Remote Sensing of Environment 270 (2022) 112860

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Simultaneous retrieval of selected optical water quality indicators from Landsat-8, Sentinel-2, and Sentinel-3

Nima Pahlevan [a,b,*], Brandon Smith [a,b], Krista Alikas [c], Janet Anstee [d], Claudio Barbosa [e], Caren Binding [f], Mariano Bresciani [g], Bruno Cremella [h], Claudia Giardino [g], Daniela Gurlin [i], Virginia Fernandez [j], Cédric Jamet [k], Kersti Kangro [c], Moritz K. Lehmann [l], Hubert Loisel [k], Bunkei Matsushita [m], Nguyên Hà [n], Leif Olmanson [o], Geneviève Potvin [h], Stefan G.H. Simis [p], Andrea VanderWoude [q], Vincent Vantrepotte [k], Antonio Ruiz-Verdù [r]

[a] NASA Goddard Space Flight Center, Greenbelt, MD, USA
[b] Science Systems and Applications, Inc. (SSAI), Lanham, MD, USA
[c] University of Tartu, Tartu, Estonia
[d] Oceans and Atmosphere, Commonwealth Scientific and Industrial Research Organization (CSIRO), Canberra, Australia
[e] Instrumentation Lab for Aquatic Systems (LabISA), National Institute for Space Research (INPE), Sao Jose dos Campos, Brazil
[f] Environment and Climate Change Canada, Burlington, ON, Canada
[g] National Research Council of Italy, Institute for Electromagnetic Sensing of the Environment, CNR-IREA, Italy
[h] University of Sherbrooke, Department of Applied Geomatics, Sherbrooke, Québec, Canada
[i] Wisconsin Department of Natural Resources, Madison, WI, USA
[j] University of the Republic, Department of Geography, Montevideo 11400, Uruguay
[k] Univ. Littoral Côte d'Opale, CNRS, Univ. Lille, IRD, UMR 8187 - LOG - Laboratoire d'Océanologie et de Géosciences, F-62930 Wimereux, France
[l] Xerra Earth Observation Institute and the University of Waikato, Hamilton, New Zealand
[m] University of Tsukuba, Ibaraki, Japan
[n] VNU University of Science, Vietnam National University, Hanoi, Viet Nam
[o] University of Minnesota, Department of Forest Resources, St. Paul, MN, USA
[p] Plymouth Marine Laboratory, Plymouth PL1 3DH, UK
[q] Great Lakes Environmental Research Laboratory, National Oceanic and Atmospheric Administration, Ann Arbor, MI, USA
[r] University of Valencia, Laboratory for Earth Observation, Valencia, Spain

ARTICLE INFO

Editor: Menghua Wang

Keywords:
Machine learning
Water quality
Inland and coastal waters
OLI
MSI
OLCI

ABSTRACT

Constructing multi-source satellite-derived water quality (WQ) products in inland and nearshore coastal waters from the past, present, and future missions is a long-standing challenge. Despite inherent differences in sensors' spectral capability, spatial sampling, and radiometric performance, research efforts focused on formulating, implementing, and validating universal WQ algorithms continue to evolve. This research extends a recently developed machine-learning (ML) model, i.e., Mixture Density Networks (MDNs) (Pahlevan et al., 2020; Smith et al., 2021), to the inverse problem of simultaneously retrieving WQ indicators, including chlorophyll-*a* (Chl*a*), Total Suspended Solids (TSS), and the absorption by Colored Dissolved Organic Matter at 440 nm ($a_{cdom}(440)$), across a wide array of aquatic ecosystems. We use a database of *in situ* measurements to train and optimize MDN models developed for the relevant spectral measurements (400–800 nm) of the Operational Land Imager (OLI), MultiSpectral Instrument (MSI), and Ocean and Land Color Instrument (OLCI) aboard the Landsat-8, Sentinel-2, and Sentinel-3 missions, respectively. Our two performance assessment approaches, namely hold-out and leave-one-out, suggest significant, albeit varying degrees of improvements with respect to second-best algorithms, depending on the sensor and WQ indicator (e.g., 68%, 75%, 117% improvements based on the hold-out method for Chl*a*, TSS, and $a_{cdom}(440)$, respectively from MSI-like spectra). Using these two assessment methods, we provide theoretical upper and lower bounds on model performance when evaluating similar and/or out-of-sample datasets. To evaluate multi-mission product consistency across broad spatial scales, map products are demonstrated for three near-concurrent OLI, MSI, and OLCI acquisitions. Overall, estimated TSS and $a_{cdom}(440)$ from these three missions are consistent within the uncertainty of the model, but Chl*a* maps from MSI and OLCI

* 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.2021.112860
Received 22 July 2021; Received in revised form 23 November 2021; Accepted 20 December 2021
Available online 4 January 2022

achieve greater accuracy than those from OLI. By applying two different atmospheric correction processors to OLI and MSI images, we also conduct matchup analyses to quantify the sensitivity of the MDN model and best-practice algorithms to uncertainties in reflectance products. Our model is less or equally sensitive to these uncertainties compared to other algorithms. Recognizing their uncertainties, MDN models can be applied as a global algorithm to enable harmonized retrievals of Chl*a*, TSS, and $a_{cdom}(440)$ in various aquatic ecosystems from multi-source satellite imagery. Local and/or regional ML models tuned with an apt data distribution (e.g., a subset of our dataset) should nevertheless be expected to outperform our global model.

1. Introduction

With the availability of free satellite images from multiple missions capable of making high fidelity radiometric measurements over water bodies, devising proper processing chains to produce harmonized global water quality (WQ) products is currently a high-priority research topic. Global WQ products are critical for addressing the pressing questions surrounding aquatic ecosystems threatened by a changing climate and ever-increasing human exploitation of natural resources. These challenges primarily pertain to aquatic ecosystem responses to shifts in climate and human activities and the disentanglement of the impacts of these two main drivers (e.g., wildfire (Kramer et al., 2020) or extended dry/wet periods (Mosley, 2015) versus landuse and landcover change (Baker, 2006)) across different eco-hydrodynamic regions. Landcover and landuse change, deforestation and reforestation (Mapulanga and Naito, 2019), or mining activities (Dethier et al., 2019; Lobo et al., 2015), for example, can alter the regional carbon cycle in aquatic ecosystems, thereby affecting aquatic life and water quality (Kutser et al., 2005). Further, operational utilities and water resource management agencies increasingly tend to rely on satellite assets for improved sampling schemes and informed decision-making (Binding et al., 2021; Schaeffer et al., 2018). At large spatial scales with seasonal variability, it is consequential for these end-users to access reliable satellite-derived WQ indicators applicable to a broad range of aquatic ecosystem conditions (e.g., various trophic states).

The optical WQ indicators considered in this study include chlorophyll-*a* (Chl*a*), total suspended solids (TSS), and Colored Dissolved Organic Matter (CDOM), measured as CDOM absorption at 440 nm ($a_{cdom}(440)$). These indicators, together with water and other non-water Inherent Optical Properties (IOPs), regulate the color of the water (Gordon et al., 1975), which is commonly quantified as water-leaving radiance ($L_w(\lambda)$; W m^{-2} sr^{-1} nm^{-1}) just above the water (Mobley, 1994). Because $L_w(\lambda)$ varies according to the illumination conditions, it is commonly normalized by the total downwelling irradiance and expressed as spectral remote sensing reflectance, i.e., $R_{rs}(\lambda)$, in units of sr^{-1} (Mobley, 1999). For conciseness, the spectral dependency is dropped hereafter.

Among the three WQ indicators targeted here, Chl*a*, the primary pigment found in all phytoplankton types, is a proxy for phytoplankton biomass whose shortage or excess leads to detrimental consequences on ecosystem function, health, and integrity. While phytoplankton absorb and scatter incoming solar irradiance, in open-ocean environments, it is the absorption by Chl*a,* with two major absorbing peaks at 443 and 670 nm, that primarily regulates the shape of R_{rs} in the visible spectrum (Kiefer and Mitchell, 1983). R_{rs} is further modulated by the absorption by other pigments, such as carotenoids and phycobiliproteins (Sathyendranath et al., 1989), leading to peaks and troughs throughout the visible spectrum (Bricaud et al., 1998). In optically complex waters, these spectral features are generally masked by the absorption and/or scattering induced by suspended particles as well as by a_{cdom}. Across global aquatic ecosystems, TSS, inclusive of all organic and inorganic suspended particles, influences turbidity. Its elevated quantities may limit the availability of light to benthic habitats, trap heat in the upper layers of the water column, thereby triggering harmful algal blooms (HABs) (Bukata et al., 1995). They can be associated with the transport of polluting substances including heavy metals and nutrients to waterways (Snyder and Carson, 1986). Its contribution to R_{rs} is through the absorption and scattering regimes and, in general, governs the overall magnitude of R_{rs} (Novo et al., 1989). CDOM may be considered as a proxy for dissolved organic carbon (Hestir et al., 2015; Kutser et al., 2015; Spencer et al., 2012), enabling regional and/or global assessments of the carbon budget and cycling in terrestrial and aquatic ecosystems (Duarte and Prairie, 2005; Roehm et al., 2009). Due to its light-absorbing characteristics, particularly in the ultraviolet (UV) and blue spectrum, an increase in CDOM concentration restricts light availability in the water column, hence aquatic life (Roulet and Moore, 2006). In coastal estuaries, a_{cdom} is regarded as a proxy for salinity, an important variable that mediates coastal biodiversity (Benlloch et al., 2002) and plays a major role in identifying aquaculture sites (Resley et al., 2006). Global quantification of changes in these optical WQ indicators over time can elucidate aquatic ecosystem responses to environmental stressors and inform the global environmental policy (Hakimdavar et al., 2020).

Global assessment of these indicators using field sampling techniques alone is not feasible. Despite decades of research and developments (Brezonik et al., 2015; Bukata et al., 1995; Gitelson, 1992; Odermatt et al., 2012), the potential of remote sensing for many aquatic ecosystems has not been fully realized, leaving behind a significant portion of global satellite observations Further, methods for generating reliable global products still deserve thorough studies. The objective of this research is to extend a previously developed machine-learning (ML) model, i.e., Mixture Density Networks (MDNs) (Pahlevan et al., 2020; Smith et al., 2021) to concurrently estimate Chl*a*, TSS, and $a_{cdom}(440)$ from the Operational Land Imager (OLI), MultiSpectral Instrument (MSI), and Ocean and Land Color Instrument (OLCI) aboard the Landsat-8, Sentinel-2, and Sentinel-3 missions, respectively. To the extent possible (owing to the differences in the spectral, spatial, and radiometric characteristics of these instruments), we aim to advance consistency among multi-mission products to deliver globally seamless WQ products. This development relies on an *in situ* database that contains hyperspectral radiometric measurements and co-located Chl*a*, TSS, and $a_{cdom}(440)$. The model performance is gauged against that of select, top-performing state-of-the-art algorithms. The model is then retrained using all the available *in situ* data for application to OLI, MSI, and OLCI images and assessment of WQ matchups.

2. The state of the practice

Optically relevant WQ indicators and inherent and apparent optical properties (IOPs, AOPs) have been collected over four decades using radiometers aboard vessels, airborne platforms, or satellites. This has generated considerable advances in our ability to estimate Chl*a* (Gitelson et al., 2007; Gower et al., 2005; Moses et al., 2012a; O'Reilly et al., 1998), TSS (Doxaran et al., 2002; Knaeps et al., 2015; Petus et al., 2010), or a_{cdom} (Kowalczuk et al., 2003; Nieke et al., 1997). Recent and upcoming launches of satellite missions continuously improve spatiotemporal coverage of aquatic environments and produce higher quality radiometric measurements. This has spurred the development of new high-quality WQ products (Balasubramanian et al., 2020; Cao and Tzortziou, 2021; Jiang et al., 2021; Smith et al., 2021) and algorithms that operate across different sensors (Pahlevan et al., 2020). Concurrently with enhancements in satellite data, the body of high-quality *in*

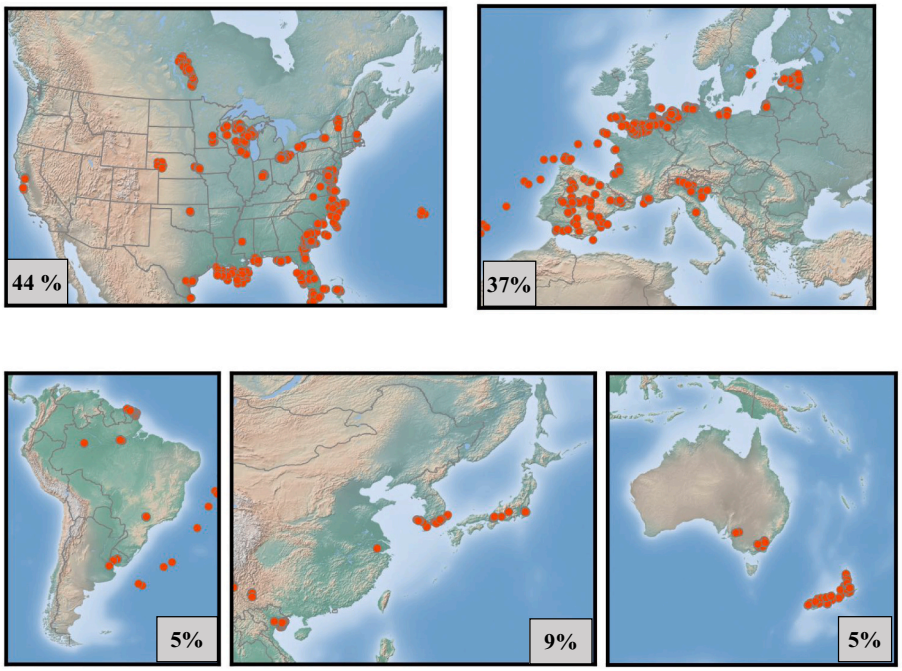

Fig. 1. Continental distribution of the development data where *in situ* R_{rs} and WQ indicators were measured. The percentage of contributed data from each continent is also provided for the data applicable to MSI (see text). Note that 63 samples from South Africa have not been illustrated (see Table A).

situ radiometric and biogeochemical data has increased (Spyrakos et al., 2018), offering more realistic representations of natural waters in comparison to synthetic data that has long served as the backbone of similar efforts in the ocean color domain (IOCCG, 2006).

Similar to other geophysical problems, estimating Chl*a*, TSS, and $a_{cdom}(440)$ from multispectral R_{rs} is by definition an inverse problem, where a function of form f^{-1} is sought to predict a vector x given R_{rs}, i.e., $x[\text{Chl}a, \text{TSS}, a_{cdom}(440)] = f^{-1}(R_{rs})$. Note that here we assume IOP components are not measured or unavailable. Given the inherent co-variations and nonlinearities of IOPs and the associated substance concentrations in global inland and coastal waters (Bukata et al., 1995; Doerffer and Fischer, 1994; IOCCG, 2000, 2006, 2018; Mobley, 1994; Zheng et al., 2015), a simultaneous retrieval scheme may exploit the fact that various natural combinations of these variables differently drive the magnitude and shape of R_{rs}. For optically complex waters, such schemes fall into three families of algorithms: a) Linear Matrix Inversion (LMI) (Brando and Dekker, 2003; Giardino et al., 2007; Hoogenboom et al., 1998), b) spectral matching based on look-up-tables (Arabi et al., 2020; Concha and Schott, 2016; Gerace et al., 2013; Raqueno, 2003), and c) machine-learning models (Doerffer and Schiller, 2007; Hieronymi et al., 2017; Keiner and Yan, 1998; Tanaka et al., 2004). All these approaches have shown promise for estimating WQ indicators and, in some cases, for IOPs in inland and nearshore coastal waters; however, most developed models have been evaluated at local or regional scales only, due to the lack of globally representative development data (e.g., the absence of specific IOPs in case of LMI). Among these methods, the Multi-Layer Perceptron (MLP) neural network (Doerffer and Schiller, 2007) trained with a Hydrolight simulated dataset (Mobley and Sundman, 2008) has been adopted as the standard processing of OLCI data over inland waters (Toming et al., 2017) and applied to OLI (Hafeez et al., 2019) and MSI (Ansper and Alikas, 2019; Brockmann et al., 2016; Soomets et al., 2020). Instead of using simulated data, our study leverages a sizable database of *in situ* data for training and testing of an MDN for the simultanous estimation of Chl*a*, TSS, and $a_{cdom}(440)$ at a global scale.

3. Datasets

3.1. Development data

A large portion of our database (R_{rs} – Chl*a* and R_{rs} – TSS) has been described in previous publications (Balasubramanian et al., 2020; Jiang et al., 2021; Pahlevan et al., 2020; Smith et al., 2021). For this study, we added new co-located samples from Australia, New Zealand, the Netherlands, Spain, Uruguay, and Brazil (Fig. 1). This dataset has been further augmented by *in situ* R_{rs} – $a_{cdom}(440)$ measurements. Our most complete dataset includes paired R_{rs} and Chl*a* samples that correspond to 35 different sources, each encompassing data from one or multiple campaigns from a single principal investigator (PI). Our TSS and $a_{cdom}(440)$ datasets came from 26 and 17 sources, respectively. The details on the geographic scope of these sources are provided in Appendix A. Note that the data from a single PI delivered at different stages were treated as separate sources, leading to a total of 42 different sources (Appendix A).

The frequency distributions of each WQ component and their respective correlations are shown in Fig. 2. It is these covariances that our algorithm uses to determine the inverse function and produce the three indicators (Section 4). For our development, the measured hyperspectral R_{rs} were resampled with the four, seven, and 12 Relative Spectral Responses (RSRs) of OLI, MSI, and OLCI within the 410–800 nm spectral range, respectively (see Appendix B for more details). The spectral coverage of our R_{rs} spectra fell into two categories; the ones covering 400–700 nm and those that had extended coverage from 400 through 800 nm. For this reason, the number of samples for the spectral band configurations of the sensors differed. OLCI's band centered near 400 nm was not considered here due to the small number of *in situ* spectra encompassing the UV range.

3.2. Satellite data and matchups

To demonstrate the quality of our WQ map products, three instances of near-simultaneous images acquired by OLI, MSI, and OLCI over two major estuaries and one hypereutrophic lake (Section 5.2) were

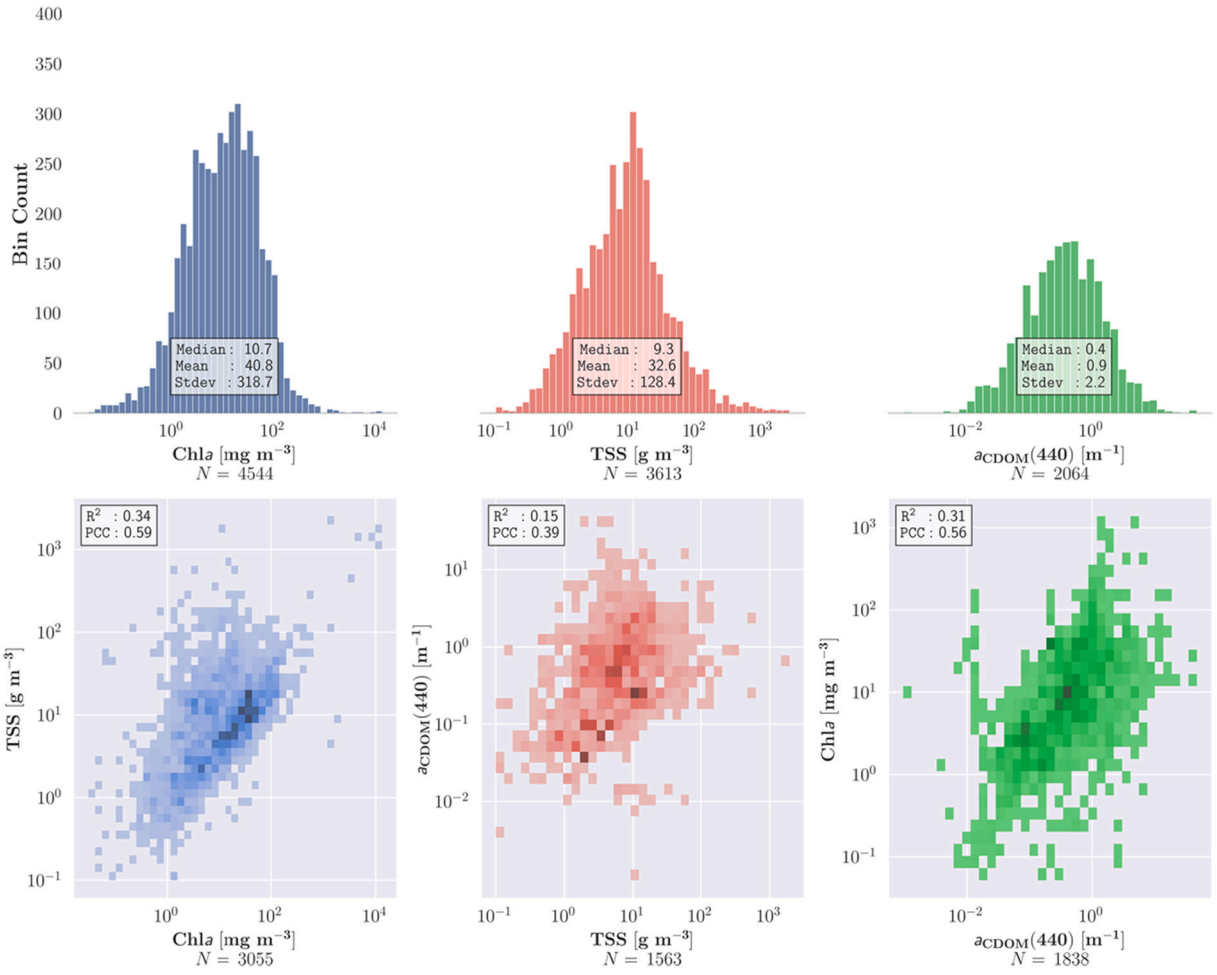

Fig. 2. Frequency distribution (top) of WQ indicators for which co-located R_{rs} spectra covering the MSI visible-Near-InfraRed spectral bands (< 800 nm) were available. 2D histograms (bottom) exhibit correlations between co-located sample pairs. The number of available data samples was the highest for OLI (410 < λ < 700 nm) and slightly lower for OLCI due to a broader spectral coverage. The R-squared (R^2) and Pearson Correlation Coefficient (PCC) are annotated to give insights into the correlations between WQ indicators.

processed via the Atmospheric Correction for OLI 'lite' (ACOLITE) (Vanhellemont and Ruddick, 2021). These scenarios were selected because they represent a wide range of WQ conditions (e.g., 0.5 < Chl*a* < 100 mg m^{-3}). While no ideal atmospheric correction (AC) processor for generating seamless R_{rs} from these missions has yet become available (Pahlevan et al., 2021a; Warren et al., 2019), ACOLITE (v20190326.0) is a convenient processor for inland and nearshore coastal waters and allows for retrieving R_{rs} under highly turbid or eutrophic optical regimes, thereby making it a suitable choice for demonstration purposes.

To understand how uncertainties in both Top-of-Atmosphere (TOA) observations and AC processing manifest in derived R_{rs}, i.e., uncertainties in R_{rs} - hereafter referred to as σ(R_{rs}), and subsequent MDN products, OLI and MSI images at WQ matchup locations were reduced to R_{rs} via two AC processors, namely the SeaWiFS Data Analysis System (SeaDAS v7.5.3) and ACOLITE. Per Pahlevan et al. (2021a), these two processors perform relatively well in nearshore coastal and inland waters, respectively. For SeaDAS, the Near-InfraRed (NIR) and ShortWave InfraRed (SWIR) band combinations were utilized to account for aerosol contribution, whereas ACOLITE was used with its default setting (dark spectrum fitting). Because of MSI's low signal-to-noise ratio in the NIR and SWIR bands (Pahlevan et al., 2017a), SeaDAS did not return as many matchups as ACOLITE when MSI imagery were processed. This issue however does not apply to OLI (Pahlevan et al., 2021a). Here, median satellite-derived WQ estimates within 3 × 3-element windows were computed.

In situ WQ data for this matchup exercise were obtained from various databases across North America. A majority of the Chl*a* matchup dataset has been employed in previous studies (Pahlevan et al., 2021a;(Smith et al., 2021)). This study added TSS and $a_{cdom}(440)$ matchups. The TSS data were mostly available through the U.S. Water Quality Portal (https://www.waterqualitydata.us) and Geological Survey's National Water Information System (https://nwis.waterdata.usgs.gov/nwis) that contain data from small inland waters, as well as through the Chesapeake Bay repository (https://www.chesapeakebay.net) with data mostly ranging from 0.1 to 20 g m^{-3}. These datasets were further augmented by the repository provided in Ross et al. (2019). Several filtering techniques were executed to exclude duplicate data records. More Chl*a* and $a_{cdom}(440)$ data were made available through the Natural Science and Engineering Research Council of Canada's (NERC) Lake Pulse Network (LPN), an extensive multi-year field campaign conducted across select drinking water supplies and lakes in Canada (Huot et al., 2019). Additional $a_{cdom}(440)$ datasets acquired in Lake Winnipeg (Manitoba, Canada) and across several lakes in Minnesota were contributed by Environment and Climate Change Canada (ECCC) and the University of Minnesota (Olmanson et al., 2020), respectively. For Chl*a* and TSS matchups, we permitted a +/− 3 h time window from the satellite overpass for coastal matchups and a same-day overpass for inland waters. Given the known stability of $a_{cdom}(440)$ over extended periods at our matchup sites, i.e., boreal lakes (Cardille et al., 2013), and to incorporate an adequate number of matchups, we adopted a +/− 3

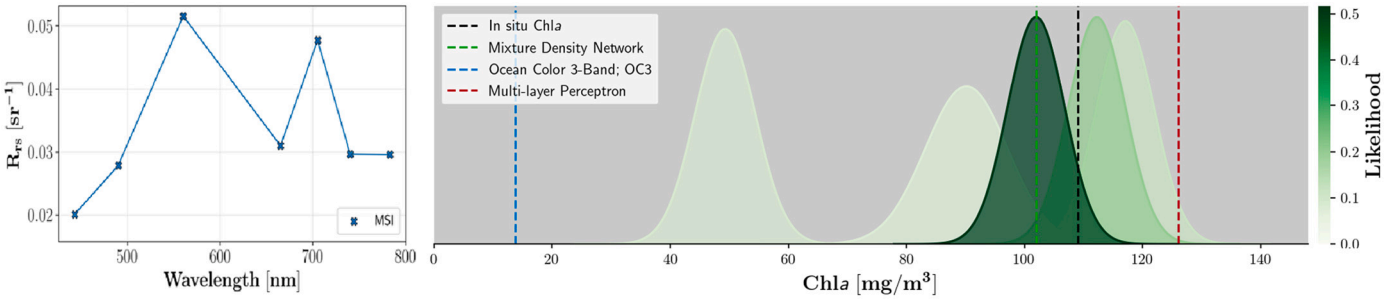

Fig. 3. A sample R_{rs} spectrum resampled with MSI RSRs (left) is shown with five Gaussian distributions output by an MDN model (Pahlevan et al., 2020) for estimating Chl*a*. The maximum likelihood (dark green curve) is utilized to make the final Chl*a* approximation through the MDN. Chl*a* estimates by an MLP (Smith et al. 2021) and the OC3, as well as the *in situ* measured Chl*a* are also denoted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

day time window for this parameter (Brezonik et al., 2015). Note that a large portion of our dataset originates from narrow ecosystems or nearshore coastal waters making it difficult to obtain enough matchups for OLCI's 300 m nominal spatial resolution.

4. Methods

MDNs differ from traditional neural networks (Bishop, 1995; Bricaud et al., 2007; Jamet et al., 2012) in that they produce the likelihoods of generated estimates (Bishop, 1994). Succinctly, MDNs learn a probability distribution over the output space, thereby accomodating multimodal target distributions. This multimodality is a fundamental characteristic of inverse problems, where a non-unique relationship exists between input and output features. Using the probabilities estimated for each individual prediction, users may choose the maximum likelihood estimate (the prediction with the highest probability) or the weighted average of all predictions (e.g., see five likelihood functions in Fig. 3). Similar to previous publications, the model input and output features encompass *in situ* R_{rs} and *in situ* WQ indicators (Section 3.1), respectively. Adopted from Pahlevan et al. (2020), an example for a Chl*a* estimation from an arbitrary MSI-like R_{rs} spectrum is illustrated in Fig. 3, with Gaussian mixtures output by the model and the maximum likelihood estimate representing the final prediction. For comparison, estimates from a conventional MLP (Smith et al., 2021) and the Ocean Color 3-band algorithm (OC3) (O'Reilly and Werdell, 2019) are also depicted. This multimodal representation within the target space allows for enhanced predictions compared to other ML models. Traditional ML models like MLPs or other empirical Chl*a* models report a single estimate without insights into the distribution of estimates.

Further, Smith et al. (2021) have shown that for a Chl*a* retrieval task from OLI observations, MDNs outperform traditional MLPs by 20 to 30%, an improvement attributable to MDNs' ability in learning conditional probabilities. MDNs inherently learn the covariances among the output features, intuitively improving the accuracy over models intended for retrieving an individual parameter (Bishop, 1994). In general, multi-parameter inversion algorithms for biogeochemical properties and/or IOPs (e.g., Kallio et al. (2001)) are anticipated to constrain the solution space given the covariances among parameters of interest in a natural environment (Fig. 2). This study leverages this characteristic to estimate optically relevant WQ indicators through MDNs.

4.1. Estimating missing values

For generating simultaneous WQ indicators, several modifications were made to the basic MDN model. First, the data itself necessitates special handling, due to hundreds of missing samples from one or two categories (Fig. 2). For instance, within the *in situ* Chl*a* dataset suitable for OLI, there are 5680 available measurements out of a total of 6475 samples. For TSS and $a_{cdom}(440)$, the number of available measurements is even smaller (3847 and 2910 samples, respectively), with only 1613 samples encompassing all three quantities - just over 20% of the total dataset. To make use of the remaining 80% of the data with missing values, a method is needed to allow the MDN model to learn from partial information.

To fill in the missing values, we apply a method referred to as multiple imputation (Rubin, 2004), which has been utilized in inverse problems related to geophysical parameter retrievals (Hudak et al., 2008; Junninen et al., 2004; Kalteh and Hjorth, 2009). Simply put, the missing values are replaced by values randomly drawn from a distribution that represents the dataset. After *m* values are drawn, each of those values is independently used to complete the analysis as if the value were not missing and subsequently pooled into a final analysis output. This process assumes the missing values are either Missing Completely At Random (MCAR), or as a slightly weaker assumption, Missing At Random (MAR) (King et al., 2001). Concretely, this means that we are assuming a missing Chl*a* is not dependent upon the Chl*a* value itself - it is, at worst, dependent only on the other *available* measurements for that sample. Suppose the missing value is instead conditionally missing upon its value (or the value of other missing variables). In that case, the data is considered Missing Not At Random (MNAR) and requires additional care in handling (King et al., 2001).

MDNs inherently model the joint probability of all output variables, leading to a multiple imputation approach which is capable of handling even MNAR variables in certain circumstances (in addition to the usual MAR or MCAR) by allowing us to draw directly from the learned posterior (Buuren and Groothuis-Oudshoorn, 2010; Galimard et al., 2018). Moreover, the multiple imputation estimates are improved continuously during the training of the MDN, i.e., as the MDN learns to better model the target variables, the accuracy of the imputed samples improves, thus leading to further improvements in the target variable estimates. Separate parts of this type of chained bootstrap learning have been documented in the literature (Ghahramani and Jordan, 1995).

4.2. Model architecture and hyperparameters

Brief experiments were performed to gauge how much improvement might be gained through hyperparameter optimization of the model, utilizing Amazon Sagemaker's Bayesian Search method to quickly search through the large space. It was found that the added complexity did not justify the relatively minor gains in performance, however, which is in line with the results shown previously (Smith et al. 2021). Instead, we used the configuration outlined in antecedent work to create models for all performance assessments: five Gaussian mixtures (Fig. 3), defined by a network using five hidden layers with 100 nodes in each layer; a learning and L2 normalization rate of 0.001 (a regularization term with model weights squared); and a bagging ensemble of ten models, over which the median is taken. See previous publications (Smith et al. 2021; Pahlevan et al., 2020) for additional details on the

Table 1
Retrieval algorithms assessed in this study.

WQ indicator	Algorithm						
Chl*a*	MDN	GI2B* (Gilerson et al., 2010)	Gons (Gons et al., 2002)	GU2B* (Gurlin et al., 2011)	Blend (Smith et al., 2018)	Moses-2B* (Moses et al., 2012b)	OC3 (O'Reilly and Werdell, 2019)
TSS		SOLID (Balasubramanian et al., 2020)	Novoa (Novoa et al., 2017)	Nechad (Nechad et al., 2010)	Petus Petus et al., 2010)	Miller & McKee (Miller and McKee, 2004)	
$a_{cdom}(440)$		Ficek (Ficek et al., 2011)	QAA-CDOM (Zhu and Yu, 2012)	Mannino (Mannino et al., 2008)			

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

model implementation.

4.3. Model evaluation

Per previous literature (Balasubramanian et al., 2020; Pahlevan et al., 2021b; Pahlevan et al., 2020;), several robust single-parameter retrieval methods applicable to OLI, MSI, and OLCI with their original form and coefficients were considered (Table 1). While >20 algorithms for Chl*a* and TSS estimations were initially examined, only a few $a_{cdom}(440)$ models viable for a wide range of CDOM content were

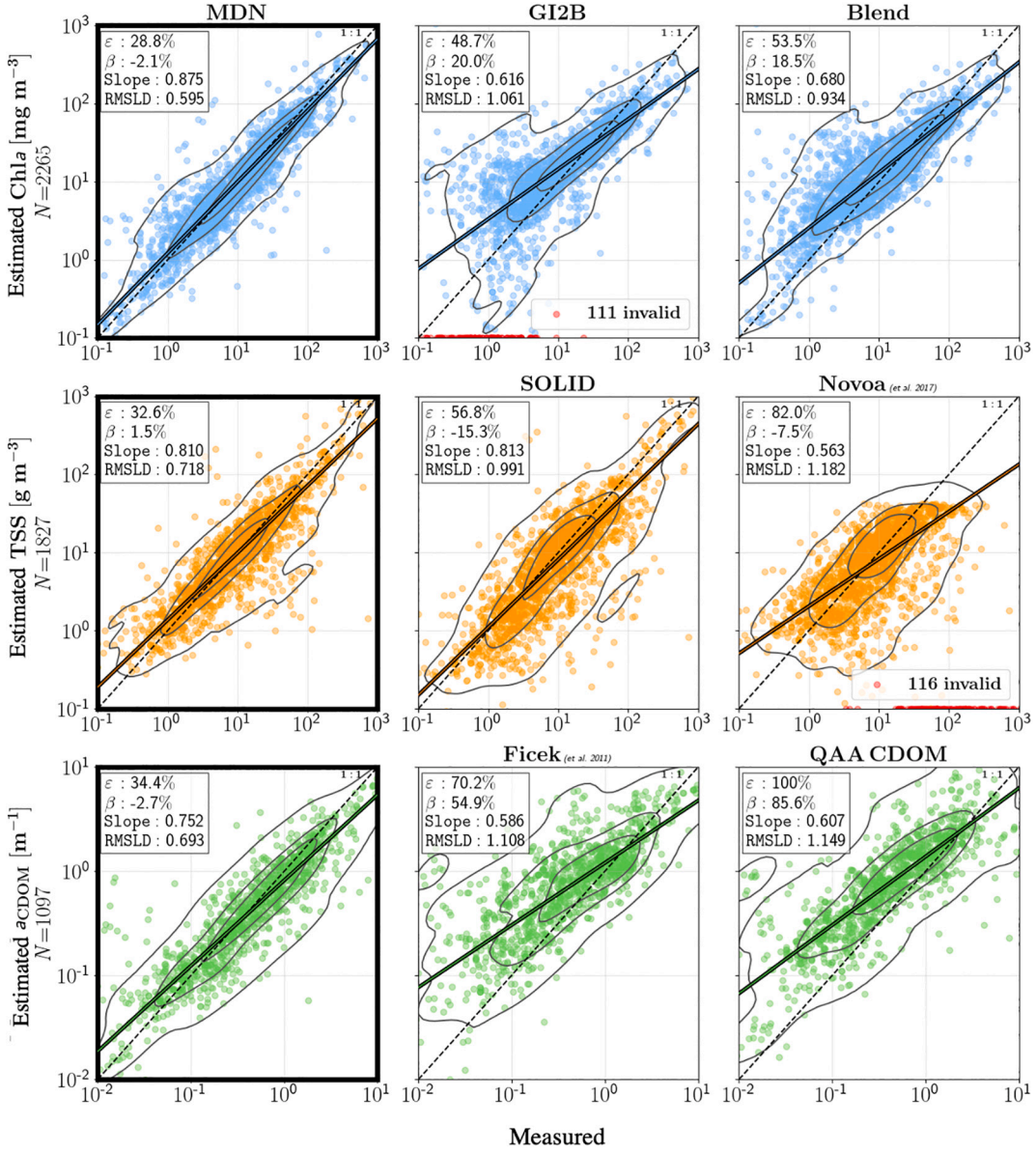

Fig. 4. Scatterplots showing the performances of MDN and best-performing state-of-the-art models for R_{rs} spectra resampled to MSI RSRs. These analyses correspond to a hold-out assessment approach with a 50–50% data split for which actual sample numbers are shown on the y-axes. See references noted in Table 1. Contour lines are presented to better illustrate data distributions. Red dots and the annotated values refer to invalid (negative) estimates. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

identified. Prior experiences with algorithm tuning of similar single-parameter models (O'Shea et al., 2021) suggest that due to non-linear covariations between R_{rs} and WQ indicators (Defoin-Platel and Chami, 2007), adjusting one or two single (fixed) coefficients does not offer significant performance improvements when a globally applicable model is sought. On the other hand, most multi-parameter retrieval algorithms suitable for inland and nearshore coastal waters (e.g., LMI) would have required further calibration data to adapt their specific IOPs unavailable for most of our dataset. When possible, equations for the empirical models listed in Table 1 are provided in Appendix C. It should be noted that generic ocean-color algorithms (e.g., default Generalized IOP model; (Werdell et al., 2013) have been attempted in prior research efforts (Pahlevan et al., 2021b) with (expectedly) little success.

The two primary metrics used to gauge the accuracy of retrievals are the median symmetric accuracy and signed systematic percentage bias (Morley et al., 2018) referred to as uncertainty (ε) and Bias (β) throughout:

$$\varepsilon = 100 \times \left(e^{median(\,|\,log(\,e/o)\,|\,)} - 1\right) \quad (1)$$

$$MR = median(\,log(\,e/o)\,) \quad (2)$$

$$\beta = 100 \times sign(MR) \times \left(e^{|\,MR\,|} - 1\right) \quad (3)$$

where e and o stand for estimated and *in situ* observed WQ indicators, respectively, and MR is median ratio. These metrics are similar to those applied in Seegers et al. (2018) with slight modifications for enhanced interpretability and robustness. Previous studies (Smith et al. 2021) have shown that these median measures, along with the slope of linear regressions (Type I), and root mean squared logarithmic difference (RMSLD), are adequate, offering a comprehensive perspective for evaluating algorithms.

Two different schemes for the performance assessment were adopted. The first method was the commonly used hold-out method that entails a random split of the entire dataset. Similar to our prior research (Pahlevan et al., 2021b), we chose a 50–50 random split to allow equal datasets for testing and training. This approach, however, may favorably impact the model performance as the test data may not be entirely independent of the training samples (e.g., data acquired in one lake through one or multiple field campaigns may be correlated). To overcome this drawback and provide a more realistic (or even pessimistic) view of the performance, our second assessment strategy was the leave-one-out (LOO) approach where the model is trained with all the data sources (Appendix A) except one against which its performance is evaluated. This process is iterated until each data source is utilized once as a test set. The ultimate performance for MDN and the rest of the benchmarking algorithms is reported as the median or weighted average of all the individual performances ($N = 35$ for Chl$a$ in Fig. 5). This approach would deduce the expected performance of the model when an out-of-sample (unseen) dataset is applied to the model. In general, these two approaches (hold-out and LOO) are anticipated to provide a practical range for algorithm uncertainties and elucidate how well the model can estimate optical WQ indicators in globally diverse aquatic ecosystems.

Due to the necessity of training an MDN for each data source in the LOO scheme, we used the Amazon SageMaker Studio (https://aws.amazon.com/sagemaker/) to train the models simultaneously. Thus, each data split was sent as a separate training job ten times, with a different random seed used to initialize network weights and training data shuffling for each of these ten trials. The results of the ten trials were then averaged, and the final performance for each data source is used in the overall LOO weighted performance assessment (Table 3). In total, this led to several hundred training jobs necessary for each sensor.

Table 2
Performance metrics for MDN based on the hold-out assessment for the spectral band configurations of OLI, MSI, and OLCI.

	WQ indicator	ε [%]	β [%]	Slope []	RMSLD []
OLI	Chl$a$	59.1	−9.5	0.77	0.86
	TSS	40.6	7.5	0.78	0.74
	$a_{cdom}(440)$	37.9	−5.4	0.80	0.53
MSI	Chl$a$	29.2	−3.4	0.88	0.62
	TSS	32.1	1.4	0.80	0.70
	$a_{cdom}(440)$	34.5	−1.6	0.73	0.72
OLCI	Chl$a$	26.2	0.2	0.90	0.58
	TSS	29.9	3.6	0.80	0.67
	$a_{cdom}(440)$	26.8	1.2	0.81	0.65

5. Results

5.1. Performance assessment

The performance assessment for MDN and two other best-performing algorithms based on the hold-out method implemented to MSI resampled R_{rs} spectra (Section 3.1) is presented in Fig. 4. The MDN model shows negligible biases and moderate uncertainties across the three WQ indicators. Among all the tested Chl$a$ and TSS algorithms, Blend (Smith et al., 2018) and SOLID (Balasubramanian et al., 2020) are the second-best performing algorithms, respectively. Notably, the MDN model predicts TSS with virtually no bias. For $a_{cdom}(440)$, MDN outperforms the other two algorithms with nearly half ε and RMSLD. Further, uncertainties in MDN estimates of $a_{cdom}(440)$ are highly uniform across the dynamic range spanning three orders of magnitude, whereas Ficek and QAA-CDOM approximations are only optimal for $a_{cdom}(440) \sim 1\ m^{-1}$. Reported Slopes slightly less than unity suggest degraded performance at the two tails of data distributions (e.g., Chl$a$ < 1 mg m^{-3} and Chl$a$ $>$ 100 mg m^{-3}) which may be alleviated with a larger dataset or by augmenting the existing dataset to fill in the gaps further. The equivalent metrics for OLI- and OLCI-like spectra are shown in Table 2. While the uncertainty measures from the MSI and OLCI spectra are similar, the estimates through OLI-like band settings contain uncertainties 30 to 100% larger for the three variables. The reason for the lower accuracy of Chl$a$ estimates is OLI's lack of spectral bands within the 700–800 nm range (Gitelson, 1992), underscoring that MSI and OLCI spectral band configurations in this region introduce more degrees of freedom for quantifying TSS and $a_{cdom}(440)$, despite $a_{cdom}(440)$ being a UV-blue absorbing component (Kirk, 1994). It should also be noted that the OLCI's spectral band configuration appears to enable 10–30% more accurate products over those from MSI, which corroborates the benefits of additional spectral bands for the optical sensing of WQ indicators (Lee et al., 2007).

The LOO assessment of MDN and other best-performing algorithms for an MSI-like instrument is provided in bar charts for all the available data sources and each WQ component in Fig. 5. Again, MDN produces the most accurate estimates for most datasets. There are exceptions to that, however, where other algorithms like GI2B, Moses-2B, SOLID, and Ficek outperform MDN - though it should be noted that some of these occasionally generate invalid retrievals (e.g., GI2B for Chl$a$ < 8 mg m^{-3}), rendering them less ideal for global applications. We note that there are datasets for which predictions are associated with considerable uncertainties (e.g., Dataset #39 for Chl$a$ and TSS). For these scenarios, MDN typically provides the best approximations. The results do not imply that the dataset is of lower quality; instead, they may signify that the range of measured values from a specific dataset is not well represented within the rest of the dataset.

A summary of the LOO performance assessment for the MDN model and the second-best algorithms in terms of a median and weighted average of ε is included in Table 3. Here, medians correspond to the median of all $\boldsymbol{\varepsilon}$ from all the datasets, and weighted averages represent sum of all $\boldsymbol{\varepsilon}$ weighed by the number of samples per dataset normalized

Fig. 5. Performance assessments based on the leave-one-out (LOO) experiments for an MSI-like mission for Chl*a*, TSS, and $a_{cdom}(440)$. See Appendix A for more details on the data sources, indices, and corresponding average R_{rs} spectra (Figure A). Equations for the state-of-the-art algorithms are provided in Appendix C.

by N. The reported uncertainties are often >50% higher than those denoted in Table 2 and bear similar conclusions. For example, the performance metrics suggest that OLCI TSS and $a_{cdom}(440)$ products are expected to be of higher quality than those from MSI and OLI. Analogous to the hold-out analyses (Table 2), it is found that OLI can produce viable TSS and $a_{cdom}(440)$ estimates - although the derived Chl*a* in highly eutrophic waters may be questionable, with more than 80% estimated uncertainties. The large differences in median and weighted average pertain to an unequal sample size of different data sources. For instance, large uncertainties (ε) for small datasets (<10 samples) do not significantly contribute to ultimate uncertainties when reporting weighted average ε, whereas median ε (computed over all datasets) is equally affected by uncertainties in each data source.

Among the three sensors and WQ indicators, the largest uncertainties ranging from 47 to 85% correspond to $a_{cdom}(440)$, which may be attributable to the availability of fewer samples and a possible presence of outliers. Moreover, $a_{cdom}(440)$ does not posses uniquely diagnostic features in R_{rs}, and so depends on how well the other components can be estimated, pooling uncertainty into $a_{cdom}(440)$ estimates. Collectively, aside from uncertainties in Chl*a* estimates from OLI, the average ε for all the sensors and WQ indicators for MDN are ~49 and ~ 65% (3rd and 4th columns in Table 3), which, together with the average ε from the hold-out assessment, provide insights into the expected range of uncertainties, i.e., 26 to 49%, or 26 to 65%, associated with the MDN predictions. Whether median or weighted average uncertainties are examined, MDN WQ estimates are more accurate than those of the second-best models by an average of 11 and 53%, approximated from the mean ratios of columns 3–5 and 4–6 in Table 3, respectively.

5.2. Visual assessments

Three near-coincident OLI, MSI, and OLCI WQ maps produced via the MDN model over the San Francisco Bay, Chesapeake Bay, and Upper Klamath Lake (Oregon, USA) are evaluated for multi-mission product consistency. To simplify this analysis, we only focus on ACOLITE R_{rs} products.

Table 3
Performance metrics (ε[%]) for MDN and second-best models corresponding to the LOO assessment for the spectral band configurations of OLI, MSI, and OLCI. N corresponds to the number of samples used for assessing MDN performance. Reported medians correspond to median of all ε from all the datasets, and weighted averages represent the sum of all ε weighed by the number of samples per dataset normalized by N.

		MDN		Second-best model (OC3/GI2B, SOLID, Ficek)			
		Median [%]	Weighted Average [%]	Median [%]	Weighted Average [%]	Number of data sources	N
OLI	Chl*a*	89.77	102.80	138.2*	126.3*	38	5577
	TSS	44.71	66.08	67.3	76.8	28	3847
	$a_{cdom}(440)$	60.75	85.83	67.9	203.4	23	2910
MSI	Chl*a*	48.30	42.79	52.2**	62.7	35	4441
	TSS	48.23	62.20	58.5	73.0	26	3613
	$a_{cdom}(440)$	55.03	91.45	54.2	184.4	19	2064
OLCI	Chl*a*	46.30	43.70	47.1**	58.6	35	4366
	TSS	45.12	59.41	58.2	80.2	26	3514
	$a_{cdom}(440)$	47.70	87.45	46.7	166.3	20	1984
Sensor-average	Chl*a*	47.30***	43.24***	47.12***	60.65***	35	4198
	TSS	46.02	62.56	58.25	76.60	26	3526
	$a_{cdom}(440)$	54.55	88.24	50.30	175.35	19.5	1769
Average		49.29	64.61	51.88	104.20	26.9	3164

* Performance is associated with OC3.
** Performance is associated with GI2B for which >850 samples (~ 28% of N) returned invalid Chl*a* estimates.
*** OLI statistics are excluded.

The images over the San Francisco Bay (Fig. 6) were acquired on March 16th, 2019, under clear sky conditions with OLI and OLCI imagery taken ~3 min apart, followed by MSI passing over the region ~30 min later. There appears to be a reasonable agreement among the WQ maps, verifying the potential of MDN for producing congruous, practical multi-mission products. Although the OLI-derived Chl*a* map resembles those of MSI and OLCI in the lower section of the bay, a closer examination of the maps suggests a higher degree of consistency in MSI and OLCI products, which aligns with the assessments in Section 5.1. A notable illustration of the advantage of the higher spatial resolution of OLI and MSI imagery is the identification of the South Bay salt ponds (Warnock et al., 2002) with very high Chl*a* estimates. We also note more resemblance in OLI and MSI TSS maps than in those of MSI-OLCI or OLI-OLCI. For instance, in the mid-bay area (east of San Rafael), a major gradient in TSS is captured by OLI and MSI, whereas the OLCI map shows a more uniform distribution across this region. The between-sensor consistency is highest for the $a_{cdom}(440)$ products in the northern section of the bay; spatial patterns show the largest concentrations (~ 2 m^{-1}) in the Sacramento River outlet region and much lower values (0.2–0.3 m^{-1}) in the southern section of the bay. In the lower bay, the three maps agree well by predicting values on the order of 1 m^{-1}. *In situ* measured Chl*a* and TSS data (Schraga et al., 2020) available through the U.S. Geological Survey's field cruise (https://sfbay.wr.usgs.gov/water-quality-database/) two days prior to the overpass date are in general in agreement with those depicted in our maps. The reported near-surface Chl*a* and TSS measurements in the South Bay at station #24 were ~10.7 mg m^{-3} and 5 g m^{-3}, respectively, while they are 9 mg m^{-3} and 7 g m^{-3}, respectively, for station #30.

The mid-section of Chesapeake Bay is mapped by the three missions ~15 min apart from each other on Nov 7th, 2016 (Fig. 7). In the main stem of the bay, OLI and MSI exhibit very similar spatial patterns of Chl*a,* with OLCI showing slightly elevated concentrations in the bay and along the Potomac River. A single point measurement by the Maryland Department of Natural Resources (MD DNR) within the Patuxent River discharge (CB5.1 W) on the same date at ~12:20 UTC (> 3 h before satellite overpasses) indicates 6.2 mg m^{-3} of Chl*a* for which OLI, MSI, and OLCI approximate 4.0, 2.9, and 9.4 mg m^{-3}, respectively. The between-sensor consistency is better in the TSS maps, where the most apparent difference appears to be that OLI and MSI resolve more subtle spatial features which are diffused by OLCI's coarser spatial sampling. For the same MD DNR matchup with a measured TSS of 7.4 g m^{-3}, TSS has been approximated as 4.0, 2.6, and 4.9 g m^{-3} by OLI, MSI, and OLCI, respectively The $a_{cdom}(440)$ products also match each other quite well with concentrations ranging from 0.2 to 0.3 m^{-1} along the main stem of the bay with OLCI maps illustrating lower CDOM levels. Of further note are the banding effects in MSI products caused by changes in view azimuth angles across odd and even focal plane modules common to pushbroom images (Pahlevan et al., 2017b).

Shown in Fig. 8 are the maps from July 29th, 2019, over the Upper Klamath Lake (Oregon), a eutrophic ecosystem known for its major summertime cyanobacteria HAB events (Bradbury et al., 2004). Here, full OLCI lake-wide maps via ACOLITE were not likely producible due to the instrument line-of-sight's turbid atmospheric conditions. The reported near-surface Chl*a* data measured the following day by the USGS Oregon Water Science Center at three locations show concentrations varying within the 27–67 mg m^{-3} range, suggesting that our estimates are a realistic representation of lake conditions. Chl*a* at sites RS, WR, and MN on July 30th were reported as 57.0, 27.4, and 66.7 mg m^{-3}, respectively. The differences in MSI and OLCI products for these three locations lie between 14 and 35%. As in previous examples, $a_{cdom}(440)$ maps appear more consistent than the other two indicators. For instance, there are major distinctions between the OLI TSS map and those produced from MSI and OLCI that agree well with the retrieved spatial patterns. This implies that OLI's lack of spectral information within the 700–800 nm range may render the derived TSS or Chl*a* products inaccurate in hypereutrophic waters. In addition, differences in estimated R_{rs} spectra can introduce large uncertainties in predicted WQ products. The elevated MSI TSS estimates, for example, can be ascribed to such inconsistencies that drive inaccurate retrievals for one or all sensors (see Section 6 for discussions).

5.3. Matchup analysis

The performance of MDN in practical terms for MSI and OLI matchups has been illustrated in Figs. 9, 10, and 11. Although the Chl*a* matchups cover a broad range of trophic states (0.1 < Chl*a* < 100 mg m^{-3}), the range and distribution of TSS and $a_{cdom}(440)$ are relatively marginal and/or sparse. Therefore, in this subsection, for Chl*a*, we only present MSI matchups. For TSS and $a_{cdom}(440)$, because of the availability of larger sample sizes since April 2013, only OLI matchups are analyzed. Although OLI's performance for Chl*a* estimation is subject to large uncertainties (Table 2), we also report OLI Chl*a* matchup analyses in Appendix D (Figure D). Recall that the purpose of this analysis is to gauge the sensitivity of MDN and selected models (Table 1) to $\sigma(R_{rs})$ from SeaDAS and ACOLITE, i.e., a direct comparison of the two processors is not intended. In addition, the number of matchups for each processor and/or sensor is different, making such direct comparisons futile. We further note that OLCI's coarser spatial resolution

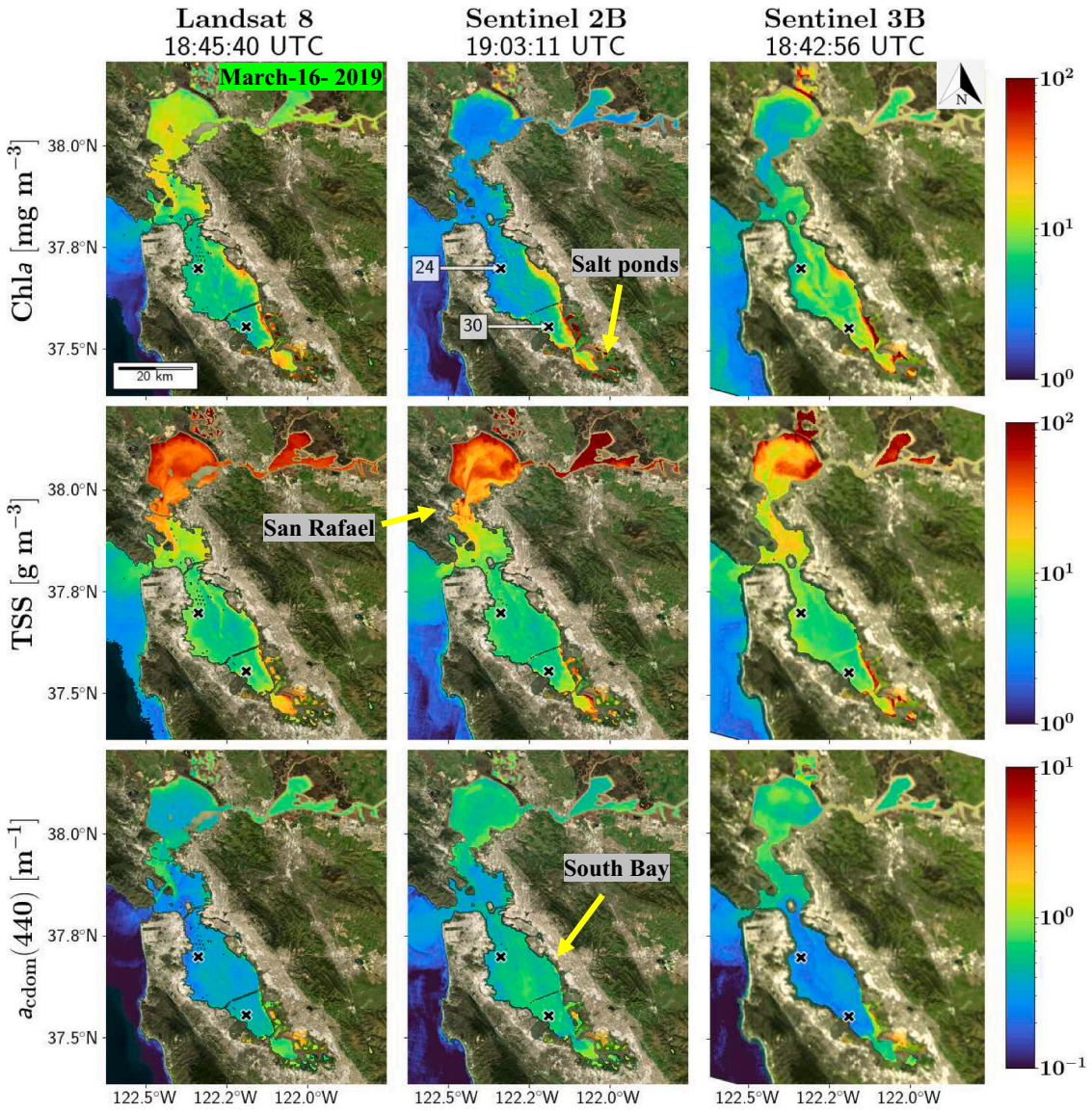

Fig. 6. Water quality maps were produced via MDN for near-concurrent images of OLI, MSI, and OLCI onboard Landsat-8, Sentinel-2B, and Sentinel-3B, respectively, from March 16th, 2019. The images were processed to R_{rs} using ACOLITE. Labeled locations correspond to the sites sampled two days prior to the satellite overpasses. The near-surface Chl*a* and TSS measurements for that day at station 24 were reported as ~ 10.7 mg m^{-3} and 5 g m^{-3}, respectively, while 9 mg m^{-3} and 7 g m^{-3} were measured for station 30.

significantly reduced the number of Chl*a* matchups (Section 3.2), preventing robust statistical analyses.

As shown in the matchups derived from ACOLITE (Fig. 9), neither MDN nor OC3 is viable for generating Chl*a* products from MSI. MDN approximations from SeaDAS-derived R_{rs}, however are highly more accurate than those estimated from OC3, although the number of matchups is much smaller than that for ACOLITE. In essence, the performance of MDN and other algorithms is likely better evaluated when a wider range of Chl*a* is utilized. It is also shown that Blend exhibits large uncertainties for Chl*a* < 10 mg m^{-3} and MDN predictions are noisy and especially less accurate than OC3 predictions for Chla <1 mg m^{-3}, indicating that $\sigma(R_{rs})$ primarily affect the retrieval accuracies within this range compared to the demonstrations in Section 5.1.

Predictions of TSS from OLI-derived R_{rs} yields the best result via MDN for almost all statistical descriptors compared to SOLID and Novoa for both AC processors (Fig. 10). The level of improvement is more discernible for SeaDAS R_{rs} implying that ACOLITE may generate biased-high R_{rs} causing varying degrees of overestimated TSS through all three algorithms. It also appears that SOLID is more sensitive to $\sigma(R_{rs})$ from ACOLITE than Novoa, while it is marginally less susceptible in the case of SeaDAS. Nonetheless, SOLID operates more uniformly across the dynamic range for both processors.

Combined $a_{cdom}(440)$ matchups acquired from Lake Winnipeg, Lake Pulse Network (LPN), and Minnesota lakes are depicted in Fig. 11. Overall, MDN outperforms Ficek, despite its parametrization based on a green-red band ratio, which is expected to reduce sensitivity to $\sigma(R_{rs})$. The Ficek model tends to overestimate absorption, and its predictions contain significant uncertainties (~ 80%) and noise as inferred by the RMSLD when SeaDAS output is examined. The noise is approximately equal for ACOLITE estimates through MDN and Ficek. It is evident that $a_{cdom}(440)$ retrievals are largely affected by $\sigma(R_{rs})$ with QAA-CDOM being the most sensitive for both AC processors (Mannino et al., 2014). The consistently negative biases in MDN predictions from SeaDAS agree well with the results of previous research (Pahlevan et al., 2021a), corroborating biased low R_{rs} retrievals, particularly in the blue bands. It should also be noted that a preliminary, independent matchup assessment of these three sources manifested more considerable uncertainties for the LPN (N = 29 and 46 for SeaDAS and ACOLITE,

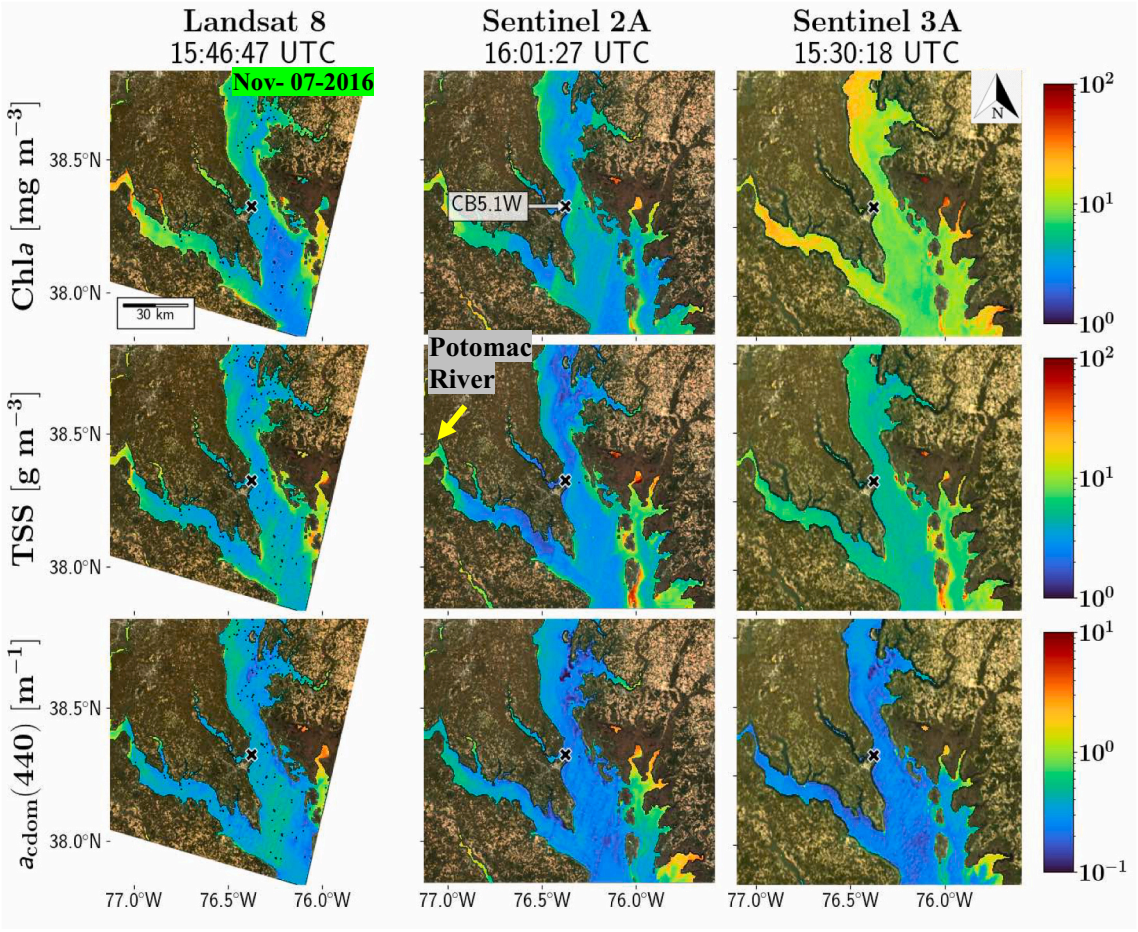

Fig. 7. Similar to Fig. 6, but over the Chesapeake Bay on Nov 7th, 2016. Near-surface Chl*a* and TSS measurements at CB5.1 W > 3 h before satellite overpasses were reported 6.2 mg m^{-3} and 7.4 g m^{-3}, respectively.

respectively) where impacts of photons emanating from adjacent land were apparent (Pahlevan et al., 2021a; Sterckx et al., 2011).

6. Discussion

We showed that the MDN model has the potential for producing consistent, high-quality WQ products from the past, present, and future satellite observations at a global scale. The model was demonstrated to be capable of generating WQ products spanning a wide range of Chl*a*, TSS, and $a_{cdom}(440)$ that conform to previous studies and expectations. It can also be readily adapted to heritage instruments by applying the respective RSRs. These missions and/or instruments include, but are not limited to, the Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), Medium Resolution Imaging Spectrometer (MERIS), and Moderate Resolution Imaging Spectroradiometer (MODIS). The extension of MDNs to these instruments will enable the construction of a long-term record of WQ variability across scales from watersheds to rivers and to nearshore and offshore areas that represent distinct optical regimes and offers a historical dataset to quantify trends and shifts in WQ indicators.

To inform a plausible range of uncertainties for the model performance, two commonly used approaches in the machine-learning domain were employed. The LOO method is possibly the preferred assessment subject to the availability of high-quality, methodically measured multi-source datasets, as well as datasets with overlapping distributions. This is not currently feasible and, owing to the diversity in *in situ* measurement methodologies and laboratory analyses likely presents the worst-case model performance per category (Fig. 5). Our reported weighted average and median statistics are, however, reliable metrics to gauge the overall performance, and in combination with the hold-out method, offer a viable range of uncertainties (e.g., 26 to 65%). This implies that if the model is tested with unseen datasets that resemble our development data, the performance would compare with the statistical descriptors reported through the hold-out method (Table 2). Nevertheless, distinct test datasets would yield performances bearing a resemblance to the LOO approach (Table 3).

We further demonstrated that depending on the AC processor, uncertainties in MDN estimates vary differently for each WQ component. We presented two instances over the San Francisco Bay and Chesapeake Bay, for which MDN generated relatively consistent WQ products across three sensors. Given the range of model uncertainties, however, the difference in products from one sensor to another may reach >100% even when a perfect AC processor is applied (e.g., when ε for a Chl*a* estimate from MSI is −50% and the corresponding ε in an OLCI-derived estimate is 60%). Hence, we surmise that the products are consistent within the algorithm uncertainty and caution users on the interpretation and application until pixel-level uncertainties are provided (see below). These circumstances may exacerbate under more complex atmospheric conditions, such as in the presence of absorbing aerosols, haze, and thin clouds, where $\sigma(R_{rs})$ are relatively high. An excellent example for this scenario is the Upper Klamath Lake demonstration (Fig. 8), where TSS estimates through MSI appear to be higher in magnitude than those of OLI and OLCI. A brief assessment of ACOLITE R_{rs} at a few selected locations verified higher MSI R_{rs} spectra, which is anticipated to lead to overestimated TSS (Woźniak et al., 2010). The extracted spectra for the MN location are illustrated in Fig. 12.

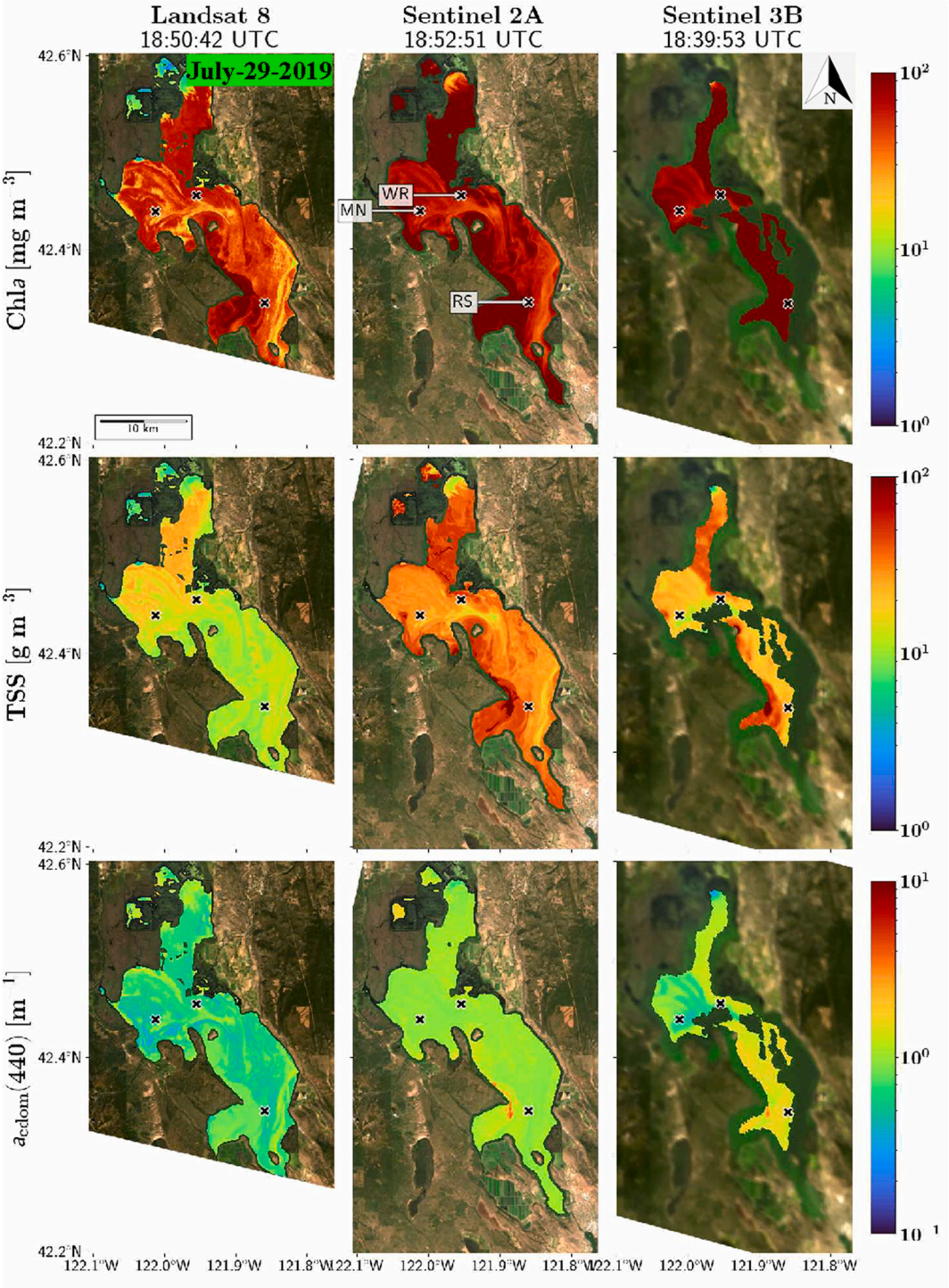

Fig. 8. Like Fig. 6, but showing the Upper Klamath Lake (Oregon) on July 29th,2019, a hypereutrophic lake with a long history of seasonal algal blooms. Near-surface Chl*a* for RS, WR, and MN on July 30th were 57.0, 27.4, and 66.7 mg m^{-3}, respectively.

In general, for Chl*a* and TSS, compared to other retrieval algorithms applicable to 0.1 to 20 mg m^{-3} (or g m^{-3}) shown in Figs. 9 and 10, MDN appears to be less sensitive to $\sigma(R_{rs})$ when SeaDAS is employed. MDN predictions of $a_{cdom}(440) > 1$ m^{-1} however, are more impacted than those of Ficek, a band ratio model, while QAA-CDOM is proven to be highly sensitive to $\sigma(R_{rs})$. Since MDN applies all the available bands to predict $a_{cdom}(440)$, an AC processor with better performance across all the bands (e.g., iCOR; Pahlevan et al. (2021a)) is expected to return more accurate retrievals. Yet, our matchup assessment using POLYMER R_{rs} (Steinmetz and Ramon, 2018) in conjunction with QAA-CDOM for Lake Winnipeg (not shown here) indicated ~50% uncertainties in $a_{cdom}(440)$, suggesting >30% improved performance compared to MDN and Ficek. This analysis (not shown here) indicates that AC processors that may be impractical for global inland and coastal waters (Pahlevan et al., 2021a) may still offer viable products when combined with algorithms that favorably offset $\sigma(R_{rs})$ for local and/or regional applications (e.g., Lake Winnipeg). More extensive matchup assessments (e.g., Chl*a* > 20 mg m^{-3}) are critical to fully characterize the effects of $\sigma(R_{rs})$ on MDN estimates. Our analysis of these two AC processors serves as a sensitivity assessment to set user expectations of model performance in

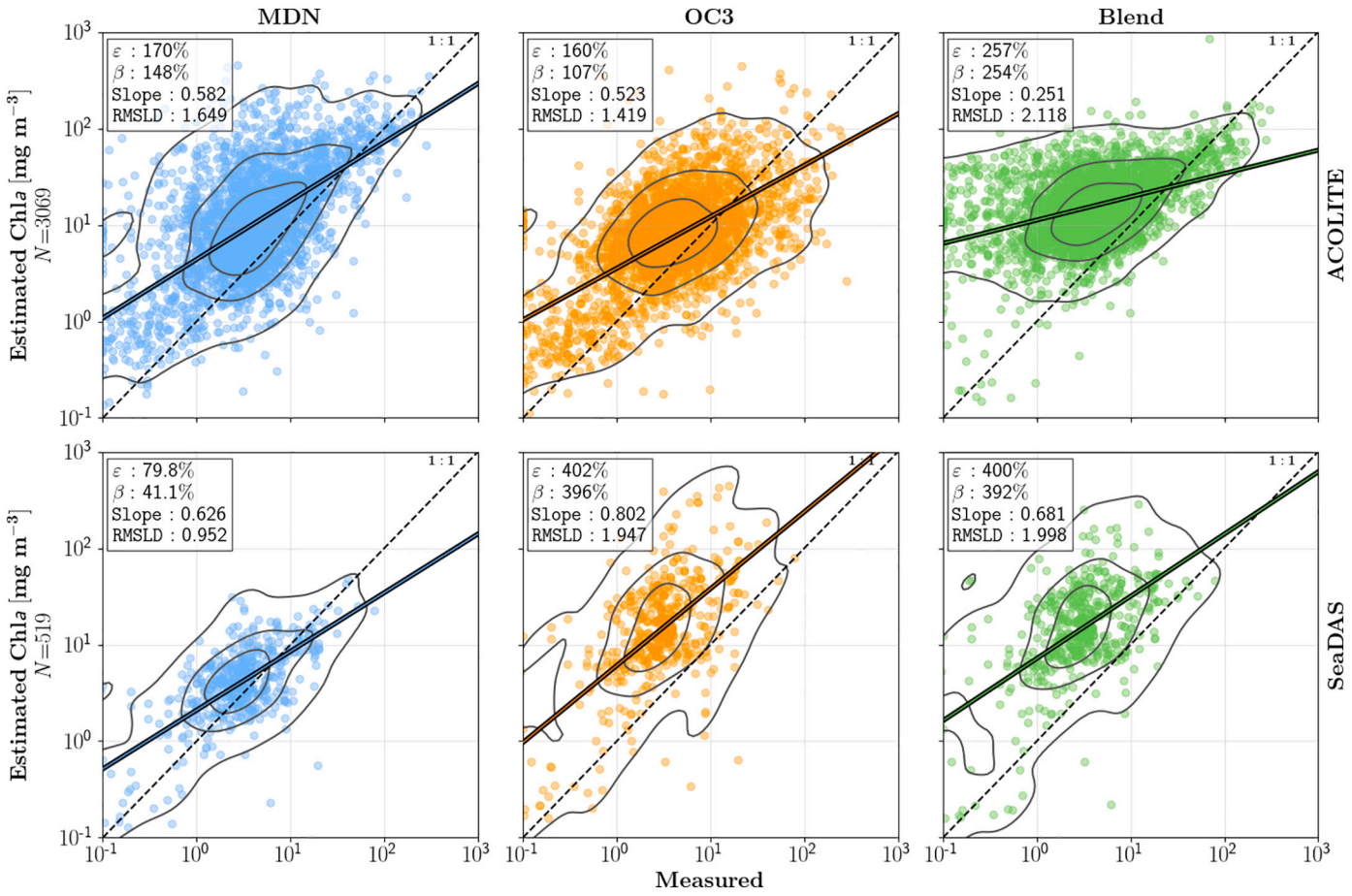

Fig. 9. MSI Chl*a* matchups for evaluating MDN approximations. Matching images were processed through ACOLITE (top row) and SeaDAS (bottom row) (Section 3.2). The number of matchups for each processor is shown on the y-axes.

practice and further verifies the intricacies in multi-mission product developments. Per recommendations in Pahlevan et al. (2021a), we anticipate that the MDN model would show the least sensitivity in coastal waters (optical water types 1, 2, and 3) when used in association with SeaDAS; whereas its application to ACOLITE R_{rs} (or iCOR; De Keukelaere et al. (2018)) in inland waters, i.e., optical water types 4, 5, 6, and 7 in Pahlevan et al. (2021), would lead to the most accurate retrievals. Further, improved handling of relative (sun-sensor) azimuth angles is necessary to characterize surface reflected skylight more accurately in the AC process and minimize banding effects present in OLI and MSI products.

In the future, one avenue for improving the model performance is to augment the input feature space with band-ratio or line-height features that are expected to diminish the algorithm sensitivity to $\sigma(R_{rs})$ (O'Shea et al., 2021), which are primarily driven by uncertainties in the AC. One should, however, note that incorporating such features may worsen WQ estimates when denominators approach zero, typically due to over-corrections for aerosol contributions. Another line of research that deserves attention is the extent to which TOA reflectance observations across these sensors are consistent. While the consistency between OLI and MSI is monitored on a routine basis (Helder et al., 2018), more efforts are required to quantify their level of agreement with OLCI's TOA products. It is also critical to provide per-pixel uncertainty estimates, which are inherently modeled by the MDN itself, i.e., the network directly parameterizes the uncertainty within the learned covariance matrices of the mixture components (Brando Guillaumes, 2017; Choi et al., 2018). This is directly applicable to the task of providing confidence levels, as the negative log-likelihood loss function used to guide MDN training has been shown to provide accurate characterizations of estimated uncertainty, especially when used in combination with a bootstrap-aggregating (bagging) ensemble (Breiman, 1996; Lakshminarayanan et al., 2017). The utility of our products in scientific studies and decision-making activities is meaningful only when pixel-level confidence intervals are generated and reported (IOCCG, 2019).

While we anticipate the MDN model to produce adequately accurate WQ products at a global scale, we note that locally or regionally tuned ML models should generally be expected to outperform our global model because of the lesser complexity in the range and magnitude of variability across smaller spatial scales (although episodic events may not be well captured in such scenarios). These developments are particularly encouraged in regions where a robust atmospheric correction may not be feasible due to frequent haze, thick aerosol layers, or complex surrounding topography and landcover types that may lead to strong adjacency effects (Cao et al., 2020; Xue et al., 2019).

7. Conclusion

Building upon previous research efforts, this study formulated, developed, and demonstrated MDNs for simultaneous retrieval of optical WQ indicators, namely Chl*a*, TSS, and $a_{cdom}(440)$, from the Landsat-8, Sentinel-2, and Sentinel-3 missions. With a goal of producing consistent multi-mission global WQ products, to the extent possible, given inherent differences in the spectral, spatial, and radiometric sampling of the sensors, we developed an MDN model using co-located *in situ* radiometry and WQ data. The model performance was comprehensively compared with that of several existing algorithms using two different methods, which proved advancements in WQ product accuracies. These analyses suggest uncertainties ranging from 26 to 62% for

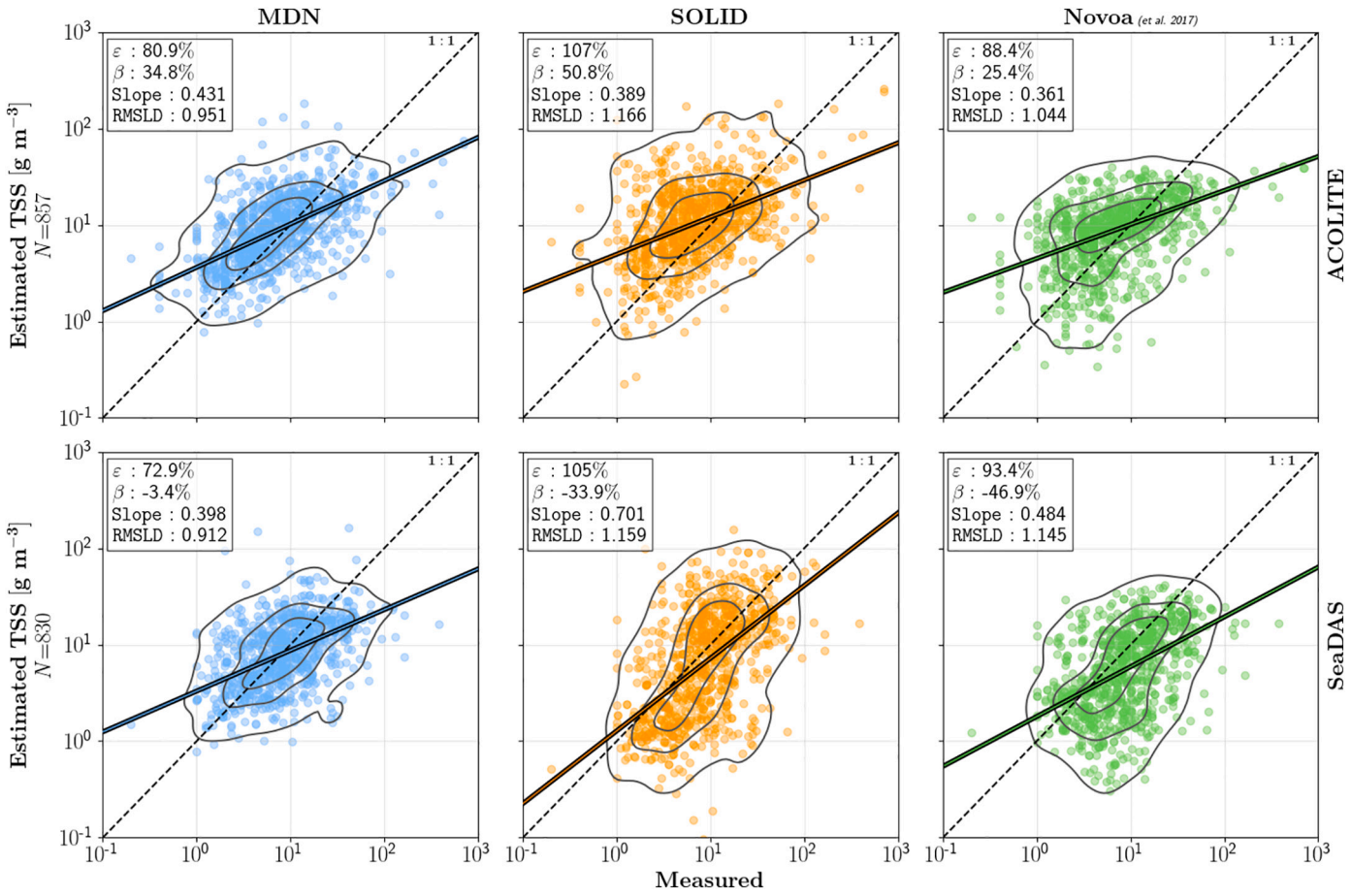

Fig. 10. OLI TSS matchups for assessing the quality of TSS products from MDN. Matching images were processed through ACOLITE (top row) and SeaDAS (bottom row) (Section 3.2). The number of matchups for each processor is shown on the y-axes.

Chl*a* and TSS, and 26 to 91% for $a_{cdom}(440)$ from MSI- and OLCI-like R_{rs} spectra through MDN. For a spectral band configuration like that of OLI, 37 to 85% uncertainties are expected for TSS and $a_{cdom}(440)$, whereas this range spans from 59 to ~100% for Chl*a*. Our demonstration maps exhibited that our model enables consistent multi-mission products within model uncertainties. Chl*a* products for less eutrophic waters (e.g., large estuaries) are anticipated to agree well across OLI, MSI, and OLCI. In contrast, across more eutrophic waters (Chl*a* >5 mg m^{-3}), MSI and OLCI should provide consistent products assuming negligible differences in the respective TOA observations. Given high-quality R_{rs} products, we expect TSS and $a_{cdom}(440)$ maps from the three missions to match very well within the uncertainty of the model across a diverse spectrum of optical regimes represented in our development database. Our matchup assessments through samples representing an acceptable range of WQ conditions (< 20 mg m^{-3} of Chl*a*, < 20 g m^{-3} of TSS and $a_{cdom}(440) < 2.5$ m^{-1}) suggested that the sensitivity of MDN to $\sigma(R_{rs})$ is subject to the quality of R_{rs} across all the bands and requires further studies at global and regional scales. Although our MDN model is anticipated to offer practical global multi-mission WQ products, local and regional ML models developed using relevant data or tied to a specific AC processor may well be advantageous over our global model; hence, such exercises, particularly in areas with challenging atmospheric and/or aquatic conditions, are immensely encouraged. These developments will likely better serve local communities and water resource managers with high demands for precise and timely WQ monitoring. Future advancements for the MDN model will include adding band-ratio and line-height features to the input feature space to maximize its resistance to $\sigma(R_{rs})$. Moreover, future models will provide accompanying per-pixel uncertainty estimates to allow users to choose fit-for-purpose products. The model is also planned to be extended to heritage instruments (e.g., MERIS) and hyperspectral sensors to demonstrate its utility for constructing long-term WQ products from the past into the future.

Code and data availability

The model can be accessed via https://github.com/STREAM-RS. The development dataset (Section 3.1) will soon be published in a data publication/repository. A copy will also be submitted to SeaBASS and other relevant open databases.

Author credit statement

Nima Pahlevan: Conceptualization; Brandon Smith, Krista Alikas, Janet Anstee, Claudio Barbosa, Caren Binding, Mariano Bresciani, Claudia Giardino, Daniela Gurlin, Virginia Fernandez, Cédric Jamet, Kersti Kangro, Moritz Lehmann, Hubert Loisel, Bunkei Matsushita, Nguyên Hà, Leif Olmanson, Bruno Palmerini, Geneviève Potvin, Stefan Simis, Andrea VanderWoude, Vincent Vantrepotte, Antonio Ruiz-Verdù; Data curation; Brandon Smith: Software and code; Brandon Smith: Formal analysis; Brandon Smith: Visualization; Nima Pahlevan and Brandon Smith: Investigation; Nima Pahlevan: Funding acquisition and supervision; Nima Pahlevan: Roles/Writing - original draft; Nima Pahlevan, Caren Binding, Daniela Gurlin, Moritz Lehmann, Hubert Loisel, Stefan Simis: Writing - review & editing.

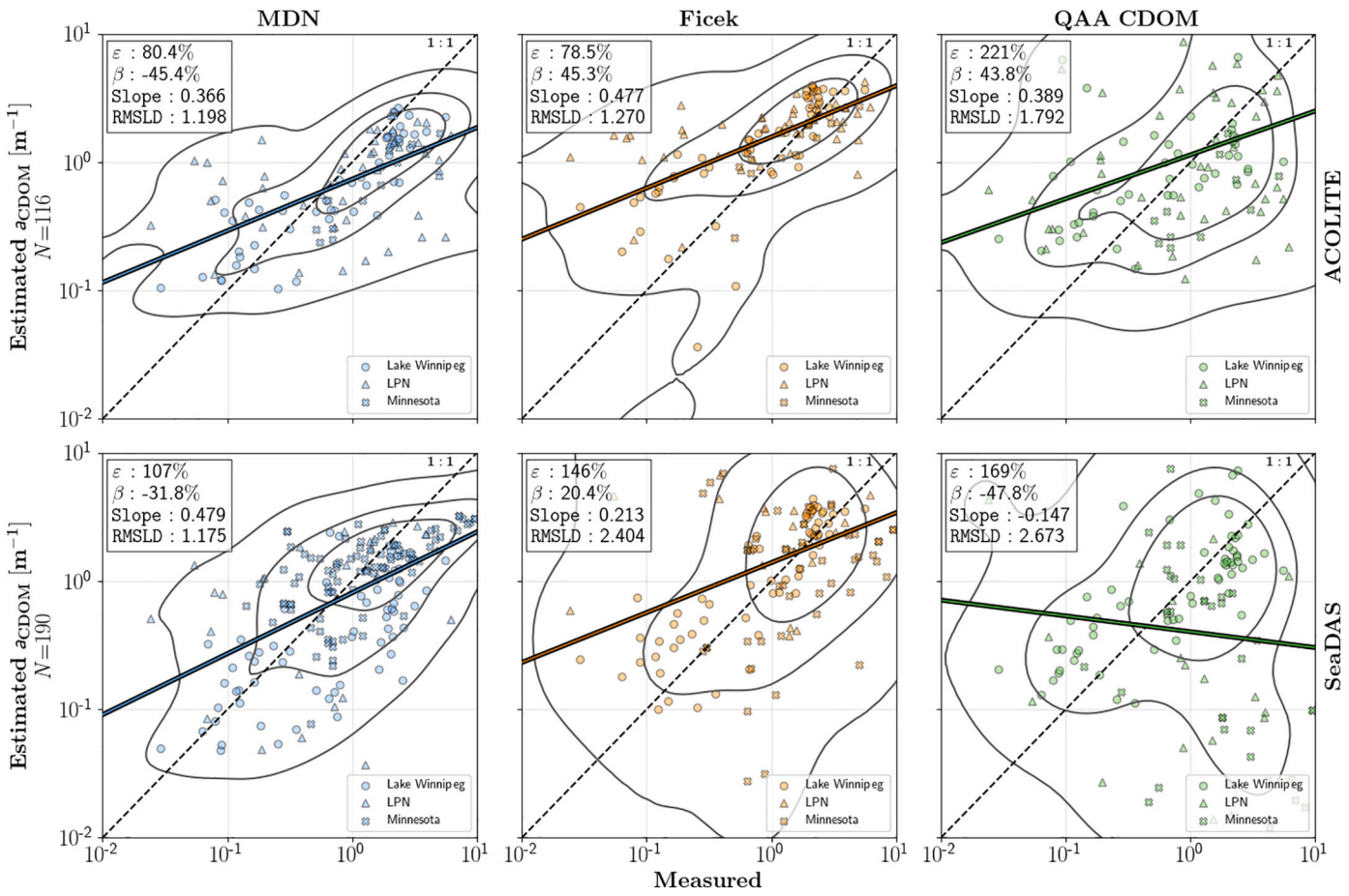

Fig. 11. Same as Fig. 10 but for $a_{cdom}(440)$. Matching images were processed through ACOLITE (top row) and SeaDAS (bottom row) (Section 3.2). Matchups were acquired in Lake Winnipeg (circles) and several small lakes (triangles) across Canada (LPN) and the state of Minnesota. The number of matchups for each data source (Lake Winnipeg, LPN, and Minnesota) is 59, 46, and 11 for ACOLITE and 68, 29, and 93 for SeaDAS, respectively.

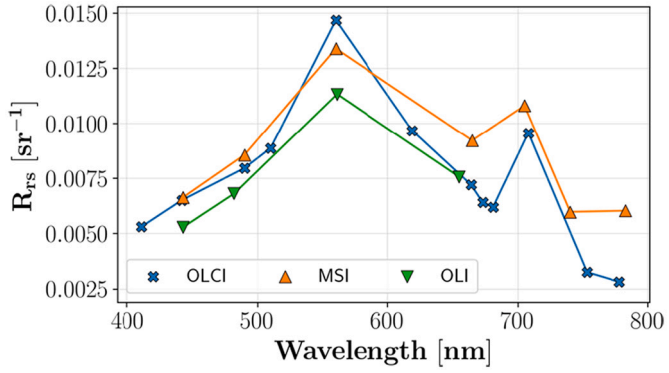

Fig. 12. OLI, MSI, and OLCI derived R_{rs} spectra (retrieved through ACOLITE) supplied to MDN for WQ indicator estimation. The spectra correspond to the MN site in the Upper Klamath Lake (Figure 8). A higher magnitude of MSI R_{rs} leads to larger estimates of TSS.

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.

Acknowledgement

The principal investigators providing the development data (Section 3.1) included Ronghua Ma, John Schalles, Anatoly Gitelson, Wesley Moses, Tim Moore, Natascha Oppelt, Mike Ondrusek, Steve Greb, Cedric Fichot, and Deepak Mishra. The data in Lake Kummerow were partly produced within the scope of the LAKESAT (grant # 50EE1340) funded by the Federal Ministry for Economic Affairs and Energy, Germany. We also acknowledge NASA's Ocean Ecology Lab for creating and maintaining SeaBASS and Amazon Web Services (AWS) for providing computing resources through SageMaker Studio as part of the Goddard Commercial Cloud (GCC) Mission Cloud Platform (MCP) and NASA Goddard Artificial Intelligence Center of Excellence (AICOE) pilot projects. This work was primarily accomplished with funding from the NASA ROSES contract #80HQTR19C0015, Remote Sensing of Water Quality element, and the USGS Landsat Science Team Award #140G0118C0011.

Appendix A

Table A

A guide to the various data sources employed in the algorithm development and performance assessments as part of the leave-one-out experiment (Fig. 5 and Table 3). NA stands for Not Available, and N is the total number of datasets (irrespective of sensors or WQ indicators) available from each set.

Data Source Index			Data Source Index			Data Source Index			Sites (examples)	N	Prior Data Usage
OLI			MSI			OLCI					
Chl*a*	TSS	$a_{cdom}(440)$	Chl*a*	TSS	$a_{cdom}(440)$	Chl*a*	TSS	$a_{cdom}(440)$			
1	1	1	1	1	1	1	1	1	Lake Erie, Lake of the Woods, Lake Winnipeg, Lake Ontario (Canada)	193	(Binding et al., 2018; Binding et al., 2019)
2	2	2	2	2	2	2	2	2	Lake Garda, Lake Iseo, Lake Como, Lake Maggiore, Lake Trasimeno (Italy)	332	(Bresciani et al., 2018; Bresciani et al., 2020; Warren et al., 2019)
3	3	3	3	3	3	3	3	3	Lake Peipsi, Lake Saadjärv, Lake Kaiavere, Lake Võrtsjärv, Lake Vän, Lake Pärnu, Lake Nohipalu, Lake Holstre järv (Estonia), Lake Vänern (Sweden)	193	(Alikas et al., 2020; Ansper and Alikas, 2019)
4	4	4	4	4	4	4	4	4	Lake Kummerow (Germany)	47	(Dörnhöfer et al., 2018)
5	5	5	5	5	5	5	5	5	Lake Hume, Wachtels Lagoon, Lake Victoria, Lake Burrinjuck, Lake Bonney Riverland, Lake Dartmouth (Australia)	67	(Botha et al., 2020)
6	6	6	6	6	6	6	6	6	Japura River, Tiete River, Lake Curuai, Lake Acai, Prana River (Brazil)	49	(Cairo et al., 2020) (Jorge et al., 2017) (Maciel et al., 2019) (da Silva et al., 2020)
7	7	7	7	7	7	7	7	7	Lake Suwa, Lake Kasumigaura, Lake Dianchi, Lake Biwa, Lake Erhai (Japan/China)	64	(Jiang et al., 2021; Matsushita et al., 2015; Yang et al., 2011)
NA	8	NA	NA	8	NA	NA	8	NA	Grizzly Bay (California, USA)	19	(Jensen et al., 2019)
NA	9	NA	NA	9	NA	NA	9	NA	Northern Gulf of Mexico (Louisiana shelf)	68	
NA	10	NA	NA	10	NA	NA	10	NA	Plum Island (Massachusetts, USA)	24	(Zhang et al., 2020)
11	NA	NA	11	NA	NA	11	NA	NA	Curonian Lagoon (Lithuania)	54	NA
12	NA	12	NA	NA	NA	NA	NA	NA	Lake IJsselmeer, Lake Loosdrecht (the Netherlands)	424	(Simis et al., 2005)
13	13	13	13	13	13	13	13	13	Fremont Lakes (Nebraska, USA)	164	(Gurlin et al., 2011; Moses et al., 2012b)
14	14	NA	14	14	NA	14	14	NA	Western Lake Erie (2014–2019) (USA)	278	NA
15	15	15	NA	15	NA	NA	NA	NA	Lake Champlain, Lake Michigan, Lake Erie, Oneida Lake (USA)	55	(Mouw et al., 2013)
16	16	16	NA	16	NA	NA	NA	NA	Green Bay (Wisconsin, USA)	15	(Lee et al., 2013)
17	17	17	17	17	17	17	17	17	Lake Koshkonong, Lake Kegonsa, Lake Waubesa (Wisconsin, USA)	188	(Pahlevan et al., 2020)
NA	18	NA	NA	18	NA	NA	18	NA	Vietnam, France, Spain, and Guyana coastlines	770	(Han et al., 2016; Loisel et al., 2017)
19	19	NA	19	19	NA	19	19	NA	Eagle Creek Reservoir, Morse Reservoir (Indiana, USA)	192	(Li et al., 2015; Song et al., 2013)
20	20	20	20	20	20	20	20	20	Lake Garda, Lake Iseo, Lake Como, Lake Maggiore, Lake Trasimeno (Italy)	17	(Bresciani et al., 2018; Bresciani et al., 2020; Warren et al., 2019)
21	21	NA	21	21	NA	21	21	NA	South Korea coastal waters	53	(Pahlevan et al., 2020)
22	NA	NA	22	NA	NA	22	NA	NA	Lake Xingyun (China)	20	NA
23	23	23	23	23	23	23	23	23	Mantua lakes (Italy)	23	(Pinardi et al., 2018)
24	NA	NA	24	NA	NA	24	NA	NA	Massachusetts Bay (Massachusetts, USA)	3	(Wei et al., 2016)
25	25	25	25	25	25	25	25	25	Cannon Lake, Sakatah Lake, Francis Lake, Minnetonka Lake (Minnesota, USA)	12	(Page et al., 2019)
26	NA	26	26	NA	26	26	NA	26	Mississippi ponds	41	(Wang et al., 2016)
27	27	27	27	27	27	27	27	27	Lake Ngaroto, Lake Rotomanuka, Lake Waikare, Lake Rotokakahi, Lake Ohau, Lake Manapouri, Lake Karaka (New Zealand)	236	(Balasubramanian et al., 2020)
28	NA	NA	28	NA	NA	28	NA	NA	Chesapeake Bay (USA)	43	(Pahlevan et al., 2020)
29	29	NA	29	29	NA	29	29	NA	Apalachicola Bay, St. Mary's River, Roat West Carib, Edisto River, Bayou Cumbest, Delaware Bay, Silver Lake (USA)	759	(Balasubramanian et al., 2020)
30	NA	30	30	NA	30	30	NA	30	West Florida shelf, Chesapeake Bay, Monterey Bay (NASA's SeaBASS, USA)	1042	(Werdell and Bailey, 2005)
NA	31	NA	NA	31	NA	NA	31	NA	Gironde and Scheldt estuaries (France/Belgium)	79	(Knaeps et al., 2018)
32	32	NA	32	32	NA	32	32	NA	Hartbeespoort, Theewaterskloof and Loskop dams (South Africa)	63	(Matthews, 2020; Matthews and Bernard, 2013)
33	NA	NA	33	NA	NA	33	NA	NA	Fifty-three reservoirs and lakes in Spain	250	(Ruiz-Verdú et al., 2008)
34	NA	34	34	NA	NA	34	NA	34	Lake Taihu (2011) (China)	45	(Wang et al., 2016)
35	NA	35	35	NA	35	35	NA	35	Lake Taihu (2008), Lake Chaohu, Lake Hongze (China)	260	(Cao et al., 2020)
36	NA	NA	36	NA	NA	36	NA	NA	Lake Trasimeno and Curonian lagoon (Italy/Lithuania)	20	(Bresciani et al., 2012; Giardino et al., 2015)
37	NA	NA	37	NA	NA	37	NA	NA	Pinto Lake (California, USA)	9	(Kudela et al., 2015)

(*continued on next page*)

Table A (*continued*)

Data Source Index			Data Source Index			Data Source Index			Sites (examples)	N	Prior Data Usage
OLI			MSI			OLCI					
Chl*a*	TSS	$a_{cdom}(440)$	Chl*a*	TSS	$a_{cdom}(440)$	Chl*a*	TSS	$a_{cdom}(440)$			
38	NA	NA	38	NA	NA	38	NA	NA	Rincon del Bonete Lake and Paso Palmar Lake on Negro river (Uruguay)	48	(Smith et al. 2021)
39	39	NA	39	39	NA	39	39	NA	Thac Ba Reservoir, Ba Be Lake, Hanoi Lake, Ha Long Bay (Vietnam)	115	(Ha et al., 2017b; Vinh et al., 2019)
40	NA	NA	40	NA	NA	40	NA	NA	Lake Hoan Kiem, Lake Linh Dam, West Lake, Lan Van Quan (Vietnam)	70	(Ha et al., 2017a)
41	41	41	41	41	41	41	41	41	Western Lake Erie (2013, 2014) (USA)	36	(Moore et al., 2019; Moore et al., 2017)
42	42	42	42	42	42	42	42	42	South_Green_Bay AERONET-OC site (Wisconsin, USA)	6	(Pahlevan et al., 2021b)

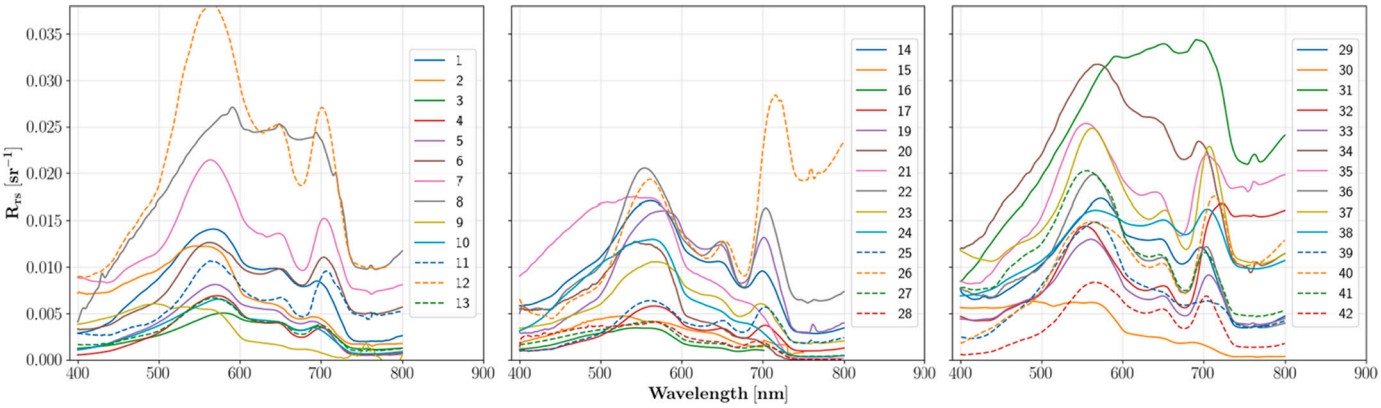

Fig. A. Average hyperspectral R_{rs} spectra associated with the 42 data sources are shown in three plots (each showing 14 samples). Note that the often-observed noise (peaks/troughs) in the ~ 760 nm region did not affect our developments/assessments (see Table B for OLI, MSI, and OLCI's spectral bands), and not all the spectra cover the full 400–800 nm range.

Appendix B

Table B
Relevant characteristics of the sensors evaluated throughout this study.

	OLI	MSI	OLCI
Nominal spatial resolution (m)	30	10, 20, 60	300
Swath (km)	185	290	1270
# of relevant bands < 800 nm	4	7	13
Spectral bands (nm)	443, 482, 560, 655	443, 492, 560, 665, 705, 740, 783	410, 443, 490, 510, 560,620, 665, 673, 681, 708, 753, 778
Full-width-half-maximum (nm)	20–60	20–65	≤ 15
Lifetime	2013 –	2015 –	2016 –
Data source	https://earthexplorer.usgs.gov	https://scihub.copernicus.eu	

Appendix C

Gilerson 2B (GI2B) - (Gilerson et al., 2010)

$$Chl = [35.75^{*}(R_{rs}(708)/R_{rs}(665)) - 19.3]^{1.124}$$

Gons (Gons et al., 2002)

$$b_b = 1.61\pi R_{rs}(778)/(0.082 - 0.6\pi R_{rs}(778))$$

$$Chl = \left[(R_{rs}(708)/R_{rs}(665))^{*}(0.7 + b_b) - 0.4 - {b_b}^{1.063}\right]/0.016$$

Gurlin 2B (GU2B) - (Gurlin et al., 2011)

$$\phi = R_{rs}(708)/R_{rs}(665)$$

$$Chl_{G2B} = 25.28\phi^2 + 14.85\phi - 15.18$$

Blend (Smith et al., 2018)

$$\phi = R_{rs}(708)/R_{rs}(665)$$

$$\alpha_1 = (\phi - 0.75)/(1.15 - 0.75)$$

$$\alpha_2 = (1.15 - \phi)/(1.15 - 0.75)$$

$$Chl = \alpha_1 Chl_{G2B} + \alpha_2 Chl_{OCx}$$

MDN (Pahlevan et al., 2020;(Smith et al., 2021)).
Codes are accessible via: https://github.com/STREAM-RS/STREAM-RS
SOLID (Balasubramanian et al., 2020)

$$b_{bp}(865) = (R_{rs}(865)\ (a_w(865) + b_{bw}(865) + 0.5\ ) - 0.105\ b_{bw}(865)\ )/(0.105 - R_{rs}(865)\ )$$

TSS = {

$$53.736\ {b_{bp}}^{MDN}(665)^{0.8559}\ if\ R_{rs}(492). < R_{rs}(665). < R_{rs}(560)$$

$$224.43\ b_{bp}(865) - 12.575\ if\ R_{rs}(560). < R_{rs}(665) \& R_{rs}(740). > 0.01$$

$$53.736\ {b_{bp}}^{QAA}(665)^{0.8559}\ if\ R_{rs}(560). < R_{rs}(492)$$

$$53.736\ {b_{bp}}^{MDN}(665)^{0.8559}\ else$$

}
$*{b_{bp}}^{MDN}$ is derived from a MDN model trained with a synthetic dataset (Balasubramanian et al., 2020).
Code is available via: https://github.com/STREAM-RS/STREAM-RS.
Nechad (Nechad et al., 2010)

$$TSS = 1.74 + (355.85 * \pi R_{rs}(665)\ )/(1 - \pi R_{rs}(665)/1728\ )$$

Novoa (Novoa et al., 2017)

$$C_{green} = 130.1 \pi R_{rs}(561)$$

$$C_{red} = 531.5 \pi R_{rs}(655)$$

$$C_{nir} = 37150 \pi R_{rs}(865)^2 + 1751 \pi R_{rs}(865)$$

$$\alpha_1 = log(0.016/(\pi R_{rs}(665)\ )\ )/log(0.016/0.007)$$

$$\beta_1 = log((\pi R_{rs}(665)\ )/0.007\ )/log(0.016/0.007)$$

$$\alpha_2 = log(0.120/(\pi R_{rs}(665)\ )\ )/log(0.120/0.080)$$

$$\beta_2 = log((\pi R_{rs}(665)\ )/0.080\ )/log(0.120/0.080)$$

TSS = {

$$C_{green}\ if\ \pi R_{rs}(655). < 0.007$$

$$\alpha_1 C_{green} + \beta_1 C_{red}\ if\ 0.007 \le \pi R_{rs}(655) \le 0.016$$

$$C_{red}\ if\ 0.016 < \pi R_{rs}(655). < 0.080$$

$$\alpha_2 C_{red} + \beta_2 C_{nir}\ if\ 0.080 \le \pi R_{rs}(655) \le 0.120$$

$$C_{nir}\ if\ 0.120 < \pi R_{rs}(655)$$

}
Ficek et al., 2011

$$a_{cdom}(440) = 3.65\ (R_{rs}(570)/R_{rs}(655)\ )^{-1.93}$$

Mannino et al. 2007:

$$a_{cdom}(440) = -0.0736\ log_e(0.408\ R_{rs}(490)/R_{rs}(555) - 0.173\ )$$

QAA-CDOM:
The QAA procedure (http://www.ioccg.org/groups/Software_OCA/QAA_v6_2014209.pdf) was implemented. From the reported IOPs, $a_{cdom}(440)$ was estimated as follows

$$a_p = 0.63\ {b_{bp}}^{0.88}$$

$$a_{cdom}(440) = a(440) - a_w(440) - a_p(440)$$

Where a_p and a_w represent pigment and pure water absorption, respectively, and b_{bp} is the particulate backscattering.

Appendix D

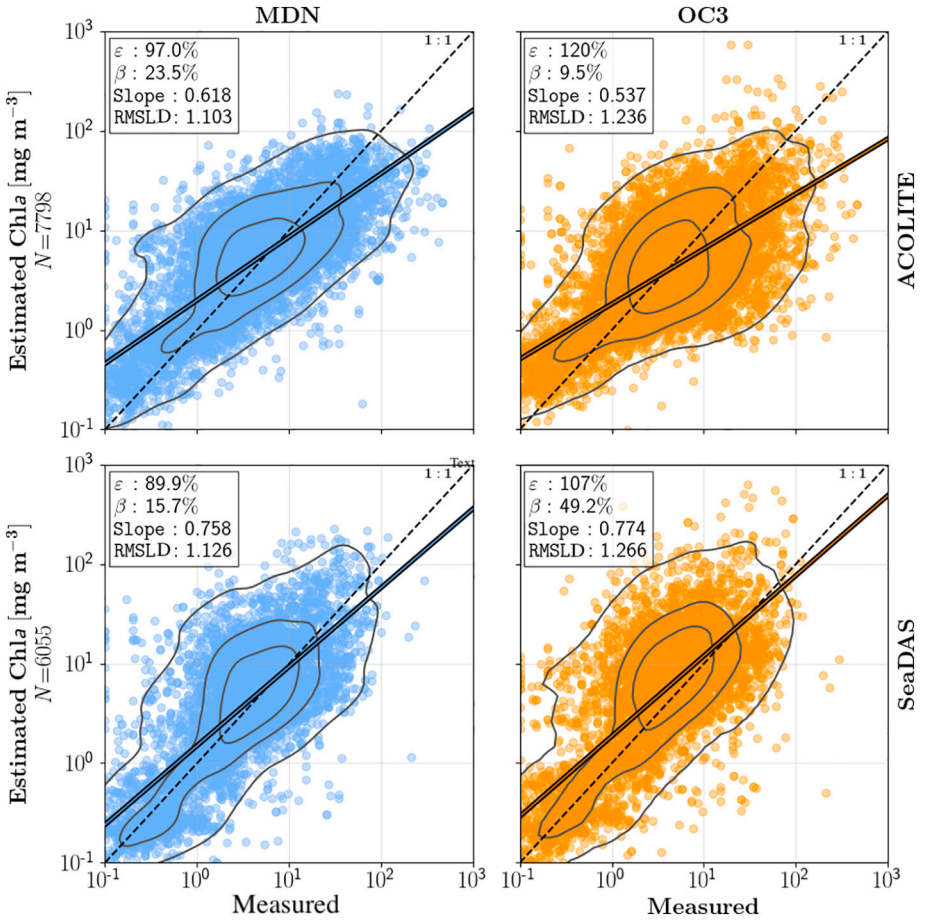

Fig. D. OLI Chl*a* matchups for evaluating MDN retrievals. Matching images were processed through ACOLITE (top row) and SeaDAS (bottom row) (Section 3.2). The number of matchups for each processor is shown on the y-axes.

References

Alikas, K., Ansko, I., Vabson, V., Ansper, A., Kangro, K., Uudeberg, K., Ligi, M., 2020. Consistency of radiometric satellite data over lakes and coastal waters with local field measurements. Remote Sens. 12, 616.

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.

Arabi, B., Salama, M.S., Pitarch, J., Verhoef, W., 2020. Integration of in-situ and multi-sensor satellite observations for long-term water quality monitoring in coastal areas. Remote Sens. Environ. 239, 111632.

Baker, A., 2006. Land Use and Water Quality (Encyclopedia of hydrological sciences).

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.

Benlloch, S., López-López, A., Casamayor, E.O., Øvreås, L., Goddard, V., Daae, F.L., Smerdon, G., Massana, R., Joint, I., Thingstad, F., 2002. Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern. Environ. Microbiol. 4, 349–360.

Binding, C., Greenberg, T., McCullough, G., Watson, S., Page, E., 2018. An analysis of satellite-derived chlorophyll and algal bloom indices on Lake Winnipeg. J. Great Lakes Res. 44, 436–446.

Binding, C., Zastepa, A., Zeng, C., 2019. The impact of phytoplankton community composition on optical properties and satellite observations of the 2017 western Lake Erie algal bloom. J. Great Lakes Res. 45, 573–586.

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.

Bishop, C.M., 1994. Mixture Density Networks. NCRG/94/004. Aston University Birmingham. http://www.ncrg.aston.ac.uk.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford university press.

Botha, E.J., Anstee, J.M., Sagar, S., Lehmann, E., Medeiros, T.A., 2020. Classification of Australian waterbodies across a wide range of optical water types. Remote Sens. 12, 3018.

Bradbury, J Platt, Colman, Steven M, Reynolds, Richard L, 2004. The history of recent limnological changes and human impact on Upper Klamath Lake, Oregon. Journal of Paleolimnology 31, 151–165.

Brando, V.E., Dekker, A.G., 2003. Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality. IEEE Trans. Geosci. Remote Sens. 41, 1378–1387.

Brando Guillaumes, A., 2017. Mixture Density Networks for Distribution and Uncertainty Estimation. Universitat Politècnica de Catalunya.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140.

Bresciani, M., Giardino, C., Stroppiana, D., Pilkaitytė, R., Zilius, M., Bartoli, M., Razinkovas, A., 2012. Retrospective analysis of spatial and temporal variability of chlorophyll-a in the Curonian lagoon. J. Coast. Conserv. 16, 511–519.

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

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.

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

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

Bricaud, A., Mejia, C., Blondeau-Patissier, D., Claustre, H., Crepon, M., Thiria, S., 2007. Retrieval of pigment concentrations and size structure of algal populations from their absorption spectra using multilayered perceptrons. Appl. Opt. 46, 1251–1260.

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.
Bukata, R.P., Jerome, J.H., Kondratyev, K.Y., Pozdnyakox, D.V., 1995. Optical Properties and Remote Sensing of Inland and Coastal Waters. CRC Press, New York.
Buuren, S.V., Groothuis-Oudshoorn, K., 2010. MICE: multivariate imputation by chained equations in R. J. Stat. Softw. 1–68.
Cairo, C., Barbosa, C., Lobo, F., Novo, E., Carlos, F., Maciel, D., Flores Júnior, R., Silva, E., Curtarelli, V., 2020. Hybrid chlorophyll-a algorithm for assessing trophic states of a tropical Brazilian reservoir based on MSI/Sentinel-2 data. Remote Sens. 12, 40.
Cao, F., Tzortziou, M., 2021. Capturing dissolved organic carbon dynamics with Landsat-8 and Sentinel-2 in tidally influenced wetland–estuarine systems. Sci. Total Environ. 777, 145910.
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.
Cardille, J.A., Leguet, J.-B., del Giorgio, P., 2013. Remote sensing of lake CDOM using noncontemporaneous field data. Can. J. Remote. Sens. 39, 118–126.
Choi, S., Lee, K., Lim, S., Oh, S., 2018. Uncertainty-aware learning from demonstration using mixture density networks with sampling-free variance modeling. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 6915–6922.
Concha, J.A., Schott, J.R., 2016. Retrieval of color producing agents in case 2 waters using Landsat 8. Remote Sens. Environ. 185, 95–107.
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. Eur. J. Remote Sens. 51, 525–542.
Defoin-Platel, M., Chami, M., 2007. How ambiguous is the inverse problem of ocean color in coastal waters? J. Geophys. Res. Oceans 112.
Dethier, E.N., Sartain, S.L., Lutz, D.A., 2019. Heightened levels and seasonal inversion of riverine suspended sediment in a tropical biodiversity hot spot due to artisanal gold mining. Proc. Natl. Acad. Sci. 116, 23936–23941.
Doerffer, R., Fischer, J., 1994. Concentrations of chlorophyll, suspended matter, and gelbstoff in case II waters derived from satellite coastal zone color scanner data with inverse modeling methods. J. Geophys. Res. Oceans 99, 7457–7466.
Doerffer, R., Schiller, H., 2007. The MERIS case 2 water algorithm. Int. J. Remote Sens. 28, 517–535.
Dörnhöfer, K., Scholze, J., Stelzer, K., Oppelt, N., 2018. Water colour analysis of Lake Kummerow using time series of remote sensing and in situ data. PFG – J. Photogramm. Remote Sens. Geoinform. Sci. 86, 103–120.
Doxaran, D., Froidefond, J.-M., Lavender, S., Castaing, P., 2002. Spectral signature of highly turbid waters: application with SPOT data to quantify suspended particulate matter concentrations. Remote Sens. Environ. 81, 149–161.
Duarte, C.M., Prairie, Y.T., 2005. Prevalence of heterotrophy and atmospheric CO 2 emissions from aquatic ecosystems. Ecosystems 8, 862–870.
Ficek, D., Zapadka, T., Dera, J., 2011. Remote sensing reflectance of Pomeranian lakes and the Baltic. Oceanologia 53, 959–970.
Galimard, J.-E., Chevret, S., Curis, E., Resche-Rigon, M., 2018. Heckman imputation models for binary or continuous MNAR outcomes and MAR predictors. BMC Med. Res. Methodol. 18, 1–13.
Gerace, A.D., Schott, J.R., Nevins, R., 2013. Increased potential to monitor water quality in the near-shore environment with Landsat's next-generation satellite. J. Appl. Remote. Sens. 7, 073558.
Ghahramani, Z., Jordan, M.I., 1995. Learning from Incomplete Data.
Giardino, C., Brando, V.E., Dekker, A.G., Strömbeck, N., Candiani, G., 2007. Assessment of water quality in Lake Garda (Italy) using Hyperion. Remote Sens. Environ. 109, 183–195.
Giardino, C., Bresciani, M., Valentini, E., Gasperini, L., Bolpagni, R., Brando, V.E., 2015. Airborne hyperspectral data to assess suspended particulate matter and aquatic vegetation in a shallow and turbid lake. Remote Sens. Environ. 157, 48–57.
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.
Gitelson, A., 1992. The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. Int. J. Remote Sens. 13, 3367–3373.
Gitelson, A.A., Schalles, J.F., Hladik, C.M., 2007. Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. Remote Sens. Environ. 109, 464–472.
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., Brown, O.B., Jacobs, M.M., 1975. Computed relationships between the inherent and apparent optical properties of a flat homogeneous ocean. Appl. Opt. 14, 417–427.
Gower, J., King, S., Borstad, G., Brown, L., 2005. Detection of intense plankton blooms using the 709 nm band of the MERIS imaging spectrometer. Int. J. Remote Sens. 26, 2005–2012.
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.
Ha, N.T.T., Koike, K., Nhuan, M.T., Canh, B.D., Thao, N.T.P., Parsons, M., 2017a. Landsat 8/OLI two bands ratio algorithm for chlorophyll-a concentration mapping in hypertrophic waters: an application to West Lake in Hanoi (Vietnam). IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 4919–4929.
Ha, N.T.T., Thao, N.T.P., Koike, K., Nhuan, M.T., 2017b. Selecting the best band ratio to estimate chlorophyll-a concentration in a tropical freshwater lake using sentinel 2A images from a case study of Lake Ba be (northern Vietnam). ISPRS Int. J. Geo Inf. 6, 290.
Hafeez, S., Wong, M.S., Ho, H.C., Nazeer, M., Nichol, J., Abbas, S., Tang, D., Lee, K.H., Pun, L., 2019. Comparison of machine learning algorithms for retrieval of water quality indicators in case-II waters: a case study of Hong Kong. Remote Sens. 11, 617.
Hakimdavar, R., Hubbard, A., Policelli, F., Pickens, A., Hansen, M., Fatoyinbo, T., Lagomasino, D., Pahlevan, N., Unninayar, S., Kavvada, A., 2020. Monitoring water-related ecosystems with earth observation data in support of sustainable development goal (SDG) 6 reporting. Remote Sens. 12, 1634.
Han, B., Loisel, H., Vantrepotte, V., Mériaux, X., Bryère, P., Ouillon, S., Dessailly, D., Xing, Q., Zhu, J., 2016. Development of a semi-analytical algorithm for the retrieval of suspended particulate matter from remote sensing over clear to very turbid waters. Remote Sens. 8, 211.
Helder, D., Markham, B., Morfitt, R., Storey, J., Barsi, J., Gascon, F., Clerc, S., LaFrance, B., Masek, J., Roy, D., 2018. Observations and recommendations for the calibration of Landsat 8 OLI and sentinel 2 MSI for improved data interoperability. Remote Sens. 10, 1340.
Hestir, E.L., Brando, V., Campbell, G., Dekker, A., Malthus, T., 2015. The relationship between dissolved organic matter absorption and dissolved organic carbon in reservoirs along a temperate to tropical gradient. Remote Sens. Environ. 156, 395–402.
Hieronymi, M., Müller, D., Doerffer, R., 2017. The OLCI neural network swarm (ONNS): a bio-geo-optical algorithm for open ocean and coastal waters. Front. Mar. Sci. 4, 140.
Hoogenboom, H., Dekker, A., De Haan, J., 1998. Retrieval of chlorophyll and suspended matter from imaging spectrometry data by matrix inversion. Can. J. Remote. Sens. 24, 144–152.
Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., Falkowski, M.J., 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sens. Environ. 112, 2232–2245.
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.
IOCCG, 2000. Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters. S. Sathyendranath International Ocean Colour Coordinating Group. IOCCG. 3.
IOCCG, 2006. Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. Z.-P. Lee International Ocean Colour Coordinating Group. IOCCG. 5.
IOCCG, 2018. Earth Observations in Support of Global Water Quality Monitoring. S. Greb, A. Dekker, & C. Binding International Ocean Colour Coordinating Group. IOCCG. 17.
IOCCG, 2019. Uncertainties in Ocean Colour Remote Sensing. F. Mélin International Ocean Colour Coordinating Group. IOCCG. 18.
Jamet, C., Loisel, H., Dessailly, D., 2012. Retrieval of the spectral diffuse attenuation coefficient Kd (λ) in open and coastal ocean waters using a neural network inversion. J. Geophys. Res. Oceans 117.
Jensen, D., Simard, M., Cavanaugh, K., Sheng, Y., Fichot, C.G., Pavelsky, T., Twilley, R., 2019. Improving the transferability of suspended Solid estimation in wetland and deltaic waters with an empirical hyperspectral approach. Remote Sens. 11, 1629.
Jiang, D., Matsushita, B., Pahlevan, N., Gurlin, D., Lehmann, M.K., Fichot, C.G., Schalles, J., Loisel, H., Binding, C., Zhang, Y., Alikas, K., Kangro, K., Uusõue, M., Ondrusek, M., Greb, S., Moses, W.J., Lohrenz, S., O'Donnell, D., 2021. Remotely estimating total suspended solids concentration in clear to extremely turbid waters using a novel semi-analytical method. Remote Sens. Environ. 258, 112386.
Jorge, D.S., Barbosa, C.C., De Carvalho, L.A., Affonso, A.G., Lobo, F.D.L., Novo, E.M.D. M., 2017. Snr (signal-to-noise ratio) impact on water constituent retrieval from simulated images of optically complex amazon lakes. Remote Sens. 9, 644.
Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., Kolehmainen, M., 2004. Methods for imputation of missing values in air quality data sets. Atmos. Environ. 38, 2895–2907.
Kallio, K., Kutser, T., Hannonen, T., Koponen, S., Pulliainen, J., Vepsäläinen, J., Pyhälahti, T., 2001. Retrieval of water quality from airborne imaging spectrometry of various lake types in different seasons. Sci. Total Environ. 268, 59–77.
Kalteh, A.M., Hjorth, P., 2009. Imputation of missing values in a precipitation–runoff process database. Hydrol. Res. 40, 420–432.
Keiner, L.E., Yan, X.-H., 1998. A neural network model for estimating sea surface chlorophyll and sediments from thematic mapper imagery. Remote Sens. Environ. 66, 153–165.
Kiefer, D.A., Mitchell, B.G., 1983. A simple, steady state description of phytoplankton growth based on absorption cross section and quantum efficiency1. Limnol. Oceanogr. 28, 770–776.
King, G., Honaker, J., Joseph, A., Scheve, K., 2001. Analyzing incomplete political science data: an alternative algorithm for multiple imputation. Am. Polit. Sci. Rev. 49–69.
Kirk, J., 1994. Optics of UVB radiation in natural waters. Ergebnisse Limnol. 43, 16.
Knaeps, E., Ruddick, K., Doxaran, D., Dogliotti, A.I., Nechad, B., Raymaekers, D., Sterckx, S., 2015. A SWIR based algorithm to retrieve total suspended matter in extremely turbid waters. Remote Sens. Environ. 168, 66–79.

Knaeps, E., Doxaran, D., Dogliotti, A., Nechad, B., Ruddick, K., Raymaekers, D., Sterckx, S., 2018. The SeaSWIR dataset. Earth Syst. Sci. Data 10, 1439–1449.
Kowalczuk, P., Cooper, W.J., Whitehead, R.F., Durako, M.J., Sheldon, W., 2003. Characterization of CDOM in an organic-rich river and surrounding coastal ocean in the South Atlantic bight. Aquat. Sci. 65, 384–401.
Kramer, S.J., Bisson, K.M., Fischer, A.D., 2020. Observations of phytoplankton community composition in the Santa Barbara Channel during the Thomas fire. J. Geophys. Res. Oceans 125 e2020JC016851.
Kudela, R.M., Palacios, S.L., Austerberry, D.C., Accorsi, E.K., Guild, L.S., Torres-Perez, J., 2015. Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. Remote Sens. Environ. 167, 196–205.
Kutser, T., Pierson, D.C., Kallio, K.Y., Reinart, A., Sobek, S., 2005. Mapping lake CDOM by satellite remote sensing. Remote Sens. Environ. 94, 535–540.
Kutser, T., Verpoorter, C., Paavel, B., Tranvik, L.J., 2015. Estimating lake carbon fractions from remote sensing data. Remote Sens. Environ. 157, 138–146.
Lakshminarayanan, B., Pritzel, A., Blundell, C., 2017. Simple and scalable predictive uncertainty estimation using deep ensembles. In: Advances in Neural Information Processing Systems, pp. 6402–6413.
Lee, Z., Carder, K., Arnone, R., He, M., 2007. Determination of primary spectral bands for remote sensing of aquatic environments. Sensors 7, 3428–3441.
Lee, Z., Pahlevan, N., Ahn, Y.-H., Greb, S., O'Donnell, D., 2013. Robust approach to directly measuring water-leaving radiance in the field. Appl. Opt. 52, 1693–1701.
Li, L., Li, L., Song, K., 2015. Remote sensing of freshwater cyanobacteria: an extended IOP inversion model of inland waters (IIMIW) for partitioning absorption coefficient and estimating phycocyanin. Remote Sens. Environ. 157, 9–23.
Lobo, F.L., Costa, M.P., Novo, E.M., 2015. Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities. Remote Sens. Environ. 157, 170–184.
Loisel, H., Vantrepotte, V., Ouillon, S., Ngoc, D.D., Herrmann, M., Tran, V., Mériaux, X., Dessailly, D., Jamet, C., Duhaut, T., 2017. Assessment and analysis of the chlorophyll-a concentration variability over the Vietnamese coastal waters from the MERIS Ocean color sensor (2002−2012). Remote Sens. Environ. 190, 217–232.
Maciel, D., Novo, E., Sander de Carvalho, L., Barbosa, C., Flores Júnior, R., de Lucia Lobo, F., 2019. Retrieving Total and inorganic suspended sediments in Amazon Floodplain Lakes: A multisensor approach. Remote Sens. 11, 1744.
Mannino, A., Russ, M.E., Hooker, S.B., 2008. Algorithm development and validation for satellite-derived distributions of DOC and CDOM in the US middle Atlantic bight. J. Geophys. Res. Oceans 113.
Mannino, A., Novak, M.G., Hooker, S.B., Hyde, K., Aurin, D., 2014. Algorithm development and validation of CDOM properties for estuarine and continental shelf waters along the northeastern US coast. Remote Sens. Environ. 152, 576–602.
Mapulanga, A.M., Naito, H., 2019. Effect of deforestation on access to clean drinking water. Proc. Natl. Acad. Sci. 116, 8249–8254.
Matsushita, B., Yang, W., Yu, G., Oyama, Y., Yoshimura, K., Fukushima, T., 2015. A hybrid algorithm for estimating the chlorophyll-a concentration across different trophic states in Asian inland waters. ISPRS J. Photogramm. Remote Sens. 102, 28–37.
Matthews, M., 2020. Data for: distinguishing cyanobacteria from algae in optically complex inland waters using a radiative transfer inversion algorithm. Mendel. Data V1.
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.
Miller, R.L., McKee, B.A., 2004. Using MODIS Terra 250 m imagery to map concentrations of total suspended matter in coastal waters. Remote Sens. Environ. 93, 259–266.
Mobley, C.D., 1994. Light and Water: Radiative Transfer in Natural Waters. Academic Press, Inc.
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.
Moore, T.S., Mouw, C.B., Sullivan, J.M., Twardowski, M.S., Burtner, A.M., Ciochetto, A. B., McFarland, M.N., Nayak, A.R., Paladino, D., Stockley, N.D., Johengen, T.H., Yu, A.W., Ruberg, S., Weidemann, A., 2017. Bio-optical properties of Cyanobacteria blooms in Western Lake Erie. Front. Mar. Sci. 4.
Moore, T.S., Churnside, J.H., Sullivan, J.M., Twardowski, M.S., Nayak, A.R., McFarland, M.N., Stockley, N.D., Gould, R.W., Johengen, T.H., Ruberg, S.A., 2019. Vertical distributions of blooming cyanobacteria populations in a freshwater lake from LIDAR observations. Remote Sens. Environ. 225, 347–367.
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., Berdnikov, S., Saprygin, V., Povazhnyi, V., 2012a. Operational MERIS-based NIR-red algorithms for estimating chlorophyll-a concentrations in coastal waters—the Azov Sea case study. Remote Sens. Environ. 121, 118–124.
Moses, W.J., Gitelson, A.A., Perk, R.L., Gurlin, D., Rundquist, D.C., Leavitt, B.C., Barrow, T.M., Brakhage, P., 2012b. Estimation of chlorophyll-a concentration in turbid productive waters using airborne hyperspectral data. Water Res. 46, 993–1004.
Mosley, L.M., 2015. Drought impacts on the water quality of freshwater systems; review and integration. Earth Sci. Rev. 140, 203–214.
Mouw, C.B., Chen, H., McKinley, G.A., Effler, S., O'Donnell, D., Perkins, M.G., Strait, C., 2013. Evaluation and optimization of bio-optical inversion algorithms for remote sensing of Lake Superior's optical properties. J. Geophys. Res. Oceans 118, 1696–1714.
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.
Nieke, B., Reuter, R., Heuermann, R., Wang, H., Babin, M., Therriault, J., 1997. Light absorption and fluorescence properties of chromophoric dissolved organic matter (CDOM), in the St. Lawrence estuary (case 2 waters). Cont. Shelf Res. 17, 235–252.
Novo, E., Hansom, J., Curran, P., 1989. The effect of sediment type on the relationship between reflectance and suspended sediment concentration. Remote Sens. 10, 1283–1289.
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.
Odermatt, D., Gitelson, A., Brando, V.E., Schaepman, M., 2012. Review of constituent retrieval in optically deep and complex waters from satellite imagery. Remote Sens. Environ. 118, 116–126.
Olmanson, L.G., Page, B.P., Finlay, J.C., Brezonik, P.L., Bauer, M.E., Griffin, C.G., Hozalski, R.M., 2020. Regional measurements and spatial/temporal analysis of CDOM in 10,000+ optically variable Minnesota lakes using Landsat 8 imagery. Sci. Total Environ. 724, 138141.
O'Reilly, J.E., Werdell, P.J., 2019. Chlorophyll algorithms for ocean color sensors-OC4, OC5 & OC6. Remote Sens. Environ. 229, 32–47.
O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., McClain, C., 1998. Ocean color chlorophyll algorithms for SeaWiFS. J. Geophys. Res. Oceans 103, 24937–24953.
O'Shea, R.E., Pahlevan, N., Smith, B., Bresciani, M., Egerton, T., Giardino, C., Li, L., Moore, T., Ruiz-Verdu, A., Ruberg, S., Simis, S.G.H., Stumpf, R., Vaičiūtė, D., 2021. Advancing cyanobacteria biomass estimation from hyperspectral observations: demonstrations with HICO and PRISMA imagery. Remote Sens. Environ. 266, 112693.
Page, B.P., Olmanson, L.G., Mishra, D.R., 2019. A harmonized image processing workflow using Sentinel-2/MSI and Landsat-8/OLI for mapping water clarity in optically variable lake systems. Remote Sens. Environ. 231, 111284.
Pahlevan, N., Sarkar, S., Franz, B.A., Balasubramanian, S.V., He, J., 2017a. 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., 2017b. Landsat 8 remote sensing reflectance (R rs) products: evaluations, intercomparisons, and enhancements. Remote Sens. Environ. 190, 289–301.
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., Mangin, A., Balasubramanian, S.V., Smith, B., Alikas, K., Arai, K., Barbosa, C., Bélanger, S., Binding, C., Bresciani, M., Giardino, C., Gurlin, D., Fan, Y., Harmel, T., Hunter, P., Ishikaza, J., Kratzer, S., Lehmann, M.K., Ligi, M., Ma, R., Martin-Lauzer, F.-R., Olmanson, L., Oppelt, N., Pan, Y., Peters, S., Reynaud, N., Sander de Carvalho, L.A., Simis, S., Spyrakos, E., Steinmetz, F., Stelzer, K., Sterckx, S., Tormos, T., Tyler, A., Vanhellemont, Q., Warren, M., 2021a. ACIX-aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. Remote Sens. Environ. 258, 112366.
Pahlevan, N., Smith, B., Binding, C., Gurlin, D., Li, L., Bresciani, M., Giardino, C., 2021b. Hyperspectral retrievals of phytoplankton absorption and chlorophyll-a in inland and nearshore coastal waters. Remote Sens. Environ. 253, 112200.
Petus, C., Chust, G., Gohin, F., Doxaran, D., Froidefond, J.-M., Sagarminaga, Y., 2010. Estimating turbidity and total suspended matter in the Adour River plume (South Bay of Biscay) using MODIS 250-m imagery. Cont. Shelf Res. 30, 379–392.
Pinardi, M., Bresciani, M., Villa, P., Cazzaniga, I., Laini, A., Tóth, V., Fadel, A., Austoni, M., Lami, A., Giardino, C., 2018. Spatial and temporal dynamics of primary producers in shallow lakes as seen from space: intra-annual observations from sentinel-2A. Limnologica 72, 32–43.
Raqueno, R.V., 2003. Hyperspectral analysis tools for multiparameter inversion of water quality factors in the Lake Ontario Rochester embayment. In: Environmental and Resource Engineering. State University of New York, Syracuse, p. 165.
Resley, M.J., Webb Jr., K.A., Holt, G.J., 2006. Growth and survival of juvenile cobia, Rachycentron canadum, at different salinities in a recirculating aquaculture system. Aquaculture 253, 398–407.
Roehm, C.L., Prairie, Y.T., Del Giorgio, P.A., 2009. The pCO2 Dynamics in Lakes in the Boreal Region of Northern Québec, Canada. Global Biogeochemical Cycles, p. 23.
Ross, M.R., Topp, S.N., Appling, A.P., Yang, X., Kuhn, C., Butman, D., Simard, M., Pavelsky, T.M., 2019. AquaSat: A data set to enable remote sensing of water quality for inland waters. Water Resour. Res. 55, 10012–10025.
Roulet, N., Moore, T.R., 2006. Browning the waters. Nature 444, 283–284.
Rubin, D.B., 2004. Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons.
Ruiz-Verdú, A., Simis, S.G., de Hoyos, C., Gons, H.J., Peña-Martínez, R., 2008. An evaluation of algorithms for the remote sensing of cyanobacterial biomass. Remote Sens. Environ. 112, 3996–4008.
Sathyendranath, S., Prieur, L., Morel, A., 1989. A three-component model of ocean colour and its application to remote sensing of phytoplankton pigments in coastal waters. Int. J. Remote Sens. 10, 1373–1394.
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.

Schraga, T., Nejad, E., Martin, C., Cloern, J., 2020. USGS Measurements of Water Quality in San Francisco Bay (CA), Beginning in 2016 (Ver. 3.0, March 2020). US Geological Survey data release.
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.
da Silva, E.F.F., de Moraes Novo, E.M.L., de Lucia Lobo, F., Barbosa, C.C.F., Noernberg, M.A., da Silva Rotta, L.H., Cairo, C.T., Maciel, D.A., Júnior, R.F., 2020. Optical water types found in Brazilian waters. Limnology 1–12.
Simis, S.G., Peters, S.W., Gons, H.J., 2005. Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. Limnol. Oceanogr. 50, 237–245.
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. Front. Remote Sens. 1.
Snyder, G.W., Carson, B., 1986. Bottom and suspended particle sizes: implications for modern sediment transport in Quinault submarine canyon. Mar. Geol. 71, 85–105.
Song, K., Li, L., Tedesco, L., Li, S., Duan, H., Liu, D., Hall, B., Du, J., Li, Z., Shi, K., 2013. Remote estimation of chlorophyll-a in turbid inland waters: three-band model versus GA-PLS model. Remote Sens. Environ. 136, 342–357.
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.
Spencer, R.G., Butler, K.D., Aiken, G.R., 2012. Dissolved organic carbon and chromophoric dissolved organic matter properties of rivers in the USA. J. Geophys. Res. Biogeosci. 117.
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.
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.
Sterckx, S., Knaeps, E., Ruddick, K., 2011. Detection and correction of adjacency effects in hyperspectral airborne data of coastal and inland waters: the use of the near infrared similarity spectrum. Int. J. Remote Sens. 32, 6479–6505.
Tanaka, A., Kishino, M., Doerffer, R., Schiller, H., Oishi, T., Kubota, T., 2004. Development of a neural network algorithm for retrieving concentrations of chlorophyll, suspended matter and yellow substance from radiance data of the ocean color and temperature scanner. J. Oceanogr. 60, 519–530.
Toming, K., Kutser, T., Uiboupin, R., Arikas, A., Vahter, K., Paavel, B., 2017. Mapping water quality parameters with sentinel-3 ocean and land colour instrument imagery in the Baltic Sea. Remote Sens. 9, 1070.
Vanhellemont, Q., Ruddick, K., 2021. Atmospheric correction of Sentinel-3/OLCI data for mapping of suspended particulate matter and chlorophyll-a concentration in Belgian turbid coastal waters. Remote Sens. Environ. 256, 112284.
Vinh, P.Q., Ha, N.T.T., Binh, N.T., Thang, N.N., Oanh, L., Thao, N., 2019. Developing algorithm for estimating chlorophyll-a concentration in the Thac Ba reservoir surface water using Landsat 8 imagery. VIETNAM J. Earth Sci. 41, 10–20.
Wang, G., Lee, Z., Mishra, D.R., Ma, R., 2016. Retrieving absorption coefficients of multiple phytoplankton pigments from hyperspectral remote sensing reflectance measured over cyanobacteria bloom waters. Limnol. Oceanogr. Methods 14, 432–447.
Warnock, N., Page, G.W., Ruhlen, T.D., Nur, N., Takekawa, J.Y., Hanson, J.T., 2002. Management and conservation of San Francisco Bay salt ponds: effects of pond salinity, area, tide, and season on Pacific flyway waterbirds. Waterbirds 79–92.
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.
Wei, J., Lee, Z., Shang, S., 2016. A system to measure the data quality of spectral remote-sensing reflectance of aquatic environments. J. Geophys. Res. Oceans 121, 8189–8207.
Werdell, P.J., Bailey, S.W., 2005. An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. Remote Sens. Environ. 98, 122–140.
Werdell, P.J., Franz, B.A., Bailey, S.W., Feldman, G.C., Boss, E., Brando, V.E., Dowell, M., Hirata, T., Lavender, S.J., Lee, Z., 2013. Generalized Ocean color inversion model for retrieving marine inherent optical properties. Appl. Opt. 52, 2019–2037.
Woźniak, S.B., Stramski, D., Stramska, M., Reynolds, R.A., Wright, V.M., Miksic, E.Y., Cichocka, M., Cieplak, A.M., 2010. Optical variability of seawater in relation to particle concentration, composition, and size distribution in the nearshore marine environment at Imperial Beach, California. J. Geophys. Res. 115.
Xue, K., Ma, R., Duan, H., Shen, M., Boss, E., Cao, Z., 2019. Inversion of inherent optical properties in optically complex waters using sentinel-3A/OLCI images: a case study using China's three largest freshwater lakes. Remote Sens. Environ. 225, 328–346.
Yang, W., Matsushita, B., Chen, J., Fukushima, T., 2011. Estimating constituent concentrations in case II waters from MERIS satellite data by semi-analytical model optimizing and look-up tables. Remote Sens. Environ. 115, 1247–1259.
Zhang, X., Fichot, C.G., Baracco, C., Guo, R., Neugebauer, S., Bengtsson, Z., Ganju, N., Fagherazzi, S., 2020. Determining the drivers of suspended sediment dynamics in tidal marsh-influenced estuaries using high-resolution ocean color remote sensing. Remote Sens. Environ. 240, 111682.
Zheng, G., Stramski, D., DiGiacomo, P.M., 2015. A model for partitioning the light absorption coefficient of natural waters into phytoplankton, nonalgal particulate, and colored dissolved organic components: a case study for the Chesapeake Bay. J. Geophys. Res. Oceans 120, 2601–2621.
Zhu, W., Yu, Q., 2012. Inversion of chromophoric dissolved organic matter from EO-1 Hyperion imagery for turbid estuarine and coastal waters. IEEE Trans. Geosci. Remote Sens. 51, 3286–3298.