Remote Sensing of Environment 282 (2022) 113233

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Global satellite water classification data products over oceanic, coastal, and inland waters

Jianwei Wei [a,b,*], Menghua Wang [a], Karlis Mikelsons [a,b], Lide Jiang [a,c], Susanne Kratzer [d], Zhongping Lee [e], Tim Moore [f], Heidi M. Sosik [g], Dimitry Van der Zande [h]

[a] NOAA/NESDIS Center for Satellite Applications and Research, College Park, MD 20740, USA
[b] Global Science & Technology, Inc., Greenbelt, MD 20770, USA
[c] Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO 80523, USA
[d] DEEP, Stockholm University, 106 91 Stockholm, Sweden
[e] School for the Environment, University of Massachusetts Boston, Boston, MA 02125, USA
[f] Harbor Branch Oceanographic Institute, Florida Atlantic University, Fort Pierce, FL 34946, USA
[g] Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA
[h] Operational Directorate Natural Environment, Royal Belgian Institute of Natural Sciences, 1000 Brussels, Belgium

ARTICLE INFO

Editor by: Marie Weiss

Keywords:
Water class
Optical water type
Remote sensing reflectance
Spectral similarity
VIIRS
Hyperspectral

ABSTRACT

Satellites have generated extensive data of remote sensing reflectance spectra ($R_{rs}(\lambda)$) covering diverse water classes or types across global waters. Spectral classification of satellite $R_{rs}(\lambda)$ data allows for the distinguishing and grouping of waters with characteristic bio-optical/biogeochemical features that may influence the productivity of a given water body. This study reports new satellite water class products (Level-2 and Level-3) from the Visible Infrared Imaging Radiometer Suite (VIIRS). We developed and implemented a hyperspectral scheme that accounts for the $R_{rs}(\lambda)$ spectral shapes and globally resolves oceanic, coastal, and inland waters into 23 water classes. We characterized the light absorption and scattering coefficients, chlorophyll-a concentration, diffuse attenuation coefficient, and suspended particulate matter for individual water classes. It is shown that the water classes are separable by their distinct bio-optical and biogeochemical properties. Furthermore, validation result suggests that the VIIRS water class products are accurate globally. Finally, we examined the spatial and temporal variability of the water classes in case studies for a demonstration of applications. The water class data in open oceans reveal that the subtropical ocean gyres have experienced dramatic expansion over the last decade. In addition, the water class data appear to be a valuable (and qualitative) indicator for water quality in coastal and inland waters with compelling evidence. We stress that this new satellite product is an excellent addition to the aquatic science database, despite the need for continuous improvement toward perfection.

1. Introduction

Satellite ocean color measurements over open oceans and coastal and inland waters represent big data covering diverse water classes or types. Global partitioning of satellite ocean color data is needed to decipher the optical complexity and unravel the optical diversity of natural waters (IOCCG, 2009). The utility of water classification is manifested in the development of class-specific algorithms to improve the estimation of water's bio-optical and biogeochemical properties (Cui et al., 2020; Jiang et al., 2020; Le et al., 2011; Moore et al., 2001; Uudeberg et al., 2020). In addition, the classification of global waters provides a valuable framework for the understanding of satellite product uncertainties (Moore et al., 2009; Wei et al., 2016b) and mechanisms controlling ocean biology, chemistry, and physics (Longhurst, 1998; Longhurst et al., 1995; Martin Traykovski and Sosik, 2003; Oliver and Irwin, 2008). Therefore, generating well-defined water classes is a priority for ocean color satellite missions.

Water classification entails the grouping of waters with similar bio-optical and biogeochemical signatures. In seminal work, Morel and Prieur (1977) proposed two abstract water classes (Case 1 and Case 2). Within Case 1 waters, phytoplankton and associated colored dissolved organic matter (CDOM) and detritus (collectively colored detrital matter (CDM)) dominate the optical properties. For Case 2, inorganic materials, such as mineral particles and CDM, significantly influence the optical

* Corresponding author at: NOAA/NESDIS Center for Satellite Applications and Research, College Park, MD 20740, USA.
E-mail address: jianwei.wei@noaa.gov (J. Wei).

https://doi.org/10.1016/j.rse.2022.113233
Received 27 December 2021; Received in revised form 15 August 2022; Accepted 22 August 2022
Available online 24 September 2022

properties. Due to a lack of precise separation criteria, this binary scheme was not systematically implemented with satellite data (Lee and Hu, 2006; Mobley et al., 2004). Others based their work on in situ or satellite bio-optical/biogeochemical data, such as the diffuse attenuation coefficient ($K_d(\lambda)$) (Jerlov, 1976), water inherent optical properties (IOPs) (Mckee and Cunningham, 2006; Shi et al., 2014), and chlorophyll-a concentration (Chl-a) (Esaias et al., 2000; Hardman-Mountford et al., 2008). The majority of the recent efforts have specifically focused on using the ocean remote sensing reflectance ($R_{rs}(\lambda)$) (or normalized water-leaving radiance, $nL_w(\lambda)$) to discriminate water classes (Botha et al., 2020; Jackson et al., 2017; Martin Traykovski and Sosik, 2003; Mélin and Vantrepotte, 2015; Moore et al., 2009; Spyrakos et al., 2011). Adopting $R_{rs}(\lambda)$ spectra for water classification has its advantages, as reflectance is the primary ocean color retrieval from satellites and contains critical information about water IOPs, i.e., the absorption and scattering properties. Two main lines of $R_{rs}(\lambda)$ classifications are practiced within the ocean color communities. The first line is to construct a set of reference classes from a training data set involving absolute $R_{rs}(\lambda)$ measurements and then assign a class to each satellite $R_{rs}(\lambda)$ based on its difference from the references. Various techniques were tested, such as the fuzzy logic and *k*-means approaches (Jackson et al., 2017; Moore et al., 2009; Spyrakos et al., 2018). Another mainstream classification is to achieve water classifications through the "spectral shape" of $R_{rs}(\lambda)$, which does not rely on the absolute magnitudes of $R_{rs}(\lambda)$ spectra (Jia et al., 2021; Lubac and Loisel, 2007; Mélin and Vantrepotte, 2015; Vantrepotte et al., 2012; Wei et al., 2016b). As demonstrated in a few studies (Pitarch et al., 2019; Wernand and van der Woerd, 2010), the Forel-Ule (FU) color scale falls in the spectral shape category. In addition, Vandermeulen et al. (2020) proposed that an apparent visible wavelength (AVW) parameter can be used for water classification.

Reflecting on these achievements, we acknowledge that some fundamental tasks and challenges are yet to be overcome with the existing water classification schemes. First and foremost, despite the efforts and improvements, evidence is needed to demonstrate that the classification schemes can divide global waters into classes with distinct optical and biogeochemical properties. Second, very few classification schemes were implemented with global Level-2 ocean color data. Thus, appraisal of the classification schemes is thus often hindered by the lack of global water class data and corresponding water bio-optical and biogeochemical products (Sathyendranath et al., 2019). Third, most historical classification schemes were designed for specific instruments, limiting their general applicability across ocean color satellites. Though projects such as the Ocean-Colour Climate Change Initiative (OC-CCI) provide water classification products across a multi-sensor record, it is achieved by converting the whole record to a "reference-sensor" with an associated water class set. Ideally, we would like to have a single optical water class set that is applicable across multiple sensors at their native wavelengths.

The objectives of this study are 1) to report the first satellite water class products generated from the Visible Infrared Imaging Radiometer Suite (VIIRS) on the Suomi National Polar-orbiting Partnership (SNPP) mission (2011–present), 2) to characterize the water classes concerning their bio-optical and biogeochemical variability, and 3) to evaluate the uncertainties of the satellite water class products. The new satellite water classification has the following three major improvements. First, the scheme is composed of hyperspectral references and is generic, i.e., readily adaptable to other satellite $R_{rs}(\lambda)$ data with different spectral band settings. Second, the new water classes are characteristic of distinct bio-optical and biogeochemical properties found globally. Third, the new water classes have provided by far the broadest coverage of global oceanic, coastal, and inland waters. In the following, Section 2 provides data and methods used in the study, and the main results of this research are given in Section 3. In Section 4, we demonstrate various applications of the water class data to aquatic ecology and water quality monitoring. We provide some detailed discussion in Section 5 and finally conclude in Section 6.

2. Data and methods

2.1. Hyperspectral reference data

The training data are composed of a large set of hyperspectral $R_{rs}(\lambda)$ spectra in the visible domain (400–700 nm). Some of them are in situ $R_{rs}(\lambda)$ measurements ($N = 953$) recovered mainly from the open ocean and coastal environments. A description of the measurement procedures and data processing is available elsewhere (Wei et al., 2016b). Others were derived from radiative transfer simulations ($N = 1000$), which are further based on synthetic IOPs adapted from IOCCG (2006). The detailed model configuration was described by Wei et al. (2019). Note that, despite the relatively small training data size, the resulting reference spectra are shown to competently represent the satellite $R_{rs}(\lambda)$ data globally (to be discussed in Section 3.2).

We used the normalized $R_{rs}(\lambda)$ spectra for subsequent analysis, which are expressed as:

$$nR_{rs}(\lambda_i) = \frac{R_{rs}(\lambda_i)}{\left[\sum R_{rs}(\lambda_i)^2\right]^{1/2}}, i = 1, 2 \ldots \mathrm{N}. \quad (1)$$

The denominator in Eq. (1) is known as the root of the sum of squares (RSS). The $nR_{rs}(\lambda)$ spectra have a unique quality, with $nR_{rs}(\lambda_1)^2 + nR_{rs}(\lambda_2)^2 + \ldots + nR_{rs}(\lambda_N)^2 = 1$. An unsupervised *k*-means technique is then evoked, which clusters the $nR_{rs}(\lambda)$ spectra by taking into account the spectral similarity among them. Specifically, the cosine distance, *d*, is calculated and utilized during spectral clustering, with

$$d = 1 - \frac{\mathrm{A} \cdot \mathrm{B}}{\|\mathrm{A}\| \|\mathrm{B}\|}, \quad (2)$$

where $\mathbf{A} \cdot \mathbf{B}$ represents a dot product between two individual $nR_{rs}(\lambda)$ spectra denoted as **A** and **B**. $\|\mathbf{A}\|$ and $\|\mathbf{B}\|$ refer to the magnitudes of **A** and **B**, respectively, where $\|\mathbf{A}\| = 1$ and $\|\mathbf{B}\| = 1$. Obviously, the normalization in Eq. (1) suppresses the spectral amplitudes of $R_{rs}(\lambda)$ while retaining and highlighting the spectral shapes. Two additional aspects of water classification are particularly worth noting. First, determining the total number of classes is crucial for data clustering. We did not attempt to derive new class numbers. Instead, we adopted the class number $k = 23$ from Wei et al. (2016b). This class number is greater than almost all other existing water classes, except for Longhurst's ocean provinces (Longhurst, 1998). This number will most likely evolve with our understanding of ocean optics in the future. Second, it is ideal for clustering analysis with hyperspectral data to best describe the variability of natural waters. However, the present $R_{rs}(\lambda)$ training data are representative of a wide range of conditions in natural waters. Therefore, as a practical alternative, we partitioned the $nR_{rs}(\lambda)$ data with an emphasis on five bands, with nominal wavelengths centered at 410, 443, 488, 551, and 670 nm. These bands are "common" for operational ocean color sensors. Finally, we derived the median $nR_{rs}(\lambda)$ spectrum for each of the 23 clusters as the corresponding reference.

In Fig. 1, the median $nR_{rs}(\lambda)$ spectra are sorted in descending order of $nR_{rs}(443)/nR_{rs}(551)$, representing 23 water classes from Class 1 to Class 23. Each water class is clustered with $nR_{rs}(\lambda)$ spectra of similar spectral shape, evidenced by the small in-class variability of the $nR_{rs}(\lambda)$ spectra. From Class 1 to Class 17, the coefficient of variation (CV) (the ratio of the variance to the median of $nR_{rs}(\lambda)$) is limited to ~15% at 400–600 nm (left panel, Fig. A1). The $nR_{rs}(\lambda)$ variance slightly increases in the blue bands for Class 18–23 primarily because the $nR_{rs}(\lambda)$ values in the blue are usually small, indicative of strong light absorption in these waters. In the red bands, the $nR_{rs}(\lambda)$ spectra tend to exhibit higher variance, especially for Class 1–15, in which the $nR_{rs}(\lambda)$ values are also usually small. Class 16–23 waters represent very turbid environments. The high $nR_{rs}(\lambda)$ values in the red bands over these turbid environments do not always imply that the corresponding $nR_{rs}(\lambda)$ variance is insignificant. As shown in earlier studies, particle backscattering, including scattering

due to abundant phytoplankton, can significantly alter the reflectance across long red and near-infrared (NIR) bands (Doxaran et al., 2002; Gower et al., 2005). In addition to the relatively high variance at the shortest blue and the longest red bands, the $nR_{rs}(\lambda)$ variance in Class 16–23 remains small in green bands. The observation of the small overall variance of in-class $nR_{rs}(\lambda)$ spectra verifies the validity of the clustering procedure and the utility of the hyperspectral classes.

2.2. Application of the water classification scheme to satellite $R_{rs}(\lambda)$ data

We applied the water classification scheme to the VIIRS-SNPP Level-2 ocean color data. VIIRS observes the Earth from daytime ascending polar orbits, with a spatial resolution of 750 m. The radiance measurements were processed with the Multi-Sensor Level-1 to Level-2 (MSL12) ocean color data processing system (Wang et al., 2017; Wang et al., 2013), which involves the NIR-, shortwave infrared (SWIR)-, and NIR-SWIR algorithms for atmospheric correction (Gordon and Wang, 1994; Wang, 2007; Wang and Shi, 2007), as well as the NIR ocean reflectance correction algorithm (Jiang and Wang, 2014). The resulting $R_{rs}(\lambda)$ images have ocean color bands centered at 410, 443, 486, 551, and 671 nm. To assign a water class to a pixel, we calculated the cosine distance between the $R_{rs}(\lambda)$ at that pixel and the reference $nR_{rs}(\lambda)$ spectrum of every water class at corresponding spectral bands as

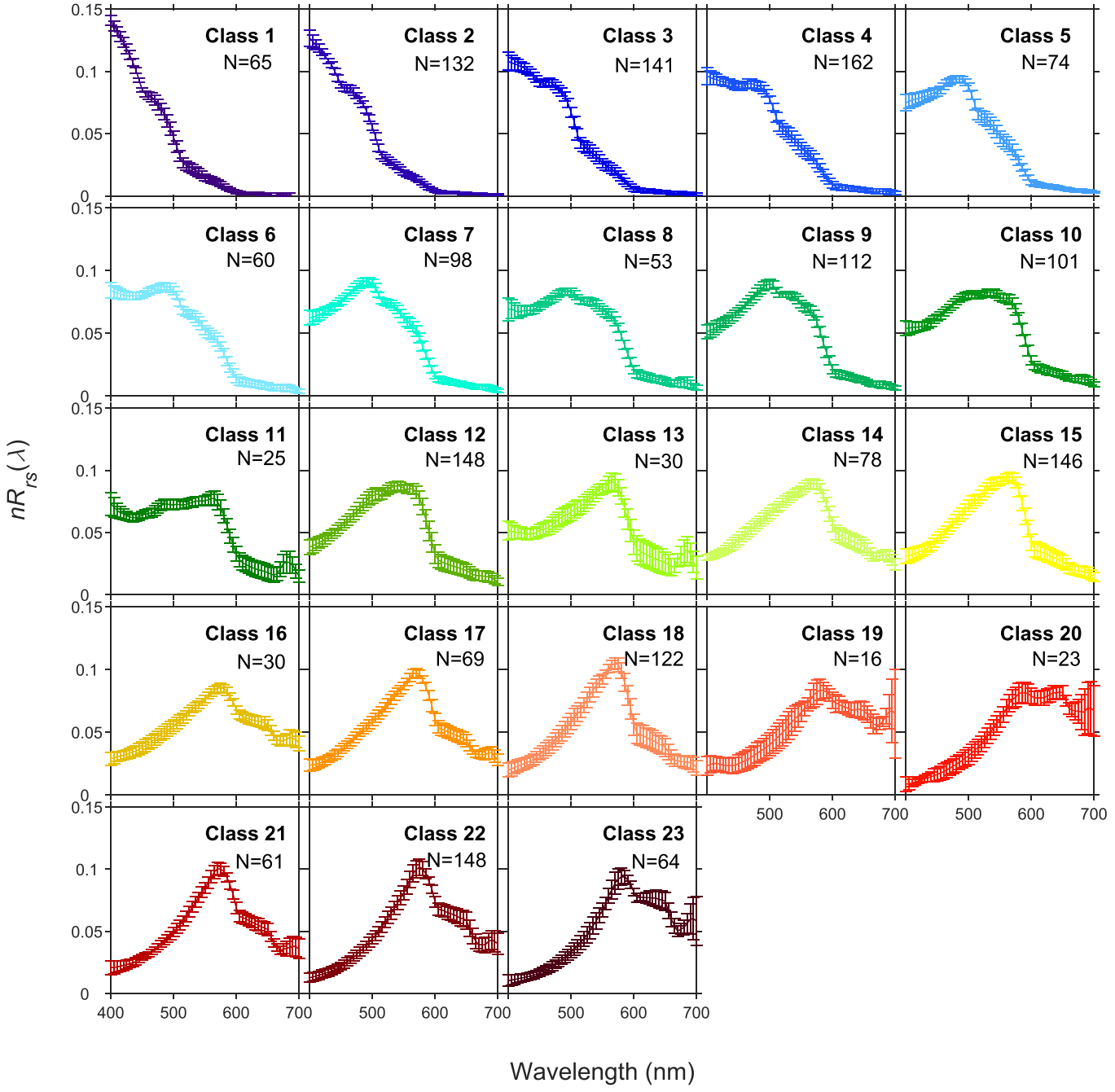

Fig. 1. Hyperspectral $nR_{rs}(\lambda)$ reference spectra for global water classes. The error bars refer to the standard deviation of $nR_{rs}(\lambda)$ within each cluster. The numbers given in each subplot are the number of the $nR_{rs}(\lambda)$ spectra assigned to every water class.

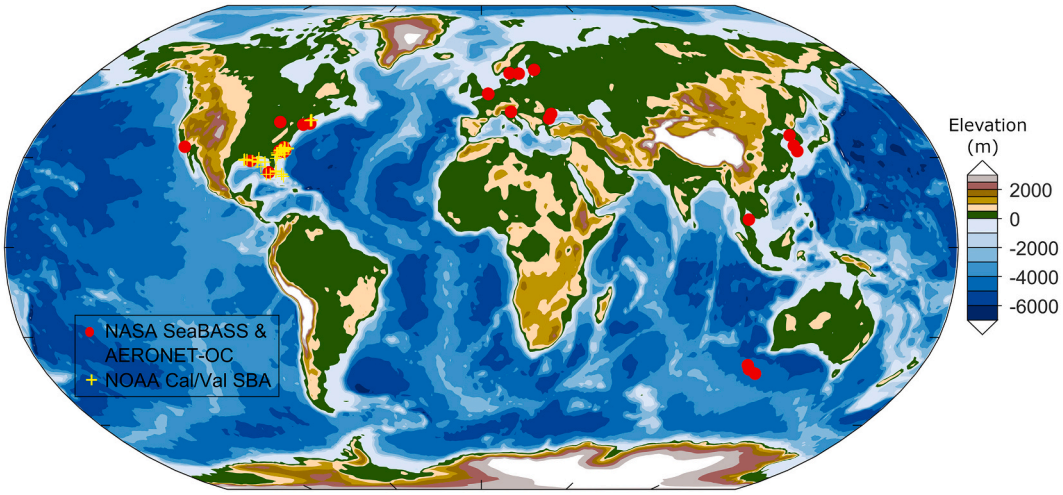

Fig. 2. Locations of the in situ $R_{rs}(\lambda)$ measurements with concurrent VIIRS-SNPP overpass.

$$d = 1 - \frac{\sum_{i=1}^{5}[nR_{rs}(\lambda_i)R_{rs}(\lambda_i)]}{\sqrt{\sum_{i=1}^{5}[nR_{rs}(\lambda_i)]^2 \sum_{i=1}^{5}[R_{rs}(\lambda_i)]^2}}, \quad (3)$$

where λ_1, λ_2, λ_3, λ_4, and λ_5 correspond to the five VIIRS spectral bands, i.e., 410, 443, 486, 551, and 671 nm, respectively. The water class was determined as the one with the smallest distance, d, to the reference. Daily, 8-day, and monthly water class products were then generated as the rounded median values of daily products within the corresponding time and/or spatial intervals. The Level-2 and Level-3 water class products are accessible through the interactive online maps or the Ocean Color Viewer (OCView) (Mikelsons and Wang, 2018).

We also applied the classification scheme to example images from the Ocean and Land Colour Instrument (OLCI) onboard the Sentinel-3A satellite and the Hyperspectral Imager for the Coastal Ocean (HICO). The OLCI image was processed with MSL12, which used two NIR bands (754 and 865 nm) to determine aerosol types and carry out atmospheric correction. The $R_{rs}(\lambda)$ spectra at the ten visible OLCI bands (400, 413, 443, 490, 510, 560, 620, 665, 674, and 681 nm) were considered for water classification. The HICO images were processed using L2GEN with full spectral coverage (380 to 960 nm sampled at 5.7 nm) (Ibrahim et al., 2018). Only the visible bands of HICO were considered in our analyses. The water classes derived from OLCI and HICO will be discussed in Section 5.2.

2.3. Global bio-optical and biogeochemical data

We examined the associated bio-optical and biogeochemical product data to characterize the water classes derived from VIIRS-SNPP. The light absorption coefficient of phytoplankton ($a_{ph}(443)$), the absorption coefficient of CDM ($a_{dg}(443)$), and the backscattering coefficient of particles ($b_{bp}(443)$) were derived from $R_{rs}(\lambda)$ data using a combination of a NIR algorithm and the quasi-analytical algorithm (NIR-QAA) (Shi and Wang, 2019). The total (non-water) absorption coefficient at 443 nm was determined as the sum of phytoplankton and CDM contributions, with $a_{pg}(443) = a_{ph}(443) + a_{dg}(443)$. The global Chl-a data were generated using the ocean color index (OCI) algorithm (Wang and Son, 2016). With $a_{ph}(443)$ and Chl-a data, we also derived the chlorophyll-specific absorption coefficient as $a_{ph}^*(443) = a_{ph}(443)/\text{Chl-a}$. The diffuse attenuation coefficient at 490 nm ($K_d(490)$) was estimated from a combination of the open ocean algorithm (empirical) and turbid water algorithm (semi-analytical) (Wang et al., 2009). SPM products were recently created with an algorithm involving the NIR, red, green, and blue bands (known as the NIR-RGB algorithm) (Wei et al., 2021a; Yu et al., 2019). This SPM switching algorithm is based on $R_{rs}(\lambda)$ at the red and NIR bands for SPM $\geq \sim 2$ mg l^{-1} (where $R_{rs}(671) \geq 0.0012$ sr^{-1}) and SPM $< \sim 2$ mg l^{-1} (where $R_{rs}(671) < 0.0012$ sr^{-1}). With the 9-km composite data (VIIRS monthly product from 2012 to 2020), we derived the class-specific means, medians, and quantiles (5%, 25%, 75%, and 95%) for each quantity. The maximum and minimum values for each water property were determined as the 97.5% and 2.5% quantiles, respectively.

2.4. Satellite and in situ matchup data

VIIRS and in situ $R_{rs}(\lambda)$ matchups were constructed to assess satellite-derived water class products (see Fig. 2). Most in situ $R_{rs}(\lambda)$ spectra were obtained from the NASA SeaWiFS Bio-optical Archive and Storage (SeaBASS). The SeaBASS database includes measurements contributed by numerous principal investigators, including ones with the Aerosol Robotic Network-Ocean Color (AERONET-OC) (Zibordi et al., 2009b). We also incorporated field measurements from the NOAA dedicated ocean color calibration/validation campaigns (Ondrusek et al., 2015, 2016; Ondrusek et al., 2017; Ondrusek et al., 2019). Specifically, the in situ hyperspectral $R_{rs}(\lambda)$ data from the skylight-blocking approach (SBA) were used (Lee et al., 2013; Wei et al., 2021b). The coincident VIIRS Level-2 $R_{rs}(\lambda)$ products were located in accordance with the in situ data and were subjected to data filtering (Bailey and Werdell, 2006). The satellite data within a 3 × 3 pixel box were checked for data quality flags. The pixels masked by land, cloud, stray light, high glint, low radiance at 551 nm, high top-of-atmosphere (TOA) radiance, and/or atmospheric correction failure were excluded from further analysis (Wang et al., 2017). The time difference between VIIRS overpass and in situ samplings was limited to within ±3 h. The maximum coefficient of variation of $R_{rs}(\lambda)$ inside the pixel box was set to 15%. The maximum difference between the measured and modeled solar irradiance at the sea surface was set to 20% to ensure sky clearness.

2.5. Error metrics

We quantified the difference between satellite- and in situ-derived water classes as

$$\Delta_i = S_i - I_i, \quad (4)$$

where S_i refers to the water class obtained from a satellite $R_{rs}(\lambda)$ spectrum and I_i is the concurrent in situ $R_{rs}(\lambda)$ measurement. Thus, the absolute error for water class estimations was calculated as $|\Delta_i| = |S_i - I_i|$.

Later in Section 4, we used two other metrics to estimate the dif-

ference between the water bio-optical properties from VIIRS and their estimations from water classes. The median absolute percentage difference (MAPD) was derived as

$$MAPD = median\left\{\left|\frac{M_i - T_i}{T_i}\right|\right\} \times 100\%, i = 1, 2, \ldots N, \quad (5)$$

where M is compared against a reference value T, and it refers to any bio-optical/biogeochemical quantity under investigation, such as Chl-a or $a_{ph}(443)$. Correspondingly, the median relative percentage difference (MRPD) or percentage bias between M and T was determined as

$$MRPD = median\left\{\frac{M_i - T_i}{T_i}\right\} \times 100\%, i = 1, 2, \ldots N. \quad (6)$$

3. Results

3.1. An overview of the new water class product

A global VIIRS climatology map (2012−2020) illustrates the spatial distribution of water classes (Fig. 3a). As indicated by the classifications, Class 1–6 waters dominate globally and occupy >90% of the surface waters (the frequency distribution of the global water classes can be found in Fig. A2). The remainder of the water classes (Class 7–23) often exist in coastal/inland water environments, which account for a relatively small fraction (~10%) of all waters globally. Examination of the global map reveals that our classifications have reproduced well-known water regimes. For instance, Class 1 prevails in the central basins of the North Atlantic, South Atlantic, North Pacific, South Pacific, and Indian Ocean, which coincides geographically with the subtropical ocean gyres (McClain et al., 2004). The Mediterranean Sea falls into two primary water classes (Class 2 and Class 3), concurring with the distinctive seasonal phytoplankton biomass patterns in its western and eastern basins (Barale et al., 2008; Bosc et al., 2004; Salgado-Hernanz et al., 2019). The turbid Yangtze River plume stands out from surrounding waters on the northwest Pacific coasts, demarcated by Class ~10–20. Along the northwest Atlantic coasts, George's Bank is separated from the Gulf of Maine and the Atlantic Ocean, and features distinct water classes (~8–11) evident on the map. The water classes depicted in the Greater North Sea are also highly comparable with other regional classifications (Van der Zande et al., 2019). As expected, waters with the highest class numbers (Class ~19–23) correspond to extremely turbid environments, such as the Yangtze River Estuary, Río de la Plata, and Amazon River Estuary. Fig. 3b further depicts the differences between maximum and minimum water classes observed in every pixel. Relatively larger differences often exist in coastal and inland water environments,

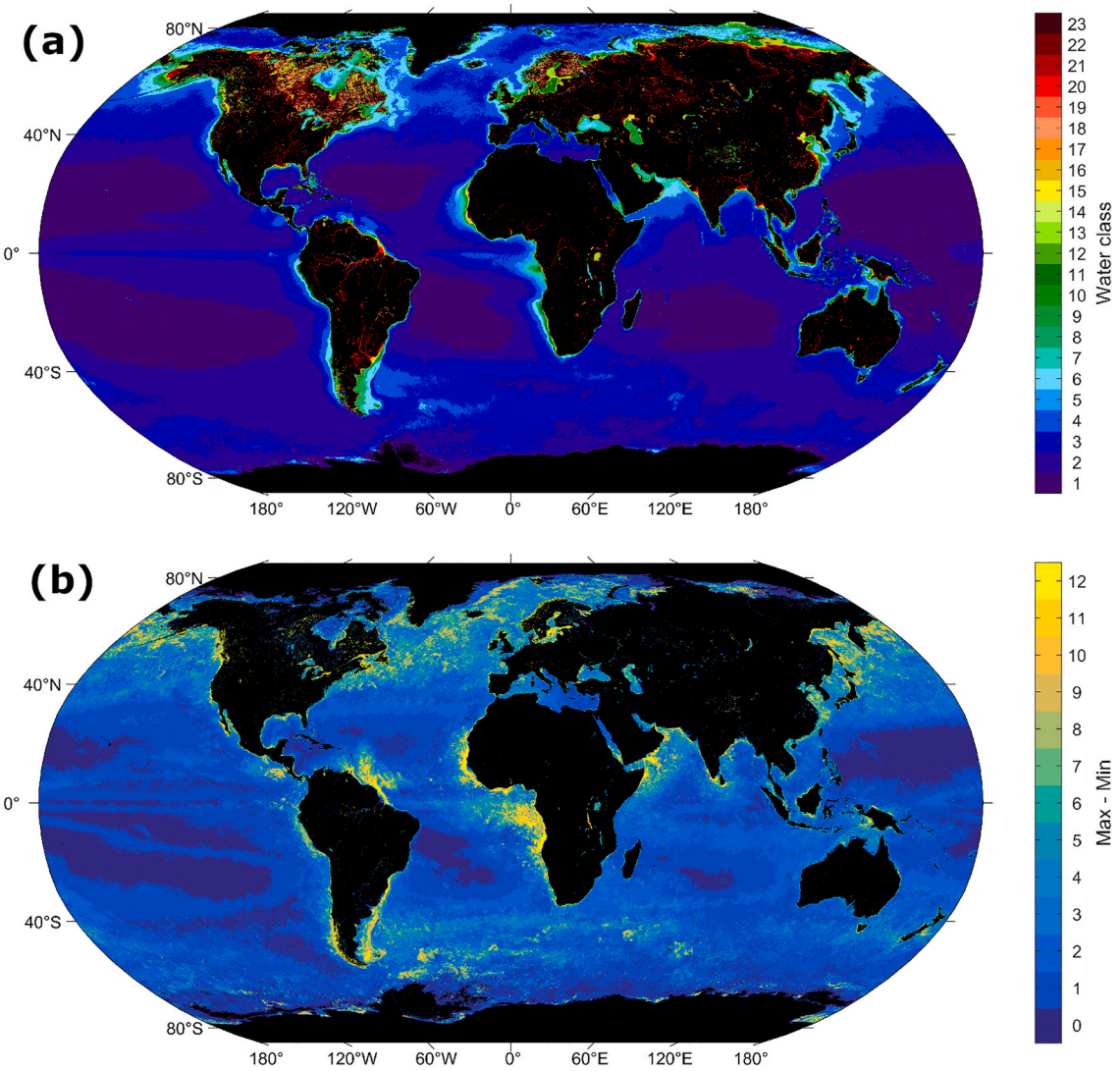

Fig. 3. (a) Climatology of the global water classes derived from VIIRS-SNPP (2012–2020, 9 km) and (b) range of variability of the monthly water classes (quantified by the differences between maximum and minimum water classes). Black pixels indicate land and/or invalid ocean color data.

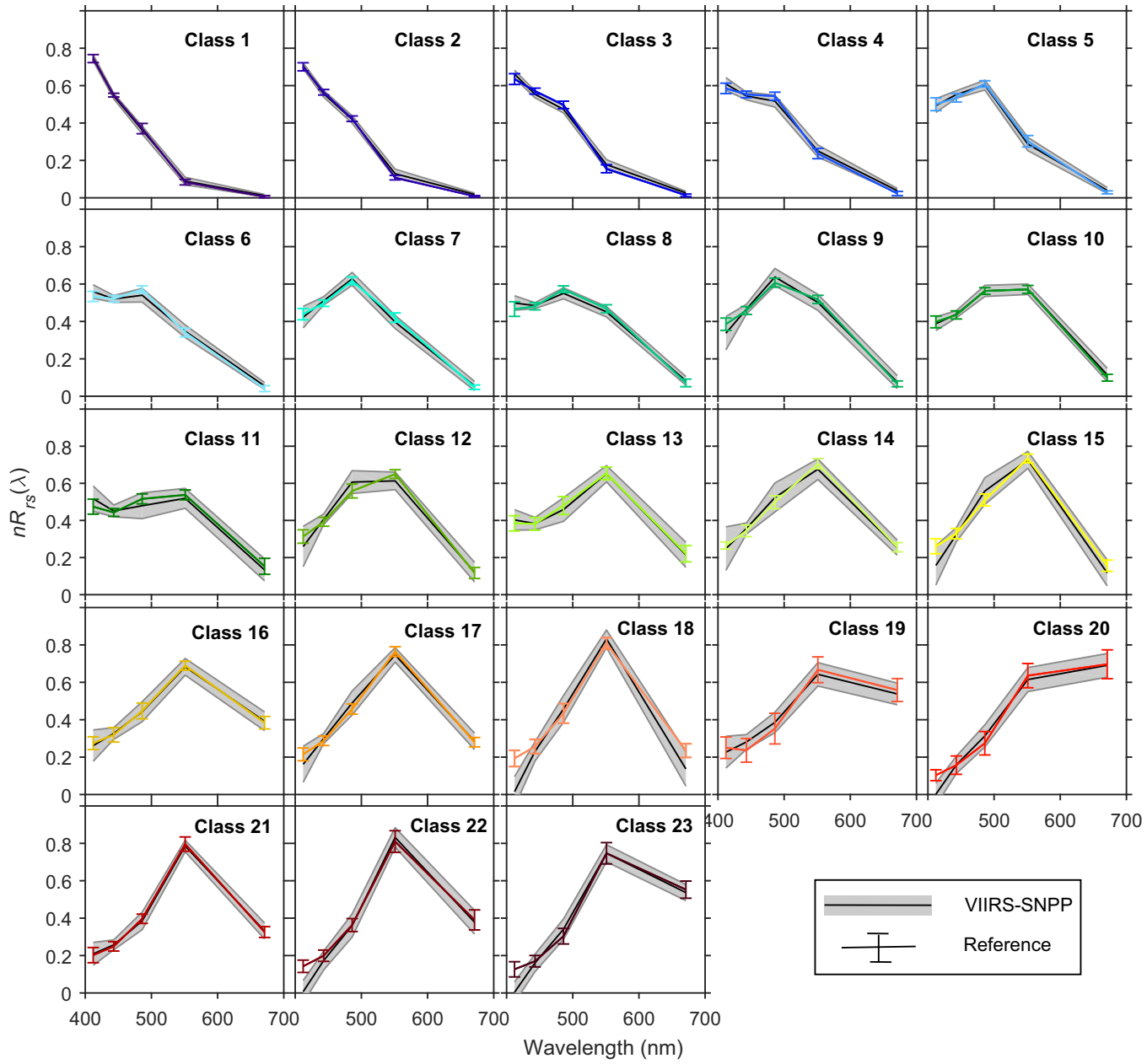

Fig. 4. Comparison of VIIRS-SNPP median $nR_{rs}(\lambda)$ spectra (the variance is denoted in the shade) with the reference spectra (with the error bars indicating the variance of the training data). VIIRS-SNPP $R_{rs}(\lambda)$ data (Level-3; 9-km; August 18, 2012) were normalized by their respective RSS values at 410, 443, 486, 551, and 670 nm. The reference spectra were extracted from the hyperspectral spectra of Fig. 1 and then rescaled by the RSS values of $nR_{rs}(\lambda)$ at the five spectral bands.

suggesting considerable variability of the water classes in these regions. The differences in open oceans are typically small in the analyzed time period.

3.2. Representativeness of the water classes

To shed light on the representativeness of the derived water classes, we compared the VIIRS reflectance spectra from global waters to the reference spectra. As shown in Fig. 4, the class-specific median VIIRS $nR_{rs}(\lambda)$ spectra are almost identical to the references. In the bands of 443 nm, 486 nm, and 551 nm, two sets of spectra show minimal differences (within ~ ±10%) (Table 1). Larger differences are occasionally present at 410 nm for Class 9–23 and at 671 nm for Class 1–11. In particular, the comparison suggests that satellite $nR_{rs}(410)$ values are often biased low in Class 15, Class 18, Class 20, Class 22, and Class 23, where the relative differences are on average equal to −40%, −92%, −99%, −95%, and −97%, respectively. Such biases will undoubtedly lead to incorrect satellite $R_{rs}(410)/R_{rs}(443)$ ratios. Underestimation of $R_{rs}(410)$ values in nearshore waters is a common problem with satellite measurements due to inadequate atmospheric correction (Wang and Jiang, 2018; Wei et al., 2020). Small $R_{rs}(671)$ values in Class 1–11 render increased percentage biases between the references and satellite measurements. Furthermore, we note that the spectral variances of VIIRS $nR_{rs}(\lambda)$ data and the references in Fig. 4 are also comparable. The similarity in spectral variance is most conspicuous in Class 1–11, representing approximately 95% of the total global water surface area (Fig. A2). On one hand, the atmospheric correction problem is a plausible explanation for such an increase (Wang and Jiang, 2018). On the other hand, the increased variability in satellite

Table 1
Relative percentage difference between the median VIIRS $nR_{rs}(\lambda)$ spectra and the classification references. The values exceeding ±10% are denoted in bold.

Water Class	nR_{rs}(410)	nR_{rs}(443)	nR_{rs}(486)	nR_{rs}(551)	nR_{rs}(670)
1	0.5%	−0.9%	−2.1%	7.0%	**103.0%**
2	0.7%	−1.5%	−0.9%	**19.8%**	**121.4%**
3	4.2%	−3.0%	−4.8%	**14.1%**	**99.8%**
4	4.0%	−1.6%	−4.8%	5.8%	**68.2%**
5	−1.4%	3.2%	−1.0%	−4.9%	**34.9%**
6	4.8%	−0.2%	−5.4%	2.1%	**34.3%**
7	−3.7%	2.6%	1.7%	−5.6%	**13.4%**
8	7.0%	0.8%	−3.8%	−3.6%	**14.9%**
9	**−12.5%**	1.5%	5.0%	−3.1%	**13.4%**
10	−2.6%	0.9%	0.1%	0.1%	**15.5%**
11	9.0%	1.9%	−7.2%	−3.5%	**−12.6%**
12	**−17.2%**	3.3%	8.6%	−5.6%	4.1%
13	4.8%	0.3%	−4.5%	−0.3%	−2.7%
14	−6.4%	3.6%	5.2%	−4.2%	−1.0%
15	**−39.6%**	−1.6%	9.4%	−1.1%	**−25.2%**
16	−4.8%	2.2%	−0.9%	−0.9%	1.9%
17	**−24.7%**	4.2%	7.0%	−2.3%	1.5%
18	**−92.1%**	**−10.1%**	5.4%	3.0%	**−41.9%**
19	−9.6%	**18.9%**	9.3%	−3.6%	−3.7%
20	**−98.9%**	1.0%	**14.2%**	−3.3%	−0.8%
21	3.3%	3.2%	−2.9%	−1.3%	1.7%
22	**−95.2%**	−13.0%	−0.7%	2.7%	−3.2%
23	**−97.2%**	−5.3%	12.3%	0.0%	−2.9%

data can result from the optical complexity of natural waters. Overall, the reference spectra in Fig. 4 are a fair representation of global VIIRS $R_{rs}(\lambda)$ spectra.

We further compared our water classification with previous efforts to highlight how our new classification might describe global waters effectively. In Table 2, six historical classifications (Jackson et al., 2017; Jia et al., 2021; Lubac and Loisel, 2007; Mélin and Vantrepotte, 2015; Pitarch et al., 2019; Spyrakos et al., 2018) (abbreviated as J17, J21, L08, M15, P19, and S18, respectively) are aligned with Class 1–23 based on their spectral similarity to Class 1–23 (recall Eq. (3)). The present classification has achieved a broad coverage of water types comparable to P19, J21, and S18. However, some water types associated with our classification are missing from P19, J21, and S18. For example, S18 lacks Class 4–5, Class 7, Class 9, Class 13, Class 15, and Class 17–18. Nevertheless, multiple instances (indicated by "M" in Table 2) from P19, J21, and S18 match with one of our water classes, especially for Class 14–23. For instance, S18 contains multiple water types that correspond to Class 23, suggesting a possibly finer resolution of this specific water

Table 2
Correspondence between the water classes developed in this study and historical classifications.[a] The letter "O" indicates that a water class type from a historical scheme matches with one of our new classes. The letter "M" refers to a situation when multiple classes from a historical scheme are found to be spectrally similar to one of the new water classes. The notation "–" indicates no corresponding water class found from other historical schemes.

Water class from this study	Numbers of correspondences					
	Jia et al. (2021)	Pitarch et al. (2019)	Jackson et al. (2017)	Moore et al. (2009)	Mélin and Vantrepotte (2015)	Spyrakos et al. (2018)
1	O	O	M		M	O
2	–	O	M	O	M	M
3	O	–	M	O	M	O
4	–	O	O	O	M	–
5	O	–	–		O	–
6	–	O	M	O	M	O
7	O	–	–	–	O	–
8	–	O	–	–	–	O
9	O	–	O	–	M	–
10	–	O	–	O	–	O
11	–	–	–	–	–	O
12	O	O	–	–	O	O
13	–	–	O	M	–	–
14	–	M	M	O	–	M
15	M	–	–	–	–	–
16	–	–	–	–	–	O
17	–	M	–	–	O	–
18	O	–	–	–	–	–
19	–	M	–	–	–	M
20	M	O	–	–	–	M
21	O	M	–	–	–	O
22	O	–	–	–	–	O
23	–	–	–	–	–	M

[a] The method described in Eq. (3) and Section 2.2 was used to establish the correspondence between these historical classes and our new water classes.

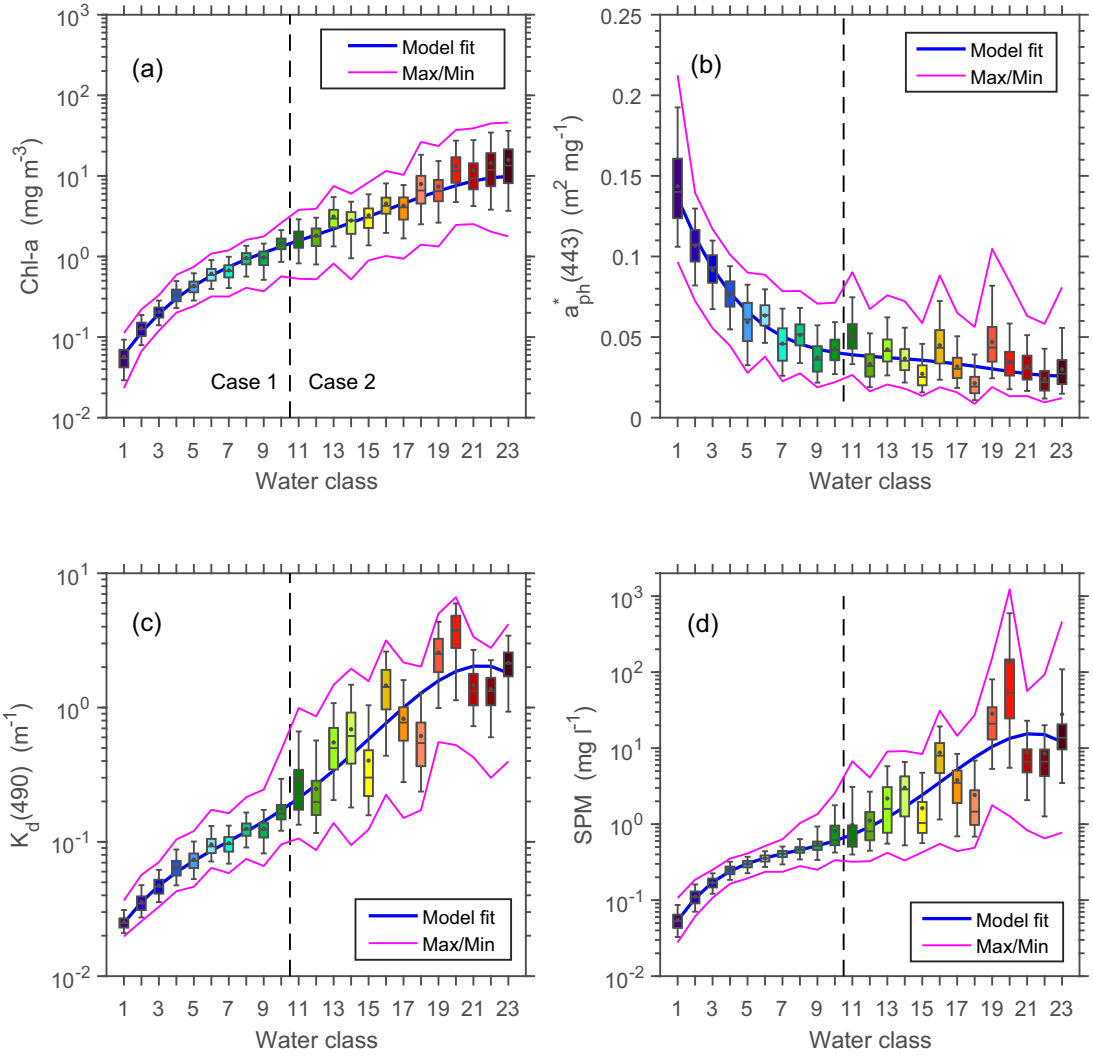

Fig. 5. Variation of Chl-a, a_{ph}*(443), K_d(490), and SPM across the water classes. Note that the dot and horizontal bar within the boxes refer to the mean and median, respectively. The upper and lower edges of the box indicate 75% and 25% percentiles, respectively, while the whiskers correspond to 95% and 5% percentiles, respectively. Beyond the boxes, the lines (in magenta) refer to the maximum and minimum values (97.5% and 2.5% percentiles, respectively). The vertical dashed line approximately separating "Case 1" and "Case 2" waters is determined with the approach of Lee and Hu (2006) (assuming ±20% deviation of $R_{rs}(410)/R_{rs}(443)$ and ±75% of $R_{rs}(551)$). Finally, a polynomial fit to the median values and the water classes is overlaid (see fitting results in Table 4). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

class. Unlike the present classification (and P19, J21, and S18), both M09 and J17 focus on open oceans and are primarily limited to Class 1–14. The latest version of M15 produces a total of 17 water types for coastal oceans (< 200 km from coastlines), which has a much-reduced resolution for Class 10–23 but may be able to reveal finer details than Class 1–7 presented here.

3.3. Bio-optical characterization

This section presents each water class's bio-optical and biogeochemical characteristics (Fig. 5 and Fig. 6). The associated statistical data for each quantity are tabulated to aid the analyses (see Table 3, Table 5, Table A1, and Table A2).

3.3.1. Chl-a concentration

The median Chl-a values, as estimated from VIIRS, increase from ~0.05 mg m^{-3} in Class 1 to >10 mg m^{-3} in Class 23 (Fig. 5a and Table 3), with a pattern closely following a 4th-order polynomial model ($R^2 = 0.99$) (Table 4). Such an observation is expected primarily because Chl-a concentration is significantly correlated with the ratio $R_{rs}(443)/R_{rs}(551)$ (O'Reilly et al., 2000), which is a partial descriptor of the $R_{rs}(\lambda)$ spectral shapes. The Chl-a data are approximately log-normally distributed for most water classes, which is consistent with earlier reports (Campbell, 1995). Some deviations occasionally exist, conspicuously in Class 18–23. As one cannot rule out the role of Chl-a uncertainties, which are usually sizeable in nearshore environments (Gons et al., 2002; Lavigne et al., 2021), the observed skewness in such turbid environments is worth further investigation. The *t*-test results

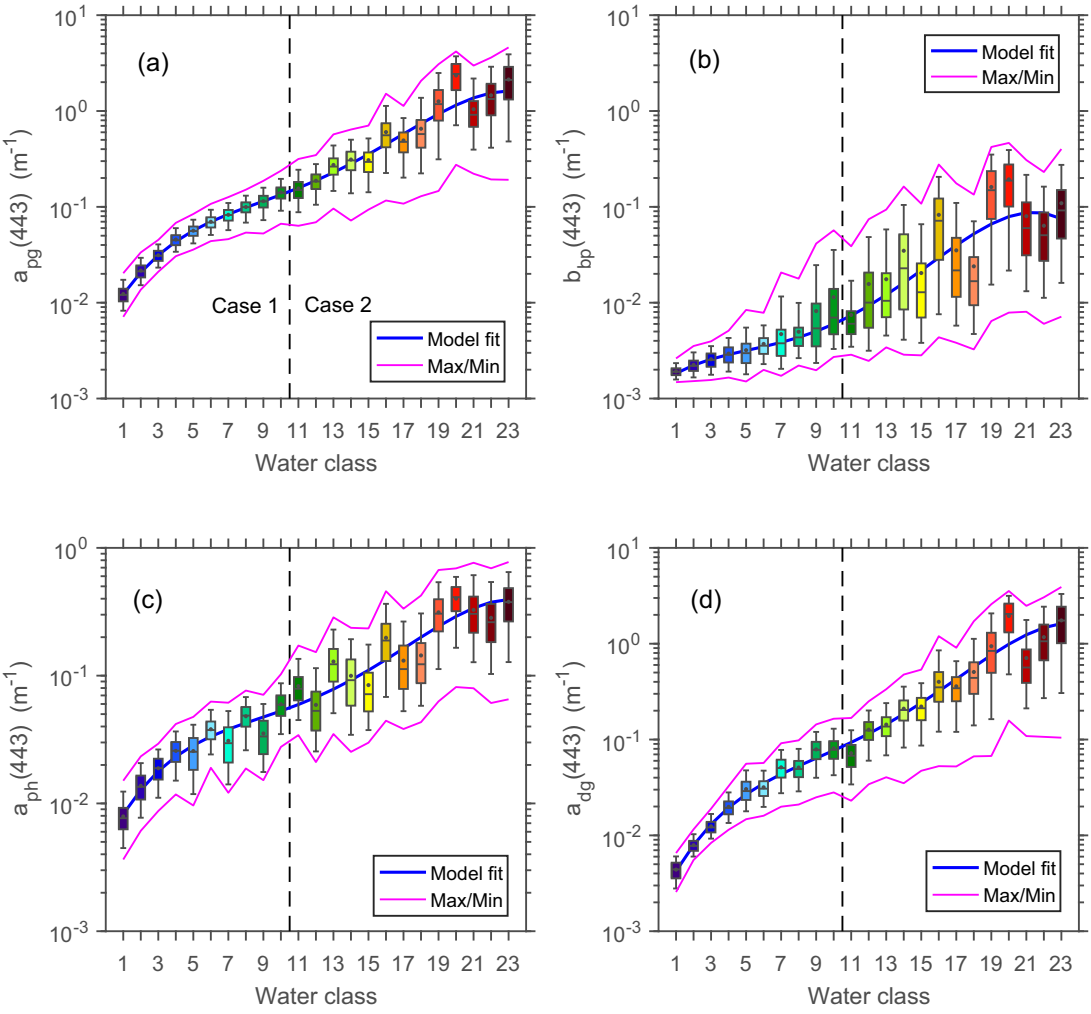

Fig. 6. Variation of light absorption and particle backscattering coefficients (all at 443 nm) with water classes (VIIRS-SNPP monthly products). Note that the dot and horizontal bar within the boxes refer to the mean and median, respectively. The upper and lower edges of the box indicate 75% and 25% percentiles, respectively, while the whiskers correspond to 95% and 5% percentiles, respectively. The lines (in magenta) beyond the boxes represent the range of variation of each quantity. The vertical dash line approximately separating "Case 1" and "Case 2" waters is determined with the approach of Lee and Hu (2006) (assuming ±20% deviation of $R_{rs}(410)/R_{rs}(551)$ and ±75% of $R_{rs}(551)$). A polynomial fit to the median values and the water classes is overlaid, with fitting results detailed in Table 4. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

indicate that Chl-a data are statistically distinctive across the water classes, especially in neighboring classes (Fig. A3). In connection with Morel and Prieur (1977), we found that Class 1–10 and Class 11–23 may, respectively, correspond to Case 1 and Case 2 waters. Note that Class 1 has the lowest Chl-a concentration among all world oceans, with ~0.025 ≤ Chl-a ≤ 0.1 mg m^{-3} (Table A1), which is representative of the subtropical ocean gyre waters (McClain et al., 2004).

3.3.2. *Chlorophyll-specific absorption coefficient*

The a_{ph}*(443) data vary between ~0.01 m^2 mg^{-1} and 0.2 m^2 mg^{-1} or about one order of magnitude globally (Fig. 5b and Table 3). This range of variation is comparable with earlier results from in situ measurements (Bricaud et al., 1995). For each water class, a_{ph}*(443) approximately follows a normal distribution, but with dispersions. The median a_{ph}*(443) values descend gradually from Class 1 toward Class 23, which can be primarily explained by the pigment packaging effect. A polynomial function fits well with the a_{ph}*(443) medians and the water classes (R^2 = 0.94, Table 4), especially for Class 1–10 or Case 1 waters. Analogous to Bricaud et al. (1995), we found that the median values of a_{ph}*(443) can relate to Chl-a concentration by a power function: a_{ph}*(443) = 0.0508 × Chl-a$^{-0.295}$, with R^2 = 0.81. Notably, the maximum and minimum a_{ph}*(443) values are remarkably different within each class. For instance, a_{ph}*(443) data can differ by two factors in Class 1 and up to seven in Class 23. Such significant variation can be partly ascribed to this quantity's dependence on regions and seasons, as well as the inability of many processors to differentiate between Chl-a and CDM absorption in the blue. Nevertheless, the *t*-test results confirm that a_{ph}*(443) is largely distinct across neighboring classes (Fig. A3).

Table 3
Class-specific Chl-a concentration (mg m^{-3}), a_{ph}*(443) (m^{-1}), K_d(490) (m^{-1}), and SPM (mg l^{-1}) in global waters. The results are based on long-term VIIRS-SNPP ocean color retrievals.

Water class	Median				Variance			
	Chl-a	a_{ph}*(443)	K_d(490)	SPM	Chl-a	a_{ph}*(443)	K_d(490)	SPM
1	0.0547	0.1401	0.0248	0.0528	0.0193	0.02598	0.0031	0.0161
2	0.1252	0.1074	0.0347	0.1098	0.0331	0.01412	0.0062	0.0273
3	0.2020	0.0932	0.0463	0.1688	0.0428	0.01270	0.0079	0.0311
4	0.3260	0.0772	0.0636	0.2466	0.0808	0.01159	0.0121	0.0403
5	0.4215	0.0609	0.0722	0.2998	0.0993	0.0154	0.0142	0.0432
6	0.5877	0.0636	0.0925	0.3543	0.151	0.0097	0.0186	0.0499
7	0.6605	0.0457	0.0963	0.4082	0.173	0.0129	0.019	0.063
8	0.9448	0.0516	0.1234	0.4598	0.232	0.0100	0.022	0.097
9	0.9653	0.0360	0.1233	0.5172	0.276	0.0109	0.028	0.174
10	1.4385	0.0420	0.1620	0.6191	0.367	0.010	0.056	0.436
11	1.6170	0.0499	0.2272	0.5934	0.613	0.012	0.169	0.979
12	1.7774	0.0320	0.1981	0.8035	0.663	0.010	0.141	0.730
13	2.9497	0.0413	0.4987	1.5881	1.228	0.011	0.269	1.747
14	2.8087	0.0353	0.6142	2.8692	1.158	0.010	0.396	1.937
15	3.0296	0.0252	0.3010	1.0352	1.368	0.009	0.278	1.387
16	4.2528	0.0433	1.4246	7.9519	1.803	0.015	0.653	5.572
17	4.0711	0.0303	0.7734	3.4954	1.841	0.010	0.386	2.477
18	6.5827	0.0193	0.5440	1.4592	4.843	0.009	0.331	2.919
19	6.4768	0.0434	2.5078	20.906	3.851	0.017	1.003	24.601
20	11.566	0.0337	3.7803	53.549	6.954	0.012	1.424	203.150
21	9.3719	0.0304	1.3166	6.5699	7.223	0.011	0.593	7.502
22	11.847	0.0220	1.3258	6.6311	9.406	0.009	0.483	9.398
23	13.435	0.0273	2.1316	13.718	9.939	0.013	0.716	53.292

Table 4
Polynomial fitting results of the water classes and median water bio-optical properties. Note that a fourth-order polynomial, $Y = a_0 + a_1X + a_2X^2 + a_3X^3 + a_4X^4$, was applied to a_{ph}*(443); for the rest, we have assumed a polynomial function of $\log_{10}Y = a_0 + a_1X + a_2X^2 + a_3X^3 + a_4X^4$.

X	Y	a_0	a_1	a_2	a_3	a_4	R^2
Water Class	a_{pg}(443)	−2.1979	0.3186	−0.03476	2.00×10^{-3}	-3.90×10^{-5}	0.98
	b_{bp}(443)	−2.8536	0.1351	−0.0206	1.77×10^{-3}	-4.30×10^{-5}	0.89
	a_{ph}(443)	−2.3254	0.2701	−0.0308	1.77×10^{-3}	-3.40×10^{-5}	0.94
	a_{dg}(443)	−2.7129	0.3772	−0.0400	2.25×10^{-3}	-4.30×10^{-5}	0.98
	Chl-a	−1.5602	0.3723	−0.0345	1.72×10^{-3}	-3.10×10^{-5}	0.99
	a_{ph}*(443)	0.1660	−0.0328	0.0032	-1.36×10^{-3}	2.12×10^{-6}	0.94
	K_d(490)	−1.8003	0.2199	−0.0257	1.81×10^{-3}	-4.10×10^{-5}	0.93
	SPM	−1.6665	0.4583	−0.0644	4.18×10^{-3}	-8.80×10^{-5}	0.88

Table 5
Class-specific light absorption and particle backscattering coefficients (all at 443 nm) in global waters. The results are based on long-term VIIRS-SNPP ocean color retrievals.

Water class	Median (unit of m^{-1} for parameters)				Variance (unit of m^{-1} for parameters)			
	a_{pg}(443)	b_{bp}(443)	a_{ph}(443)	a_{dg}(443)	a_{pg}(443)	b_{bp}(443)	a_{ph}(443)	a_{dg}(443)
1	0.0121	0.0019	0.0077	0.0044	0.0027	0.00023	0.0023	0.0010
2	0.0212	0.0022	0.0135	0.0078	0.0043	0.00041	0.0039	0.0013
3	0.0309	0.0025	0.0189	0.0121	0.0052	0.00054	0.0047	0.0022
4	0.0449	0.0029	0.0258	0.0193	0.0078	0.00073	0.0064	0.0044
5	0.0559	0.0030	0.0254	0.0290	0.0094	0.0012	0.0091	0.0091
6	0.0687	0.0036	0.0375	0.0307	0.013	0.0011	0.0088	0.0082
7	0.0822	0.0038	0.0296	0.0504	0.016	0.0031	0.012	0.015
8	0.0994	0.0043	0.0487	0.0492	0.018	0.0024	0.012	0.015
9	0.1141	0.0054	0.0334	0.0780	0.025	0.0073	0.013	0.024
10	0.1401	0.0070	0.0591	0.0782	0.031	0.011	0.016	0.026
11	0.1499	0.0060	0.0789	0.0678	0.046	0.005	0.027	0.028
12	0.1841	0.0100	0.0530	0.1242	0.051	0.015	0.028	0.041
13	0.2632	0.0105	0.1228	0.1366	0.086	0.017	0.053	0.051
14	0.3054	0.0228	0.0921	0.2032	0.107	0.033	0.050	0.080
15	0.2943	0.0129	0.0716	0.2122	0.112	0.020	0.043	0.090
16	0.5585	0.0716	0.1883	0.3505	0.268	0.063	0.089	0.219
17	0.4767	0.0217	0.1125	0.3431	0.188	0.034	0.067	0.159
18	0.5747	0.0168	0.1228	0.4399	0.351	0.022	0.077	0.301
19	1.1800	0.1492	0.3055	0.8409	0.644	0.105	0.126	0.578
20	2.4363	0.1879	0.4112	2.0425	0.929	0.115	0.127	0.832
21	0.9131	0.0600	0.3045	0.5643	0.533	0.064	0.146	0.471
22	1.3354	0.0506	0.2620	1.0603	0.744	0.047	0.132	0.653
23	2.1111	0.0918	0.3760	1.7439	1.037	0.080	0.154	0.918

3.3.3. Water turbidity

We characterize the water turbidity based on two commonly used parameters: $K_d(490)$ and SPM. Globally, $K_d(490)$ varies within three orders of magnitude (Fig. 5c and Table 3). The lowest $K_d(490)$ values are affiliated with Class 1, where $0.02 \leq K_d(490) \leq 0.037$ m^{-1}, while the highest is present in Class 20, where $0.6 \leq K_d(490) \leq 6.6$ m^{-1} (Table A1). A similar dependency is also observable in the SPM concentration, which spans a broader dynamic range (Fig. 5d and Table 3). Class 20 thus appears to be an optically dynamic environment, with the intra-class SPM varying over three orders of magnitude. Both $K_d(490)$ and SPM demonstrate an increasing trend from Class 1 to Class 20 and a slightly decreasing trend from Class 20 toward Class 23. A polynomial model is fitted to the class median $K_d(490)$ or SPM values, resulting in a high coefficient of determination (Table 4). In comparison, the fits are more accurate for Class 1–10 than Class 11–23. Both $K_d(490)$ and SPM data are roughly log-normally distributed for Class 1–9 but are more skewed for Class 10–23. The *t*-test estimation suggests that the $K_d(490)$ or SPM data are separable for almost every pair of neighboring classes (see Fig. A3).

3.3.4. Light absorption and backscattering coefficients

Fig. 6 shows the variation of light absorption and backscattering coefficients with water classes. Every quantity varies over about three orders of magnitude globally. The medians for each bio-optical property ($a_{pg}(443)$, $b_{bp}(443)$, $a_{ph}(443)$, and $a_{dg}(443)$) follow an ascending pattern with water classes. The median $a_{ph}(443)$ increases from 0.004 m^{-1} in Class 1 to 0.77 m^{-1} in Class 23, corresponding to a 200-fold difference (Table 5). The $a_{dg}(443)$ data vary between 0.003 m^{-1} and 3.9 m^{-1} (median), a prominent 1300-fold difference. A 4th-order polynomial function can fit the class-averaged bio-optical property and the water classes well (Table 4). The best fit is determined for $a_{pg}(443)$. This is mainly because our classification used normalized reflectance spectra and worked better in grouping waters of comparable absorption coefficients. Further, log-normal distributions can describe the absorption coefficients within most water classes. The $b_{bp}(443)$ data are severely skewed toward positive directions; only the Class 1–6 waters are closer to log-normal distributions. Nonetheless, log-normal distributions may approximate the global absorption and backscattering data, recalling the predominance of clear ocean waters around the globe (Fig. A2). Lastly, based on the *t*-test results, the absorption coefficients are generally distinctive between neighboring classes ($p < 0.05$, Fig. A3). For $b_{bp}(443)$, we cannot reach a consensus on data distinction, considering the relatively large deviations from an ideal probability distribution.

3.4. Validation

Fig. 7a gives the validation results from the VIIRS and in situ matchup data. The data represent the majority of the water classes quite well, except for Class 1, Class 19, Class 20, and Class 23. Among the rest, Class 2, Class 11, and Class 16 only have a small number of matchups ($N < 20$).

According to our analyses, the water classes were accurately generated from VIIRS. From Class 2 to Class 4, the median absolute difference, $|\Delta|$, is equal to zero. The $|\Delta|$ value increases to 1 or 2 through Class 5–15. From Class 17, the differences between VIIRS and in situ derivations may increase up to $|\Delta| = 4$–6. Fig. 7b–d further shows the

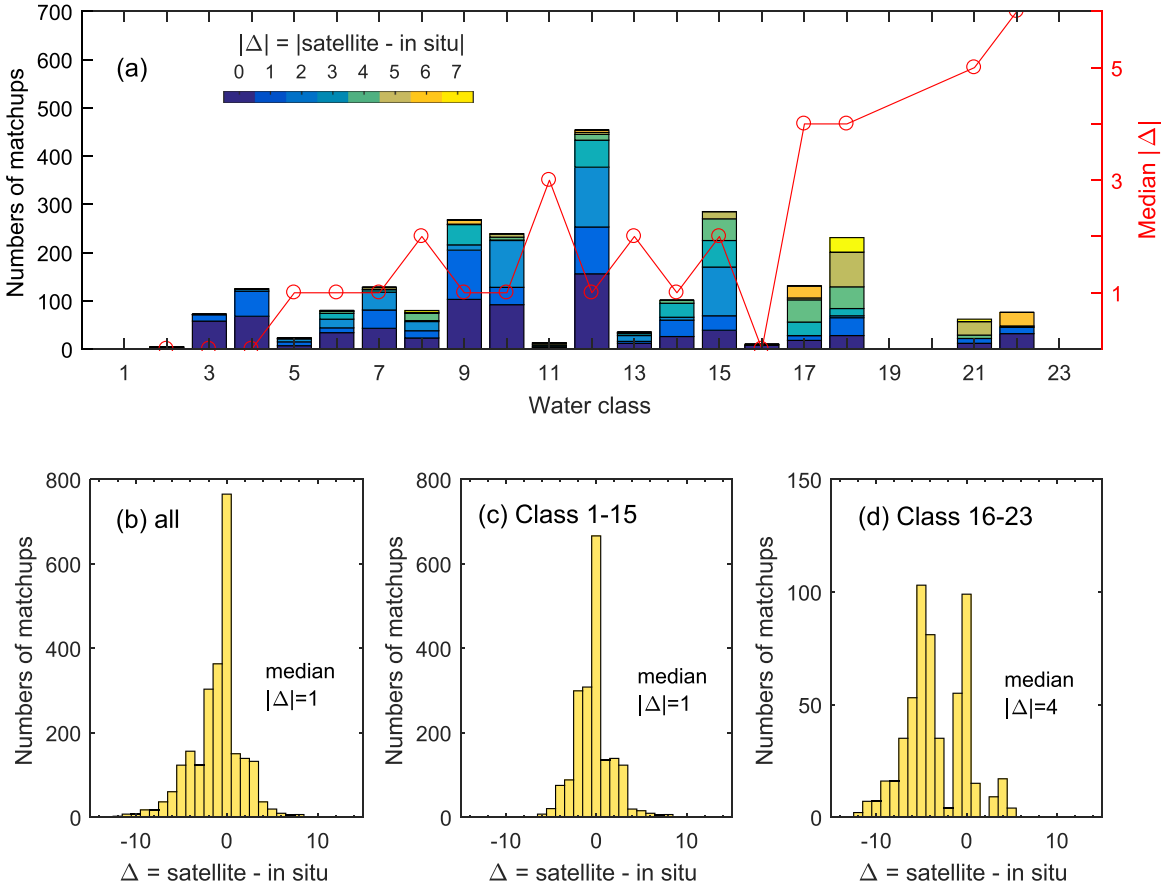

Fig. 7. Comparison of the water classes derived from concurrent VIIRS and in situ $R_{rs}(\lambda)$ measurements: (a) class-specific results, (b) frequency distribution on global scales, (c) frequency distribution for Class 1–15, and (d) frequency distribution for Class 16–23.

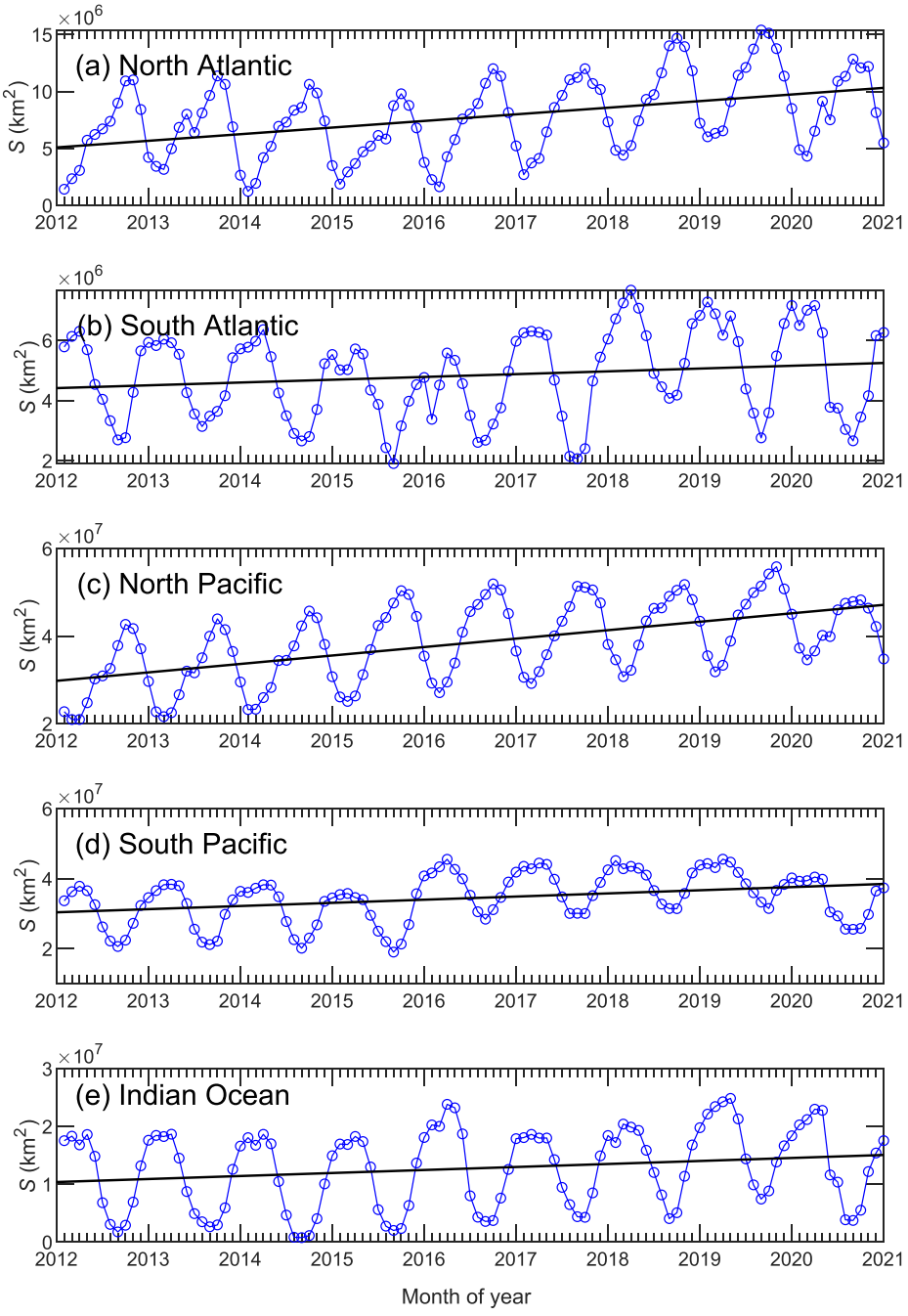

Fig. 8. Time series of the monthly mean surface area (S, unit: km^2) of Class 1 waters in the major ocean basins for the (a) North Atlantic Ocean, (b) South Atlantic Ocean, (c) North Pacific Ocean, (d) South Pacific Ocean, and (e) Indian Ocean. The straight lines are the linear fits, with statistics given in Table 6.

frequency distributions of Δ values for the entire matchup data and two subsets of matchup data. A global median $|\Delta|$ of 1 is recorded for VIIRS retrievals, and the same result is also observed for Class 1–15. In contrast, the satellite water class products are found with markedly increased differences for Class 16–23 (with median $|\Delta|$ equal to 4). We stress that the Class 16–23 waters are usually representative of turbid nearshore environments (recall Fig. 3). Accurate satellite retrieval of $R_{rs}(\lambda)$ in such waters remains a challenge (Wang and Jiang, 2018; Wei

Table 6
Long-term trends of the surface areas of Class 1 waters in the ocean basins (2012–2020). Bold numbers indicate significant results ($p < 0.05$).

	Mean area in 2012, km^2	Increase in km^2/yr (%/yr)	p-value
North Atlantic Ocean	6.4×10^6	5.8×10^5 (11)	$\mathbf{2.3 \times 10^{-6}}$
South Atlantic Ocean	4.8×10^6	9.2×10^4 (2.1)	9.0×10^{-2}
North Pacific Ocean	3.1×10^7	1.9×10^6 (6.5)	$\mathbf{4.3 \times 10^{-10}}$
South Pacific Ocean	3.0×10^7	9.0×10^5 (3.0)	$\mathbf{4.5 \times 10^{-4}}$
Indian Ocean	1.2×10^7	5.2×10^5 (5.0)	$\mathbf{4.1 \times 10^{-2}}$

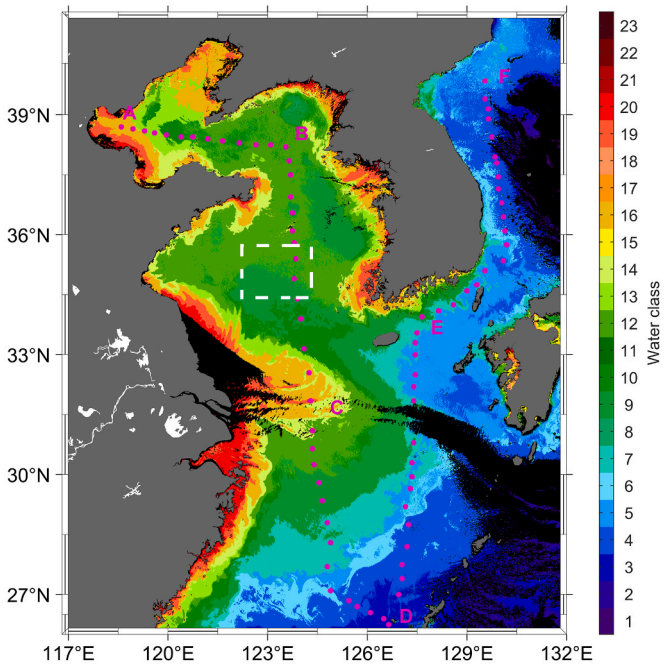

Fig. 9. Satellite-derived water classes in the northwest Pacific region (VIIRS-SNPP snapshot; February 9, 2016; 750 m resolution). We extracted the water classes and bio-optical data along the transect indicated by solid pink circles. The square area delimited by the white dashed box is the region for extracting monthly time series data. Black pixels represent invalid ocean color measurements.

et al., 2020; Zibordi et al., 2009a). Uncertainties in satellite $R_{rs}(\lambda)$ data will be primarily responsible for the increased uncertainties in water class products (see further discussion in Section 5.3).

4. Case analyses of the satellite water class retrievals

4.1. Long-term trends in ocean subtropical gyres

A promising application in satellite water classification is the potential of geographical zonation of oceans for global ecological research. Especially, the similarity with Longhurst's ocean provinces (recall Fig. 3) gives credit to our new classifications. Here, we examine the time series of the surface areas occupied by Class 1 to understand its spatial and temporal variability. Note that Class 1 waters are typical of the lowest bio-optical/biogeochemical properties, where Chl-a $\leq$ 0.1 mg m^{-3}, SPM $\leq$ 0.1 mg l^{-1}, $K_d(490) \leq 0.035$ m^{-1}, $a_{ph}(443) \leq 0.015$ m^{-1}, $a_{dg}(443) \leq 0.007$ m^{-1}, and $b_{bp}(443) \leq 0.0025$ m^{-1}. These features clearly suggest that Class 1 can represent the subtropical ocean gyres. Indeed, earlier exercises often identified the subtropical gyres for Chl-a ≤ 0.07 mg m^{-3} (McClain et al., 2004; Polovina et al., 2008). In Fig. 8, we give the time series of the surface areas of the Class 1 waters based on the VIIRS monthly products. Five ocean gyres unanimously experienced seasonal oscillations. In the northern hemisphere, they shrink during springtime and expand in the fall. In the southern hemisphere, their sizes increase in the spring and decrease in the fall. Also, all five subtropical gyres increased in size from 2012 to 2020 (see statistics in Table 6). The most dramatic change occurred in the North Atlantic Gyre, which expanded by $\sim 5.8 \times 10^5$ km^2/yr or $\sim$ 11%/yr. The North Pacific Gyre, South Pacific Gyre, and Indian Ocean Gyre also grew in surface size, albeit slower (3%/yr–6.5%/yr). The rate of change for the South Atlantic Gyre was the smallest and statistically insignificant. These annual trends are comparable to earlier chlorophyll-based analyses (McClain et al., 2004; Polovina et al., 2008; Signorini et al., 2015).

4.2. Covariation of the water classes and water bio-optical/biogeochemical properties

Benefiting from the bio-optical/biogeochemical characteristics, the water classes may serve as an indicator of water quality in coastal oceans and inland waters. With examples, we demonstrate below that the water bio-optical/biogeochemical properties covary with the water classes. In Fig. 9, the coastal ocean in the northwest Pacific is an optically dynamic environment. River discharge and sediment resuspension due to tides (Shi and Wang, 2012; Shi and Wang, 2014) lead to turbid waters near the Chinese and the Korean coasts (Class ~15–23). Moderately turbid waters (Class ~8–14) dominate the Yellow Sea and East China Sea, while clear waters (approximately Class 1–7) are not unusual in the northeast (Sea of Japan/East Sea) and the south. Fig. 10 illustrates the spatial variation of the water classes along a set of pre-defined locations starting from A and ending at F (indicated in Fig. 9) for this particular day. The VIIRS water classes exhibit a spatial pattern highly comparable to Chl-a, SPM, $K_d(490)$, $a_{dg}(443)$, and $b_{bp}(443)$. Further in Fig. 11, we compare the time series data for the water classes and bio-optical/biogeochemical properties within the Yellow Sea (the region of interest is highlighted in Fig. 9). These time series data exhibit nearly synchronous variability, with seasonal maxima occurring around early spring and minima around early fall. Furthermore, we compared the bio-optical and biogeochemical properties (from both the spatial transect and the time series station) with the corresponding global class-specific median values. These median values nicely replicated the spatial and

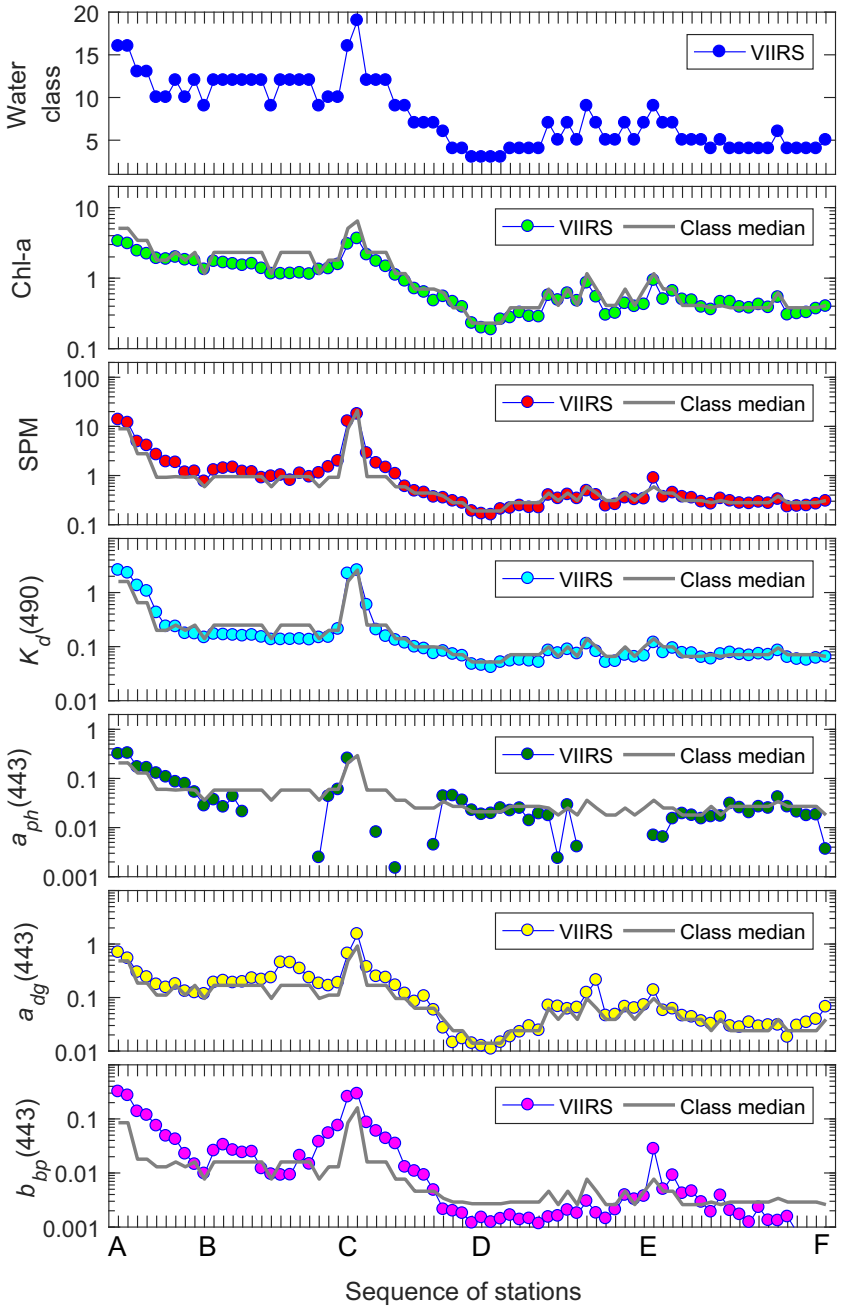

Fig. 10. Spatial variation of the water classes and related bio-optical and biogeochemical properties in the northwest Pacific region. The class-specific median values for each bio-optical or biogeochemical property are overlaid for comparison. Letters A, B, C, D, E, and F indicate the sampling locations (shown in Fig. 9). Note that units for Chl-a, SPM, $K_d(490)$, $a_{ph}(443)$, $a_{dg}(443)$, and $b_{bp}(443)$ are mg m^{-3}, mg l^{-1}, m^{-1}, m^{-1}, m^{-1}, and m^{-1}, respectively.

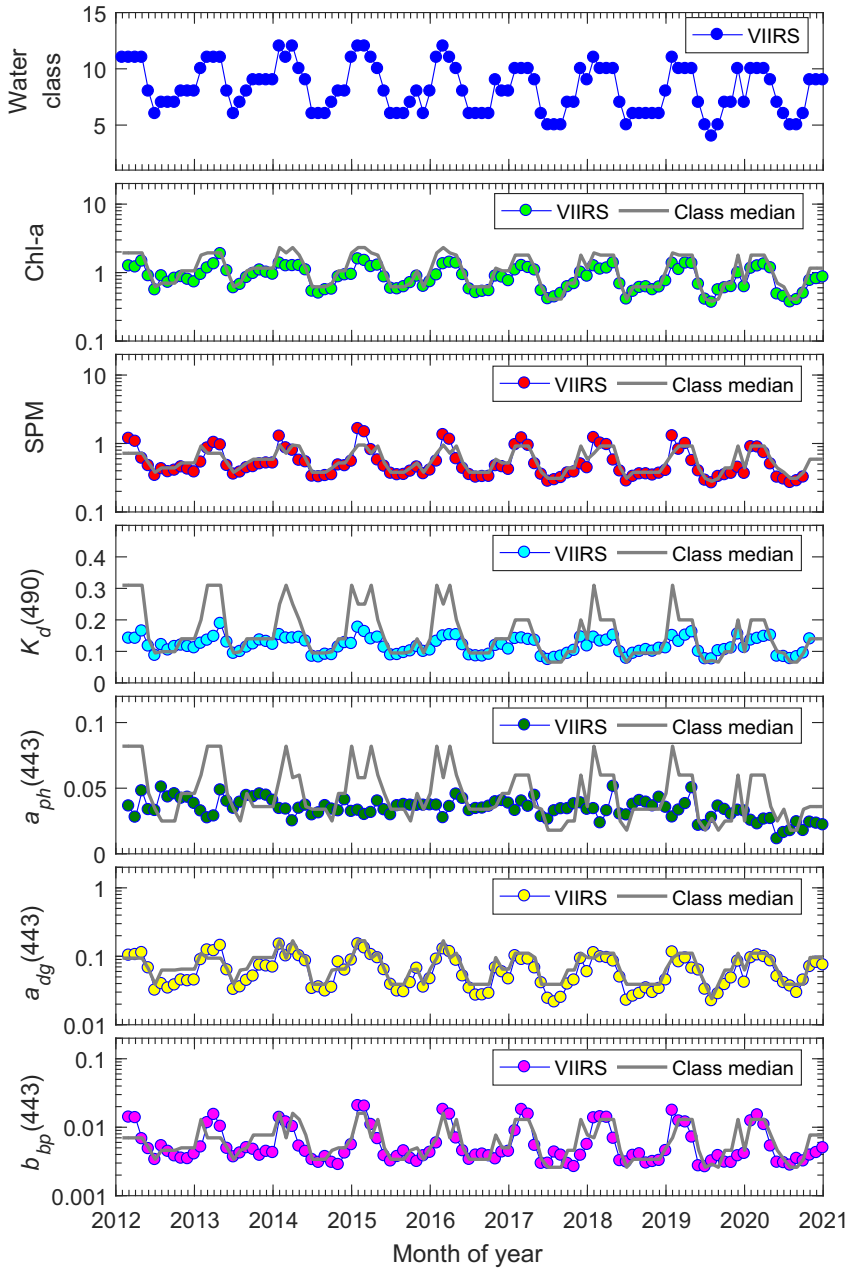

Fig. 11. Time series of the water classes and related bio-optical and biogeochemical properties in the northwest Pacific region. The class-specific median values for each bio-optical or biogeochemical property are overlaid for comparison. The study area is highlighted (white dashed box) in Fig. 9. Note that units for Chl-a, SPM, $K_d(490)$, $a_{ph}(443)$, $a_{dg}(443)$, and $b_{bp}(443)$ are mg m^{-3}, mg l^{-1}, m^{-1}, m^{-1}, m^{-1}, and m^{-1}, respectively.

Table 7
Comparison results of estimated bio-optical quantities with satellite-measured quantities in the northwest Pacific region and Lake Erie. The error statistics were calculated for the transect data and time-series data, respectively.

Parameter	Coastal ocean (Northwest Pacific region)				Inland water (Lake Erie)			
	Transect		Time series		Transect		Time Series	
	MAPD	MRPD	MAPD	MRPD	MAPD	MRPD	MAPD	MRPD
Chl-a	18%	18%	33%	31%	20%	−20%	30%	18%
SPM	19%	−2%	18%	10%	33%	13%	47%	−12%
$K_d(490)$	20%	13%	29%	19%	20%	18%	42%	−4%
$a_{ph}(443)$	33%	13%	41%	22%	45%	−44%	30%	−15%
$a_{dg}(443)$	25%	−18%	33%	29%	215%	215%	50%	48%
$a_{pg}(443)$	18%	−18%	25%	18%	21%	14%	33%	17%
$b_{bp}(443)$	67%	10%	32%	9%	40%	−40%	47%	−21%

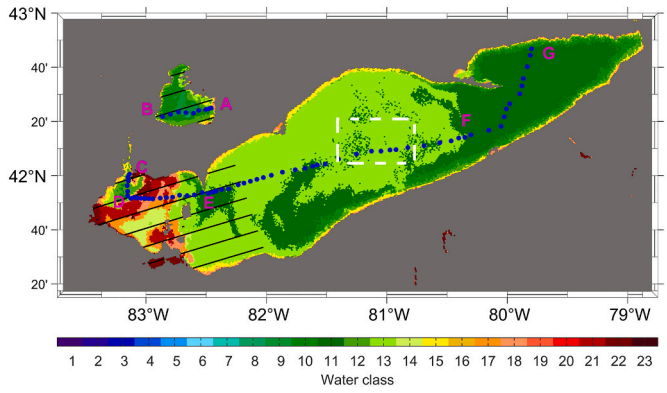

Fig. 12. Satellite-derived water classes in Lake Erie (VIIRS-SNPP snapshot; September 24, 2017; 750 m resolution). The blue solid circles indicate the stations for extraction of the water classes and other bio-optical and biogeochemical properties. The square area delimited by the white dashed box shows where the monthly time series data were extracted. Black pixels represent invalid measurements.

temporal variability for most bio-optical/biogeochemical properties (see quantitative statistics in Table 7).

Fig. 12 gives another snapshot of the water class distribution in Lake Erie. The western basin of Lake Erie is predominantly Class 11–23, reflecting a relatively turbid water environment. The rest of the lake is much clearer and more homogeneous, with water classes dominantly varying between 10 and 13. We demonstrate the spatial and temporal variation of the water classes in Fig. 13 and Fig. 14, respectively. Noticeably, the water classes mirror the spatial and temporal variation of the water bio-optical/biogeochemical properties to varying degrees. Together with Fig. 10 and Fig. 11, these observations suggest that the water class product itself may be adopted as a proxy for water quality, at least qualitatively. We further compared the global class-specific median values with the long-term or long-range bio-optical/biogeochemical properties. Despite occasional large deviations (e.g., $a_{dg}(443)$ in Fig. 13), the global median values appear to well describe the periodical ups and downs of almost every bio-optical/biogeochemical property from satellites, where MAPD <50% (see statistics in Table 7).

5. Discussion

5.1. Comparison with Longhurst's biogeographic provinces

There are many points of agreement between what is developed here and the biogeographical provinces of Longhurst (1998). In Fig. 15, we compared the climatological distribution of water classes with Longhurst's provinces. It is prominent that the new water classes reflect many significant spatial structures described by Longhurst but are somewhat more naturally delineated based on the optical properties and spectral shapes. For instance, the ocean subtropical gyre provinces are primarily associated with Class 1. The North Atlantic Drift Province (NADR; located west of Ireland) is dominated by Class 4 and stands out from the surrounding oceans. The Indian Monsoon Gyres Province (MONS; located to the north of the subtropical gyre) and the Indian South Subtropical Gyre Province (ISSG) are distinguished by Class 2 in the north and Class 1 in the south, respectively. According to the comparison, Longhurst's scheme does not give sufficient details about the optical diversity of coastal oceans and excludes inland waters. Each Longhurst ocean province can encompass a handful of optical water classes. For example, the Northwest Atlantic Shelves Province (NWCS) comprises at least ten water classes (~2–12). Such differences are mainly related to the objectives of individual classifications: the new classification in this study is targeted at all global water types, whereas Longhurst focuses on open oceans. In addition, our classification is entirely dependent on the ocean's optical features, while Longhurst considers geography. Lastly, both our classification and Longhurst's scheme separate water types with a dynamic boundary (Devred et al., 2007; Fay and McKinley, 2014; Reygondeau et al., 2020; Reygondeau et al., 2013).

5.2. Extension to other satellites

As manifested by in situ and satellite $R_{rs}(\lambda)$ matchup analyses (Cui et al., 2010; Wei et al., 2020; Zibordi et al., 2009a), satellite $R_{rs}(\lambda)$ band ratios are often subject to relatively minor errors. Thus, the satellite calibration and atmospheric correction will have less impact on the water classes generated with a spectral shape-based approach. A recent study has shown that spectral shape-based classification is relatively insensitive to the spectral bands used for classification (Jia et al., 2021). This feature makes it possible to generate highly comparable water classes across different satellite sensors. In Fig. 16, we give example water classes in the northeast U.S. coastal waters (with OLCI and VIIRS data) and Lake Erie (with HICO and VIIRS data). Overall, the derived OLCI water classes are highly comparable with the VIIRS product (Fig. 16a and b). The differences in the Gulf of St. Lawrence can be related to the bio-optical changes in water masses between the OLCI and the VIIRS overpass (> 3 h). With about the same overpass time, the HICO and VIIRS water classes are expectedly similar as well (Fig. 16c and d), and the difference, $|\Delta|$, between the two water class data is within 1. The satellite $R_{rs}(\lambda)$ uncertainties and different spatial resolutions (300 m for OLCI, 90 m for HICO, and 750 m for VIIRS) may have contributed to the observed differences, which will be investigated in future efforts. Because our classification scheme is relatively insensitive to sensor characteristics (such as calibration uncertainty, atmospheric correction errors, and spectral bands), it is promising to build decades-long time series products by merging the data across different satellite missions. Creating merged water class products among peer missions should also be feasible to mitigate spatial gaps inherent with satellite data (Liu and Wang, 2019).

5.3. Product uncertainties associated with atmospheric correction

Atmospheric correction is key to retrieving satellite $R_{rs}(\lambda)$ data. The uncertainties associated with the modeling of atmospheric and surface properties will translate to $R_{rs}(\lambda)$ (IOCCG, 2019) and eventually to the water class products. To help understand the water class uncertainty, we performed a sensitivity analysis with the $nR_{rs}(\lambda)$ spectra described in Section 2.1. In our evaluations, an error (ε) was added to $nR_{rs}(\lambda)$ at one wavelength only for each simulation. As shown in Fig. 17, minor errors in $nR_{rs}(\lambda)$ (i.e., ±10% in this study) exert minimal influence on the resulting water classes. When $nR_{rs}(\lambda)$ errors reach high levels, for example, $\varepsilon = \pm 30\%$ and ± 50%, the uncertainties in the water class products can increase substantially. Among all wavelengths, the blue and green bands exhibit dominant roles for Class 1–10, while the blue, green, and red bands are more critical for Class 11–23. In extremely clear waters, such as Class 1–3, the blue bands play a major role in the water class uncertainties. In contrast, the green and red bands are relatively more important in the opposite end of the water classes, such as Class 19–23. Fig. 17g refers to the water class uncertainties resulting from excessive negative biases in $nR_{rs}(\lambda)$. Although the uncertainty may

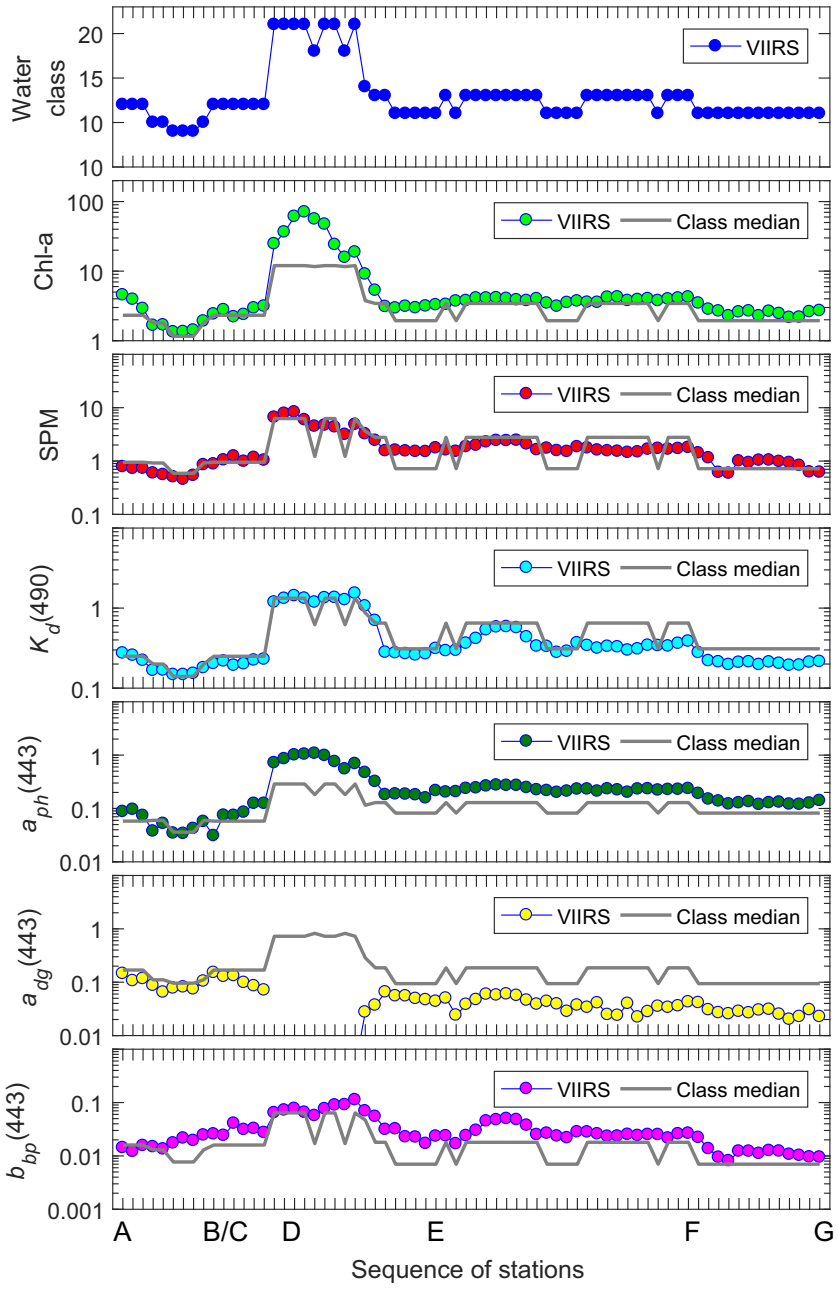

Fig. 13. Spatial variation of the water classes and corresponding bio-optical properties in Lake Erie. The class-specific median values for each bio-optical or biogeochemical property are overlaid for comparison. Letters A, B, C, D, E, and F indicate the sampling locations (shown in Fig. 12). Note that units for Chl-a, SPM, $K_d(490)$, $a_{ph}(443)$, $a_{dg}(443)$, and $b_{bp}(443)$ are mg m^{-3}, mg l^{-1}, m^{-1}, m^{-1}, m^{-1}, and m^{-1}, respectively.

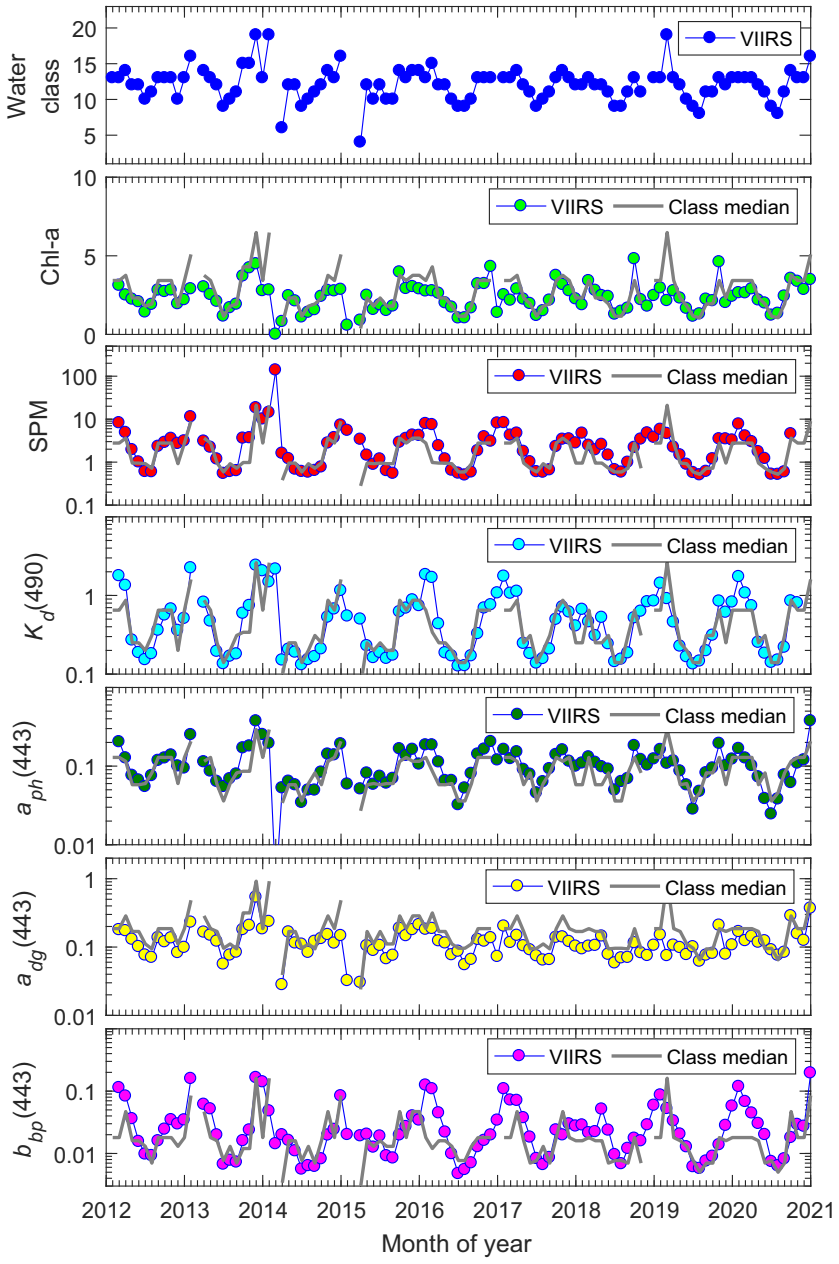

Fig. 14. Time series of the water classes and corresponding bio-optical properties in Lake Erie. The class-specific median values for each bio-optical or biogeochemical property are also overlaid for comparison. The study area is highlighted (white dashed box) in Fig. 12. Note that units for Chl-a, SPM, $K_d(490)$, $a_{ph}(443)$, $a_{dg}(443)$, and $b_{bp}(443)$ are mg m^{-3}, mg l^{-1}, m^{-1}, m^{-1}, m^{-1}, and m^{-1}, respectively.

increase markedly with an input error of −100%, we emphasize that such large-scale negative biases in $nR_{rs}(\lambda)$ are most likely limited to certain water classes. Recall that in Fig. 4, we reported an average bias of about −100% in VIIRS $nR_{rs}(410)$ data for Class 18, Class 20, and Class 22–23. According to Fig. 17g, however, the significant negative biases in $nR_{rs}(410)$ do not significantly increase the water class uncertainty. One explanation is that the $nR_{rs}(410)$ values are small in these waters due to strong light absorption of CDM, and an error-disturbed $nR_{rs}(410)$ does not substantially alter the overall spectral shape of $nR_{rs}(\lambda)$ (also see Wei et al., 2020).

5.4. Future directions

It is important to note that the present classification is built on the reflectance spectra in visible bands. Under certain circumstances, $R_{rs}(\lambda)$ spectra can be more sensitive to the variation of specific water constituents outside the visible domain, such as CDM in the ultraviolet (Wei et al., 2016a) and algal blooms in the NIR bands (Gower et al., 2005; Qi et al., 2020). Thus, expanding the spectral coverage in the reference database is warranted. In addition, the hyperspectral reference data in Fig. 1 are unlikely to encompass every situation in nature. Particularly, the bio-optical variability in inland water environments can be more complex and distinct from the open ocean and coastal waters. With few reference spectra and validation data in our analyses, further efforts are recommended for inland water applications. The presence of floating matters, such as floating algae (Hu, 2022), may introduce uncertainties to the satellite water class products. Therefore, it is imperative to accumulate in situ hyperspectral $R_{rs}(\lambda)$ measurements over optically distinct targets to enrich the reference database.

The bio-optical characterizations provide a valuable perspective of

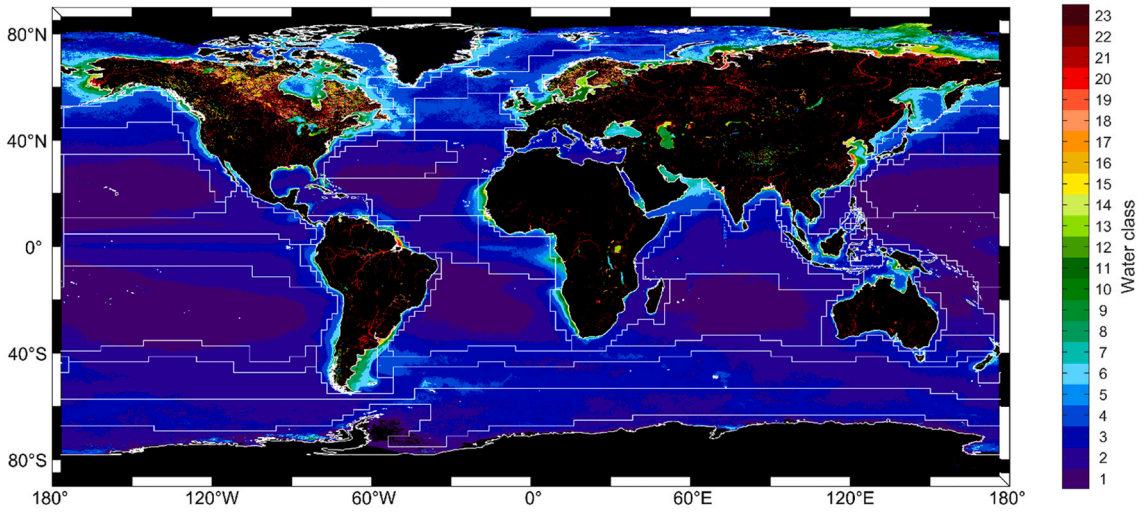

Fig. 15. Climatologic distribution of the water classes in open oceans, coastal, and inland waters (2012–2020). The biogeographical provinces of Longhurst (1998) (delineated in white lines) are overlaid on the top of the water classes.

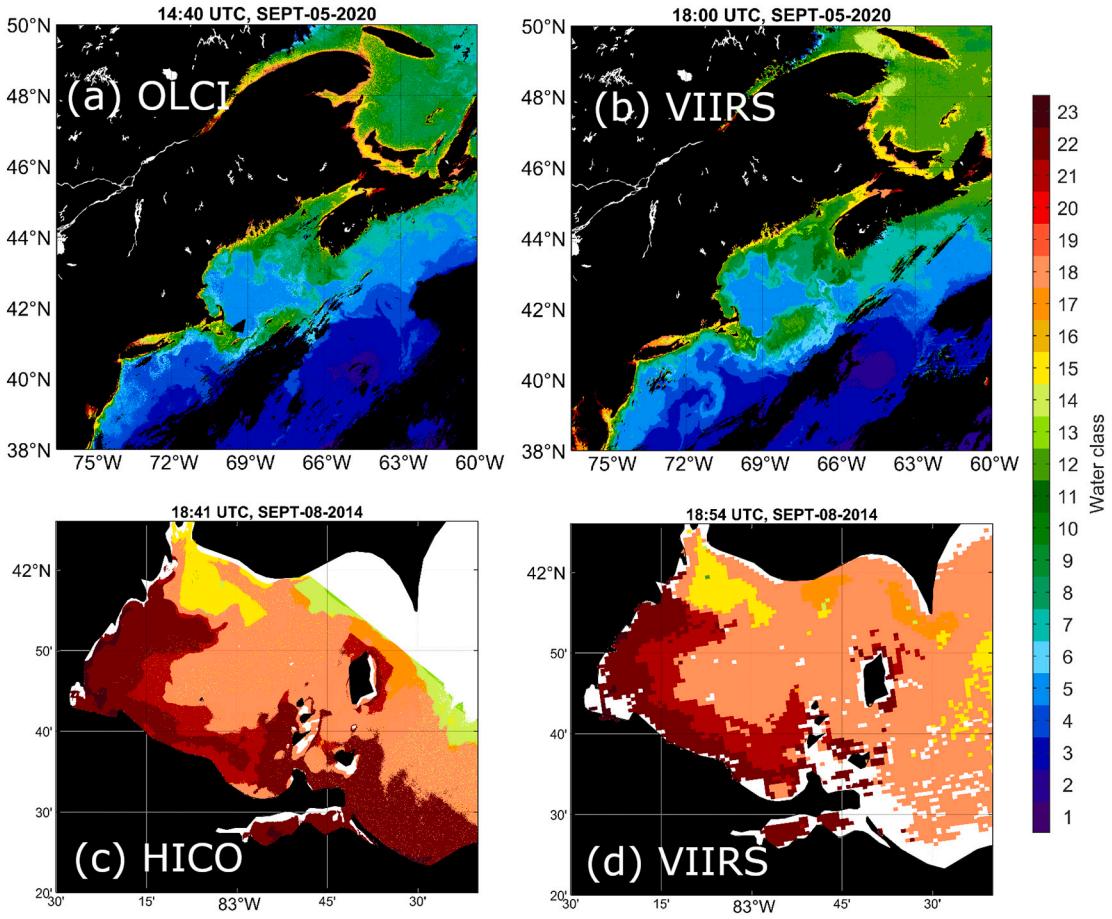

Fig. 16. Satellite-derived water classes over the northeast U.S. coasts (in top panel) and over the western basin of Lake Erie (in bottom panel). All images are given in the native resolutions of the satellite measurements.

the water classes (Fig. 5 and Fig. 6). Other descriptors are also useful to define the water classes, such as the ratios among various bio-optical and biogeochemical variables of interest, including $a_{dg}(443)/a_{pg}(443)$, $b_{bp}(443)$/Chl-a (Siegel et al., 2005), and particulate backscattering ratio (Lubac and Loisel, 2007). In addition, the seasonal and geographic dependency of each quantity mentioned above has been neglected in the present study. Thus, our characterizations are representative of a global mean situation. Recent studies have suggested that the temporal variations of the chlorophyll-specific absorption coefficient can be significant in clear oceanic waters and remains a critical determinator for the Chl-a retrieval from satellites (e.g., Lee et al., 2020). All the factors omitted in the present study should be considered in the future for a complete

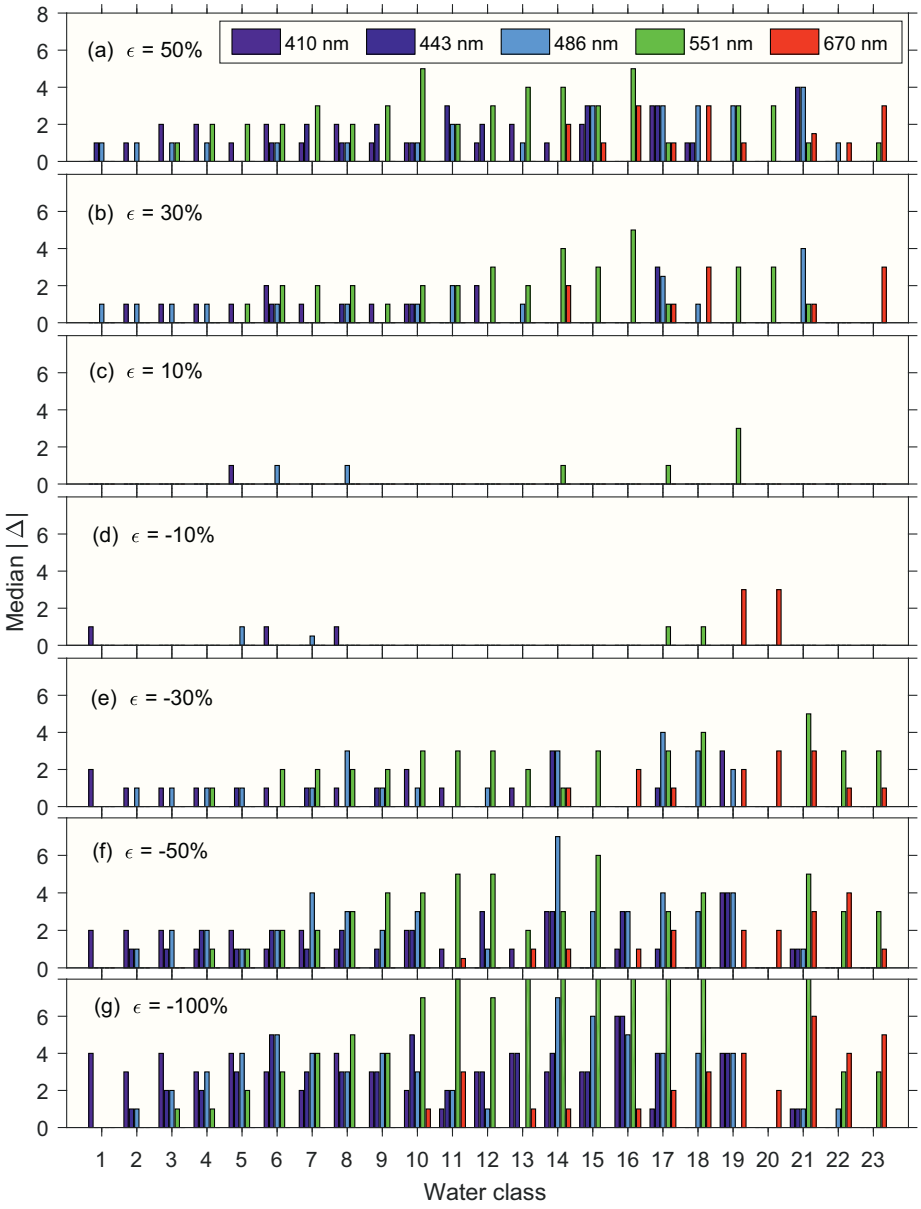

Fig. 17. Variation of the water class uncertainties (quantified as the median $|\Delta|$) with errors added to $nR_{rs}(\lambda)$. Note that the five wavelengths (410, 443, 486, 551, and 670 nm) are considered independently, where the error ε is added to $nR_{rs}(\lambda)$ at only one wavelength for each case.

picture of the water classes.

6. Conclusions

This study reports new global satellite water class products that can be routinely generated with an $R_{rs}(\lambda)$ spectral shape similarity algorithm. The global water classes presented herein represent the first of their kind for mission-long satellite observations (Level-2 and Level-3). The water classification scheme has resolved global waters into 23 water classes, with a reasonable representation of oceanic waters and complex coastal and inland waters. The classification is successful primarily due to the $R_{rs}(\lambda)$ spectral shape-preserving classification scheme. As discussed in this study, the reflectance shape conveys rich information on water IOPs, especially regarding the light absorption coefficients. Therefore, our classification is optimal for separating water bio-optical properties across water classes. Among many quantities, $a_{pg}(443)$, $a_{ph}(443)$, $a_{dg}(443)$, and Chl-a increase progressively with water classes, while $a_{ph}^{*}(443)$ exhibits a decreasing trend with water classes. Others, including $b_{bp}(443)$, $K_d(490)$, and SPM, also increase with water classes but are less structured. The characteristic variation of bio-optical and biogeochemical properties with water classes is significant, especially when they are generally separable across various water classes.

We demonstrate that these water class products are a new source of information and foresee ample opportunities for exploration and extensive applications. Our case studies provide compelling examples of the applicability of optical water classes to aquatic ecology, e.g., significant expansions of subtropical gyres. Analyses also suggest that the water class products can be used as indicators for water quality in coastal and inland water environments. The spectral shape-based classification generates reliable products because they are less affected by $R_{rs}(\lambda)$ spectral amplitudes, which are sensitive to satellite calibration and atmospheric correction. Moreover, the algorithm is generic and enables the generation of highly comparable water classes among satellites of different spectral band settings. Therefore, it is feasible to create merged products from multiple satellites and long-term time series by applying the algorithm to peer satellite missions.

It is also recognized that VIIRS water class products are limited to the five spectral bands, which lack the capability to resolve certain yet important water bio-optical features, including the phycocyanin absorption band at 625 nm. Future directions include extending the classification scheme to other sensors such as OLCI with more spectral bands. Finally, we acknowledge that the problems arising from the VIIRS water classification may be solved with emerging hyperspectral ocean color sensors, including the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission and the Geosynchronous Littoral Imaging and Monitoring Radiometer (GLIMR).

CRediT author statement

Jianwei Wei: Conceptualization, Methodology, Formal Analysis, Writing- Original draft.

Menghua Wang: Funding acquisition, Conceptualization, Reviewing and Editing.

Karlis Mikelsons: Data curation, Software.

Lide Jiang: Data curation, Software.

Susanne Kratzer: Reviewing and Editing.

Zhongping Lee: Reviewing and Editing.

Tim Moore: Reviewing and Editing.

Heidi M. Sosik: Reviewing and Editing.

Dimitry Van der Zande: Reviewing and Editing.

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.

Data availability

Data will be made available on request.

Acknowledgments

We thank many individual investigators for sharing in situ $R_{rs}(\lambda)$ data. The AERONET-OC in situ measurements are accessible from NASA SeaBASS and are also available from Guiseppe Zibordi (Venice, Galata, Gloria, Gustav, and Helsinki stations), Dimitry Van der Zande (Thornton C station), Susanne Kratzer (Palgrunden station), Alex Gilerson (LISCO station), Hui Feng and Heidi Sosik (MVCO station), Young-Je Park (Ieodo station), and Matthew Ragan (USC station). We also thank three anonymous reviewers for their thoughtful comments. The hyperspectral training data are available upon request. The satellite water class products are currently accessible from the NOAA/STAR Ocean Color Team website (https://www.star.nesdis.noaa.gov/socd/mecb/color/). This study was supported by the Joint Polar Satellite System (JPSS) funding, the NASA Carbon Monitoring System (CMS) project (80NSSC20K0014), and the Swedish National Space Agency Grant (Dnr. 2021-00050). The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce.

Appendix 1. Variance of $R_{rs}(\lambda)$ spectra

The spectral variance of the in situ hyperspectral references is presented as the ratio of the variance to the median, which is represented as the coefficient of variation (CV). The VIIRS-SNPP statistics were based on a global image on the daily 9 km image on August 18, 2012.

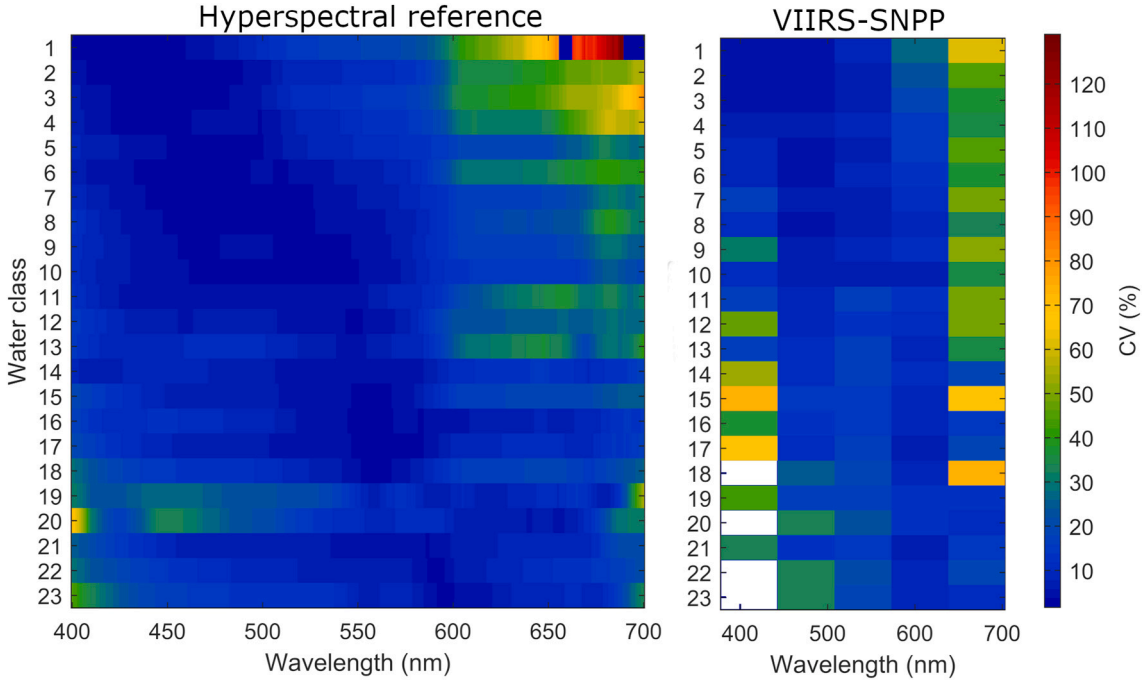

Fig. A1. Spectral variance (normalized by the median values, i.e., the coefficient of variation) of the hyperspectral reference spectra and VIIRS-SNPP satellite spectra (August 18, 2012). The white blanks in VIIRS-SNPP data indicate the CV values of the $nR_{rs}(\lambda)$ spectra have exceeded 100%.

Appendix 2. Distribution of the water classes in global surface waters

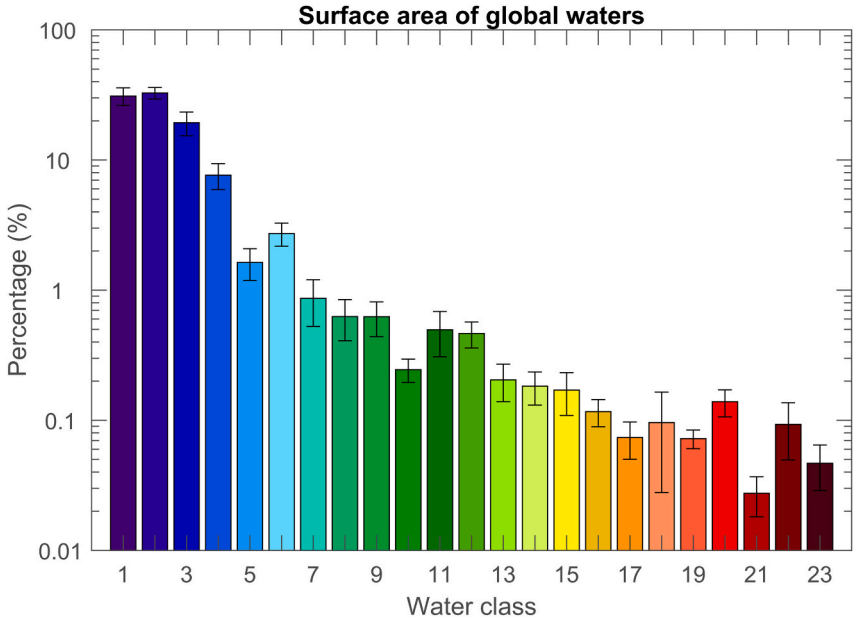

Fig. A2. Frequency distribution of the surface areas of the global water classes. The error bars indicate the standard deviation of the surface areas.

Appendix 3. Distinction of bio-optical and biogeochemical properties between water classes

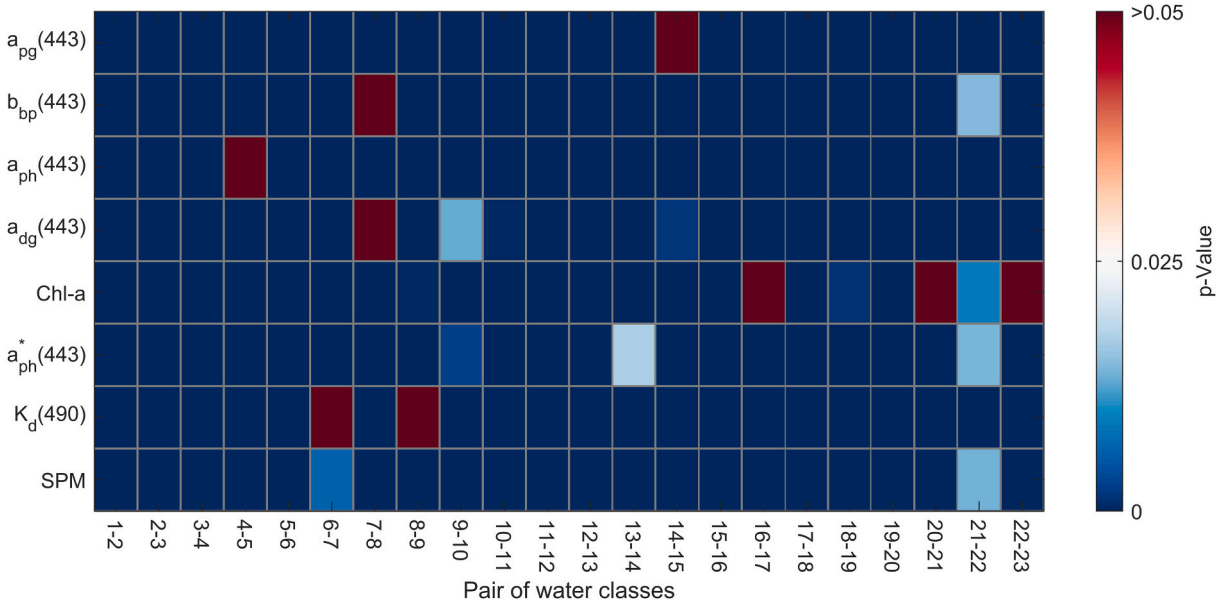

Fig. A3. Level of significance of two-tailed *t*-tests for (log-transformed) bio-optical and biogeochemical properties. The *x*-axis indicates the neighboring pairs of water classes from Class 1 to Class 23.

Appendix 4. Range of variation of biogeochemical and bio-optical properties in global surface waters

Table A1
Range of chlorophyll-a (Chl-a) concentration, chlorophyll-specific absorption coefficient at 443 nm (a_{ph}*(443)), diffuse attenuation coefficient at 490 nm (K_d(490)), and suspended particulate matter (SPM) in global waters. The results were obtained from mission-long VIIRS-SNPP monthly products (9 km).

Class	Chl-a	a_{ph}*(443)	K_d(490)	SPM
1	0.0233–0.1122	0.0966–0.2124	0.0198–0.0366	0.0278–0.1069
2	0.0667–0.2208	0.0724–0.1395	0.0256–0.0567	0.0611–0.1860
3	0.1201–0.3312	0.0555–0.1170	0.0328–0.0706	0.1073–0.2505

(*continued on next page*)

Table A1 (*continued*)

Class	Chl-a	$a_{ph}^*(443)$	$K_d(490)$	SPM
4	0.2008–0.5906	0.0446–0.1011	0.0429–0.1040	0.1645–0.3526
5	0.2424–0.7456	0.0278–0.0900	0.0463–0.1203	0.1934–0.4113
6	0.3194–1.0861	0.0379–0.0886	0.0640–0.1735	0.2356–0.5151
7	0.3188–1.1922	0.0226–0.0785	0.0585–0.1633	0.2363–0.6322
8	0.4089–1.6088	0.0275–0.0786	0.0746–0.2151	0.2810–1.0315
9	0.3697–1.7692	0.0188–0.0707	0.0660–0.2450	0.2514–1.3551
10	0.5668–2.6150	0.0220–0.0713	0.0956–0.4818	0.3358–2.5567
11	0.5286–3.7867	0.0265–0.0902	0.1061–0.9908	0.3200–6.6985
12	0.5217–3.9059	0.0163–0.0674	0.0869–0.8601	0.3254–4.0952
13	0.8167–7.4743	0.0206–0.0760	0.1382–1.4769	0.4198–9.0257
14	0.5200–6.0213	0.0180–0.0723	0.0946–1.9462	0.3329–9.1442
15	0.8947–8.3119	0.0135–0.0587	0.1234–1.5707	0.4270–8.3679
16	1.0148–11.525	0.0189–0.0882	0.2241–3.1639	0.5497–31.104
17	0.9386–10.286	0.0156–0.0651	0.1508–2.1665	0.4415–14.533
18	1.4034–26.387	0.0086–0.0563	0.1722–2.0161	0.4893–27.188
19	1.3308–23.512	0.0190–0.1047	0.5529–4.9987	1.7672–154.48
20	2.4621–37.111	0.0133–0.0843	0.5259–6.6335	1.2741–1222.4
21	2.5248–38.924	0.0134–0.0632	0.4267–3.3610	0.8317–56.458
22	2.0333–44.753	0.0095–0.0582	0.2997–2.7765	0.6508–92.902
23	1.7769–46.055	0.0121–0.0809	0.3969–4.1711	0.7747–461.71

Table A2
Range of the absorption and particle backscattering coefficients (all at 443 nm) in global waters. The results were obtained from mission-long VIIRS-SNPP monthly products (9 km).

Class	$a_{pg}(443)$	$b_{bp}(443)$	$a_{ph}(443)$	$a_{dg}(443)$
1	0.0071–0.0203	0.0015–0.0026	0.0036–0.0151	0.0026–0.0065
2	0.0137–0.0335	0.0015–0.0035	0.0061–0.0233	0.0055–0.0116
3	0.0210–0.0452	0.0016–0.0040	0.0087–0.0293	0.0083–0.0190
4	0.0306–0.0681	0.0017–0.0051	0.0118–0.0418	0.0114–0.0326
5	0.0358–0.0842	0.0015–0.0084	0.0096–0.0476	0.0147–0.0559
6	0.0439–0.1074	0.0020–0.0078	0.0190–0.0627	0.0160–0.0569
7	0.0460–0.1264	0.0017–0.0207	0.0121–0.0614	0.0199–0.0915
8	0.0539–0.1510	0.0022–0.0178	0.0188–0.0764	0.0210–0.0981
9	0.0527–0.1868	0.0020–0.0414	0.0152–0.0708	0.0250–0.1438
10	0.0667–0.2379	0.0027–0.0569	0.0277–0.1025	0.0281–0.1648
11	0.0635–0.3158	0.0029–0.0388	0.0342–0.1720	0.0230–0.1679
12	0.0691–0.3469	0.0025–0.0742	0.0211–0.1528	0.0341–0.2505
13	0.0960–0.5703	0.0034–0.0936	0.0349–0.2858	0.0404–0.3348
14	0.0719–0.6413	0.0029–0.1629	0.0253–0.2368	0.0350–0.4763
15	0.0936–0.7047	0.0028–0.1079	0.0299–0.2340	0.0472–0.5365
16	0.1165–1.5153	0.0044–0.2765	0.0444–0.4576	0.0528–1.1975
17	0.1080–1.1302	0.0038–0.1759	0.0382–0.3342	0.0524–0.9082
18	0.1291–2.0633	0.0033–0.1343	0.0433–0.4237	0.0668–1.7134
19	0.1464–3.0771	0.0064–0.4215	0.0629–0.6710	0.0673–2.5834
20	0.2748–4.1827	0.0079–0.4645	0.0814–0.6909	0.1578–3.5472
21	0.2217–2.9785	0.0080–0.3079	0.0797–0.7644	0.1087–2.4821
22	0.1933–3.6118	0.0060–0.2310	0.0611–0.6921	0.1067–3.0480
23	0.1916–4.6077	0.0071–0.4012	0.0653–0.7744	0.1046–3.9048

References

Bailey, S.W., Werdell, P.J., 2006. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. Remote Sens. Environ. 102, 12–23. https://doi.org/10.1016/j.rse.2006.01.015.

Barale, V., Jaquet, J.-M., Ndiaye, M., 2008. Algal blooming patterns and anomalies in the Mediterranean Sea as derived from the SeaWiFS data set (1998–2003). Remote Sens. Environ. 112, 3300–3313. https://doi.org/10.1016/j.rse.2007.10.014.

Bosc, E., Bricaud, A., Antoine, D., 2004. Seasonal and interannual variability in algal biomass and primary production in the Mediterranean Sea, as derived from 4 years of SeaWiFS observations. Glob. Biogeochem. Cycles 18, GB1005. https://doi.org/10.1029/2003GB002034.

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

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

Campbell, J.W., 1995. The lognormal distribution as a model for bio-optical variability in the sea. J. Geophys. Res. 100, 13237–13254. https://doi.org/10.1029/95JC00458.

Cui, T.W., Zhang, J., Groom, S., Sun, L., Smyth, T., Sathyendranath, S., 2010. Validation of MERIS Ocean-color products in the Bohai Sea: a case study for turbid coastal waters. Remote Sens. Environ. 114, 2326–2336.

Cui, T.W., Zhang, J., Wang, K., Wei, J.W., Mu, B., Ma, Y., et al., 2020. Remote sensing of chlorophyll a concentration in turbid coastal waters based on a global optical water classification system. ISPRS J. Photogramm. Remote Sens. 163, 187–201. https://doi.org/10.1016/j.isprsjprs.2020.02.017.

Devred, E., Sathyendranath, S., Platt, T., 2007. Delineation of ecological provinces using ocean colour radiometry. Mar. Ecol. Prog. Ser. 346, 1–13.

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.

Esaias, W.E., Iverson, R.L., Turpie, K., 2000. Ocean province classification using ocean colour data: observing biological signatures of variations in physical dynamics. Glob. Chang. Biol. 6, 39–55.

Fay, A.R., McKinley, G.A., 2014. Global open-ocean biomes: mean and temporal variability. Earth Syst. Sci. Data 6, 273–284. https://doi.org/10.5194/essd-6-273-2014.

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. https://doi.org/10.1093/plankt/24.9.947.

Gordon, H.R., Wang, M., 1994. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. Appl. Opt. 33, 443–452.
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.
Hardman-Mountford, N.J., Hirata, T., Richardson, K.A., Aiken, J., 2008. An objective methodology for the classification of ecological pattern into biomes and provinces for the pelagic ocean. Remote Sens. Environ. 112, 3341–3352. https://doi.org/10.1016/j.rse.2008.02.016.
Hu, C., 2022. Hyperspectral reflectance spectra of floating matters derived from hyperspectral imager for the Coastal Ocean (HICO) observations. Earth Syst. Sci. Data 14, 1183–1192. https://doi.org/10.5194/essd-14-1183-2022.
Ibrahim, A., Franz, B., Ahmad, Z., Healy, R., Knobelspiesse, K., Gao, B.-C., et al., 2018. Atmospheric correction for hyperspectral ocean color retrieval with application to the hyperspectral imager for the Coastal Ocean (HICO). Remote Sens. Environ. 204, 60–75. https://doi.org/10.1016/j.rse.2017.10.041.
IOCCG, 2006. In: Lee, Z.P. (Ed.), Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. International Ocean Color Coordinating Group, Dartmouth, Nova Scotia, Canada, p. 126. https://doi.org/10.25607/OBP-96.
IOCCG, 2009. Partition of the ocean into ecological provinces: role of ocean-color. In: Dowell, M., Platt, T. (Eds.), Reports of the International Ocean-Color Coordinating Group. International Ocean Color Coordinating Group, Dartmouth, Canada, p. 99.
IOCCG, 2019. In: Mélin, F. (Ed.), Uncertainty in Ocean Color Remote Sensing. International Ocean Color Coordinating Group, Dartmouth, NS, Canada, p. 164. https://doi.org/10.25607/OBP-696.
Jackson, T., Sathyendranath, S., Mélin, F., 2017. An improved optical classification scheme for the ocean colour essential climate variable and its applications. Remote Sens. Environ. 203, 152–161. https://doi.org/10.1016/j.rse.2017.03.036.
Jerlov, N.G., 1976. Marine Optics. Elsevier Scientific Publishing Company, Amsterdam.
Jia, T., Zhang, Y., Dong, R., 2021. A universal fuzzy logic optical water type scheme for the global oceans. Remote Sens. 13, 4018.
Jiang, L., Wang, M., 2014. Improved near-infrared ocean reflectance correction algorithm for satellite ocean color data processing. Opt. Express 22, 21657–21678. https://doi.org/10.1364/OE.22.021657.
Jiang, G., Loiselle, S.A., Yang, D., Ma, R., Su, W., Gao, C., 2020. Remote estimation of chlorophyll a concentrations over a wide range of optical conditions based on water classification from VIIRS observations. Remote Sens. Environ. 241, 111735 https://doi.org/10.1016/j.rse.2020.111735.
Lavigne, H., Van der Zande, D., Ruddick, K., Cardoso Dos Santos, J.F., Gohin, F., Brotas, V., et al., 2021. Quality-control tests for OC4, OC5 and NIR-red satellite chlorophyll-a algorithms applied to coastal waters. Remote Sens. Environ. 255, 112237 https://doi.org/10.1016/j.rse.2020.112237.
Le, C., Li, Y., Zha, Y., Sun, D., Huang, C., Zhang, H., 2011. Remote estimation of chlorophyll a in optically complex waters based on optical classification. Remote Sens. Environ. 115, 725–737.
Lee, Z.P., Hu, C., 2006. Global distribution of case 1 waters: an analysis from SeaWiFS measurements. Remote Sens. Environ. 101, 270–276.
Lee, Z.P., 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.
Lee, Z., Shang, S., Zhang, S., Wu, J., Wei, G., Wu, X., 2020. Impact of temporal variation of chlorophyll-specific absorption on phytoplankton phenology observed from ocean color satellite: a numerical experiment. J. Geophys. Res. 125, e2020JC016382 https://doi.org/10.1029/2020JC016382.
Liu, X., Wang, M., 2019. Filling the gaps of missing data in the merged VIIRS SNPP/NOAA-20 ocean color product using the DINEOF method. Remote Sens. 11, 178.
Longhurst, A., 1998. Ecological Geography of the Sea. Academic Press, San Diego, California.
Longhurst, A., Sathyendranath, S., Platt, T., Caverhill, C., 1995. An estimate of global primary production in the ocean from satellite radiometer data. J. Plankton Res. 17, 1245–1271.
Lubac, B., Loisel, H., 2007. Variability and classification of remote sensing reflectance spectra in the eastern English Channel and southern North Sea. Remote Sens. Environ. 110, 45–58.
Martin Traykovski, L.V., Sosik, H.M., 2003. Feature-based classification of optical water types in the Northwest Atlantic based on satellite ocean color data. J. Geophys. Res. 108 https://doi.org/10.1029/2001JC001172.
McClain, C.R., Signorini, S.R., Christian, J.R., 2004. Subtropical gyre variability observed by ocean-color satellites. Deep-Sea Res. II 51, 281–301.
Mckee, D., Cunningham, A., 2006. Identification and characterisation of two optical water types in the Irish Sea from in situ inherent optical properties and seawater constituents. Estuar. Coast. Shelf Sci. 68, 305–316.
Mélin, F., Vantrepotte, V., 2015. How optically diverse is the coastal ocean? Remote Sens. Environ. 160, 235–251. https://doi.org/10.1016/j.rse.2015.01.023.
Mikelsons, K., Wang, M., 2018. Interactive online maps make satellite ocean data accessible (https://www.Star.Nesdis.Noaa.gov/socd/mecb/color/ocview/ocview.html). Eos 99. https://doi.org/10.1029/2018EO096563.
Mobley, C.D., Stramski, D., Bissett, W.P., Boss, E., 2004. Optical modeling of ocean waters: is the case 1 - case 2 classification still useful? Oceanography 17, 60–67.
Moore, T.S., Campbell, J.W., Feng, H., 2001. A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. IEEE Trans. Geosci. Remote Sens. 39, 1764–1776.
Moore, T.S., Campbell, J.W., Dowell, M.D., 2009. A class-based approach to characterizing and mapping the uncertainty of the MODIS Ocean chlorophyll product. Remote Sens. Environ. 113, 2424–2430.
Morel, A., Prieur, L., 1977. Analysis of variations in ocean color. Limnol. Oceanogr. 22, 709–722.
O'Reilly, J.E., Maritorena, S., O'Brien, M.C., Siegel, D.A., Toole, D.A., Menzies, D.W., 2000. In: Hooker, S.B., Firestone, E.R. (Eds.), SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3, SeaWiFS Postlaunch Technical Report Series. NASA Goddard Space Flight Center, Greenbelt, MD, p. 58).
Oliver, M.J., Irwin, A.J., 2008. Objective global ocean biogeographic provinces. Geophys. Res. Lett. 35.
Ondrusek, M., Lance, V.P., Wang, M., Arnone, R.A., Ladner, S., Goode, W., 2015. Report for dedicated JPSS VIIRS ocean color calibration/validation cruise. In: Lance, V.P. (Ed.), NOAA Technical Report NESDIS 146. NOAA National Environmental Satellite, Data, and Information Service, Washington, D.C., p. 60. https://doi.org/10.7289/V52B8W0Z
Ondrusek, M., Lance, V.P., Wang, M., Arnone, R.A., Ladner, S., Goode, W., 2016. Report for dedicated JPSS VIIRS ocean color December 2015 calibration/validation cruise. In: Lance, V.P. (Ed.), NOAA Technical Report NESDIS 148. National Environmental Satellite, Data, and Information Service, Silver Spring, Maryland: United States, p. 66. https://doi.org/10.7289/V5/TR-NESDIS-148.
Ondrusek, M., Lance, V.P., Wang, M., Stengel, E., Kovach, C., Arnone, R.A., 2017. Report for dedicated JPSS VIIRS ocean color calibration/validation cruise, October 2016. In: Lance, V.P. (Ed.), NOAA Technical Report NESDIS 151. National Environmental Satellite, Data, and Information Service, Washington, D.C.: United States, p. 65. https://doi.org/10.7289/V5/TR-NESDIS-151.
Ondrusek, M., Lance, V.P., Wang, M., Stengel, E., Kovach, C., Arnone, R.A., 2019. Report for dedicated JPSS VIIRS ocean color calibration/validation cruise May 2018. In: Lance, V.P. (Ed.), NOAA Technical Report NESDIS 152. National Environmental Satellite, Data, and Information Service, Washington, D.C.: United States, p. 91. https://doi.org/10.25923/scyb-qf42.
Pitarch, J., van der Woerd, H.J., Brewin, R.J.W., Zielinski, O., 2019. Optical properties of forel-ule water types deduced from 15 years of global satellite ocean color observations. Remote Sens. Environ. 231, 111249 https://doi.org/10.1016/j.rse.2019.111249.
Polovina, J.J., Howell, E.A., Abecassis, M., 2008. Ocean's least productive waters are expanding. Geophys. Res. Lett. 35, 1–5.
Qi, L., Hu, C., Mikelsons, K., Wang, M., Lance, V., Sun, S., et al., 2020. In search of floating algae and other organisms in global oceans and lakes. Remote Sens. Environ. 239, 111659 https://doi.org/10.1016/j.rse.2020.111659.
Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., Maury, O., 2013. Dynamic biogeochemical provinces in the global ocean. Glob. Biogeochem. Cycles 27, 1046–1058.
Reygondeau, G., Cheung, W.W.L., Wabnitz, C.C.C., Lam, V.W.Y., Frölicher, T., Maury, O., 2020. Climate change-induced emergence of novel biogeochemical provinces. Front. Mar. Sci. 7 https://doi.org/10.3389/fmars.2020.00657.
Salgado-Hernanz, P.M., Racault, M.F., Font-Muñoz, J.S., Basterretxea, G., 2019. Trends in phytoplankton phenology in the Mediterranean Sea based on ocean-colour remote sensing. Remote Sens. Environ. 221, 50–64. https://doi.org/10.1016/j.rse.2018.10.036.
Sathyendranath, S., Brewin, R.J.W., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., et al., 2019. An ocean-colour time series for use in climate studies: the experience of the ocean-colour climate change initiative (OC-CCI). Sensors 19, 4285.
Shi, W., Wang, M., 2012. Satellite views of the Bohai Sea, Yellow Sea, and East China Sea. Prog. Oceanogr. 104, 30–45. https://doi.org/10.1016/j.pocean.2012.05.001.
Shi, W., Wang, M., 2014. Ocean reflectance spectra at the red, near-infrared, and shortwave infrared from highly turbid waters: a study in the Bohai Sea, Yellow Sea, and East China Sea. Limnol. Oceanogr. 59, 427–444. https://doi.org/10.4319/lo.2014.59.2.0427.
Shi, W., Wang, M., 2019. A blended inherent optical property algorithm for global satellite ocean color observations. Limnol. Oceanogr. Methods 17, 377–394. https://doi.org/10.1002/lom3.10320.
Shi, K., Li, Y., Zhang, Y., Li, L., Lv, H., Song, K., 2014. Classification of inland waters based on bio-optical properties. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 543–561. https://doi.org/10.1109/JSTARS.2013.2290744.
Siegel, D.A., Maritorena, S., Nelson, N.B., Behrenfeld, M.J., 2005. Independence and interdependencies among global ocean color properties: reassessing the bio-optical assumption. J. Geophys. Res. 110 https://doi.org/10.1029/2004JC002527.
Signorini, S.R., Franz, B.A., McClain, C.R., 2015. Chlorophyll variability in the oligotrophic gyres: mechanisms, seasonality and trends. Front. Mar. Sci. 2 https://doi.org/10.3389/fmars.2015.00001.
Spyrakos, E., Gonzalez Vilas, L., Torres Palenzuela, J.M., Desmond Barton, E., 2011. Remote sensing chlorophyll a of optically complex waters (rias baixas, NW Spain): application of a regionally specific chlorophyll a algorithm for MERIS full resolution data during an upwelling cycle. Remote Sens. Environ. 115, 2471–2485. https://doi.org/10.1016/j.rse.2011.05.008.
Spyrakos, E., O'Donnell, R., Hunter, P.D., Miller, C., Scott, M., Simis, S.G.H., et al., 2018. Optical types of inland and coastal waters. Limnol. Oceanogr. 63, 846–870. https://doi.org/10.1002/lno.10674.
Uudeberg, K., Aavaste, A., Kõks, K.-L., Ansper, A., Uusõue, M., Kangro, K., et al., 2020. Optical water type guided approach to estimate optical water quality parameters. Remote Sens. 12, 931.
Van der Zande, D., Lavigne, H., Blauw, A., Prins, T., Desmit, X., Eleveld, M., 2019. In: Coherence in Assessment Framework of Chlorophyll a and Nutrients as Part of the EU Project 'Joint Monitoring Programme of the Eutrophication of the North Sea with Satellite Data' (Ref: DG ENV/MSFD Second Cycle/2016), Activity 2 Report, p. 106.
Vandermeulen, R.A., Mannino, A., Craig, S.E., Werdell, P.J., 2020. 150 shades of green: using the full spectrum of remote sensing reflectance to elucidate color shifts in the

ocean. Remote Sens. Environ. 247, 111900 https://doi.org/10.1016/j.rse.2020.111900.
Vantrepotte, V., Loisel, H., Dessailly, D., Mériaux, X., 2012. Optical classification of contrasted coastal waters. Remote Sens. Environ. 123, 306–323.
Wang, M., 2007. Remote sensing of the ocean contributions from ultraviolet to near-infrared using the shortwave infrared bands: simulations. Appl. Opt. 46, 1535–1547. https://doi.org/10.1364/ao.46.001535.
Wang, M., Jiang, L., 2018. Atmospheric correction using the information from the short blue band. IEEE Trans. Geosci. Remote Sens. 56, 6224–6237. https://doi.org/10.1109/TGRS.2018.2833839.
Wang, M., Shi, W., 2007. The NIR-SWIR combined atmospheric correction approach for MODIS Ocean color data processing. Opt. Express 15, 15722–15733. https://doi.org/10.1364/oe.15.015722.
Wang, M., Son, S., 2016. VIIRS-derived chlorophyll-a using the ocean color index method. Remote Sens. Environ. 182, 141–149. https://doi.org/10.1016/j.rse.2016.05.001.
Wang, M., Son, S., Harding, L.W., 2009. Retrieval of diffuse attenuation coefficient in the Chesapeake Bay and turbid ocean regions for satellite ocean color applications. J. Geophys. Res. 114 https://doi.org/10.1029/2009JC005286.
Wang, M., Liu, X., Tan, L., Jiang, L., Son, S., Shi, W., et al., 2013. Impacts of VIIRS SDR performance on ocean color products. J. Geophys. Res. 118, 10,347–310,360. https://doi.org/10.1002/jgrd.50793.
Wang, M., Liu, X., Jiang, L., Son, S., 2017. The VIIRS ocean color product algorithm theoretical basis document version 1.0. In: NOAA NESDIS STAR Algorithm Theoretical Basis Document. NOAA/NESDIS/STAR, College Park, MD, p. 68.
Wei, J., Lee, Z.P., Ondrusek, M., Mannino, A., Tzortziou, M., Armstrong, R., 2016a. Spectral slopes of the absorption coefficient of colored dissolved and detrital material inverted from UV-visible remote sensing reflectance. J. Geophys. Res. 121, 1953–1969. https://doi.org/10.1002/2015JC011415.
Wei, J., Lee, Z.P., Shang, S., 2016b. A system to measure the data quality of spectral remote sensing reflectance of aquatic environments. J. Geophys. Res. 121, 8189–8207. https://doi.org/10.1002/2016JC012126.
Wei, J., Lee, Z.P., Shang, S., Yu, X., 2019. Semianalytical derivation of phytoplankton, CDOM, and detritus absorption coefficients from the landsat 8/OLI reflectance in coastal waters. J. Geophys. Res. 124, 3682–3699. https://doi.org/10.1029/2019jc015125.
Wei, J., Yu, X., Lee, Z.P., Wang, M., Jiang, L., 2020. Improving low-quality satellite remote sensing reflectance at blue bands over coastal and inland waters. Remote Sens. Environ. 250, 112029 https://doi.org/10.1016/j.rse.2020.112029.
Wei, J., Wang, M., Jiang, L., Yu, X., Mikelsons, K., Shen, F., 2021a. Global estimation of suspended particulate matter from satellite ocean color imagery. J. Geophys. Res. 126 https://doi.org/10.1029/2021JC017303.
Wei, J., Wang, M., Lee, Z.P., Ondrusek, M., Zhang, S., Ladner, S., 2021b. Experimental analysis of the measurement precision in spectral water-leaving radiance in different water types. Opt. Express 29, 2780–2797. https://doi.org/10.1364/OE.413784.
Wernand, M.R., van der Woerd, H.J., 2010. Spectral analysis of the Forel-Ule Ocean colour comparator scale. J. Eur. Opt. Soc. 5, 10014S https://doi.org/10.2971/jeos.2010.10014s.
Yu, X., Lee, Z.P., Shen, F., Wang, M., Wei, J., Jiang, L., et al., 2019. An empirical algorithm to seamlessly retrieve the concentration of suspended particulate matter from water color across ocean to turbid river mouths. Remote Sens. Environ. 235, 111491 https://doi.org/10.1016/j.rse.2019.111491.
Zibordi, G., Berthon, J.-F., Mélin, F., D'Alimonte, D., Kaitala, S., 2009a. Validation of satellite ocean color primary products at optically complex coastal sites: northern Adriatic Sea, northern Baltic proper and gulf of Finland. Remote Sens. Environ. 113, 2574–2591.
Zibordi, G., Mélin, F., Berthon, J.-F., Holben, B., Slutsker, I., Giles, D., et al., 2009b. AERONET-OC: a network for the validation of ocean color primary products. J. Atmos. Ocean. Technol. 26, 1634–1651. https://doi.org/10.1175/2009JTECHO654.1.