

* Corresponding author at: NASA Goddard Space Flight Center, Greenbelt, MD, USA.
E-mail address: nima.pahlevan@nasa.gov (N. Pahlevan).

https://doi.org/10.1016/j.rse.2021.112366
Received 6 August 2020; Received in revised form 16 February 2021; Accepted 20 February 2021
Available online 9 March 2021

Landsat-8 and Sentinel-2 data processing. Over 1000 radiometric matchups from both freshwaters (rivers, lakes, reservoirs) and coastal waters were utilized to examine the quality of derived aquatic reflectances ($\widehat{\rho}_w$). This dataset originated from two sources: Data gathered from the international scientific community (henceforth called Community Validation Database, CVD), which captured predominantly inland water observations, and the Ocean Color component of AERONET measurements (AERONET-OC), representing primarily coastal ocean environments. This volume of data permitted the evaluation of the AC processors individually (using all the matchups) and comparatively (across seven different Optical Water Types, OWTs) using common matchups. We found that the performance of the AC processors differed for CVD and AERONET-OC matchups, likely reflecting inherent variability in aquatic and atmospheric properties between the two datasets. For the former, the median errors in $\widehat{\rho}_w(560)$ and $\widehat{\rho}_w(664)$ were found to range from 20 to 30% for best-performing processors. Using the AERONET-OC matchups, our performance assessments showed that median errors within the 15–30% range in these spectral bands may be achieved. The largest uncertainties were associated with the blue bands (25 to 60%) for best-performing processors considering both CVD and AERONET-OC assessments. We further assessed uncertainty propagation to the downstream products such as near-surface concentration of chlorophyll-*a* (Chl*a*) and Total Suspended Solids (TSS). Using satellite matchups from the CVD along with *in situ* Chl*a* and TSS, we found that 20–30% uncertainties in $\widehat{\rho}_w(490 \leq \lambda \leq 743\ nm)$ yielded 25–70% uncertainties in derived Chl*a* and TSS products for top-performing AC processors. We summarize our results using performance matrices guiding the satellite user community through the OWT-specific relative performance of AC processors. Our analysis stresses the need for better representation of aerosols, particularly absorbing ones, and improvements in corrections for sky- (or sun-) glint and adjacency effects, in order to achieve higher quality downstream products in freshwater and coastal ecosystems.

1. Introduction

Compensation for atmospheric scattering and absorption and for surface reflection at the air-water interface (i.e., sky-glint and sun-glint) from the signal measured at the Top of Atmosphere (TOA) is referred to as the process of atmospheric correction (AC). Robust AC is essential for the accurate retrieval of aquatic reflectance and downstream science products (e.g., near-surface concentration of chlorophyll-a; Chl*a*, and Total Suspended Solids; TSS) from remotely sensed observations. AC over the open ocean is carried out adequately, as reported by the International Ocean Color Coordinating Group (IOCCG) (IOCCG, 2010), but over inland and coastal waters inaccurate AC still leads to large uncertainties in satellite data products, thus limiting the detection of subtle variability in aquatic ecosystems (Pahlevan et al., 2020). As a result, some satellite-based methods for the detection of harmful algal blooms (HABs), for instance, rely on Level 1 TOA or simple Rayleigh-corrected quantities in order to avoid large uncertainties in Level 2 products introduced by poor AC performance over eutrophic waters (Binding et al., 2021; Matthews and Bernard, 2013; Schaeffer et al., 2018; Stumpf et al., 2016).

The availability of image data at resolutions on the order of tens of meters, such as those acquired by the joint National Aeronautics and Space Administration (NASA) and U.S. Geological Survey (USGS) Landsat program and the Copernicus Sentinel-2 Services, has spurred development of applications for smaller water bodies such as lakes, rivers and estuaries. This has led to the development of a number of novel AC processors to obtain accurate satellite-derived aquatic reflectances ($\widehat{\rho}_w$) for downstream products (e.g., Chl*a*). These processors show significant differences in methods and previous *ad hoc* comparisons were limited in geographic scope, or in the number or types of matchups (i.e., coastal versus freshwater), and did not often adhere to an identical matchup analysis approach (Ansper and Alikas, 2019; De Keukelaere et al., 2018; Ilori et al., 2019; Pahlevan et al., 2017b; Pahlevan et al., 2017c; Pereira-Sandoval et al., 2019; Renosh et al., 2020; Soomets et al., 2020; Vanhellemont, 2019; Warren et al., 2019). Thus, important questions remain unaddressed, for example, whether processors meet the currently defined 30% threshold requirements (Global Climate Observing System; GCOS) across all bands, how they compare to heritage approaches (e.g., Franz et al., 2015), and specifying which processor(s) can provide more reliable downstream products given the uncertainties in $\widehat{\rho}_w$ under various atmospheric conditions and/or across distinct aquatic ecosystems. A list of widely used notations and acronyms throughout this article is provided in Table 1.

To address important questions relating to atmospheric correction, NASA and the European Space Agency (ESA) jointly conducted an Atmospheric Correction Intercomparison eXercise (ACIX) in coordination with the Committee on Earth Observation Satellites (CEOS) validation group. The initial exercise (ACIX-I; Doxani et al. (2018)) focused on diverse land-cover types and continental/coastal atmospheric conditions and mostly evaluated aerosol optical thickness (AOT) retrieval to gauge performances. The scope of the first exercise was however limited for aquatic studies due to the absence of assessments of ρ_w retrieval accuracy. Therefore, a second Atmospheric Correction Intercomparison eXercise (ACIX-II), hereafter referred to as ACIX-Aqua, commenced in October 2018. To fulfill the objective of ACIX-Aqua, a large number of high-quality observations from a representative range of aquatic environments were desirable. Automated and systematically processed radiometric observations through the Ocean Color component of the Aerosol Robotic Network (AERONET-OC; Zibordi et al., 2009a; Zibordi et al., 2006) are an invaluable asset for such analyses (Hlaing et al., 2013; Jamet et al., 2011; Mélin et al., 2010). They mostly represent coastal ecosystems, but also include a number of freshwater sites (Philipson et al., 2016). To complement the AERONET-OC database and to add more data representing freshwater ecosystems, a community-wide data sharing initiative was undertaken. The goal was to include optically diverse water bodies where, for example, AC may be more challenging due to environmental factors, such as the presence of absorbing aerosols or land adjacency effects (AE) (IOCCG, 2018). This second data set is henceforth referred to as the Community Validation Database (CVD).

The purpose of this article is to provide an objective assessment of state-of-the-art atmospheric correction using a global dataset representing a wide array of atmospheric and aquatic conditions. We evaluated eight different AC processors applied to Landsat-8 Operational Land Imager (OLI) data and Sentinel-2A/B MultiSpectal Instrument (MSI) images over inland and coastal waters. Owing to their inherent differences in measurement techniques and representations of aquatic environments, the AERONET-OC and CVD matchups were primarily treated independently. Additionally, our entire dataset (AERONET-OC and CVD aggregated) was divided into Optical Water Types (OWTs, Section 3.4), which allowed an assessment of processors across widely variable coastal and inland water conditions. Furthermore, we aimed to investigate how uncertainties in $\widehat{\rho}_w$ manifest in satellite retrieved Chl*a* and TSS for each processor. To that end, we provide a brief overview of

Table 1
List of widely used notations, acronyms, and symbols.

Symbol	Description
ρ_w	*In situ* aquatic reflectance
$\widehat{\rho}_w$	Satellite derived aquatic reflectance
ρ_t	Top-Of-Atmosphere (TOA) reflectance
ρ_{rc}	Rayleigh-corrected reflectance
Chl*a*	*In situ* chlorophyll-*a*
TSS	*In situ* Total Suspended Solids
Chl*a*^{p}	Pseudo Chl*a* derived from ρ_w
TSS^{p}	Pseudo TSS derived from ρ_w
Chl*a*^{r}	Remotely sensed Chl*a* (estimated from $\widehat{\rho}_w$)
TSS^{r}	Remotely sensed TSS (estimated from $\widehat{\rho}_w$)

atmospheric correction followed by a description of the *in situ* data employed in the analysis. Subsequently, succinct descriptions of the AC processors, matchup selection, OWT classification, as well as Chl*a* and TSS retrieval algorithms, and statistical metrics used for the performance assessments are presented. The results and their implications for scientific studies and monitoring applications are further elaborated upon in Sections 4 and 5. Finally, we offer recommendations (Section 6) on the viability of the AC processors for freshwater and coastal ecosystems with due attention to the quality of downstream products, such as Chl*a* and TSS.

2. Background: Atmospheric correction

During the early ocean color era, prior to the launch of the Coastal Zone Color Scanner (CZCS), radiative transfer analyses had shown that most of the "blue" photons reaching TOA over the ocean arise from within the atmosphere (Gordon, 1976) through the scattering and absorption processes induced by gas molecules and aerosols. The purpose of AC for satellite observations is, therefore, to compensate for the photons that do not originate from the water column. The goal is to estimate non-dimensional aquatic reflectance $\widehat{\rho}_w$,which may also be denoted as water-leaving reflectance in the literature (Ruddick et al., 2006). In the absence of sun-glint (i.e., direct solar beam specular reflection), surface whitecaps and AE, the total signal expressed as TOA reflectance, $\rho_t(\lambda)$, can be simplified as

$$\rho_t(\lambda) = t\widehat{\rho}_w\left(\lambda\right) + [\rho_r(\lambda) + \rho_a(\lambda) + \rho_{ar}(\lambda)] \quad (1)$$

where t is the diffuse transmission (Gordon and Wang, 1994), ρ_r is the Rayleigh reflectance in the absence of aerosol, ρ_a is the aerosol reflectance, and ρ_{ar} is the radiance arising from Rayleigh-aerosol multiple-scattering. Depending on the methodology adopted for AC, the three components within the brackets are computed either as one unknown parameter, i.e., ρ_{path}, or by separating the Rayleigh and aerosol contributions (Antoine and Morel, 1999; Deschamps et al., 1983; Gordon, 1978). The diffuse transmission is also computed knowing the transmissions due to air and gas molecules as well as aerosols (Yang and Gordon, 1997). Among all the unknown components, estimating the aerosol contribution is the most intractable undertaking, and small errors may lead to high uncertainties in $\widehat{\rho}_w$ (Gordon and Wang, 1994).

Several approaches have been proposed for the removal of atmospheric effects over open oceans for missions like the Moderate Resolution Imaging Spectroradiometer (MODIS), Medium-Spectral Resolution Imaging Spectrometer (MERIS), Global Imager (GLI), and POLarization and Directionality of the Earth's Reflectances (POLDER) (IOCCG, 2010). These approaches differ primarily in the methodology used for the removal of the aerosol contribution (Antoine and Morel, 1999; Chomko and Gordon, 1998; Fukushima et al., 1998; Gao et al., 2000; Gordon, 1997; Gordon et al., 1997; Gordon and Wang, 1994; Harmel and Chami, 2011; Nicolas et al., 2002). This body of research collectively asserts that the performance of the existing AC processors are, in general, acceptable over clear ocean waters, where water-leaving radiance in the near-infrared (NIR) is negligible and maritime (non-absorbing) aerosols are the dominant aerosol type (IOCCG, 2010).

In both inland and coastal waters, water-leaving radiance in the NIR is often not negligible, and AE due to multiple-scattering from neighboring terrestrial terrain can contribute to ρ_{path} (Bulgarelli et al., 2014; Santer and Schmechtig, 2000; Sterckx et al., 2015). Moses et al. (2017) reported how the removal of atmospheric effects over inland and coastal waters might be further complicated by other factors, such as the proximity to terrestrial sources of aerosols, which result in an optically heterogeneous atmosphere. In coastal waters, terrestrial- and marine-source aerosols can also produce spatially variable and mixed conditions (Pahlevan et al., 2017a) that may not be fully represented in existing aerosol models (Ahmad et al., 2010). Some variants of the open-ocean algorithms attempt to account for non-negligible water-leaving radiances in the NIR in inland and coastal waters (Bailey et al., 2010; Moore et al., 1999; Siegel et al., 2000; Stumpf, 2004). Alternatively, the use of the short-wave infrared (SWIR) bands has been demonstrated to improve retrievals in sediment-dominated waters (Gao et al., 2000; Vanhellemont and Ruddick, 2014; Wang, 2007; Wang and Shi, 2007). However, AC over inland and coastal waters is still an important area of research and development.

3. Methods

3.1. Dataset

The radiometric quantity commonly utilized in remote sensing studies is the *in situ* aquatic reflectance, ρ_w, defined as

$$\rho_w = \pi \times L_w(0^+)/E_d(0^+) = \pi \times R_{rs} \quad (2)$$

where $L_w(0^+)$ and $E_d(0^+)$ are the water-leaving radiance and downwelling irradiance just above the water surface, respectively, and R_{rs} is remote-sensing reflectance (Mobley 1999). To avoid confusion with nomenclature used in the terrestrial remote sensing community, we adopt the term "aquatic reflectance" for ρ_w.

The standard AERONET-OC product (Level 2), the normalized $L_w(0^+)$ corrected for bidirectional effects (Morel et al., 2002), was divided by the solar irradiance spectrum (Thuillier et al., 2003) resampled by 11 nm square filters (Zibordi et al., 2006), and then multiplied by π to yield ρ_w. The initial pool of matchups for both missions included >1200 samples. To account for the difference in the spectral sampling of AERONET-OC and $\widehat{\rho}_w$, we applied the deep neural network approach[1] proposed in Pahlevan et al. (2017d), which converts AERONET-OC ρ_w to OLI (443, 482, 561, and 655 nm) and MSI (443, 490, 560, and 664 nm) broadband observations. Throughout this research, however, we use MSI band centers to refer to these four visible bands. The spectral band adjustments amounted to $<15\%$ change in the magnitude of ρ_w at the 490, 560, and 664 nm bands, subject to the shape and magnitude of the spectra (Fig. 3). Note that only $<17\%$ of our final AERONET-OC dataset corresponded to inland waters; thus, the associated analysis is primarily representative of coastal regions (Fig. S1).

In order to expand the representation of inland watersinland waters, a total of 2679 hyperspectral ρ_w records were compiled in the CVD. These spectra had been acquired using field radiometers that were assembled and calibrated by six different manufacturers. These instruments were utilized for field measurements following 10 different techniques, all of which are listed in Table S1 and Fig. S2. A visual inspection of ρ_w was carried out to detect data showing abnormal spectral features (e.g., noisy spectra, negative values in any spectral bands). This led to the exclusion of less $<1\%$ of the initial database. Such a small percentage of excluded spectra was a result of the data quality screening that had already been undertaken by the data providers. No further adjustments (e.g., corrections for the Bidirectional Reflectance Distribution Function; Hlaing et al., 2012) were applied to this dataset.

The *in situ* radiometric data were provided at various spectral resolutions (between 1 and 3.3 nm) and ranges (mostly within the 350–900 nm range). The spectra were convolved with the relative spectral responses of OLI and MSI to obtain band-equivalent ρ_w. Here, MSI NIR bands at 705, 743, and 783 nm bands were also simulated. Overall, $<$ 10% of the entire CVD were measured in coastal (e.g., near Crete, Greece) or brackish waters (e.g., Baltic Sea), and the rest represent a diverse range of freshwater ecosystems, from hypereutrophic lakes (e.g., Lake Taihu, China), to oligotrophic lakes (Lake Garda, Italy), and to rivers (e.g., Japurá River, Brazil) (Table S1). The CVD is thus a significant complement to AERONET-OC. The continental distributions of

[1] Code can be accessed via https://github.com/STREAM-RS/STREAM-RS

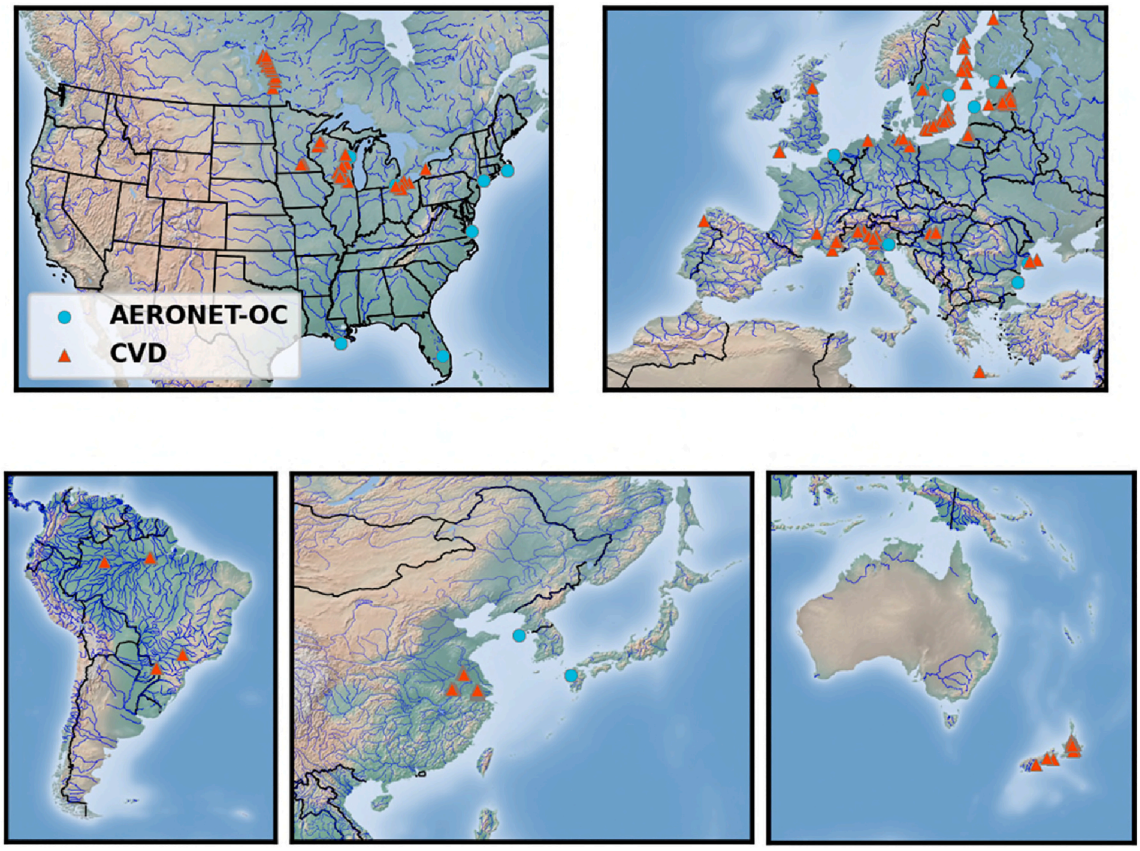

Fig. 1. Locations of valid *in situ* radiometric matchups acquired near-coincident with Landsat-8 and Sentinel-2 overpasses (see Section 3.3 for more details). These matchups correspond to diverse aquatic ecosystems, including lakes, rivers, and coastal waters (see Tables S1 and S3). The Community Validation Database (CVD) contains data mostly representing inland waters. Background map source: https://www.shadedrelief.com/

valid OLI and MSI matchups, following the implementation of the matchup criteria (Section 3.3) are illustrated in Fig. 1.

The assessment of the effect of AC on the quality of downstream products requires matchups accompanied by *in situ* measurements of water constituents. The AERONET-OC data do not include measurements of Chl*a* and TSS while the CVD has some direct measurements with 123 matchups (Fig. 2). Therefore, for the CVD, we primarily focused on a quantitative analysis (Section 4.2.1) using this subset of Chl*a* and TSS matchups and compared them against Chl*a* and TSS estimated from $\widehat{\rho}_w$ (hereafter referred to as Chla^r and TSS^r) via best-practice retrieval algorithms (Section 3.5). For the AERONET-OC data, we applied select retrieval algorithms to ρ_w to estimate Chl*a* and TSS, hereafter referred to as pseudo *in situ* estimates (Chla^p and TSS^p) (Section 4.2.2). For an ideal AC processor, Chla^p and Chla^r, for instance, are expected to agree if produced via *the same* algorithm. It is therefore plausible to assume that the differences in products are attributable to

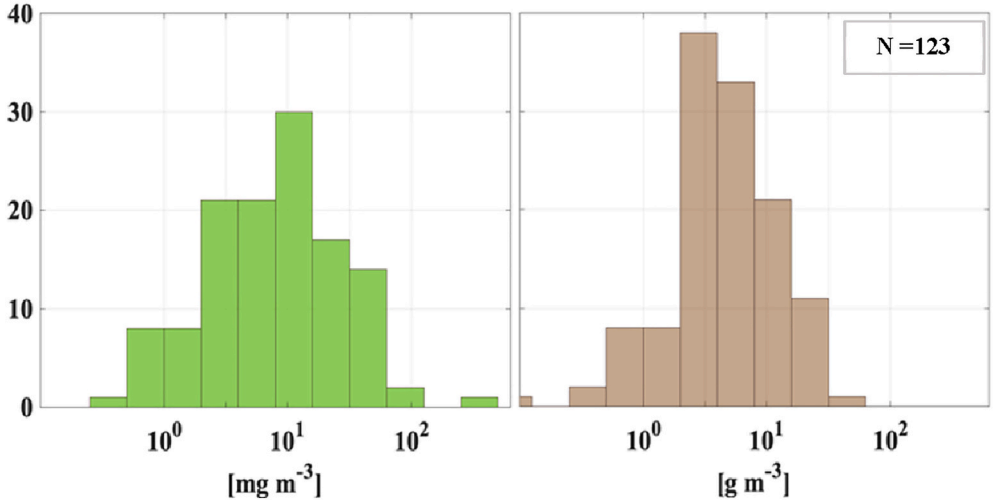

Fig. 2. Frequency distributions of available Chl*a* (left) and TSS in the CVD matchups (Section 4.2.1). Mean and median values are 15.8 and 8.3 mg m^{-3} for Chl*a*, and 6.8 and 4.25 g m^{-3} for TSS.

Table 2
Attributes of AC processors used to process OLI and MSI matchups for this exercise (ACIX-Aqua).

Two-step models						
Correction	ACOLITE	GRS	MEETC2	POLYMER	SeaDAS	iCOR
Gaseous	O2, O3 (OMI)	CH4, CO, CO2, NO2, O2, O3 (CAMS)	O2, O3, NO2 (Gasteiger et al., 2014)	O3 and NO2	O2, O3, NO2 (NCEP/OMI)	O2, O3, NO2, CO, CO2, CH4 (O3 climatology)
Water vapor	NCEP	CAMS (Rahman and Dedieu, 1994)	ECMWF	NA	NA	ECMWF (OLI) (Schläpfer, 1998)
Sun-glint	Fit to residuals at ρ_{rc}(1609) & ρ_{rc}(2200)	Treated spectrally dependent	Cox and Munk, 1954	Treated as bulk signal	Optional	Subtraction of minimum
Sky-glint	Gordon et al., 1988	OSOAA LUTs	OSOAA LUTs		Ahmad and Fraser, 1982	Haan and Kokke, 1996
Adjacency effects	NA	NA	NA		NA	SIMEC (Sterckx et al., 2015)
Aerosol	Dark target approach (area-based)	Fitted to CAMS (area-based)	Dark target and BPC (per-pixel)	Polynomial fitting (per-pixel)	NIR-SWIR band ratio (per-pixel)	Dark target and AOT multi-parameter inversion (area-based)
Rayleigh LUT	6SV (Kotchenova et al., 2006)	OSOAA (Chami et al., 2015)	OSOAA (Chami et al., 2015)	SOS (Lenoble et al., 2007)	Ahmad and Fraser, 1982	MODTRAN 5.0 (Berk et al., 2006)
Other Considerations						
Geometry	Scene center for OLI and 5-km grids for MSI	Per-pixel	multi-scale	Per-pixel	Per-pixel	Per-pixel
Aerosol model	Continental/maritime aerosols (Kotchenova et al., 2006)	Mixture of fine and coarse modes (Harmel and Chami, 2011)	(Shettle and Fenn, 1979) (Moulin et al., 2001)	NA	Coastal/Ocean (Ahmad et al., 2010)	MODTRAN rural models (Berk et al., 2006)
Cloud masking	ρ_t(1609) > 0.0215	ρ_t (865) > 0.0275 & 0 < NDWI <1.1	NA	ρ_t (865) > 0.2	ρ_t (1609) > 0.018	Cloud mask layers are provided (De Keukelaere et al., 2018)
Output grid cell size (m)	10	20	10	10/20/60	20/30	60
Assumptions on bio-optical conditions	No	No	Yes (Saulquin et al., 2016)	Yes (Park and Ruddick, 2005)	Yes (Bailey et al., 2010)	No
Version	20,181,210	1.1.4	2	4.12	7.5	2.5
Open source access	Yes (ACOLITE)	Yes (GRS)	No	Yes (POLYMER)	Yes (SeaDAS)	No
Organizations	RBINS	INRAE	ACRI-ST	HYGEOS	NASA	VITO
References	Vanhellemont, 2019 (Vanhellemont and Ruddick, 2018	Harmel et al., 2018	Saulquin et al., 2016	Steinmetz et al., 2011 Steinmetz and Ramon, 2018	Franz et al., 2015 Pahlevan et al., 2017c	De Keukelaere et al., 2018

Machine-learning models		
Model components	C2X	OC-SMART
Atmospheric modeling	SOS (Lenoble et al., 2007)	AccuRT (Stamnes et al., 2018)
Atmospheric parameterization	O3, WV, NO2, O2, P_0	O3, WV, NO2, O2, P_0
Adjacency effects	No	No
Aquatic modeling	Hydrolight (Mobley and Sundman, 2008)	AccuRT (Stamnes et al., 2018)
Aquatic parameterization	NOMAD data (Bailey and Werdell, 2006) jnun	Ocean color climatology (Fan et al., 2021)
Retrieval approach	ρ_t –> ρ_w	ρ_{rc} –> ρ_w
Architecture	MLP, sigmoid activation function; six hidden layers; 77 neurons each	MLP, hyperbolic tangent activation function; 3 layers; 100 × 75 × 50 neurons
	Other Considerations	
Geometry	Per-pixel	Per-pixel
Aerosol models	Coastal – AERONET-OC (Holben et al., 1998)	Coastal/Ocean - (Ahmad et al., 2010)
Cloud masking	IdePix	Nordkvist et al., 2009
Output grid cell size (m)	10/20/30	30/60
Version	1.0 (C2X NN)	1.0
Assumption on bio-optical conditions	Yes (Brockmann et al., 2016)	Yes (Fan et al., 2021)

(*continued on next page*)

Table 2 (*continued*)

	Machine-learning models	
Model components	C2X	OC-SMART
Open source access	Yes (SNAP)	Yes (OC-SMART)
Organization	Brockman Consult	Stevens Institute of Technology
References	Brockmann et al., 2016	Fan et al., 2021; Fan et al., 2017

ACOLITE: Atmospheric Correction for OLI lite; C2X: Case-2 Extreme Waters; GRS: Glint Removal for Sentinel-2; MEETC2: Meet Case 2 waters; POLYMER: POLYnomial based algorithm applied to MERIS; OC-SMART: Ocean Color Simultaneous Marine and Aerosol Retrieval Tool; SeaDAS: SeaWiFS Data Analysis System; iCOR: image CORrection for atmospheric effects; WV: Water vapor; MLP: Multilayer perceptron; P_0: surface pressure; ECMWF: European Centre for Medium-Range Weather Forecasts; NCEP: National Centers for Environmental Prediction; NDWI; Normalized difference water index; OSOAA:: Ordres Successifs Océan Atmosphère Avancé; TOMS: Total Ozone Mapping Spectrometer; OMI: Ozone Monitoring Instrument; CAMS: Copernicus Atmospheric Monitoring System; BPC: Bayesian Predictive Classification.

individual processors and how they influence $\widehat{\rho}_w$ in spectral bands contributing to Chl*a* or TSS retrievals. Moreover, this assessment uses all the valid matchups ($N > 400$; Section 3.3). The analysis also demonstrates the sensitivity of retrieval algorithms (Table 4) to uncertainties associated with each AC processor, providing insights into the choice of retrieval algorithm for a respective processor. For completeness, this sensitivity assessment was also repeated for the CVD matchups (Appendix C).

3.2. AC processors

Eight AC methods were evaluated (Table 2). According to their underlying mechanisms, the processors fall into two broad categories, namely two-step and machine-learning schemes. In the two-step procedure, the effects of Rayleigh and gaseous absorption are first removed and then aerosol contribution is approximated (Eq. 1). ACOLITE and iCOR are the two processors that follow heritage terrestrial approaches for removing aerosol contributions (Vermote et al., 1997) while SeaDAS applies the heritage ocean color approach (Mobley et al., 2016).

POLYMER, on the other hand, fits a second-order polynomial function to ρ_{rc} to simultaneously correct for aerosol and sun-glint signals. Similarly, GRS applies a spectral fitting approach to the observed ρ_{path} signal to approximate aerosol radiance. The two machine-learning models, C2X and OC-SMART, are both based on multilayer perceptron neural networks, trained with synthetic datasets that were generated using in-water and atmospheric radiative transfer models. MEETC2 utilizes a Bayesian Predictive Classification (BPC) method using Gaussian Mixture Model *prior* distributions on ρ_{rc} and dissociates the effects of hydrosols and aerosols. Further details on the processors and relevant citations are provided in Table 2.

3.3. Matchup selection

While each processor has specific exclusion criteria to mask out clouds and/or haze and handle AE or sun-glint, we further employed an additional masking strategy to retain or remove matchups that could better inform our analysis. Prior to the matchup filtering, all potential matchups were inspected for outliers where either no valid retrievals were derived (e.g., due to clouds) or major disparities between ρ_w and $\widehat{\rho}_w$ were identified. If the difference between $\widehat{\rho}_w$ and ρ_w in any band for at least four processors was >100%, the retrieval was flagged as an outlier and the matchup was subsequently discarded from the entire assessment. These two exclusion criteria eliminated 1392 CVD matchups leaving a total of 1287 samples. Among the excluded measurements were *in situ* measurements in small lakes (e.g., Methods #6 and #9 in Table S1), close to shorelines, or in proximity of man-made objects where AE may have impeded a reliable AC. A similar preliminary cloud-masking led to ~ 600 matchups for the AERONET-OC dataset.

For image data, $\widehat{\rho}_w$ were extracted from within 150 m × 150 m square boxes surrounding the matchup location. This is equivalent to 5 × 5-element and 15 × 15-element windows for OLI and MSI, respectively, in the visible bands. For the NIR bands of MSI, this window size represents a 7 × 7-element window (Drusch et al., 2012). An area of approximately 90 m × 90 m centered around the AERONET-OC sites was discarded to minimize contamination from the respective platforms (e.g., lighthouses), where instruments are deployed. Slight differences in the window sizes between OLI and MSI matchups were assumed to introduce negligible uncertainties. The median value of the remaining valid pixels was chosen to best represent $\widehat{\rho}_w$. Additional (conservative) masking criteria (Eq. 3 & 4) were developed (Pahlevan et al., 2014) and modified via experimenting with the available *in situ* data and visual inspection:

$$\rho_t(1610) > 0.025 \; or \; std\,[\rho_t(1610)] > 0.01 \tag{3}$$

$$NIR_{in} < 0.01 \& std\,[\rho_t(1610)] > 0.001 \tag{4}$$

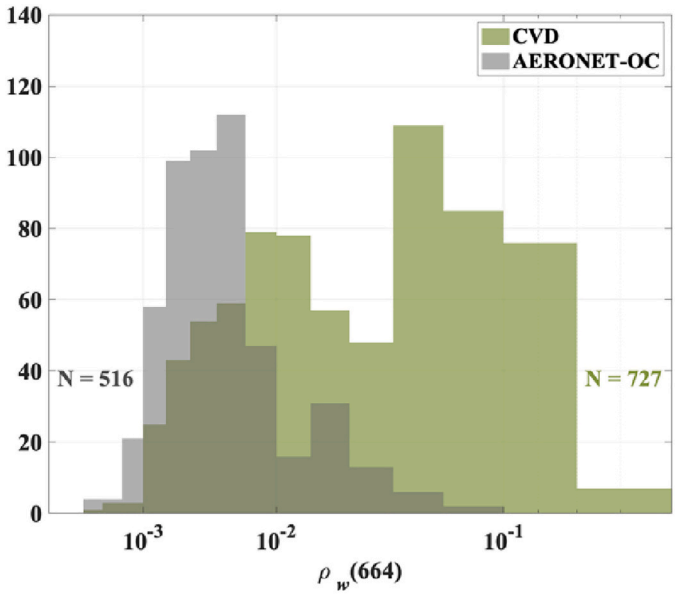

Fig. 3. Frequency distribution of (log-scale) $\rho_w(664)$ drawn from valid matchups obtained via POLYMER. The darker area corresponds to the overlap between the two distributions.

where NIR_{in} is the line height at the 865 nm band computed using ρ_t at 664 and 1609 nm (Abbott and Letelier, 1999). This index was defined to identify intense phytoplankton blooms and avoid their masking (e.g., Lake Taihu images). *Std* [ρ_t(1610)] was evaluated over the 150 m × 150 m matchup areas to evaluate local spatial variability. For instance, clouds and cloud shadows are expected to result in high standard deviations. The acceptable time difference for matchups was limited to +/− 30 h (Warren et al., 2019) for the CVD matchups and to +/− 3 h for AERONET-OC matchups (Bailey and Werdell, 2006).

Retrievals exhibiting negative values in any band were labeled as failure. Because of frequent negative retrievals in the NIR at the AERONET-OC sites, this criterion was applied only in the visible bands for the respective matchups. For the CVD matchups, this criterion was implemented independently for the visible and NIR bands (see 4.1). Fig. 3 shows the frequency distribution of $\rho_w(664)$ derived with POLYMER, which yielded the largest number of valid matchups for both groups of data ($N = 1243$; 516 and 727 for the AERONET-OC and CVD matchups, respectively). The number of valid matchups aggregated from both OLI and MSI images is provided for each processor in Table 3. The difference in the number of matchups stems from each processor adopting different theoretical assumptions and applying various thresholding schemes, or practical considerations, for retrievals. The detailed matchup statistics for other processors are also included in Table S3.

3.4. Optical water types

To enable a performance analysis per OWT, a classification scheme and set of reference OWTs were chosen. For the classification method, we used the spectral angle mapping technique that takes input spectra normalized by the area under the curve within the range of 400 to 750 nm (Yuhas et al., 1992). A given spectrum is assigned to one of the reference OWTs with which it constructs the minimum cosine angle. Among the various OWT references (Eleveld et al., 2017; Moore et al., 2014; Spyrakos et al., 2018), we found that the specific six and seven OWTs recommended in Eleveld et al. (2017) and Moore et al. (2014), respectively, do not fully encompass the optical variability found across coastal waters. As a result, the 21 OWTs of Spyrakos et al. (2018) were initially considered. The first assignment of ρ_w to these OWTs for our valid matchups (Table 3) led to under-represented OWTs that would render the interpretation of the results across a continuum of aquatic environments a challenge. Following a few experimental analyses by adding and removing classes from the original OWT set, a subset of seven OWTs was found to provide both a near-uniform distribution of

Table 3
Number of valid matchups for each processor.

	ACOLITE	C2X	GRS	MEETC2	OC-SMART	POLYMER	SeaDAS	iCOR
CVD (Visible)*	655	522	155	412	492	727	352	707
CVD (NIR)	163	130	111	78	NA	136	68	179
AERONET-OC*	455	479	366	340	460	516	450	NA

* These correspond to the number of valid matchups for the combined OLI and MSI data in the visible bands.

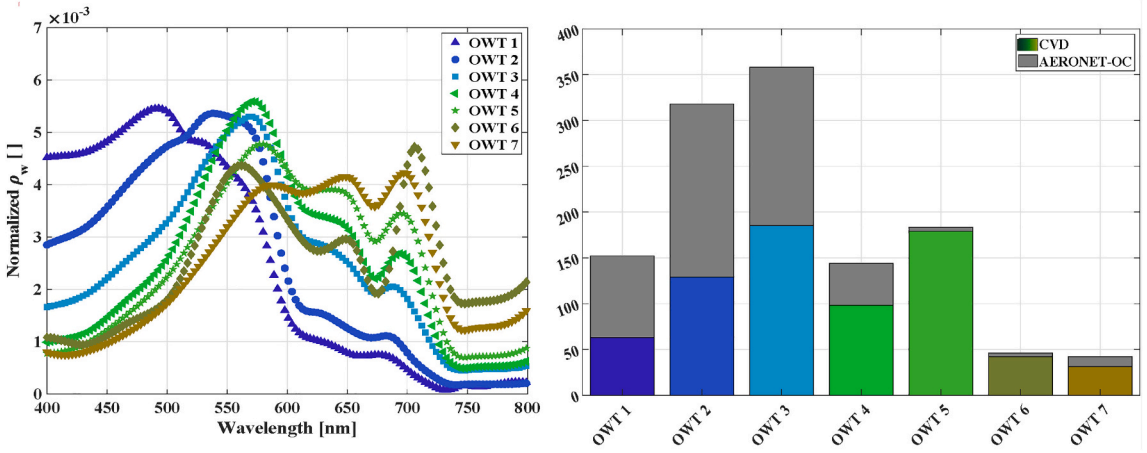

Fig. 4. Our seven OWTs (left) selected from Spyrakos et al. (2018) for the interpretation of the results (Section 4.1.2). Normalized ρ_w refers to ρ_w divided by area under the curve (Spyrakos et al., 2018). The number of matchups per OWT for POLYMER is also depicted (Fig. 3). See Table S2 for the correspondence of our OWTs with the original set (Spyrakos et al., 2018). For a detailed description of OWTs see Table 4.

Table 4
Statistics for water constituents associated with each OWT derived by taking weighted averages of data in Spyrakos et al. (2018) (N ~ 700) and all the relevant data available in CVD (N ~ 800). Units are mg m^{-3}, g m^{-3}, and m^{-1} for Chl*a*, TSS, and a_{cdom}, respectively. Note that a_{cdom} is reported at 440 nm.

	OWT1			OWT2			OWT3			OWT4			OWT5			OWT6			OWT7		
	Chl*a*	TSS	a_{cdom}	Chl*a*	TSS	a_{cdom}	Chl*a*	TSS	a_{cdom}	Chl*a*	TSS	a_{cdom}	Chl*a*	TSS	a_{cdom}	Chl*a*	TSS	a_{cdom}	Chl*a*	TSS	a_{cdom}
Median	1.2	1.1	0.06	2.3	1.99	0.09	7.1	6.6	0.44	13.2	7.2	0.65	15.7	15.3	2.38	44.5	11.8	1.14	13.7	66.9	1.75
Mean	2.7	1.2	2.7	2.4	2.1	0.09	12.1	7.7	0.64	21.1	7.9	0.70	20.1	23.1	2.60	52.6	14.1	1.56	26.2	78.9	2.56
N	41	37	37	277	224	224	254	212	212	141	140	138	267	248	243	427	242	264	294	306	211

OWTs for each processor's valid matchups and a fair coverage for various OWTs encountered in inland and coastal waters. The correspondence of our selected OWTs with the original set (Spyrakos et al., 2018) is provided in Table S2. Fig. 4. illustrates our selected seven OWTs and the distribution of OWTs, as an example, for the valid matchups of POLYMER. OWTs 1 and 2 are commonly found in the coastal waters and/or oligotrophic lakes whereas OWT3 is attributed to moderately eutrophic waters. Lakes or coastal estuaries with various degrees of phytoplankton blooms are represented by OWTs 4, 5, and 6. Lastly, OWT7 ensures that sediment-rich waters are present in our spectral library. Table 4 provides more specific statistics for the water constituents per OWT.

For the OWT-specific evaluation of AC processors (Section 4.1.2), AERONET-OC and CVD matchups were aggregated to permit a statistically robust inference in each OWT. To enable direct intercomparisons among the AC processors for each OWT, we pursued two approaches to identify common matchups: a) a pairwise matchup identification, and b) a universal matchup analysis. The pairwise matchups are determined between two AC processors while universal matchups are determined among multiple processors. The results associated with the universal matchup approach are illustrated in Appendix B (Figs. B1 and B2), but the number of matchups per OWT was considered insufficient. Therefore, we adopted the pairwise intercomparison as the primary method for ranking the AC processors for which CVD and AERONET-OC matchups were combined. The best-performing AC processor per OWT and spectral band was then ascertained by computing win rates (Section 3.6).

3.5. Water constituent retrieval

To examine how uncertainties in $\widehat{\rho}_w$ propagate to Chla^r and TSS^r products, two families of algorithms, namely Type I and Type II, were implemented. Type I algorithms are those that utilize the visible bands for retrievals, whereas Type II are those that require both visible and NIR, or only NIR, bands. Type I algorithms were applied to the combined OLI and MSI matchups to enhance the statistical significance of the analysis. For this category, widely used empirical algorithms, including Nechad (Nechad et al., 2010) for TSS and OCx (O'Reilly and Werdell, 2019) for Chl*a*, were utilized (Table 5). For Chl*a* assessments, the recently developed Mixture Density Networks (MDN) model was also employed (Pahlevan et al., 2020; Smith et al., 2021), while SOLID (Balasubramanian et al., 2020) and Novoa (Novoa et al., 2017) models were also applied for TSS analyses. While the MDN model uses all the available bands across the visible and NIR region, it is sensitive to the most relevant spectral bands (e.g., 490 and 560 nm; Smith et al., 2021). Note that the switching schemes available for SOLID and Novoa in sediment-rich waters were disabled as only OLI-MSI visible bands were utilized. Type II algorithms are only applicable for Chl*a* retrievals from MSI matchups for which 13 different algorithms, including MDN and variations of three-band models (e.g., Moses et al. (2012)), were implemented. For conciseness, outcomes of only the five top-performing models are presented (Table 7). The formulations for the retrieval algorithms listed in Table 5 are provided in the Equations section of the supplementary material.

3.6. Performance metrics

We mainly focused on investigating two metrics representing the overall error and bias in target quantities. These two norms first transform data into log space and then convert them back to linear space to assess the quality of the retrieved quantity (e.g., $\widehat{\rho}_w$) against that measured *in situ* (e.g., ρ_w). These metrics are computed as follows:

Table 5
Algorithms utilized in this study to examine the impacts of uncertainties in the AC processors on Chl*a* and TSS retrieval.

Models	Product	CVD					AERONET-OC		
Type I	Chl*a*	MDN (Smith et al., 2021)		OCx (O'Reilly and Werdell, 2019)			MDN (Pahlevan et al., 2020)	OCx (O'Reilly and Werdell, 2019)	
	TSS	SOLID (Balasubramanian et al., 2020)		Nechad (Nechad et al., 2010)			SOLID (Balasubramanian et al., 2020)	Nechad (Nechad et al., 2010)	Novoa (Novoa et al., 2017)
Type II	Chl*a*	MDN (Pahlevan et al., 2020)	GI-2B* (Gilerson et al., 2010)	Gons (Gons et al., 2002)	GU-2B* (Gurlin et al., 2011)	Blend (Smith et al., 2018)		NA	

* 2B refers to 2-band red-NIR family of algorithms.

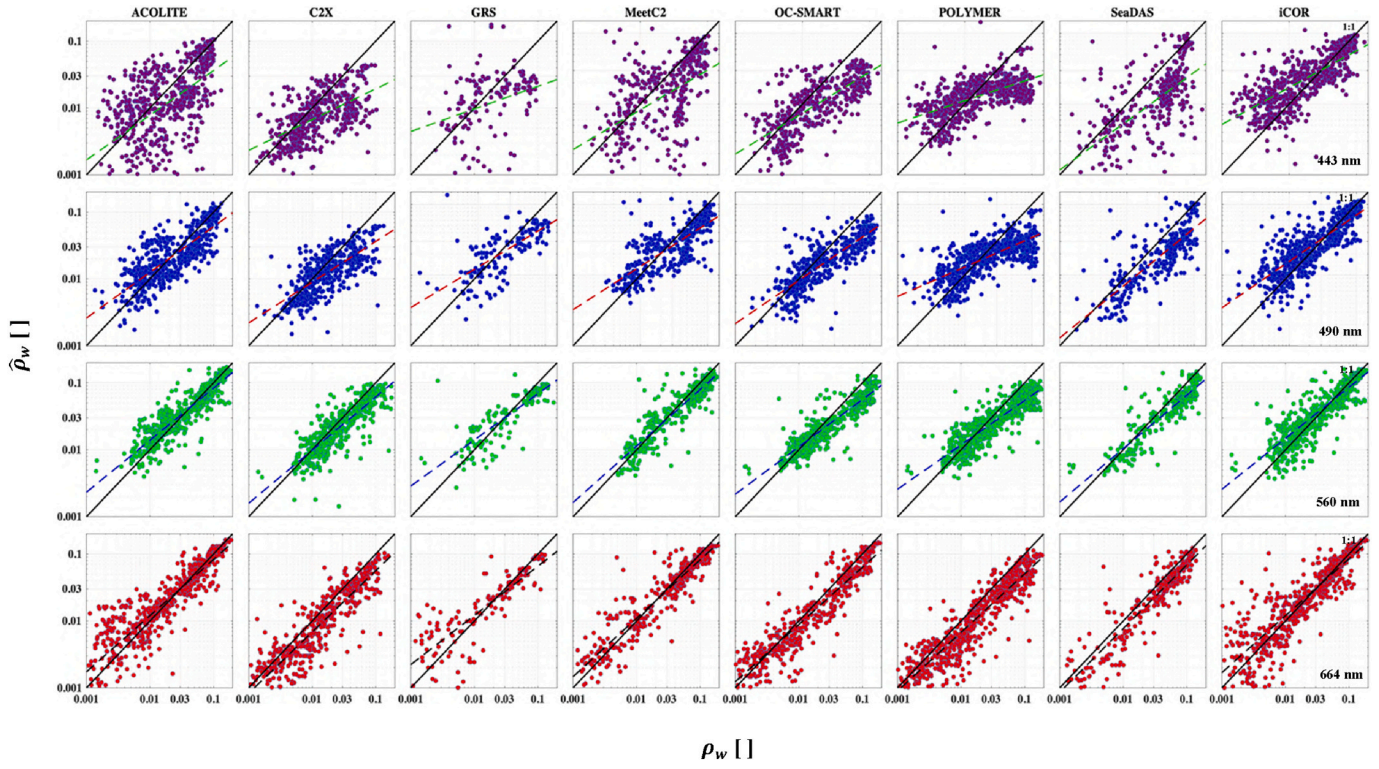

Fig. 5. Overall performance of AC processors using the CVD matchups with all the OLI and MSI data combined. Scatterplots are shown in log-log scale and the number of matchups per processor is reported in Table 3. Detailed statistical metrics are available in Table S4. The solid black lines refer to the 1:1 line.

$$\beta = 100 \times sign(z)\left(10^{|Z|} - 1\right)\ [\%] \text{ where } Z = Median\left(log_{10}\left(\widehat{\rho}_w(\lambda_i)\Big/\rho_w(\lambda_i)\right)\right) \quad (5)$$

$$\epsilon = 100 \times \left(10^{Y} - 1\right)\ [\%] \text{ where } Y = Median\left|log_{10}\left(\widehat{\rho}_w(\lambda_i)\Big/\rho_w(\lambda_i)\right)\right| \quad (6)$$

where λ_i refers to spectral band i, Median is the median operator, β is the symmetric signed percentage bias and ϵ represents the median symmetric accuracy (MdSA; Morley et al. (2018)). These metrics, expressed in %, are simple for interpretation, reasonably resistant to outliers, and zero-centered compared to those in Seegers et al. (2018). In the definition of the above metrics, we assume no uncertainties in *insitu* measurements, because to a large extent these uncertainties affect all the AC processors equally and were unknown at the time of this research (see Section 5.1). The quantities $\widehat{\rho}_w$ and ρ_w may be replaced by other quantities, such as Chla^r (or TSS^r) and Chl*a* (or TSS), respectively. Where Chla^p and TSS^p are treated as reference products, the estimated error and bias metrics are denoted $\widehat{\epsilon}$ and $\widehat{\beta}$ (Section 4.2.).

We also computed the Root Mean Squared Log Error (RMSLE), Root Mean Squared Error (RMSE), Median Absolute Percentage error (MAPE), slope of linear regression in log-space (S), and Mean Symmetric Accuracy (MSA) computed analogous to Eq. 6, except the average of the log ratios was employed to infer the overall impact of noisy retrievals (Seegers et al., 2018). Note that RMSE and MAPE are computed to merely provide traceability to previous studies (Pahlevan et al., 2017b; Pahlevan et al., 2017c; Steinmetz and Ramon, 2018; Warren et al., 2019) and is not anticipated to serve as a reliable measure given the log-normal distribution of our matchup datasets (Fig. 3).

In Section 4.1.2, where pairwise intercomparisons are presented, win rates (Seegers et al., 2018) were calculated per OWT and band for each AC pair. A winning processor for each pair, the one with lower ϵ, was assigned with unity. As a result, for each AC pair, a 4 (band) by 7 (OWT)-element binary array (Fig. 9) filled with 0 s or 1 s was formed. This pairwise intercomparison was repeated for all the processors to generate seven binary arrays, which were added and normalized by the total number of pairwise comparisons (N=7). Having these normalized heatmaps created for all processors, the AC processors likely to produce the most accurate $\widehat{\rho}_w$ for any given band and OWT were determined. The description of the metric used for the universally common matchups is provided in Appendix B.

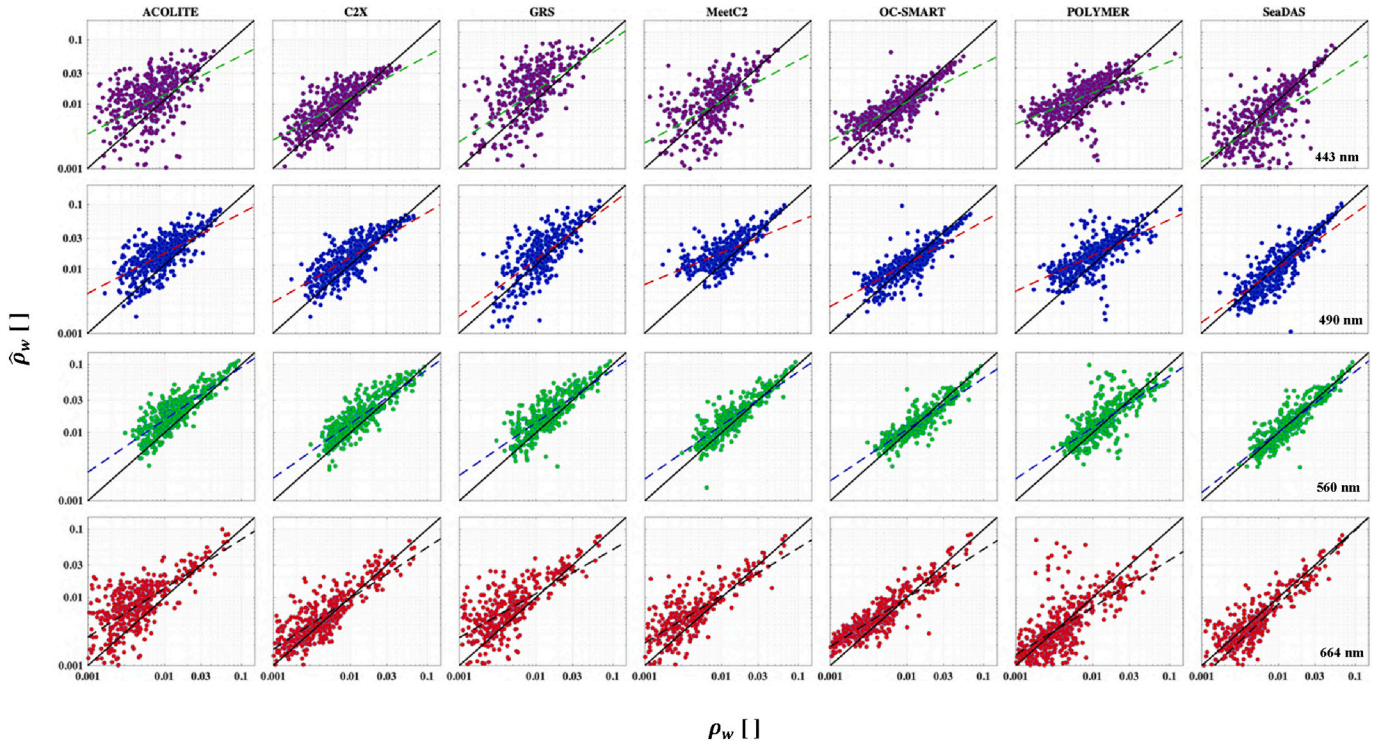

Fig. 6. Same as Fig. 5, but for the AERONET-OC matchups. Statistical descriptors are provided in Table S5. The solid black lines refer to the 1:1 line.

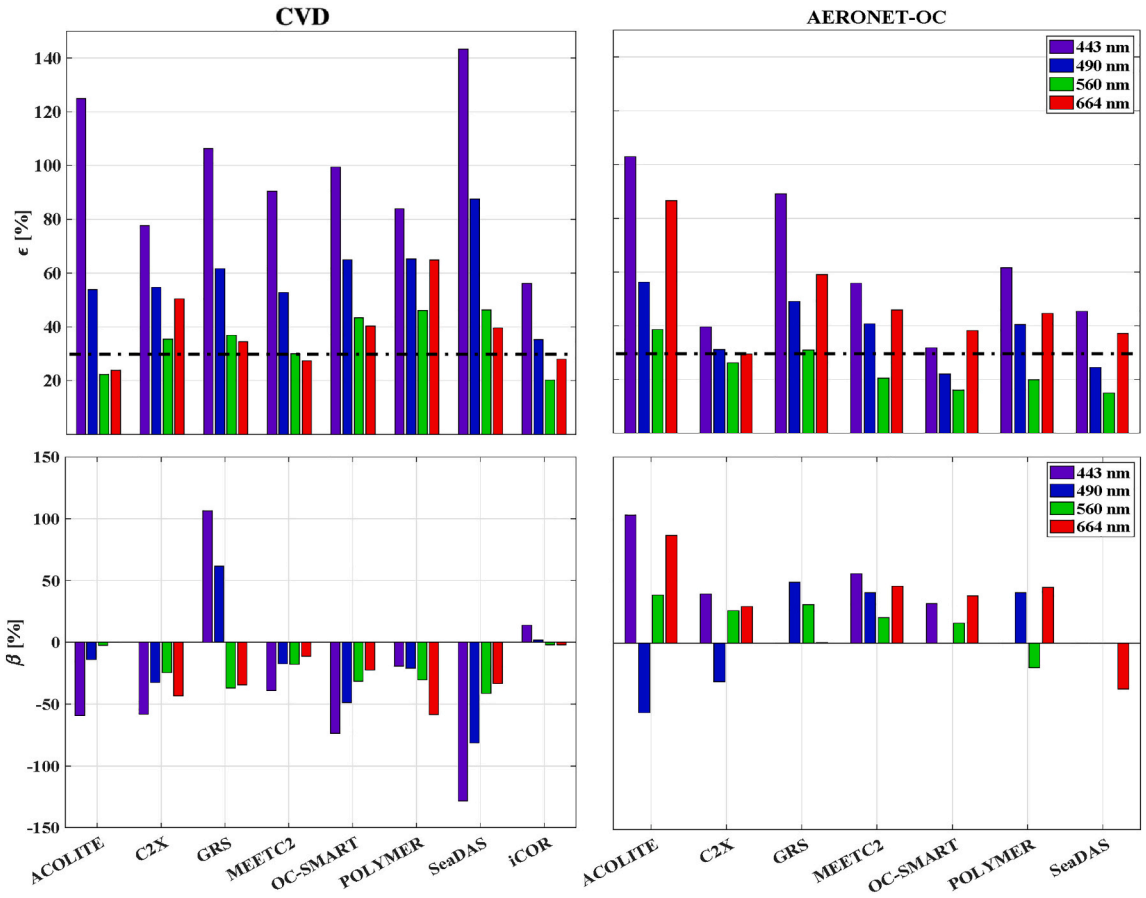

Fig. 7. Performance assessments as determined by the median symmetric accuracy (ϵ; Eq. 6) and median symmetric bias (β; Eq. 5) for OLI and MSI matchups combined. The dashed lines in the top row correspond to a 30% threshold (GCOS).

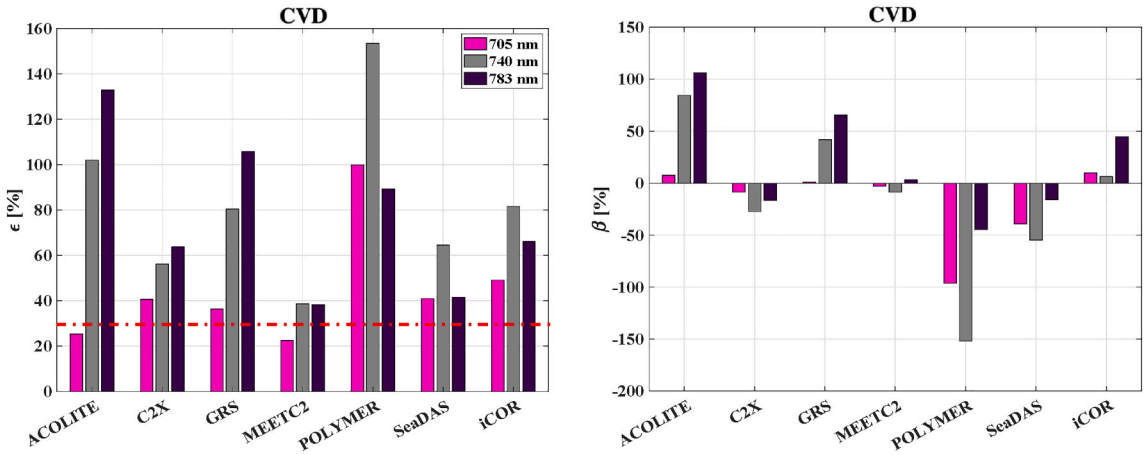

Fig. 8. Similar to Fig. 7, but computed only for the MSI's NIR bands. See Table 3 for the number of matchups. OC-SMART is excluded as it does not retrieve in the NIR bands. The dashed line on the left panel corresponds to a 30% threshold (GCOS). Statistical metrics are provided in Table S4.

4. Results

4.1. Aquatic reflectance products

The performance analyses are described in two main subsections: a) an assessment of *all* the valid matchups for each processor (Section 4.1.1) and b) an evaluation of *common* matchups between processors (Section 4.1.2). The latter enables a fair performance intercomparison using identical matchups while the former evaluates each individual processor independently, with an adequate number of matchups, permitting a full assessment of their practicality.

4.1.1. All matchups: Individual performance

To provide a straightforward and qualitative assessment of individual performance, the scatterplots for the CVD and AERONET-OC matchups for OLI and MSI combined are illustrated in Figs. 5 and 6. The full statistical metrics (ϵ, β, MSA, MAPE, RMSE, RMSLE, and S) are reported in Tables S4 and S5. Examining these two figures points to two primary observations. First, each processor performs differently when evaluated against CVD and AERONET-OC matchups, with some performing better in

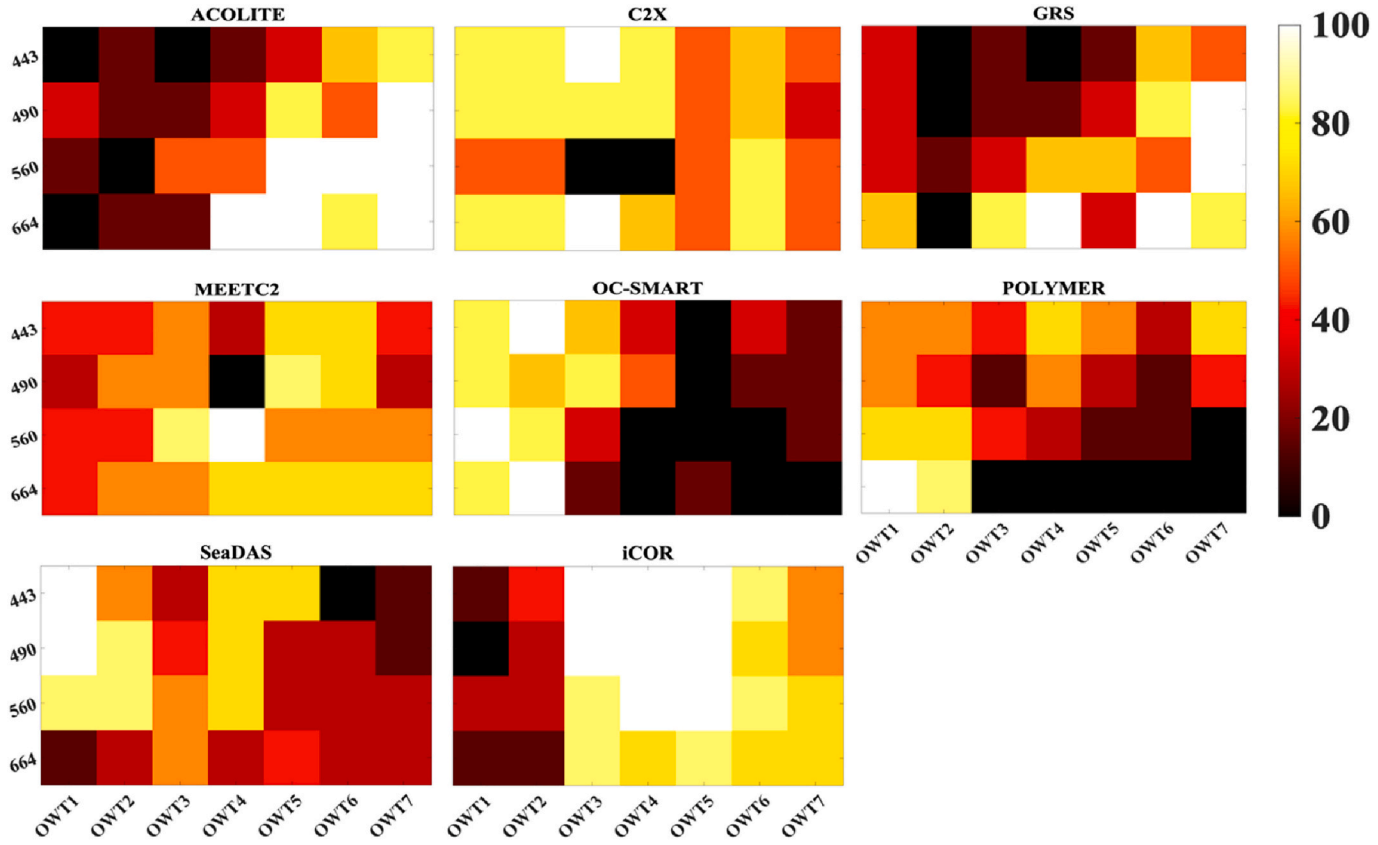

Fig. 9. Relative performance assessments [%] determined via aggregating pairwise intercomparisons (Section 3.6). Processors with brighter colors (white or yellow) are likely to generate high-quality $\widehat{\rho}_w$ for a given OWT and band. CVD and AERONET-OC matchups were combined to carry out this analysis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

locations represented in the CVD and some showing promise in AERONET-OC regions. Second, the poorest performance consistently corresponds to the blue bands, particularly when CVD matchups are examined. This likely suggests the limited performance of processors to remove aerosol contributions (Pahlevan et al., 2017a). Despite differences in the number of matchups, similarities and discrepancies among the processors can further be highlighted. For instance, SeaDAS and OC-SMART appear to exhibit similar distributions in the 443 and 490 nm bands when assessed using AERONET-OC matchups. For CVD matchups in these bands, POLYMER-derived $\widehat{\rho}_w$ tend to saturate for $\widehat{\rho}_w > 0.03$ whereas estimates from iCOR appear to be fairly reliable in this range. Interestingly, for the AERONET-OC matchups, the performance in $\widehat{\rho}_w(443)$ degrades for values $< \sim 0.02$ where the data is widely dispersed. This signal-dependent performance is relatively less pronounced in the green and red bands, particularly for the CVD matchups.

A summary of processor performances in terms of retrieval errors (ϵ) and biases (β) in the visible bands is illustrated in Fig. 7. Overall, ϵ further corroborates the limited performance in the blue bands and better performance at AERONET-OC locations compared to those represented in CVD. It is also noticeable that none of the processors appear to uniformly meet the 30% retrieval accuracy requirements (GCOS) across all four bands. The performance is superior in the 560 and 664 nm bands for the CVD matchups and in the 560 nm band for the AERONET-OC matchups. This may partially be attributed to higher signal levels typically found in these spectral regions for the majority of our data. On the other hand, β shows that nearly all the processors underestimate $\widehat{\rho}_w$ for the CVD matchups with iCOR and ACOLITE returning minimal biases in the 490, 560, and 664 nm bands. This is opposite to the performance inferred from the AERONET-OC matchups, with SeaDAS exhibiting near-zero biases in the 443, 490, and 560 nm bands. In addition, band-dependent performance (improving or degrading from the blue to red) is noted for nearly all the processors when considering CVD matchups. The performance analyses of individual processors for OLI and MSI visible bands are included in Appendix A. The assessment gives insights into the extent to which the processors are dependent on the radiometric performance of the two instruments (Pahlevan et al., 2019). The errors and biases in the MSI NIR bands determined via the CVD matchups are provided in Fig. 8. The larger uncertainties are typically found in the 740 or 783 nm bands, and no consistent positive or negative biases are evident across the processors. Assuming a 30% threshold requirement for $\widehat{\rho}_w(700 < \lambda < 800\,nm)$, retrievals from ACOLITE and MEETC2 appear to yield valid retrievals for the 705 nm band, although only 78 valid matchups (Table 3) were identified for MEETC2. Further, ϵ tends typically to rise from the 705 to 783 nm band for ACOLITE, C2X, and GRS while the 740 nm band carries larger uncertainties when retrieved from the other processors.

Table 7
Errors in estimated Chl*a* from MSI matchups for select processors using CVD matchups. Least sensitive (Type II) retrieval algorithms per processor are boldfaced. The 2B models yield four invalid retrievals.

			MDN	Blend	GI-2B	Gons	GU-2B
			$N = 19$			$N = 15$	
δChla^p	ρ_w	ϵ	**15.2**	27.9	41.3	38.9	41.3
		β	**−5.6**	−12.8	−3.5	−27.5	2.4
δChla^r	ACOLITE	ϵ	40.2	61.9	46.4	56.1	**33.8**
		β	−3.4	−10.7	−0.28	−5.3	**10.4**
	C2X	ϵ	34.7	34.4	33.2	**25.3**	39.1
		β	3.92	−8.1	5.5	**−0.8**	23.5
	POLYMER	ϵ	117.4	49.8	**25.6**	41.5	55.5
		β	−45.6	−8.4	**−1.4**	−6.5	10.3
	iCOR	ϵ	79.9	48.3	**26.6**	35.7	37.2
		β	2.1	−26.1	**−0.38**	−24.5	11.5

4.1.2. Common matchups: Performance intercomparison

The pairwise intercomparison strategy, on average, resulted in 60 samples for OWTs 1 and 4, 175 for OWTs 2 and 3, 85 for OWT5, and 20 for OWTs 6 and 7, a total of ~600 matchups. The number of matchups was higher for some AC pairs, such as iCOR and POLYMER, while SeaDAS and GRS intercomparisons led to smaller (but adequate) samples (e.g., 150 for OWT3). The heatmaps in Fig. 9 illustrate per band, per OWT performances estimated via win rates (Section 3.6). A paler color in each cell suggests an AC processor with higher chance of producing reliable $\widehat{\rho}_w$ products. For OWTs 1 and 2, OC-SMART and SeaDAS perform best with OC-SMART exhibiting a more consistent performance across the four bands. Among the processors, iCOR outperforms the rest of the processors in OWTs 3, 4, and 5, though it was assessed using CVD matchups only. The performance is superior in the blue and green bands in these OWTs. C2X is also found to perform well in OWT3 in the blue and red bands. For OWT6, iCOR together with ACOLITE, C2X, GRS, and MEETC2 appear to perform better than the rest of the schemes. ACOLITE is the obvious choice in OWT7 for which a superior performance is achieved in the 490, 560, and 664 nm bands. For this water type, GRS seems to offer a reasonable performance in particular in the 490 and 560 nm bands. It should also be noted that some of the processors are found to perform consistently across three or all four bands, a signature that underscores a processor's capability to preserve the spectral shape of ρ_w. This applies to, for example, iCOR (OWTs 3, 4, 5), SeaDAS (OWTs 1 and 4), MEETC2 (OWTs 5 and 6), or C2X (OWTs 1, 2, and 3) - see Section 5.2 for further discussion. The secondary approach to our OWT analysis conducted using universally common matchups is provided in Appendix B.

4.2. Downstream products

4.2.1. Community validation database (CVD)

For the common matchup analysis, we focus only on the AC

Table 6
Errors in Chla^r and TSS^r estimated through Type I retrieval algorithms using common CVD matchups (Fig. 2; $N = 123$). Performance of each retrieval algorithm (δChla^p and δTSS^p) is also provided in the 2nd and 3rd columns. Note that the Novoa model is excluded here as it requires $\boldsymbol{\rho_w(865)}$, which was not assessed in this exercise. Best-performing retrieval algorithms per processor are boldfaced.

	δChla^p		δChla^r							
	ρ_w		ACOLITE		C2X		POLYMER		iCOR	
Algorithm	MDN	OCx	MDN	OCx	MDN	OCx	MDN	OCx	MDN	OCx
ϵ [%]	17.7	61.2	77.1	142	82.7	**67.4**	83.6	102.7	89.5	151.2
β [%]	0.35	−13.2	30.1	21.3	−15.2	**10.9**	−37.4	−71.0	74.2	−89.4
S [ ]	0.67	0.47	0.81	0.45	0.75	0.58	0.67	0.52	**0.86**	0.45

	δTSS^p		δTSS^r							
Algorithm	Nechad	SOLID	Nechad	SOLID	Nechad	SOLID	Nechad	SOLID	Nechad	SOLID
ϵ [%]	69.8	58.9	**58.1**	95.1	**58.5**	85.5	71.2	107.8	62.6	72.9
β [%]	57.6	45.1	37.7	78.2	**6.6**	−20.3	9.6	−52.1	35.1	−2.1
S [ ]	0.31	0.75	0.30	0.59	0.41	**0.89**	0.31	0.67	0.40	0.85

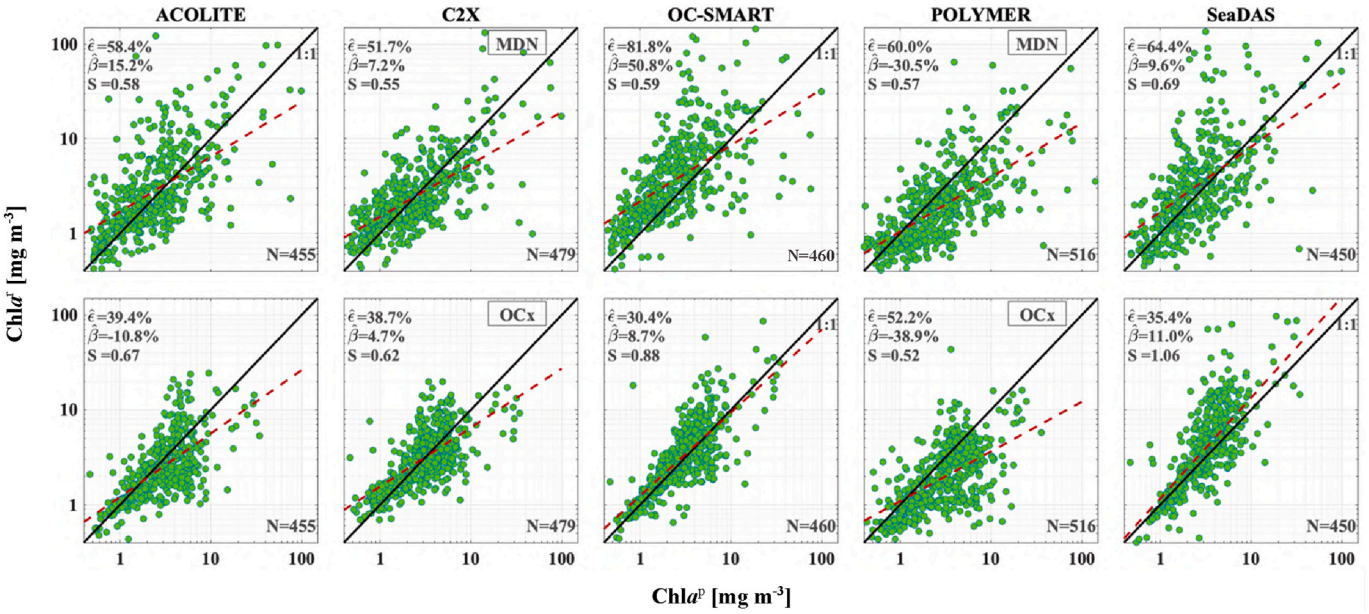

Fig. 10. Sensitivity of Chl*a* retrieval algorithms (Type I) to uncertainties in $\hat{\rho}_w$ obtained from five AC processors. Chla^p were derived from *in situ* AERONET-OC radiometric matchups using MDNs (top) (Smith et al., 2021) and OCx (bottom) (O'Reilly and Werdell, 2019). Chla^r were derived using the same algorithms. This analysis is not intended for comparing the absolute performance of MDN or OCx. Note that combined OLI and MSI's visible bands are applied here and that the axes are truncated to 120 mg m^{-3} for enhanced visualizations. Number of valid matchups for each processor has been annotated.

processors that returned adequate valid matchups to permit statistically robust inference (N=123). The results pertaining to Type I algorithms (Section 3.5) are included in Table 6. The errors (ϵ) in Chla^r and TSS^r (e.g., δChla^r) as well as in Chla^p and TSS^p (e.g., δChla^p) are presented to allow for comparisons between the theoritical limits and practical performance of the retrieval algorithms.

Considering the MDN retrieval algorithm, δChla^r are, on average, four times larger than δChla^p. Despite its superior performance in predicting Chla^p (Smith et al., 2021), MDN does not always outperform OCx for a given $\hat{\rho}_w$, suggesting its higher sensitivity to uncertainties in AC. Overall, C2X appears to yield the lowest error (67.4%) and bias (10.9%).

Fig. C1 further illustrates Chla^r – Chla^p plots, allowing to gauge the sensitivity of Type I algorithms to uncertainties in $\hat{\rho}_w$ across a wider concentration range for all the processors. For TSS retrievals, C2X, ACOLITE, and iCOR are found to return more accurate products with ϵ within the 58–63% range, although the biases are appreciably high for ACOLITE and iCOR. The Nechad model outperforms SOLID when applied to $\hat{\rho}_w$, implying that SOLID is more sensitive to uncertainties in AC. Despite this practical limitation, the iCOR-SOLID combination allows reasonable TSS estimates, which may be attributed to iCOR's ability to better preserve the shape of ρ_w as discussed in Section 4.1.2 (Fig. 9). Moreover, as shown in Balasubramanian et al. (2020), the

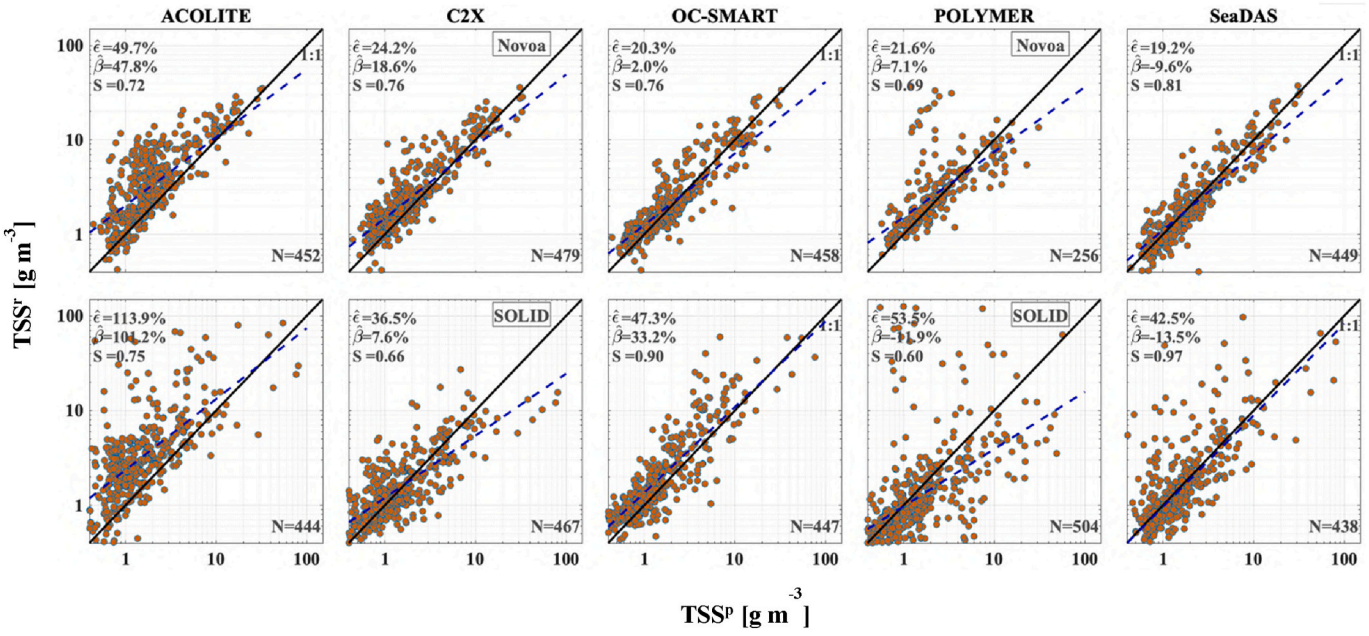

Fig. 11. Same as Fig. 10, but for TSS. TSS^p and TSS^r were derived from *in situ* AERONET-OC and satellite derived spectra via Novoa (top) and SOLID (bottom) models (Balasubramanian et al., 2020; Novoa et al., 2017).

performance of the Nechad model is limited in clear to moderately turbid waters where TSS < 4 g m^{-3} (Fig. 2), which is why the slopes are much lower than unity for all the processors. The sensitivity of the Nechad and SOLID models to uncertainties in $\widehat{\rho}_w$ for all the processors is provided in Fig. C2.

We further investigated the performances of the five best-performing Chl*a* algorithms (Table 5) using a small subset of universally common MSI matchups (N ~ 19). From the results provided in Table 7, one may infer that although $\widehat{\rho}_w(705)$ and $\widehat{\rho}_w(743)$ carry relatively large uncertainties (Fig. 8), the critical, but perturbed, spectral information within this wavelength range still results in improvements in Chla^r when compared to that shown in Table 6. Further, we note that the propagation of uncertainities works in favor of the family of 2B algorithms such that the estimated errors in Chla^r are less than that in Chla^p. For instance, when applying Gons, Chla^r from C2X shows ~20% error whereas Chla^p contains ~39% retrieval errors. However, the 2B models (GU-2B and GI-2B in Table 5) are mostly suited to eutrophic waters (e.g., Chl$a > 8$ mg m^{-3}) and are typically insensitive to the Chl*a* ranges found in oligotrophic lakes or coastal waters. This is in contrast to MDN and Blend which provide valid retrievals across a broader range. An extensive analysis of the sensitivity of Type II algorithms to uncertainties in $\widehat{\rho}_w$ computed through all the CVD matchups is illustrated in Fig. C3.

4.2.2. AERONET-OC

For this sensitivity exercise, we present apparent error and bias metrics ($\widehat{\epsilon}$ and $\widehat{\beta}$), as well as slopes of linear regressions (S) in Fig. 10. For an ideal processor that introduces no uncertainties in $\widehat{\rho}_w$, the data distributions are expected to align with the 1:1 line. The sensitivity to uncertainties in $\widehat{\rho}_w$ across AERONET-OC sites (with mean Chl*a* value of 2 to 3 mg m^{-3}) varies from 30 to 82%. In general, OCx derived Chla^r estimated through SeaDAS and OC-SMART appear to adequately reproduce Chla^p suggesting that the blue-green band ratios are preserved fairly well for these two processors. When applying MDN, OC-SMART and POLYMER are shown to contain major biases, which have implications in uncertainties introduced in $\widehat{\rho}_w$ across all the visible bands used for MDN estimations. For POLYMER, this also applies to retrievals from the OCx algorithm. We caution readers that this assessment merely points to the sensitivity of the AC processors and should not be interpreted as a verification for the choice of a Chl*a* algorithm.

To examine TSS retrievals (with a mean TSS^p of 2.6 g m^{-3}), we applied the Novoa and SOLID models (Table 5) since Nechad was found to perform poorly for this range of TSS (>80% with TSS < 4 g m^{-3}) (Balasubramanian et al., 2020). From Fig. 11, one may infer that the Novoa model is less perturbed by the uncertainties introduced via the AC processors, with SeaDAS and OC-SMART introducing the least amount of noise and bias to $\widehat{\rho}_w$. Similar to the Chl*a* analysis, when applying SOLID, OC-SMART tends to add major biases to TSS^r. ACOLITE is shown to yield noisy and biased TSS^r for both algorithms, while POLYMER appears to introduce noise in SOLID estimates relative to that of Novoa.

5. Discussion

5.1. In situ radiometry

In situ radiometric measurements acquired by various instruments, field setups and processing procedures added uncertainty to our matchup statistics which remains uncharacterized in this analysis. For example, recent studies have shown that above-water radiometric measurements (Method 7 in Table S1) may carry uncertainties ranging from 5 to 50% across the visible bands depending on environmental conditions (Alikas et al., 2020). The same measurement technique with well-calibrated instruments was demonstrated to differ from AERONET-OC measurements by up to 10% (Tilstone et al., 2020). Uncertainties associated with other measurement techniques largely vary according to instrument calibration, environmental conditions, measurement protocols, and data processing schemes. Challenges of in-water radiometric measurements may be particularly acute in highly turbid waters or during intense phytoplankton blooms particularly for spectral bands with low penetration depths. In general, compiling a large pool of globally representative matchups, re-used from previous calibration/validation research (Table S1), was a prerequisite for this effort and thus necessitated acceptance of some underlying uncertainty. We believe that such uncertainties did not alter the comparative assessment of the AC processors and the downstream products for three main reasons. First, use of the median symmetric accuracy metric (ϵ) reduced the impact of any highly uncertain *in situ* measurements. We carried out a few experiments by varying the matchup filtering criteria (Eqs. (3) & (4)) to evaluate the effect of the number of matchups on the metric. This analysis found no major changes in the apparent performances (e.g., Fig. 7). In contrast, we found the bias metric (β) to be sensitive to changes in the number of matchups, suggesting its sensitivity to outliers, although different matchup samples did not affect whether a processor was generally under- or over-estimating (i.e., bias signs were unaffected). Thus, the majority of the analysis presented was based upon the error metric (ϵ). Second, our measurements were made by hyperspectral radiometers built and/or assembled by different manufacturers, with varying calibration histories and measurement approaches. Creating a large pool of data using a combination of various methods likely minimizes any systematic errors in our performance assessments. Finally, the results from the pairwise common matchup datasets with the CVD and AERONET-OC matchups aggregated seem to converge with those from the individual performance assessments in Section 4.1.1, supporting negligible effects due to potential uncertainties within the CVD.

Regardless, more coordinated data acquisition approaches, such as those offered by AERONET-OC (Zibordi et al., 2009a) or the WATER-HYPERNET (Vansteenwegen et al., 2019) currently under development) will be crucial for better quantifying uncertainties and their sources in space and time. The recently established WISPstation network (Bresciani et al., 2020) is another example of a system with a major contribution to this exercise. Although the WISPstation instrumentation may not meet the highest level of standards required for ocean color validation practices (uncertainties <5%) (Zibordi et al., 2009b), the availability of these measurements extended our valid matchup datasets by ~10%. As automated networks are currently expanding, a harmonization among field radiometric observations will become important to facilitate the evolution of globally applicable AC processors and achieve improved retrieval accuracies (<<30%). Research cruises and field campaigns targeting wide ranges of coastal and inland water types to collect data for calibration/validation will remain essential to increase the environmental diversity of the CVD, and we strongly advocate for their continued support by space or operational agencies. The availability of a web-based platform like the Coastal Thematic Exploitation Platform (CTEP) was instrumental in expediting the coordination, processing, storage, and distribution of the products for matchup assessments. However, for the CVD matchups, a significant amount of time and effort had to be invested to assemble and re-format the *in situ* radiometric data. Hence, a community curated protocol for data preparation, formatting and ingestion for such analyses would greatly facilitate future AC assessments (IOCCG, 2019). Examples of databases containing *in situ* radiometric data for inland and coastal waters already exist (e.g., the SeaWiFS Bio-optical Archive and Storage System; SeaBASS, Lake Bio-optical Measurements and Matchup Data for Remote Sensing data; LIMNADES), but need to expand. In addition, harmonizing data input from various sensors, including low-cost sensors, into central data systems is currently a subject under development in the MONOCLE project.

5.2. Matchup assessments

The performance assessments of AC processors determined using all available matchups, in general, agree well with those from the stricter common-matchup analysis (Sections 4.1.2). For example, both analyses

suggest that iCOR and ACOLITE outperform other schemes for the diverse, mostly freshwater observations contained within the CVD or in turbid and/or eutrophic ecosystems represented by OWTs 3 through 7. In clearer waters (OWTs 1 and 2), OC-SMART and SeaDAS are the best performers. These two processors, along with C2X, also performed well using only AERONET-OC matchups (Fig. B2), confirming their utility for coastal waters. The iCOR processor performed well in OWTs 3, 4, and 5, whereas ACOLITE performed poorly in coastal waters represented in AERONET-OC matchups. An examination of Fig. 9 and Fig. B1 points to POLYMER and MEETC2's potential for producing high-quality $\widehat{\rho}_w$ in OWTs 1 and 2, but their overall performance may not exceed those of SeaDAS and OC-SMART. Moreover, Figs. 5 and 7 suggest that iCOR produces better quality $\widehat{\rho}_w$ in the 443 nm band than that of other processors, an observation congruous with the analyses offered in Fig. 9 and Fig. B1.

There is also agreement between our assessments of the reflectance and the downstream water-quality products (Sections 4.1 and 4.2). For example, C2X returns $\widehat{\rho}_w(664)$ adequately (Figs. 6 and 7) which results in better quality TSS estimates using the Nechad algorithm (Table 6) in water types represented by the AERONET-OC matchups. The Novoa model was also demonstrated to exhibit slight sensitivities to uncertainties in $\widehat{\rho}_w$ generated via C2X (Fig. 11). It may therefore be advantageous to seek a fit-for-purpose AC that estimates satisfactory $\widehat{\rho}_w$ in only one or two individual bands for specific applications (e.g., TSS estimation using a single-band algorithm). ACOLITE, for instance, performs noticeably well in the 664 nm band for OWTs 4,5, 6, and 7 (Fig. 9); or the blue-green Chl*a* algorithm shows a minimal sensitivity to uncertainties in $\widehat{\rho}_w$ derived from SeaDAS in OWTs 1, 2, and 4 (Figs. 6, 7, 9, and 11). Using the CVD matchups and two-band (2B) retrieval algorithms (Table 7), we also demonstrated that C2X, POLYMER, and iCOR yield Chl*a* with accuracies ranging from 25 to 30%, which potentially support their utility in eutrophic ecosystems. These improvements likely stem from counteracting effects of biases in $\widehat{\rho}_w(705)$ and $\widehat{\rho}_w(664)$ (Fig. 7). In contrast, Chl*a* retrievals using these algorithms with ACOLITE performed poorly, indicating that the $\widehat{\rho}_w(705)/\widehat{\rho}_w(664)$ value is adversely affected by this processor.

Overall, the processors like iCOR and ACOLITE that often outperform other schemes for regions represented in CVD achieve 20 to 30% errors in the green and red bands and return Chla^r and TSS^r estimates with 25–70% uncertainties (Tables 6 and 7). Top-performing processors in coastal waters such as OC-SMART, SeaDAS, and C2X yield errors within the range of 15 to 30% in these spectral bands but the uncertainties in Chla^r and TSS^r are expected to remain between 50 and 70% given the assessments provided in Table 6 and Figs. 10 and 11. For these processors, the larger errors in Chla^r stem from higher sensitivity to errors in $\widehat{\rho}_w(560)$ and $\widehat{\rho}_w(664)$, the spectral bands that contain the most relevant information for Chl*a* estimation in OWTs 1, 2, and 3. In the blue bands, the errors tend to be larger (25–60%) for best-performing schemes, limiting their utility in scientific studies (see next section). It is also worth noting that the larger sensitivity of MDN and SOLID retrieval algorithms to uncertainties in $\widehat{\rho}_w$ are attributable to the fact that these models employ all the available spectral bands instead of a band ratio or a single band. For instance, SOLID uses spectral information in the visible bands to estimate spectral particulate backscattering. This observation likely favors retrieval algorithms that exhibit minimal sensitivities to uncertainties in AC even though they may not work adequately on high-quality ρ_w.

5.3. Potential pathways towards enhanced performances

We find that the limited success (Fig. 7) in the removal of the contributing aerosol signal remains a major source of uncertainty in $\widehat{\rho}_w$. Regardless of their approach, the AC processors utilize aerosol models that may not accurately represent aerosol types over land and coastal waters. The large inaccuracies (25–70%) in Chla^r and TSS^r, even from processors that generally meet the 30% threshold requirements for $\widehat{\rho}_w(560)$ and $\widehat{\rho}_w(664)$ (GCOS), further corroborate the demand for more precise $\widehat{\rho}_w$ across all the visible and NIR bands. Should improved downstream products ($\sim$ 10% uncertainties) be desired, the requirements on ϵ must be lowered to <10% (Cetinic et al., 2019). We caution readers that the largest uncertainties in $\widehat{\rho}_w$ were commonly found for the 443 nm band whose spectral signature makes little contribution to Chl*a* and TSS predictions in most water types examined in this study (Smith et al., 2021). This situation applies to MDN and SOLID, but other algorithms considered here do not utilize information in this spectral band. Therefore, a thorough impact assessment of downstream products should encompass a wider array of products (e.g., a_{cdom}). For instance, given the large uncertainties in the blue bands, retrieving variables like the spectral diffuse attenuation coefficients (K_d) necessary for semi-analytical estimates of Secchi-disk depth (Lee et al., 2016), may be extremely challenging in most inland and coastal waters given the large uncertainties in the blue bands. Another potential reason for the limited performances in the 443 and 490 nm bands is likely the lack of rigor in handling the sky-glint or observations in areas with enhanced backscatter from haze in portions of OLI and MSI images east of nadir. A more robust approach in accounting for the residual signal in the blue may be achieved through the explicit inclusion of sun-sensor geometry, multiple-scattering near the water surface (Gilerson et al., 2018), and cloud fraction information.

On average, the opposite signs in biases in the CVD and AERONET-OC matchup assessments (Fig. 7) for nearly all the processors appear to suggest under-correction and overcorrection in inland and coastal areas, respectively. This reversal pattern rules out significant cross-sensor biases between OLI and/or MSI ρ_t products. The OLI- and MSI-specific performance analyses (Appendix A) further underscore the sensitivity of each processor to instruments-specific radiometric performance (biases and signal-to-noise-ratio; SNR). For instance, SeaDAS is found to be sensitive to the low SNRs in the MSI 865 and 1610 nm bands employed in approximating aerosol contributions (Pahlevan et al., 2019; Pahlevan et al., 2017b). The discrepancies found for other processors' performances require further investigations but, in general, a processor capable of producing harmonized $\widehat{\rho}_w$ products from OLI and MSI is anticipated to produce similar ϵ, which is minimally sensitive to the number of matchups. Note that this assessment could not be made on an OWT-specific basis (Fig. 9) due to insufficient matchups per OWT for each sensor. Another research area is identifying the source of uncertainties in the 740 and 783 nm bands, which also demand more matchup datasets collected with well-calibrated instruments.

AE continues to be a confounding phenomenon encountered in inland and nearshore coastal waters. It varies as a function of aerosol type, aerosol height, landcover type and its seasonal variability, topography, and other environmental conditions. Hence, geomorphological factors (extent, width, shape) of an ecosystem are not the only determinants for the presence or absence of AE. Interestingly, iCOR is the only processor considered in this study attempting to correct for AE and, overall, the best-performing algorithm across OWTs 3 through 6. Further, in this exercise, all the AC processors frequently failed to retrieve realistic $\widehat{\rho}_w$ in small Canadian lakes (e.g., Professor's Lake, a 0.26 km^2 reservoir located in Brampton, Ontario; Huot et al., 2019), whereas acceptable performance was observed in other fairly small ecosystems, such as Elistvere Järv (a 1.29 km^2 lake in Tartu County, Estonia) or tributaries of the Paraná River (Brazil-Paraguay border). Very low water-leaving radiance across the whole visible-NIR range in Canadian lakes dominated by colored dissolved organic matter, combined with a surrounding environment of dense vegetation (e.g., boreal forest) with high NIR reflectance may explain this issue. Focused regional performance assessments will aid in addressing such phenomena. Future global intercomparison exercises may additionally consider relevant environmental, physical, topographical attributes, and their spatiotemporal variability, to improve characterization and correction of AE. For the AERONET-OC sites, we further analyzed the retrieval

Table 8
Ranking of AC processors specified using band-average (column-wise) performances in Fig. 9. Note that iCOR was only analyzed using CVD matchups.

Order	OWT1	OWT2	OWT3	OWT4	OWT5	OWT6	OWT7
1	OC-SMART	OC-SMART	iCOR	iCOR	iCOR	iCOR	ACOLITE
2	SeaDAS	SeaDAS	C2X	SeaDAS	ACOLITE	MEETC2	GRS
3	C2X	POLYMER	MEETC2	C2X	MEETC2	ACOLITE	iCOR
4	POLYMER	C2X	OC-SMART	OC-SMART	SeaDAS	C2X	MEETC2

accuracies with respect to AOT, wind direction, water vapor, and imaging geometry and found very weak or no correlations with ϵ and/or β. With expected improvements in the performance of the AC schemes, higher degrees of correlations with environmental and physical variables are anticipated. Further, dependencies on aerosol types and AOTs may be analyzed using observations at nearby AERONET sites.

Since both C2X and OC-SMART exhibited promising performances, we encourage continued research into novel machine-learning modeling solutions which could capture some of the more subtle influences on retrieval performance mentioned above. These may include improved forward radiative transfer modeling, enhanced representation of aerosols and in-water optical properties, leveraging latest architectures (e.g., activation functions) or models that are adept in identifying better solutions to the inverse problem. Similarly, the *prior* and posterior probabilities learned and applied by MEETC2 also showed promising results for retrieving $\widehat{\rho}_w(700 < \lambda < 800\,nm)$; hence, future investments in applying Bayesian approaches would be worthwhile (Frouin et al., 2019; Thompson et al., 2019).

6. Recommendations

This international, collaborative effort is one of the first major community-wide research activities to address a critical component of satellite data processing for monitoring and scientific studies of aquatic biogeochemical cycling and water quality at a global scale. Among the objectives of ACIX-Aqua was to make recommendations to the user community, satellite practitioners, water resource managers, as well as space and operational agencies for the implementation of an AC approach suitable for their applications. The present work demonstrates that research in this area is highly topical and needs to continue to achieve a reduction of uncertainties in $\widehat{\rho}_w$ and consequently in downstream quantities, such as the inherent optical properties and water constituents (e.g., Chl*a* and TSS). The problem addressed is highly multi-dimensional, and we discussed uncertainity metrics specific to AC processors, specific to spectral bands of satellite instruments, specific to OWTs and specific to two commonly estimated downstream products. In order to simplify these results and to provide a guideline for end-users to consider when processing or re-processing OLI and MSI images for coastal and inland aquatic applications, we offer the following recommendations and considerations.

- For global studies of inland and coastal waters, there is *no single* solution, and a preferred AC processor may be chosen according to the specific scientific objective and application. To facilitate this choice, we provide a ranking of AC processors per OWT based on the relative pairwise performance presented in Section 4.2.1 in Table 8. We note that Table 8 is a simplification and higher granularity in the decision process may be achieved by consulting our results with respect to the spectral bands of interest, the desired downstream products, coastal versus inland water applications, or specific satellite sensor (Appendix A). For example, in nearshore coastal waters (OWTs 1, 2, and 3), OC-SMART, SeaDAS, and C2X outperform the rest of the processors (Fig. B2). In these regions, while both SeaDAS and C2X show better performance for the OLI data, OC-SMART is found to perform equally for both OLI and MSI (Fig. A1).
- We demonstrated that each processor has different degrees of sensitivity to varying choices of constituent retrieval algorithms (Section 4.2 and Appendix C may be consulted). This suggests that a switching scheme to select the optimal AC based on OWTs may be a promising approach. Therefore, we recommend further studies dedicated to identifying optimal retrieval algorithms for each processor.
- The uncertainties associated with our Chl*a* and TSS matchups demonstrate that a cautious approach is needed where products are intended for scientific studies where high accuracies are required. For example, $\epsilon < 10\%$ in $\widehat{\rho}_w$ is encouraged for the identification of subtle climate-change signals, estimation of the absorption of colored dissolved organic matter (Cao et al., 2018) relevant to carbon budget assessments, and analysis of the variability of in-water particulate backscattering (Zawada et al., 2007). We speculate that this uncertainty requirement should also yield uncertainties <20% in global Chl*a* and TSS products and allow robust assessments of seasonal variability across a wide range of aquatic ecosystems. One should note that this requirement shall *not* preclude existing products (e.g., Tables 6 and 7) obtained from current versions of the AC processors for global and/or regional water quality monitoring applications. With future hyperspectral missions in sight and their potential ability to better address scientific questions related to phytoplankton properties (Pahlevan et al., 2021), more stringent uncertainty requirements in $\widehat{\rho}_w$, similar to those adopted for the Plankton, Aerosol, Clouds, and Ecosystem (PACE) mission (Cetinic et al., 2019) are expected.
- While AC processors are evolving, for some of the application areas (e.g., HAB detection) novel regional techniques that fully or partially bypass the AC process and the associated uncertainties are viable alternatives (Binding et al., 2013; Cao et al., 2020; Matthews and Odermatt, 2015; Smith et al., 2021; Stumpf et al., 2016). For a global applicability of such an approach, a further global data drive is needed to compile existing optical proxies of water quality indicators (e.g., Chl*a*, TSS, turbidity, Secchi-disk depth).
- Further research dedicated to enhancing the representativeness of aerosol models integrated into the AC processors is required. Future mission designs should consider the inclusion of observation modalities (e.g., polarimetric, hyperspectral, and multi-angular radiometry as well as ranging) to improve the discrimination and/or characterization of aerosol types, heights, and optical thickness (Frouin et al., 2019). Moreover, additions of high-fidelity radiometric measurements in the deeper blue bands (He et al., 2012) and/or within the ultraviolet region should further constrain the solution space for aerosol retrievals, especially, in dystrophic ecosystems (e.g., boreal lakes) where negligible ρ_w is expected in this spectral range.

7. Conclusion

The aquatic subgroup of the second Atmospheric Correction Intercomparison eXercise (ACIX-Aqua), a joint ESA and NASA initiative under the CEOS direction, was specifically developed to carry out a comprehensive assessment of the existing AC processors for Landsat-8 and Sentinel-2 data processing over inland and coastal waters. This required a community-wide data sharing effort from field campaigns in freshwaters around the globe and utilizing the observation records available through the primary coastal ocean AERONET-OC sites.

Through a considerable effort of *in situ* data collation, we were able to assess comprehensively the performance of eight different AC processors. We found marked performance differences between inland and coastal waters, where some processors performed satisfactorily in freshwater bodies, while others showed superior performance in coastal waters. Overall, the uncertainties were lower in the coastal environments. By combining our freshwater and AERONET-OC matchups, we produced performance matrices to determine the AC processors considered most capable of generating reliable products for specific OWTs. Despite the processors failing to meet a 30% error threshold across all the visible bands in freshwater environments, we showed (for best-performing AC processors) that the median errors in Chl*a* and TSS range from 25 to 70%. The derived products from best-performing AC processors should be suitable for some water quality monitoring activities, such as hot-spot identification and assessing impacts of episodic events. However, accuracies demonstrated here may limit their suitability for long-term trend studies targeting incremental shifts in climate, biogeochemical cycling, organic and/or inorganic particle discrimination, or phytoplankton property identification. It is anticipated that near-future research will lead to advancements in the performance of AC processors by employing more representative aerosol types and/or bio-optical models depending on the underlying mechanisms of the AC processors. Further, characterizing and quantifying adjacency effects deserve major advancements to ensure practical products for scientific studies or water quality assessments in small or hydrologically complex water bodies.

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.

Acknowledgment

We acknowledge NASA's AERONET team for maintaining the network. Our greatest appreciation is extended to the AERONET-OC Principal Investigators (Giuseppe Zibordi, Brent Holben, Alex Gilerson, Samir Ahmed, Burton Jones, Hui Feng, Young-Je Park, Heidi Sosik, Sherwin Ladner, Timothy Moore, Menghua Wang, Steven Greb, Sarah Bartlett, Dimitry Van der Zande) and their corresponding funding agencies (e.g., JAXA's GCOM-C project). We are thankful to Yannick Huot and Giuseppe Zibordi for their general comments as well as to three anonymous reviewers for providing critical reviews and constructive comments that improved the presentation of the results. The field campaigns conducted in the Brazilian territory were funded by the São Paulo Research Foundation (FAPESP) Project 2014/23903-9. Steef Peters, Evangelos Spyrakos, Peter Hunter, Andrew Tyler, Martin Ligi, and Mark Warren were funded under the European Union's Horizon 2020 research and innovation program under grant agreements No. 776480 (MONOCLE) and No. 730066 (EOMORES). Krista Alikas and Martin Ligi were funded by EOMORES. Evangelos Spyrakos, Peter Hunter, and Andrew Tyler were also funded under the UK Natural Environment Research Council (NERC) projects GloboLakes (NE/J024279/1) and INCIS-3IVE (NE/L013312/1). BONUS FerryScope for *in situ* data collection in European waters is also acknowledged. The consistency in image processing and data extraction among all the processors were made possible via the CTEP platform under ESA support. Nima Pahlevan was funded under NASA ROSES contract # 80HQTR19C0015, Remote Sensing of Water Quality element, and the USGS Landsat Science Team Award # 140G0118C0011.

Appendix A

A sensor-specific assessment of the AC processors is shown in Fig. A1, where the number of valid matchups per processor is also provided. For the 443 nm band analyzed via the AERONET-OC matchups, C2X and SeaDAS exhibit distinct performances with higher uncertainties associated with MSI $\widehat{\rho}_w$. Similar differences also exist for ACOLITE and GRS as revealed through CVD matchups. When considering the 664 nm band, C2X, OC-SMART, and iCOR show different performances, i.e., $>50\%$, for OLI and MSI determined via both CVD and AERONET-OC matchups. Another example is ACOLITE which appears to output relatively consistent $\widehat{\rho}_w(664)$ in freshwaters well represented in the CVD matchup pool.ool

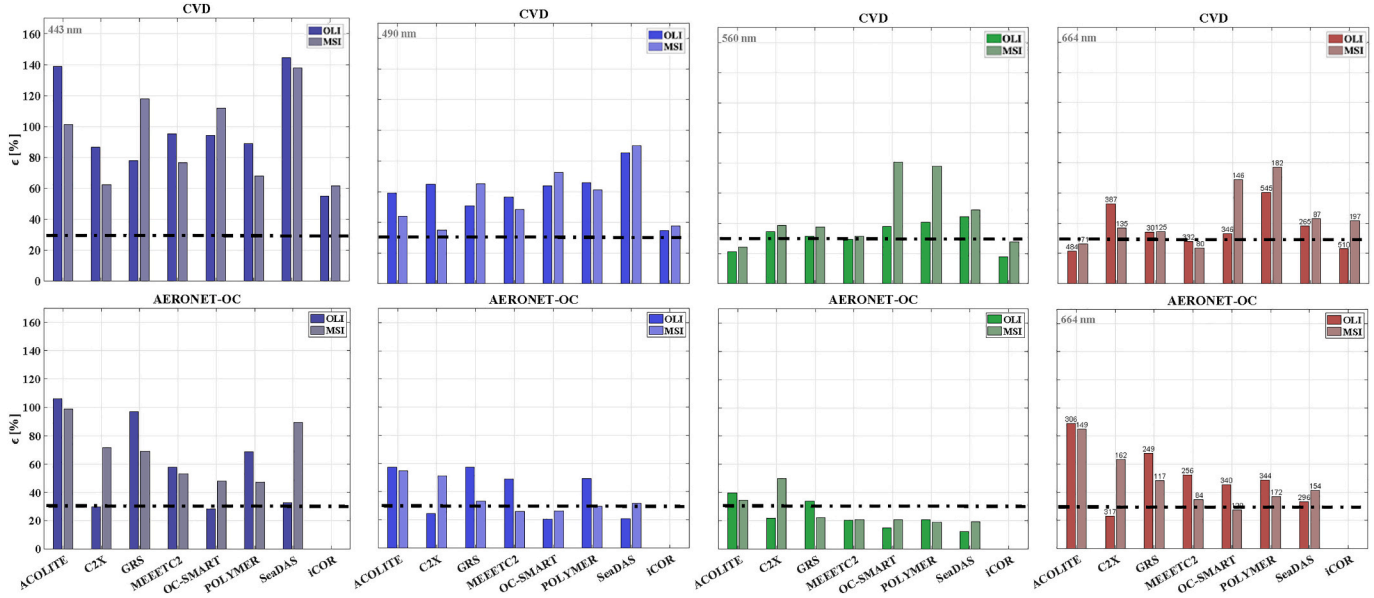

Fig. A1. Performance assessments demonstrating OLI-MSI interconsistency for the visible bands using valid matchups for each individual processor. Bars associated with the 664 nm chart are labeled with the number of valid matchups for each sensor. The dashed lines correspond to a 30% threshold (GCOS).

Appendix B

As the secondary performance assessment approach for each OWT (Section 3.4), we identified universally common matchups among all the processors that return adequate numbers of matchups for the CVD and AERONET-OC datasets. To rank the processors, a composite metric was formulated using the five (unsigned) log-based figures of merit, i.e., ϵ, $|\beta|$, RMSLE, $|1 - S|$, and MSA (Section 3.6). The metric values for each OWT were stored in a 4 x n array where n is the number of processors evaluated for the four visible bands. These values were then normalized within the [0,1] interval for each OWT and band. This strategy allowed for a uniform scaling of all the metrics. Assuming equal weights for each, the metrics for each processor/OWT/band were added, yielding a value ranging from 0 to 5 – the best-performing AC scheme takes on a value close to zero. This assessment is virtually equivalent of performance analyses based on radar (or spider) diagrams. Note that the analysis has been carried out independently for the CVD and AERONET-OC matchups. Here, *only* processors returning larger numbers of valid $\widehat{\rho}_w$ matchups were evaluated. Therefore, GRS and SeaDAS from the CVD, and GRS from the AERONET-OC analyses, were eliminated. Further analyses suggested that including OC-SMART reduces the number of valid matchups for the CVD assessment; thus, this processor was also eliminated

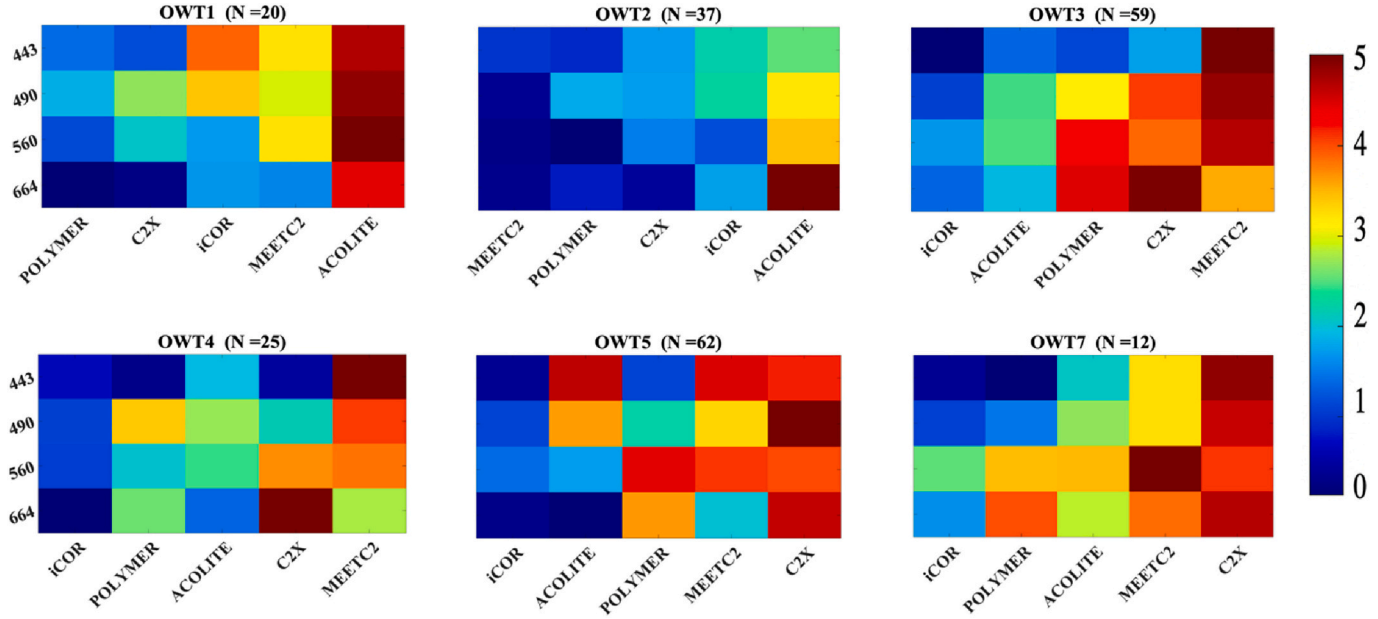

Fig. B1. Relative performance assessments determined via a composite metric derived from the universally common CVD matchups. The number of common and valid matchups is denoted by N. Cooler colors indicate a better performance. Five log-based metrics, i.e., RMSLE, ϵ, $|\beta|$, MSA, and $|1\text{-}S|$ (Section 3.6), were utilized to produce the heatmaps. Band-average metrics were applied to rank the processors from left to right. Note that no common matchups were identified for OWT6.

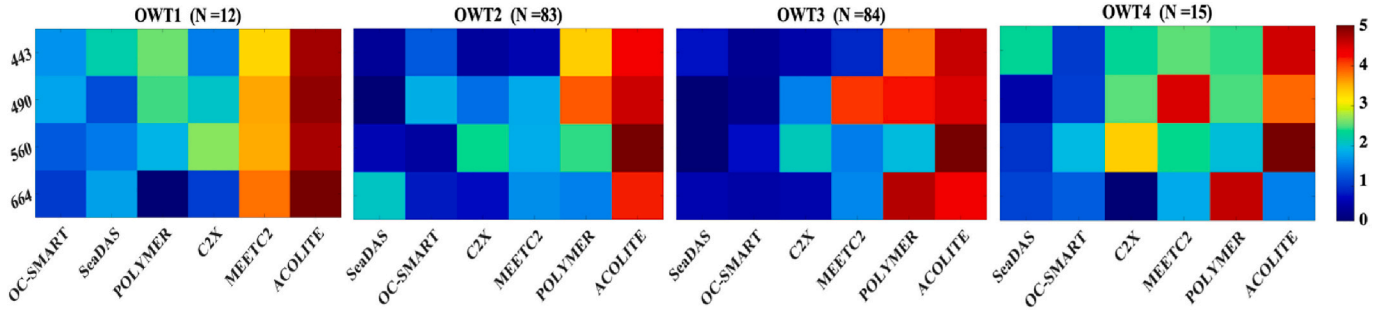

Fig. B2. Same as Fig. B1, but derived from AERONET-OC data analysis. Less than five common matchups were identified for OWTs 5,6, and 7, and hence they are not illustrated.

Appendix C

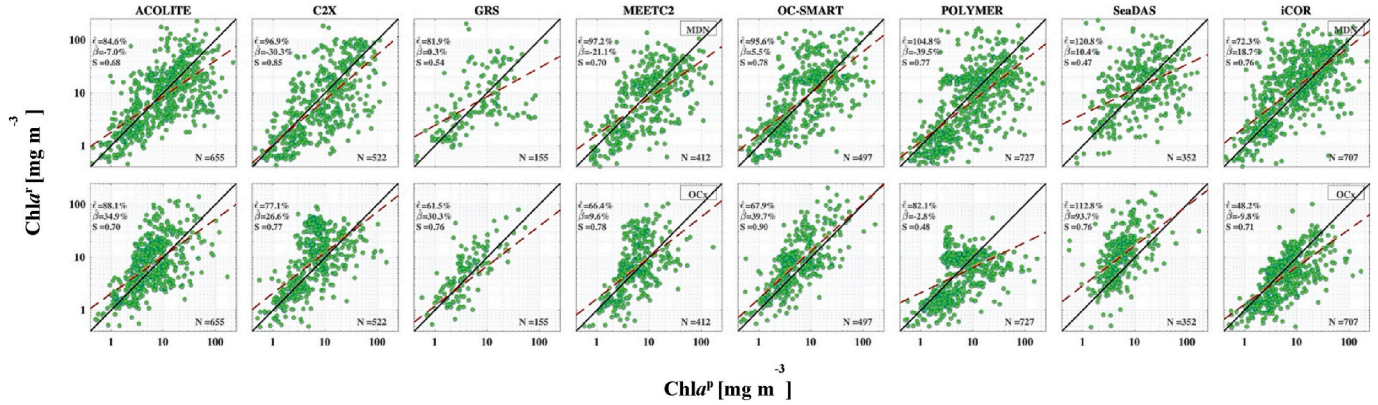

Fig. C1. Satellite-derived Chl*a* (Chl*a*^{r}) against pseudo Chl*a* products (Chl*a*^{p}) estimated from the CVD radiometric matchups. The plots show sensitivity of two (Type I) Chl*a* algorithms to uncertainties in $\widehat{\rho}_w$. Note that combined OLI and MSI's visible bands are applied here, and that the same Chl*a* algorithms are used to create each scatterplot (e.g., MDN for top row). For an ideal AC processor, data distributions align with 1:1 lines. This assessment does not provide insights into the absolute performance of Chl*a* algorithms (see Sections 3.1 and 3.5).

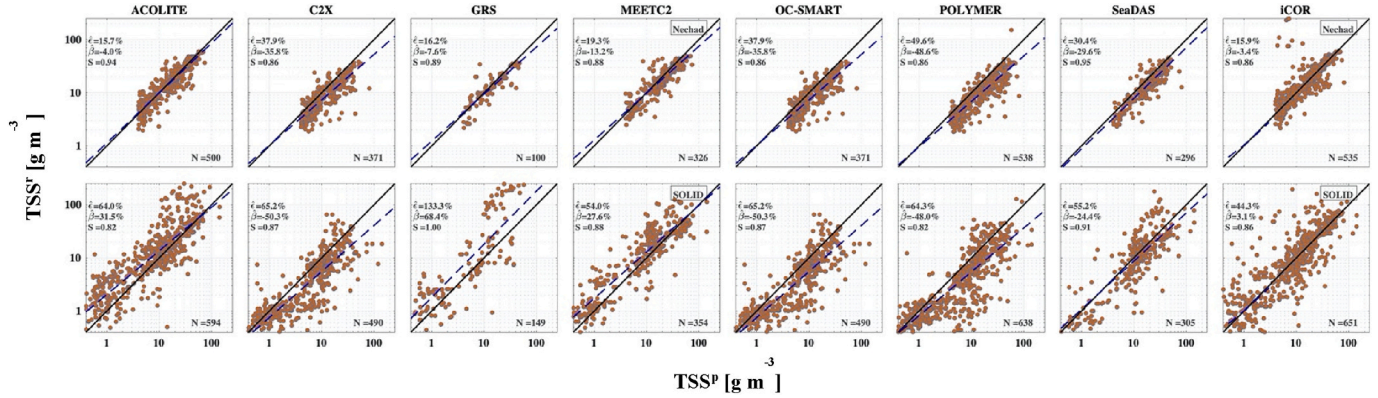

Fig. C2. Same as C.1, but for TSS products. Top row shows sensitivity of the Nechad model to uncertainties in $\widehat{\rho}_w$ derived from different processors. For this algorithm, data with TSS $<$ 4 g m $^{-3}$ were removed from the matchups leading to fewer matchup samples.

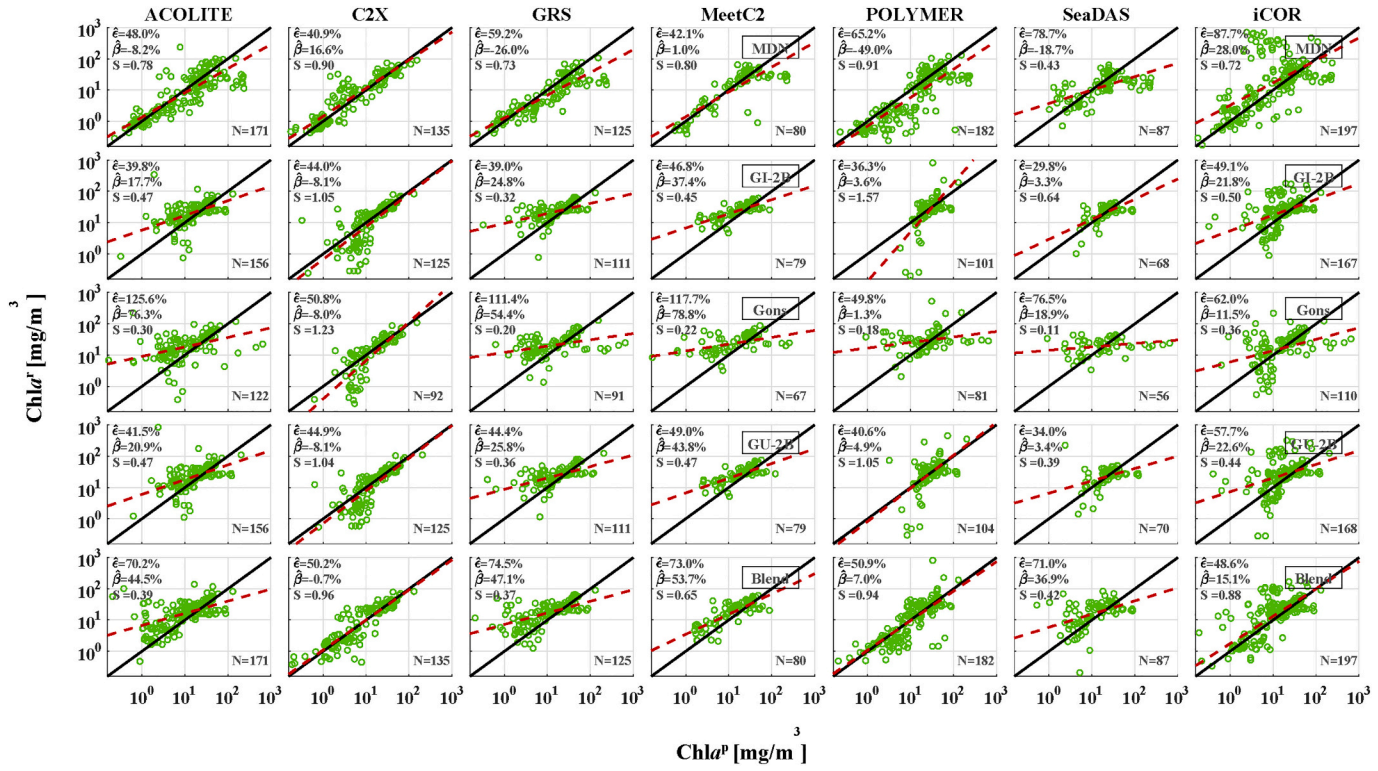

Fig. C3. Same as C.1, but using Type II algorithms for MSI matchups only. The algorithms include MDN (Pahlevan et al., 2020), GU-2B (Gurlin et al., 2011), GI-2B (Gilerson et al., 2010), Gons (Gons et al., 2002), and Blend (Smith et al., 2018). GI-2B appears to be least sensitive to uncertainties in $\widehat{\rho}_w$.

Appendix D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2021.112366.

References

Abbott, M.R., Letelier, R.M., 1999. Algorithm theoretical basis document chlorophyll fluorescence (MODIS product number 20). *NASA*. *http://www.modis.gsfc.nasa.gov/data/atbd*.

Ahmad, Z., Fraser, R.S., 1982. An iterative radiative transfer code for ocean-atmosphere systems. J. Atmos. Sci. 39, 656–665.

Ahmad, Z., Franz, B.A., McClain, C.R., Kwiatkowska, E.J., Werdell, J., Shettle, E.P., Holben, B.N., 2010. New aerosol models for the retrieval of aerosol optical thickness and normalized water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open oceans. Appl. Opt. 49, 5545–5560.

Alikas, K., Vabson, V., Ansko, I., Tilstone, G.H., Dall'Olmo, G., Nencioli, F., Vendt, R., Donlon, C., Casal, T., 2020. Comparison of above-water Seabird and TriOS radiometers along an Atlantic Meridional Transect. Remote Sens. 12, 1669.

Ansper, A., Alikas, K., 2019. Retrieval of chlorophyll a from Sentinel-2 MSI data for the European Union water framework directive reporting purposes. Remote Sens. 11, 64.

Antoine, D., Morel, A., 1999. A multiple scattering algorithm for atmospheric correction of remotely sensed ocean colour (MERIS instrument): principle and implementation for atmospheres carrying various aerosols including absorbing ones. Int. J. Remote Sens. 20, 1875–1916.

Bailey, S.W., Werdell, P.J., 2006. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. Remote Sens. Environ. 102, 12–23.

Bailey, S.W., Franz, B.A., Werdell, P.J., 2010. Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. Opt. Express 18, 7521–7527.

Balasubramanian, S.V., Pahlevan, N., Smith, B., Binding, C., Schalles, J., Loisel, H., Gurlin, D., Greb, S., Alikas, K., Randla, M., Bunkei, M., Moses, W., Nguyễn, H., Lehmann, M.K., O'Donnell, D., Ondrusek, M., Han, T.-H., Fichot, C.G., Moore, T., Boss, E., 2020. Robust algorithm for estimating total suspended solids (TSS) in inland and nearshore coastal waters. Remote Sens. Environ. 246, 111768.

Berk, A., Anderson, G.P., Acharya, P.K., Bernstein, L.S., Muratov, L., Lee, J., Fox, M., Adler-Golden, S.M., Chetwynd, J.J.H., Hoke, M.L., Lockwood, R.B., Gardner, J.A., Cooley, T.W., Borel, C.C., Lewis, P.E., Shettle, E.P., 2006. MODTRAN5: 2006 update, 62331F–62331F.

Binding, C., Greenberg, T., Bukata, R., 2013. The MERIS maximum chlorophyll index; its merits and limitations for inland water algal bloom monitoring. J. Great Lakes Res. 39, 100–107.

Binding, C.E., Pizzolato, L., Zeng, C., 2021. EOLakeWatch; delivering a comprehensive suite of remote sensing algal bloom indices for enhanced monitoring of Canadian eutrophic lakes. Ecol. Indic. 121, 106999.

Bresciani, M., Pinardi, M., Free, G., Luciani, G., Ghebrehiwot, S., Laanen, M., Peters, S., Della Bella, V., Padula, R., Giardino, C., 2020. The use of multisource optical sensors to study phytoplankton spatio-temporal variation in a shallow turbid Lake. Water 12, 284.

Brockmann, C., Doerffer, R., Peters, M., Kerstin, S., Embacher, S., Ruescas, A., 2016. Evolution of the C2RCC neural network for sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters. ESASP 740, 54.

Bulgarelli, B., Kiselev, V., Zibordi, G., 2014. Simulation and analysis of adjacency effects in coastal waters: a case study. Appl. Opt. 53, 1523–1545.

Cao, F., Tzortziou, M., Hu, C., Mannino, A., Fichot, C.G., Del Vecchio, R., Najjar, R.G., Novak, M., 2018. Remote sensing retrievals of colored dissolved organic matter and dissolved organic carbon dynamics in north American estuaries and their margins. Remote Sens. Environ. 205, 151–165.

Cao, Z., Ma, R., Duan, H., Pahlevan, N., Melack, J., Shen, M., Xue, K., 2020. A machine learning approach to estimate chlorophyll-a from Landsat-8 measurements in inland lakes. Remote Sens. Environ. 248, 111974.

Cetinic, I., McClain, C.R., Werdell, P.J., Ahmad, Z., Franz, B.A., Karakoylu, E.M., McKinna, L.I., Patt, F.S., 2019. PACE Technical Report Series, Volume 6: Data Product Requirements and Error Budgets Consensus Document.

Chami, M., Lafrance, B., Fougnie, B., Chowdhary, J., Harmel, T., Waquet, F., 2015. OSOAA: a vector radiative transfer model of coupled atmosphere-ocean system for a rough sea surface application to the estimates of the directional variations of the water leaving reflectance to better process multi-angular satellite sensors data over the ocean. Opt. Express 23, 27829–27852.

Chomko, R.M., Gordon, H.R., 1998. Atmospheric correction of ocean color imagery: use of the Junge power-law aerosol size distribution with variable refractive index to handle aerosol absorption. Appl. Opt. 37, 5560–5572.

Cox, C., Munk, W., 1954. Measurement of the roughness of the sea surface from photographs of the sun's glitter. JOSA 44, 838–850.

De Keukelaere, L., Sterckx, S., Adriaensen, S., Knaeps, E., Reusen, I., Giardino, C., Bresciani, M., Hunter, P., Neil, C., Van der Zande, D., 2018. Atmospheric correction of Landsat-8/OLI and Sentinel-2/MSI data using iCOR algorithm: validation for coastal and inland waters. European Journal of Remote Sensing 51, 525–542.

Deschamps, P., Herman, M., Tanre, D., 1983. Modeling of the atmospheric effects and its application to the remote sensing of ocean color. Appl. Opt. 22, 3751–3758.

Doxani, G., Vermote, E., Roger, J.-C., Gascon, F., Adriaensen, S., Frantz, D., Hagolle, O., Hollstein, A., Kirches, G., Li, F., 2018. Atmospheric correction inter-comparison exercise. Remote Sens. 10, 352.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ. 120, 25–36.

Eleveld, M.A., Ruescas, A.B., Hommersom, A., Moore, T.S., Peters, S.W., Brockmann, C., 2017. An optical classification tool for global lake waters. Remote Sens. 9, 420.

Fan, Y., Li, W., Gatebe, C.K., Jamet, C., Zibordi, G., Schroeder, T., Stamnes, K., 2017. Atmospheric correction over coastal waters using multilayer neural networks. Remote Sens. Environ. 199, 218–240.

Fan, Y., Li, W., Chen, N., Ahn, J.-H., Park, Y.-J., Kratzer, S., Schroeder, T., Ishizaka, J., Chang, R., Stamnes, K., 2021. OC-SMART: a machine learning based data analysis platform for satellite ocean color sensors. Remote Sens. Environ. 253, 112236.

Franz, B.A., Bailey, S.W., Kuring, N., Werdell, P.J., 2015. Ocean color measurements with the operational land imager on Landsat-8: implementation and evaluation in SeaDAS. J. Appl. Remote. Sens. 9, 096070.

Frouin, R.J., Franz, B.A., Ibrahim, A., Knobelspiesse, K., Ahmad, Z., Cairns, B., Chowdhary, J., Dierssen, H.M., Tan, J., Dubovik, O., 2019. Atmospheric correction of satellite ocean-color imagery during the PACE era. Front. Earth Sci. 7, 145.

Fukushima, H., Higurashi, A., Mitomi, Y., Nakajima, T., Noguchi, T., Tanaka, T., Toratani, M., 1998. Correction of atmospheric effect on ADEOS/OCTS ocean color data: algorithm description and evaluation of its performance. J. Oceanogr. 54, 417–430.

Gao, B.-C., Montes, M.J., Ahmad, Z., Davis, C.O., 2000. Atmospheric correction algorithm for hyperspectral remote sensing of ocean color from space. Appl. Opt. 39, 887–896.

Gasteiger, J., Emde, C., Mayer, B., Buras, R., Buehler, S., Lemke, O., 2014. Representative wavelengths absorption parameterization applied to satellite channels and spectral bands. J. Quant. Spectrosc. Radiat. Transf. 148, 99–115.

Gilerson, A.A., Gitelson, A.A., Zhou, J., Gurlin, D., Moses, W., Ioannou, I., Ahmed, S.A., 2010. Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. Opt. Express 18, 24109–24125.

Gilerson, A., Carrizo, C., Foster, R., Harmel, T., 2018. Variability of the reflectance coefficient of skylight from the ocean surface and its implications to ocean color. Opt. Express 26, 9615–9633.

Gons, H.J., Rijkeboer, M., Ruddick, K.G., 2002. A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. J. Plankton Res. 24, 947–951.

Gordon, H.R., 1976. Radiative transfer: a technique for simulating the ocean in satellite remote sensing calculations. Appl. Opt. 15, 1974–1979.

Gordon, H.R., 1978. Removal of atmospheric effects from satellite imagery of the oceans. Appl. Opt. 17, 1631–1636.

Gordon, H.R., 1997. Atmospheric correction of ocean color imagery in the earth observing system era. J. Geophys. Res. Atmos. 102, 17081–17106.

Gordon, H.R., Wang, M., 1994. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. Appl. Opt. 33, 443–452.

Gordon, H.R., Brown, J.W., Evans, R.H., 1988. Exact Rayleigh scattering calculations for use with the Nimbus-7 coastal zone color scanner. Appl. Opt. 27, 862–871.

Gordon, H.R., Du, T., Zhang, T., 1997. Remote sensing of ocean color and aerosol properties: resolving the issue of aerosol absorption. Appl. Opt. 36, 8670–8684.

Gurlin, D., Gitelson, A.A., Moses, W.J., 2011. Remote estimation of chl-a concentration in turbid productive waters—return to a simple two-band NIR-red model? Remote Sens. Environ. 115, 3479–3490.

Haan, J.F., Kokke, J., 1996. Remote sensing algorithm development: toolkit I: operationalization of atmospheric correction methods for tidal and inland waters(pp. 96-16). Netherlands Remote Sensing Board (BCRS).

Harmel, T., Chami, M., 2011. Influence of polarimetric satellite data measured in the visible region on aerosol detection and on the performance of atmospheric correction procedure over open ocean waters. Opt. Express 19, 20960–20983.

Harmel, T., Chami, M., Tormos, T., Reynaud, N., Danis, P.-A., 2018. Sunglint correction of the multi-spectral instrument (MSI)-SENTINEL-2 imagery over inland and sea waters from SWIR bands. Remote Sens. Environ. 204, 308–321.

He, X., Bai, Y., Pan, D., Tang, J., Wang, D., 2012. Atmospheric correction of satellite ocean color imagery using the ultraviolet wavelength for highly turbid waters. Opt. Express 20, 20754–20770.

Hlaing, S., Gilerson, A., Harmel, T., Tonizzo, A., Weidemann, A., Arnone, R., Ahmed, S., 2012. Assessment of a bidirectional reflectance distribution correction of above-water and satellite water-leaving radiance in coastal waters. Appl. Opt. 51, 220–237.

Hlaing, S., Harmel, T., Gilerson, A., Foster, R., Weidemann, A., Arnone, R., Wang, M., Ahmed, S., 2013. Evaluation of the VIIRS Ocean color monitoring performance in coastal regions. Remote Sens. Environ. 139, 398–414.

Holben, B.N., Eck, T.F., Slutsker, I., Tanré, D., Buis, J.P., Setzer, A., Vermote, E., Reagan, J.A., Kaufman, Y.J., Nakajima, T., Lavenu, F., Jankowiak, I., Smirnov, A., 1998. AERONET—A federated instrument network and data archive for aerosol characterization. Remote Sens. Environ. 66, 1–16.

Huot, Y., Brown, C.A., Potvin, G., Antoniades, D., Baulch, H.M., Beisner, B.E., Bélanger, S., Brazeau, S., Cabana, H., Cardille, J.A., del Giorgio, P.A., Gregory-Eaves, I., Fortin, M.-J., Lang, A.S., Laurion, I., Maranger, R., Prairie, Y.T., Rusak, J. A., Segura, P.A., Siron, R., Smol, J.P., Vinebrooke, R.D., Walsh, D.A., 2019. The NSERC Canadian Lake pulse network: a national assessment of lake health providing

science for water management in a changing climate. Sci. Total Environ. 695, 133668.
Ilori, C.O., Pahlevan, N., & Knudby, A. (2019). Analyzing performances of different atmospheric correction techniques for Landsat 8: application for coastal remote sensing. *Remote Sensing*, 11, 469.
IOCCG, 2010. Atmospheric Correction for Remotely-Sensed Ocean-Colour Products. M. Wang I.O.C.C. Group.
IOCCG, 2018. Earth observations in support of global water quality monitoring. In: Greb, S., Dekker, A. (Eds.), Binding International Ocean Colour Coordinating Group, 17. I.O.C.C. Group.
IOCCG, 2019. Protocols for Satellite Ocean Colour Data Validation: In Situ Optical Radiometry, vol. 3.0.
Jamet, C., Loisel, H., Kuchinke, C.P., Ruddick, K., Zibordi, G., Feng, H., 2011. Comparison of three SeaWiFS atmospheric correction algorithms for turbid waters using AERONET-OC measurements. Remote Sens. Environ. 115, 1955–1965.
Kotchenova, S.Y., Vermote, E.F., Matarrese, R., Klemm Jr., F.J., 2006. Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: path radiance. Appl. Opt. 45, 6762–6774.
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.
Lenoble, J., Herman, M., Deuzé, J., Lafrance, B., Santer, R., Tanré, D., 2007. A successive order of scattering code for solving the vector equation of transfer in the earth's atmosphere with aerosols. J. Quant. Spectrosc. Radiat. Transf. 107, 479–507.
Matthews, M., Bernard, S., 2013. Characterizing the absorption properties for remote sensing of three small optically-diverse south African reservoirs. Remote Sens. 5, 4370–4404.
Matthews, M.W., Odermatt, D., 2015. Improved algorithm for routine monitoring of cyanobacteria and eutrophication in inland and near-coastal waters. Remote Sens. Environ. 156, 374–382.
Mélin, F., Clerici, M., Zibordi, G., Holben, B., Smirnov, A., 2010. Validation of SeaWiFS and MODIS aerosol products with globally distributed AERONET data. Remote Sens. Environ. 114, 230–250.
Mobley, C.D., 1999. Estimation of the remote-sensing reflectance from above-surface measurements. Appl Opt 38, 7442–7455.
Mobley, C.D., Sundman, L.K., 2008. Hydrolight 5, Ecolight5 User Guide. Sequoia Scientific, Inc.
Mobley, C.D., Werdell, J., Franz, B., Ahmad, Z., & Bailey, S. (2016). Atmospheric correction for satellite ocean color radiometry. NASA/TM-2016-217551, GSFC-E-DAA-TN35509.
Moore, G., Aiken, J., Lavender, S., 1999. The atmospheric correction of water colour and the quantitative retrieval of suspended particulate matter in case II waters: application to MERIS. Int. J. Remote Sens. 20, 1713–1733.
Moore, T.S., Dowell, M.D., Bradt, S., Verdu, A.R., 2014. An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. Remote Sens. Environ. 143, 97–111.
Morel, A., Antoine, D., Gentili, B., 2002. Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function. Appl. Opt. 41, 6289–6306.
Morley, S.K., Brito, T.V., Welling, D.T., 2018. Measures of model performance based on the log accuracy ratio. Space Weather 16, 69–88.
Moses, W.J., Gitelson, A.A., Perk, R.L., Gurlin, D., Rundquist, D.C., Leavitt, B.C., Barrow, T.M., Brakhage, P., 2012. Estimation of chlorophyll-a concentration in turbid productive waters using airborne hyperspectral data. Water Res. 46, 993–1004.
Moses, W.J., Sterckx, S., Montes, M.J., De Keukelaere, L., Knaeps, E., 2017. Atmospheric correction for inland waters. In: Bio-optical Modeling and Remote Sensing of Inland Waters. Elsevier, pp. 69–100.
Moulin, C., Gordon, H.R., Banzon, V.F., Evans, R.H., 2001. Assessment of Saharan dust absorption in the visible from SeaWiFS imagery. Journal of Geophysical Research: Atmospheres 106, 18239–18249.
Nechad, B., Ruddick, K., Park, Y., 2010. Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters. Remote Sens. Environ. 114, 854–866.
Nicolas, J.-M., Deschamps, P.-Y., Loisel, H., & Moulin, C. (2002). POLDER-2 / Ocean Color: Algorithm Theoretical Basis Document for Atmospheric correction Algorithm. Laboratoire d'Optique Atmosphérique (LOA) Laboratoire des Sciences du Climat et de l'Environnement (LSCE).
Nordkvist, K., Loisel, H., Gaurier, L.D., 2009. Cloud masking of SeaWiFS images over coastal waters using spectral variability. Opt. Express 17, 12246–12258.
Novoa, S., Doxaran, D., Ody, A., Vanhellemont, Q., Lafon, V., Lubac, B., Gernez, P., 2017. Atmospheric corrections and multi-conditional algorithm for multi-sensor remote sensing of suspended particulate matter in low-to-high turbidity levels coastal waters. Remote Sens. 9, 61.
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., Lee, Z., Wei, J., Schaff, C., Schott, J., Berk, A., 2014. On-orbit radiometric characterization of OLI (Landsat-8) for applications in aquatic remote sensing. Remote Sens. Environ. 154, 272–284.
Pahlevan, N., Roger, J.-C., Ahmad, Z., 2017a. Revisiting short-wave-infrared (SWIR) bands for atmospheric correction in coastal waters. Opt. Express 25, 6015–6035.
Pahlevan, N., Sarkar, S., Franz, B.A., Balasubramanian, S.V., He, J., 2017b. Sentinel-2 MultiSpectral instrument (MSI) data processing for aquatic science applications: demonstrations and validations. Remote Sens. Environ. 201, 47–56.
Pahlevan, N., Schott, J.R., Franz, B.A., Zibordi, G., Markham, B., Bailey, S., Schaaf, C.B., Ondrusek, M., Greb, S., Strait, C.M., 2017c. Landsat 8 remote sensing reflectance (R rs) products: evaluations, intercomparisons, and enhancements. Remote Sens. Environ. 190, 289–301.
Pahlevan, N., Smith, B., Binding, C., O'Donnell, D.M., 2017d. Spectral band adjustments for remote sensing reflectance spectra in coastal/inland waters. Opt. Express 25, 28650–28667.
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.
Pahlevan, N., Smith, B., Schalles, J., Binding, C., Cao, Z., Ma, R., Alikas, K., Kangro, K., Gurlin, D., Hà, N., Matsushita, B., Moses, W., Greb, S., Lehmann, M.K., Ondrusek, M., Oppelt, N., Stumpf, R., 2020. Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: a machine-learning approach. Remote Sens. Environ. 240, 111604.
Pahlevan, N., Smith, B., Binding, C., Gurlin, D., Li, L., Bresciani, M., Giardino, C., 2021. Hyperspectral retrievals of phytoplankton absorption and chlorophyll-a in inland and nearshore coastal waters. Remote Sens. Environ. 253, 112200.
Park, Y.-J., Ruddick, K., 2005. Model of remote-sensing reflectance including bidirectional effects for case 1 and case 2 waters. Appl. Opt. 44, 1236–1249.
Pereira-Sandoval, M., Ruescas, A., Urrego, P., Ruiz-Verdú, A., Delegido, J., Tenjo, C., Soria-Perpinyà, X., Vicente, E., Soria, J., Moreno, J., 2019. Evaluation of atmospheric correction algorithms over Spanish inland waters for sentinel-2 multi spectral imagery data. Remote Sens. 11, 1469.
Philipson, P., Kratzer, S., Ben Mustapha, S., Strömbeck, N., Stelzer, K., 2016. Satellite-based water quality monitoring in Lake Vänern, Sweden. Int. J. Remote Sens. 37, 3938–3960.
Rahman, H., Dedieu, G., 1994. SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. Remote Sens. 15, 123–143.
Renosh, P.R., Doxaran, D., Keukelaere, L.D., Gossn, J.I., 2020. Evaluation of atmospheric correction algorithms for Sentinel-2-MSI and Sentinel-3-OLCI in highly turbid estuarine waters. Remote Sens. 12, 1285.
Ruddick, K.G., De Cauwer, V., Park, Y.-J., Moore, G., 2006. Seaborne measurements of near infrared water-leaving reflectance: the similarity spectrum for turbid waters. Limnol. Oceanogr. 51, 1167–1179.
Santer, R., Schmechtig, C., 2000. Adjacency effects on water surfaces: primary scattering approximation and sensitivity study. Appl. Opt. 39, 361–375.
Saulquin, B., Fablet, R., Bourg, L., Mercier, G., d'Andon, O.F., 2016. MEETC2: ocean color atmospheric corrections in coastal complex waters using a Bayesian latent class model and potential for the incoming sentinel 3—OLCI mission. Remote Sens. Environ. 172, 39–49.
Schaeffer, B.A., Bailey, S.W., Conmy, R.N., Galvin, M., Ignatius, A.R., Johnston, J.M., Keith, D.J., Lunetta, R.S., Parmar, R., Stumpf, R.P., 2018. Mobile device application for monitoring cyanobacteria harmful algal blooms using Sentinel-3 satellite ocean and land colour instruments. Environ. Model Softw. 109, 93–103.
Schläpfer, D., 1998. Differential Absorption Methodology for Imaging Spectroscopy of Atmospheric Water Vapor.
Seegers, B.N., Stumpf, R.P., Schaeffer, B.A., Loftin, K.A., Werdell, P.J., 2018. Performance metrics for the assessment of satellite data products: an ocean color case study. Opt. Express 26, 7404–7422.
Shettle, E.P., Fenn, R.W., 1979. Models for the Aerosols of the Lower Atmosphere and the Effects of Humidity Variations on their Optical Properties. DTIC Document.
Siegel, D.A., Wang, M., Maritorena, S., Robinson, W., 2000. Atmospheric correction of satellite ocean color imagery: the black pixel assumption. Appl. Opt. 39, 3582–3591.
Smith, M.E., Lain, L.R., Bernard, S., 2018. An optimized chlorophyll a switching algorithm for MERIS and OLCI in phytoplankton-dominated waters. Remote Sens. Environ. 215, 217–227.
Smith, B., Pahlevan, N., Schalles, J., Ruberg, S., Errera, R., Ma, R., Giardino, C., Bresciani, M., Barbosa, C., Moore, T., Fernandez, V., Alikas, K., & Kangaro, K. (2021). A chlorophyll-a algorithm for Landsat-8 based on mixture density networks. *Frontiers in Remote Sensing, 1*.
Soomets, T., Uudeberg, K., Jakovels, D., Brauns, A., Zagars, M., Kutser, T., 2020. Validation and comparison of water quality products in Baltic Lakes using Sentinel-2 MSI and Sentinel-3 OLCI data. Sensors 20, 742.
Spyrakos, E., O'Donnell, R., Hunter, P.D., Miller, C., Scott, M., Simis, S.G., Neil, C., Barbosa, C.C., Binding, C.E., Bradt, S., 2018. Optical types of inland and coastal waters. Limnol. Oceanogr. 63, 846–870.
Stamnes, K., Hamre, B., Stamnes, S., Chen, N., Fan, Y., Li, W., Lin, Z., Stamnes, J., 2018. Progress in forward-inverse modeling based on radiative transfer tools for coupled atmosphere-snow/ice-ocean systems: a review and description of the accurt model. Appl. Sci. 8, 2682.
Steinmetz, F., Ramon, D., 2018. Sentinel-2 MSI and Sentinel-3 OLCI consistent ocean colour products using POLYMER. In: Remote Sensing of the Open and Coastal Ocean and Inland Waters. International Society for Optics and Photonics, p. 107780E.
Steinmetz, F., Deschamps, P.-Y., Ramon, D., 2011. Atmospheric correction in presence of sun glint: application to MERIS. Opt. Express 19, 9783–9800.
Sterckx, S., Knaeps, S., Kratzer, S., Ruddick, K., 2015. SIMilarity environment correction (SIMEC) applied to MERIS data over inland and coastal waters. Remote Sens. Environ. 157, 96–110.
Stumpf R. P., A.R.A., Gould J. R. , Martinolich P. M., Ransibrahmanakul V. (2004). A Partially Coupled Ocean-Atmosphere Model for Retrieval of Water-Leaving Radiance from SeaWiFS in Coastal Waters. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD.
Stumpf, R.P., Davis, T.W., Wynne, T.T., Graham, J.L., Loftin, K.A., Johengen, T.H., Gossiaux, D., Palladino, D., Burtner, A., 2016. Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. Harmful Algae 54, 160–173.
Thompson, D.R., Cawse-Nicholson, K., Erickson, Z., Fichot, C.G., Frankenberg, C., Gao, B.-C., Gierach, M.M., Green, R.O., Jensen, D., Natraj, V., 2019. A unified

approach to estimate land and water reflectances with uncertainties for coastal imaging spectroscopy. Remote Sens. Environ. 231, 111198.
Thuillier, G., Hersé, M., Labs, D., Foujols, T., Peetermans, W., Gillotay, D., Simon, P.C., Mandel, H., 2003. The solar spectral irradiance from 200 to 2400 nm as measured by the SOLSPEC spectrometer from the atlas and Eureca missions. Sol. Phys. 214, 1–22.
Tilstone, G., Dall'Olmo, G., Hieronymi, M., Ruddick, K., Beck, M., Ligi, M., Costa, M., D'alimonte, D., Vellucci, V., Vansteenwegen, D., 2020. Field intercomparison of radiometer measurements for ocean colour validation. Remote Sens. 12, 1587.
Vanhellemont, Q., 2019. 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., Ruddick, K., 2014. Turbid wakes associated with offshore wind turbines observed with Landsat 8. Remote Sens. Environ. 145, 105–115.
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.
Vansteenwegen, D., Ruddick, K., Cattrijsse, A., Vanhellemont, Q., Beck, M., 2019. The pan-and-tilt hyperspectral radiometer system (PANTHYR) for autonomous satellite validation measurements—prototype design and testing. Remote Sens. 11, 1360.
Vermote, E.F., El Saleous, N., Justice, C.O., Kaufman, Y.J., Privette, J.L., Remer, L., Roger, J.C., Tanr, D., 1997. Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation. Journal of Geophysical Research: Atmospheres 102, 17131–17141.
Wang, M., 2007. Remote sensing of the ocean contributions from ultraviolet to near-infrared using the shortwave infrared bands: simulations. Appl. Opt. 46, 1535–1547.
Wang, M., Shi, W., 2007. The NIR-SWIR combined atmospheric correction approach for MODIS Ocean color data processing. Opt. Express 15, 15722–15733.
Warren, M.A., Simis, S.G., Martinez-Vicente, V., Poser, K., Bresciani, M., Alikas, K., Spyrakos, E., Giardino, C., Ansper, A., 2019. Assessment of atmospheric correction algorithms for the sentinel-2A MultiSpectral imager over coastal and inland waters. Remote Sens. Environ. 225, 267–289.
Yang, H., Gordon, H.R., 1997. Remote sensing of ocean color: assessment of water-leaving radiance bidirectional effects on atmospheric diffuse transmittance. Appl. Opt. 36, 7887–7897.
Yuhas, R.H., Goetz, A.F., Boardman, J.W., 1992. Discrimination among Semi-Arid Landscape Endmembers Using the Spectral Angle Mapper (SAM) Algorithm.
Zawada, D.G., Hu, C., Clayton, T., Chen, Z., Brock, J.C., Muller-Karger, F.E., 2007. Remote sensing of particle backscattering in Chesapeake Bay: a 6-year SeaWiFS retrospective view. Estuar. Coast. Shelf Sci. 73, 792–806.
Zibordi, G., Holben, B., Hooker, S.B., Mélin, F., Berthon, J.-F., Slutsker, I., Giles, D., Vandemark, D., Feng, H., Rutledge, K., Schuster, G., Al Mandoos, A., 2006. A network for standardized ocean color validation measurements. Eos, Transactions American Geophysical Union 87, 293–297.
Zibordi, G., Berthon, J.-F., Mélin, F., D'Alimonte, D., Kaitala, S., 2009a. Validation of satellite ocean color primary products at optically complex coastal sites: northern Adriatic Sea, northern Baltic proper and gulf of Finland. Remote Sens. Environ. 113, 2574–2591.
Zibordi, G., Mélin, F., Berthon, J.-F., Holben, B., Slutsker, I., Giles, D., D'Alimonte, D., Vandemark, D., Feng, H., Schuster, G., Fabbri, B.E., Kaitala, S., Seppälä, J., 2009b. AERONET-OC: a network for the validation of ocean color primary products. J. Atmos. Ocean. Technol. 26, 1634–1651.