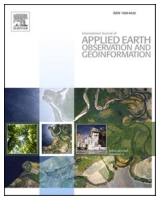
International Journal of Applied Earth Observations and Geoinformation 116 (2023) 103173

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

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



Mapping *Ulva prolifera* green tides from space: A revisit on algorithm design and data products

Chuanmin Hu [a,*], Lin Qi [b,c], Lianbo Hu [d], Tingwei Cui [e], Qianguo Xing [f], Mingxia He [d], Ning Wang [g], Yanfang Xiao [h], Deyong Sun [i], Yingcheng Lu [j], Chao Yuan [h], Mengquan Wu [k], Changying Wang [l], Yanlong Chen [m], Haipeng Xu [n], Li'e Sun [o], Maohua Guo [p], Menghua Wang [b]

[a] College of Marine Science, University of South Florida, St. Petersburg, FL, USA
[b] NOAA Center for Satellite Applications and Research, College Park, MD, USA
[c] Global Science & Technology Inc., Greenbelt, MD, USA
[d] Ocean Remote Sensing Institute, Ocean University of China, Qingdao, Shandong, China
[e] School of Atmospheric Sciences, Sun Yat-sen University, Key Laboratory of Tropical Atmosphere-Ocean System, Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai, Guangdong, China
[f] Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, Shandong, China
[g] North China Sea Marine Forecasting Center of State Oceanic Administration, Qingdao, Shandong, China
[h] First Institute of Oceanography, Ministry of Natural Resources, Qingdao, Shandong, China
[i] School of Marine Sciences, Nanjing University of Information Science & Technology, Nanjing, China
[j] International Institute for Earth System Science, Nanjing University, Nanjing, China
[k] College of Resources and Environmental Engineering, Ludong University, Yantai, Shandong, China
[l] College of Computer Science and Technology, Qingdao University, Qingdao, China
[m] National Marine Environmental Monitoring Center, Dalian, China
[n] Lianyungang Sea Area Use and Protection Dynamic Management Center, Lianyugang, Jiangsu, China
[o] Qingdao Ecological and Environmental Monitoring Centre of Shandong Province, Qingdao, Shandong, China
[p] National Satellite Ocean Application Service, Haidian District, Beijing, China

ARTICLE INFO

Keywords:
Ulva prolifera
Coverage
Biomass
Remote sensing
MODIS
VIIRS
OLCI
GOCI
MSI
OLI
GaoFen
CZI
FAI
AFAI
NDVI
DVI
VB-FAH
EVI

ABSTRACT

Since the first report in 2008, macroalgal blooms of *Ulva prolifera* (often called green tides) in the Yellow Sea have occurred every year, with their origins, transport pathways, temporal changes, as well as causes and consequences studied extensively. Of these studies, satellite remote sensing has been used widely to detect the bloom presence and quantify the bloom size (i.e., *U. prolifera* coverage in km^2 or biomass in kilotons). However, substantial variability has been found in the refereed literature in the remote sensing methodology, results, and interpretation of the *U. prolifera* coverage, especially in the attempts to study inter-annual changes or long-term trends. There are often inconsistent or contradicting results even from the same satellite sensor. Such inconsistencies or contradictions create difficulty not only within the remote sensing community when presenting new methodology or results, but also to researchers when attempting to use the remote sensing results to make predictions or perform impact assessments. Here, we review the literature on the remote sensing methodology to detect and quantify *U. prolifera* blooms, and make recommendations based on physical principles. Specifically, we propose the following conceptual guidelines: 1) a reliable index or algorithm should be relatively tolerant to perturbations by non-optimal observing conditions (thick aerosols, thin clouds, moderate sun glint, cloud-adjacent straylight, which can all be found frequently in the study region) for presence/absence detection, as well as to small errors in the selected thresholds to quantify *U. prolifera*; 2) a reliable index or algorithm should also make it relatively easy to account for variability in subpixel coverage of *U. prolifera* (i.e., through pixel unmixing) in order to obtain an accurate estimate of total *U. prolifera* coverage from an image; 3) a reliable data product (i.e., *U. prolifera* maps) should be able to account for the variable clouds when interpreting spatial patterns or temporal changes, with uncertainty estimates provided whenever possible; and 4) both the algorithm and the data product should minimize manual work in order to make them more objective and repeatable by other researchers. Finally, we show different types of time series of *U. prolifera* amounts in the Yellow Sea using

* Corresponding author.
E-mail address: huc@usf.edu (C. Hu).

https://doi.org/10.1016/j.jag.2022.103173
Received 22 October 2022; Received in revised form 29 November 2022; Accepted 27 December 2022
Available online 3 January 2023

the approaches based on these guidelines and Moderate Resolution Imaging Spectroradiometer (MODIS) observations, and discuss their implications on the interpretation of annual changes in interdisciplinary studies.

Notations

R	Reflectance (dimensionless), wavelength (band) dependent and often refers to R_{rc}
R_{rc}	Rayleigh-corrected reflectance (dimensionless), wavelength dependent
R_w	Reflectance of water (dimensionless), wavelength dependent
R_T	Reflectance of the target pixel, which is often a mixture of *U. prolifera* and water
R_U	Endmember reflectance of *U. prolifera* (100% subpixel coverage)
NIR	Near-infrared wavelengths, 700–1100 nm
SWIR	Shortwave infrared wavelengths, 1100–3000 nm
α	sub-pixel fraction (0.0–1.0, or 0%–100%) of algae coverage. Some earlier studies used the symbol χ instead
FAI	Floating Algae Index, a scalar quantity derived from a linear combination of reflectance at three bands (R_{b1}, R_{b2}, R_{b3}) to measure the enhanced NIR reflectance due to floating algae
AFAI	Alternative Floating Algae Index, similar to *FAI* but three other bands are used for the linear index
NDVI	Normalized Difference Vegetation Index, defined as $(R_{NIR} - R_{red})/(R_{NIR} + R_{red})$
EVI	Enhanced Vegetation Index, an index to use one blue band, one red band, and one NIR band to reduce the impact of aerosols
DVI	Difference Vegetation Index, defined as $(R_{NIR} - R_{red})$
VB-FAH	Virtual Baseline Floating macroAlgae Height, an index to use a green band as a surrogate of SWIR band in the *FAI* calculation
RGB	Red-Green-Blue composite, often called true color
FRGB	False-color Red-Green-Blue composite, where the Green channel uses a NIR band instead of a green band in the RGB composite.
SeaDAS	SeaWiFS Data Analysis System, originally developed by NASA to process SeaWiFS data but further developed to have the capacity to process data collected by most ocean color sensors
OCView	Ocean Color View, an online tool developed by NOAA to visualize ocean color imagery in near real-time
SNR	signal-to-noise ratio ($=L_{typ} / \sigma$), wavelength dependent
CZI	Coastal Zone Imager onboard the HY-1C (2018 – present) and HY-1D (2020 – present) satellites
DOVE	High-resolution (3–4 m) multi-band sensors available on a constellation of small satellites operated by Planet Labs since 2014
GF	Gao-Fen (high resolution in English translation), a series of Chinese satellite missions starting 2013 to map Earth surface
GOCI	Geostationary Ocean Color Imager on COMS satellite (2011–2021) operated by the Korea Institute of Ocean Science & Technology
HJ	Huan-Jing (environment in English translation), Chinese satellite missions starting 2008 to map Earth surface with high-resolution sensors (3–100 m) onboard
Landsat	A series of Earth-observing satellite missions since 1972 operately jointly by the United States Geological Survey and NASA, with most missions equipped with multi-band sensors
MERIS	Medium Resolution Imaging Spectrometer (2002–2012), one of the main instruments onboard the Envisat platform of the European Space Agency
MSI	Multi-Spectral Instrument on Sentinel-2A and Sentinel-2B satellites operated by the European Space Agency
MODIS	Moderate Resolution Imaging Spectroradiometer on Terra (2000 – present) and Aqua satellites (2002 – present) operated by the U.S. NASA
OLI	Operational Land Imager on Landsat-8 satellite (2013 – present)
OLCI	Ocean and Land Colour Instrument on Sentinel-3A and Sentinel-3B satellites operated by the European Space Agency
SAR	Synthetic Aperture Radar
VIIRS	Visible Infrared Imaging Radiometer Suite, ocean color sensor on the Suomi NPP (2012 – present) and NOAA-20 satellites (2018 – present)

1. Background

Green macroalgae blooms of *Ulva prolifera* (often called green tides) in the western Yellow Sea (Fig. 1) have captured the attention of researchers and public media since June 2008, when the bloom became known internationally because of its impact on the Olympic sailing game (Hu and He, 2008). Since then, these annually recurrent blooms have been studied extensively, including the bloom origin, transport pathways, reasons behind inter-annual variations, and implications to ocean ecology. A general consensus is that the blooms originate from the Subei Shoal (Huo et al., 2013; Lee et al., 2011; Liu et al., 2010; Liu et al., 2009), where aquaculture of *Porphyra yezoensis* (the *nori* seaweed used to make sushi) is an important source to provide the initial seed population of *U. prolifera* (Hu et al., 2010; S. Hu et al., 2014; Liu et al., 2010; Liu et al., 2013; Liu et al., 2009; Liu et al., 2015; Wang et al., 2015; Xing et al., 2019; Zhang et al., 2017; Y. Zhang et al., 2019). This is because *U. prolifera* could grow on the aquaculture rafts, which was usually discarded during seaweed harvest. Such discarded *U. prolifera* was transported to offshore waters by ocean currents, where optimal light, temperature, and nutrient conditions promote rapid growth, resulting in large-scale blooms in the western YS (Hu et al., 2010; Liu et al., 2016).

Satellite remote sensing has been used in many studies to show bloom location and to estimate bloom size in terms of *U. prolifera* coverage or biomass (Table 1). Some of these studies also used several environmental factors (winds, currents, water temperature, light, nutrients) to explain spatial patterns and temporal changes of the blooms. Various satellite sensors with different spectral bands and spatial resolutions have been used, including MODIS, VIIRS, OLCI, OLI, MSI, GOCI, GaoFen, CZI, among others. Likewise, different indexes or algorithms have been developed and used to detect and quantify the blooms. These include the Normalized Difference Vegetation Index (NDVI) (Hu and He 2008; Cui et al., 2012), Normalized Difference Algae Index (NDAI) (Garcia et al., 2013; Keesing et al., 2011; Shi and Wang, 2009) Floating Algae Index (FAI) (He et al., 2011; Hu, 2009; Hu et al., 2010; Hu et al., 2017; Qi et al., 2016; Xu et al., 2014), alternative FAI (AFAI) (Qi et al.,

2022a), Enhanced Vegetation Index (Xiao et al., 2019), Virtual-Baseline Floating macroAlgae Height (VB-FAH) (Xing and Hu, 2016), Difference Vegetation Index (DVI) (Xing et al., 2019), and machine learning algorithms (Qiu et al., 2018; Wan et al., 2021; Gao et al., 2022). A list of references is provided in Table 1. After the detection of the algae-containing pixels from individual images, some papers took further steps to account for variable sub-pixel fractional cover through pixel unmixing, while most of the papers simply counted the total number of algae-containing pixels without pixel unmixing (Table 1). Furthermore, to study inter-annual changes in bloom size, most of the studies used the daily maximum bloom size within a year to represent that year (e.g., Xing et al., 2015, 2019; Hu L et al., 2019; Xiao et al., 2019; Li et al., 2022; Yuan et al., 2022), while Qi et al. (2016 & 2022a) used image composites to remove variable cloud cover when estimating the mean monthly biomass or bloom size.

Such diversified approaches to detect *U. prolifera* and estimate bloom size led to different and sometimes contradicting results. For example, for the same MODIS image, the estimated bloom size could differ by more than one order of magnitude (Table 1 of Hu L et al., 2019); for inter-annual changes during the period of 2008–2016, some studies reported maximum bloom size in 2009 (e.g., Zhang et al., 2020; J. Zhang et al., 2019; Y. Zhang et al., 2019) while others showed 2009 as one of the minimum-bloom years with 2016 being the maximum year (Hu L et al., 2019). Likewise, for the period of 2007–2021, several studies reported maximum year of 2021 with *U. prolifera* coverage being several times higher than that in 2019 (Li et al., 2022; Zheng et al., 2022), while another study (Qi et al., 2022a) showed 2019 being the maximum year.

These different and contradicting results make it extremely difficult to explain or predict bloom patterns in either space or time using either multi-variant statistical analysis or numeric models. After all, if a time series of bloom patterns is incorrect to begin with in the first place, what reliable findings or conclusions can be drawn through analyzing environmental factors?

Recognizing such a difficulty, Hu L et al. (2019) explained in detail how to apply a linear index to detect and unmix algae-containing pixels to estimate fractional cover and biomass, and Qi et al. (2016) and Qi et al. (2022a) further combined multi-day detection results to generate monthly image composites to estimate the monthly mean biomass. Unfortunately, as more data have been collected and made available from multiple satellite missions, more recent publications still applied some of the unreliable methods to generate long-term time-series images, and, as before, these results continued to differ and sometimes contradict each other. Such results create difficulty not only within the remote sensing community when presenting new methodology or results, but also to researchers when attempting to use the remote sensing results to make predictions or perform impact assessments.

The objectives of this paper are three folds: 1) to compile and comment on the published methodology and results on the remote sensing of *U. prolifera* blooms; 2) to propose several general guidelines on algorithm design and on the approach to generate time series of *U. prolifera* maps; and 3) to demonstrate time-series of *U. prolifera* data and discuss their implications on interdisciplinary research. Within this text, for simplicity, the words "algae", "macroalgae", "*Ulva*" and "*U. prolifera*" are used interchangeably.

2. Data and methods

2.1. Literature search

The official literature database available at webofscience.com was searched in August 2022 with the keyword combination of "*Ulva*" and "remote sensing", resulting in 69 references. Of these references, those not directly relevant (e.g., *Ulva* in other regions, or remote sensing data are only used superficially for illustrations) are excluded. The remaining references were further searched through Google Scholar to obtain additional papers that cited these references. In the end, only those papers that show representative methodology on optical remote sensing or long-term time-series were compiled.

2.2. Satellite data

Although all optical satellite sensors can be used for the purpose of detecting and quantifying *U. prolifera* blooms, in this study, for demonstration purpose, we use MODIS data only. This is because MODIS has appropriate spectral bands, high signal-to-noise ratios (SNRs), and near daily revisit frequency since 2000.

MODIS Level-0 data were obtained from the NASA OB.DAAC (https://oceancolor.gsfc.nasa.gov) and processed to generate Level-1 data, and then processed to generate Rayleigh corrected reflectance ($R_{rc}(\lambda)$, dimensionless) using the SeaDAS software (version 8.0). The use

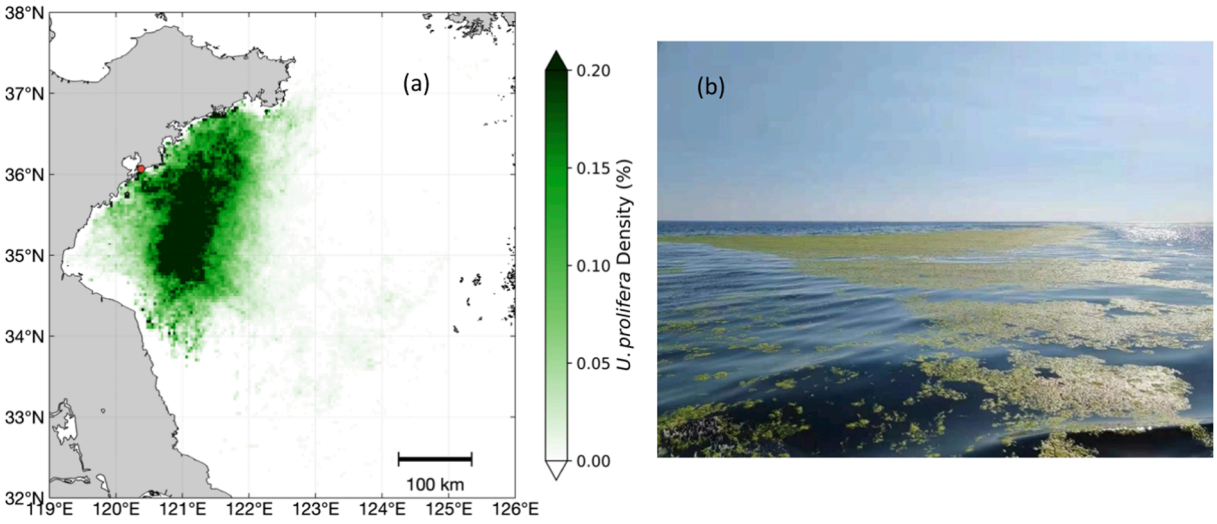


Fig. 1. (a) Mean surface cover density (in % of each 4-km grid) of *U. prolifera* in the western YS, based on MODIS (2000–2021) satellite observations between May and August (figure redrawn from Qi et al., 2022a). The red dot denotes the location of Qingdao, China. The low-density features between 33 and 35°N and 122 – 125°E mainly came from July 2011, June 2015, and June 2016, which were confirmed to be *U. prolifera* instead of *Sargassum horneri* based on analysis of spectral shapes (Qi and Hu, 2021); (b) Digital photo taken from the YS on 23 June 2021 showing *U. prolifera* mats on the ocean surface. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Refereed papers on optical remote sensing of *U. prolifera*, together with their brief descriptions of attributes. The papers are arranged in chronological order.

Author	Year	Sensor(s)	Algorithm/ Index	Pixel unmixing	Daily maximum	Monthly mean	Multi-year	Max year (range)	Max area (km^2)	Max day
Hu & He	2008	MODIS	NDVI	No	No	No	No	2008	N/A	N/A
Liu et al.	2009	MODIS	RGB	No	No	No	No	2008	N/A	N/A
Shi & Wang	2009	MODIS	NDAI	No	No	No	No	2008	N/A	N/A
Hu et al.	2010	MODIS, Landsat	FAI	Yes	Yes	No	Yes	2008 (2000–2008)	1940	N/A
Keesing et al.	2011	MODIS	NDVI	No	Yes	No	Yes	2009 (2004–2009)	4994	15-07-2009
Son et al.	2012	GOCI	NDVI, EVI, etc.	No	Yes	No	No	2011	N/A	N/A
Garcia et al.	2013	MODIS	NDVI, SAI	Yes	Yes	No	Yes	2008 (2008–2009)	481	31-05-2008
Liu et al.	2013	MODIS	NDVI	No	Yes	No	Yes	2009 (2007–2012)	4994	15-07-2009
Zhang et al.	2013	HJ	FRGB	No	Yes	No	No	2012	220	22-05-2012
Xu et al.	2014	MODIS	FAI	No	Yes	No	Yes	2013 (2008–2013)	N/A	N/A
Son et al.	2015	GOCI	IGAC	No	No	No	No	2011	N/A	N/A
Xing et al.	2015	MODIS	NDVI	Yes	Yes	No	No	2008 (2007–2013)	1200	31-05-2008
Qi et al.	2016	MODIS	FAI	Yes	Yes	Yes	Yes	2015 (2007–2015)	1160	21-06-2015
Xing & Hu	2016	HJ, Landsat	VB-FAH	No	No	No	Yes	2012 (1995–2014)	N/A	N/A
Xu et al.	2016	MODIS, HJ, Landsat	NDVI	No	No	No	No	N/A	N/A	N/A
Hu et al.	2017	MODIS	FAI	Yes	Yes	No	Yes	2015 (2008–2015)	895	N/A
Wang et al.	2017	Landsat	FAI, NDVI, etc.	No	No	No	No	N/A	N/A	N/A
Xiao et al.	2017	MODIS, HJ	NDVI	Yes	Yes	No	Yes	2015 (2013–2015)	1065	21-06-2015
Xu et al.	2017	MODIS, others	DVI	No	Yes	No	Yes	2015 (2007–2018)	3800	N/A
Cui et al.	2018	MODIS, GOCI, HJ	NDVI, DVI	Yes	Yes	No	Yes	2016 (2007–2016)	739	25-06-2016
Harun-Al-Rashid	2018	Landsat8	FAI	No	Yes	No	Yes	N/A	N/A	N/A
Jin et al.	2018	GOCI	AFAI	Yes	Yes	No	Yes	2016 (2011–2016)	908	N/A
Li et al.	2018	GF, HJ	NDVI, DVI, etc.	Yes	No	No	No	N/A	N/A	N/A
Qiu et al.	2018	GOCI, Landsat8	NDVI, AFAI, etc.	Yes	Yes	No	Yes	2016 (2016–2017)	885	25-06-2016
Sun et al.	2018	MODIS, HJ	NDVI	No	Yes	No	No	only 2015	1752	N/A
Xing et al.	2018	MODIS, GF, CBERS	DVI	Yes	Yes	No	No	only 2016	539	25-06-2016
Zheng et al.	2020	MODIS, GOCI, etc.	RVI, NDVI, EVI	Yes	Yes	No	No	only 2017	N/A	N/A
Cao et al.	2019	MODIS, HJ, OLI	NDVI	No	Yes	No	Yes	2016 (2016–2018)	2906	25-06-2016
Hu et al.	2019	MODIS, OLI, GF, WV2	FAI, DVI	Yes	Yes	No	Yes	2016 (2008–2016)	1350	25-06-2016
Kim et al.	2019	MODIS, GOCI, OLI	NDVI	No	Yes	No	Yes	2016 (2008–2016)	13353	N/A
Xiao et al.	2019	MODIS	EVI	Yes	Yes	No	Yes	2016 (2007–2016)	1.19 M tons	N/A
Xing et al.	2019	MODIS, others	DVI	No	Yes	No	Yes	2015 (2007–2018)	3800	N/A
Zhang J et al.	2019	HJ-1A & HJ-1B CCD	N/A	N/A	Yes	No	Yes	2009 (2008–2017)	2100	N/A
Zhang Y et al.	2019	MODIS, others	NDVI	No	Yes	No	Yes	2009 (2007–2018)	2100	N/A
Zhang H et al.	2019	GF, HJ, ETM+, GOCI	NDVI, FAI, etc.	No	No	No	No	N/A	N/A	N/A
Cui et al.	2020	GF4	NDVI	No	No	No	No	N/A	N/A	N/A
Zhang G et al.	2020	MODIS, HJ	NDVI	No	Yes	No	Yes	2015 (2011–2018)	1714	N/A
Zhang H et al.	2020	N/A	manual draw	N/A	Yes	No	Yes	2009 (2008–2019)	2100	N/A
An et al.	2021	MODIS, HJ, GF, MSI	DVI	Yes	No	No	Yes	N/A (2007–2020)	N/A	N/A
Li et al.	2021	MODIS, MSI	NDVI	No	Yes	No	Yes	2016 (2007–2020)	1700	N/A
Sun et al.	2021	OLI	FAI	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Wan et al.	2021	GOCI		N/A	N/A	N/A	N/A	N/A	N/A	N/A

(*continued on next page*)

Table 1 (*continued*)

Author	Year	Sensor(s)	Algorithm/ Index	Pixel unmixing	Daily maximum	Monthly mean	Multi-year	Max year (range)	Max area (km^2)	Max day
			Unet, SegNet, etc.							
Wang et al.	2021	MODIS, GF, CZI, etc.	DVI	No	Yes	No	Yes	2019 (2009, 18, 19)	3644	23-06-2019
Zhang G et al.	2021	MSI, OLI, MODIS	NDVI, FAI	No	Yes	No	Yes	2016 (2016–2020)	1582	25-06-2016
Zhang H et al.	2021	HJ, OLI	VB_FAH, FAI	Yes	No	No	No	N/A	N/A	N/A
Gao et al.	2022	MODIS, GF5/ SAR	FAI, NDVI, etc.	N/A	No	No	No	N/A	N/A	N/A
Li et al.	2022	MODIS, others	NDVI	No	Yes	No	Yes	2021 (2007–2021)	2600	N/A
Ma et al.	2022	MODIS. Sentinel-1/SAR	DVI	No	Yes	No	No	only 2021	42384	18-06-2021
Qi et al.	2022	MODIS, Landsat	AFAI, FAI	Yes	No	Yes	Yes	2019 (1984–2021)	N/A	N/A
Wang et al.	2022	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Yuan et al.	2022	MODIS	DVI	Yes	Yes	No	Yes	2021 (2015–2021)	585	19-06-2021
Zhang G et al.	2022	MODIS, HJ, GF1	NDVI	N/A	Yes	No	Yes	2015 (2012–2016)	1752	21-06-2015
Zhang B et al.	2022	MODIS	NDVI	No	Yes	No	Yes	2015 (2015–2019)	2600	18-06-2015
Zheng L et al.	2022	MODIS, Landsat, others	NDVI	No	Yes	No	Yes	2021 (2013–2021)	2000	N/A

of Level-0 is to merge consecutive granules (MODIS data were provided in 5-minute granules) seamlessly before processing to Level-1, as the study region sometimes requires two consecutive granules to cover. In contrast, merging two consecutive Level-1 granules would result in a data gap of 20 scan lines. When SeaDAS is not available, other software (e.g., ENVI) or in-house codes may also be used to generate $R_{rc}(\lambda)$. The use of $R_{rc}(\lambda)$ instead of the fully-corrected remote sensing reflectance ($R_{rs}(\lambda)$, sr^{-1}, a standard product of SeaDAS) is because aerosol or other corrections often fail over pixels containing *U. prolifera* due to their contributions to the atmospheric correction bands (near-infrared (NIR) or shortwave infrared (SWIR)) (Hu, 2009). The use of $R_{rc}(\lambda)$ also follows the tradition on remote detection of floating matters (e.g., Qi et al., 2020; others). Therefore, $R_{rc}(\lambda)$ was used for all subsequent calculations and analysis, including generation of false-color Red-Green-Blue (FRGB) composite images to visualize the *U. prolifera* features. In such images, *U. prolifera* features appear greenish because a NIR MODIS band at 859 nm (250-m resolution) was used to represent the green channel in the FRGB composites (the 645-nm and 469-nm bands were used as the red and blue channels, respectively). *U. prolifera* has enhanced NIR reflectance, thus making the features appear greenish in the FRGB images (Qi et al., 2020).

2.3. Field data

Surface reflectance of *U. prolifera* has been measured in the laboratory and in the field (Hu L et al., 2017; Xiao et al., 2019; others). These measurements were from mixtures of *U. prolifera* and water with different mixing ratio (i.e., subpixel fractional cover in % or biomass density in kg m^{-2}). In this study, the reflectance spectra corresponding to the full *U. prolifera* coverage (~2 kg m^{-2}) from Hu L et al. (2017) were used as the endmember spectra to determine the upper bound threshold used in pixel unmixing.

3. Results

3.1. Literature search

Table 1 lists all papers published in English on the remote detection of *U. prolifera* blooms using optical satellite sensors. Most of these papers used MODIS, while some used higher-resolution sensors for demonstration purposes (e.g., Li et al., 2018; Xing and Hu, 2016). Most papers did not perform pixel unmixing, while only two papers (Qi et al., 2016, 2022a) applied image composition to remove clouds and other artifacts in order to estimate monthly mean biomass and bloom size. Similar to what was found in Hu L et al. (2019), there is substantial variability in the estimated *U. prolifera* coverage (in km^2) even from the same image, and there is also substantial variability in the multi-year annual patterns even for the same reported period. For example, for the period of 2008–2016, Hu L et al. (2019) reported maximum daily coverage of 1,350 km^2 in 2016, as compared to 13,353 km^2 estimated by Kim et al. (2019). Zhang et al. (2020) reported the maximum year of 2009 (daily maximum of 2,100 km^2) for the period of 2008–2019, while Wang et al. (2021) showed higher daily maximum of 3,644 km^2 in 2019 rather than in 2009.

Clearly, these mixed results cannot be all correct, and when the incorrect multi-year patterns are used together with other environmental factors (temperature, light, nutrients, winds, etc.) to infer causality, the inference is likely meaningless. Below, through clarifying several mapping concepts and using examples, we reveal the reasons behind such discrepancy and propose the appropriate methods, in both principle and practice, to map *U. prolifera* from space.

3.2. Requirements and general guidelines on U. prolifera mapping

In computer programming, before developing the computer code in a certain language, pseudo code is often used to first outline the logical steps, and then seek practical ways to realize these logical steps, i.e., implementing in a certain language. Here, we follow the same approach to first outline the algorithm conceptual design, and then demonstrate practical ways to implement such a design.

Conceptually, similar to detecting and quantifying other floating matters (Hu, 2021), detecting and quantifying *U. prolifera* from individual images also require three steps:

I. Is there something floating on the surface? This requires the detection of a spatial anomaly;
II. Is that "something" *U. prolifera* or something else? This requires spectral discrimination and/or knowledge of other ancillary information;
III. How much *U. prolifera* is in each image pixel and each image?

With some *a priori* information (e.g., in the western YS the only floating matter in May – July is likely *U. prolifera*), the first two steps become one, making all three steps into two, as illustrated in the conceptual diagram of Fig. 2:

1) Step 1: classification of image pixels into three categories: algae-containing, algae-free, and non-observable (i.e., invalid). This step is called "detection" or "classification." This step can often be separated in three sub-steps:

1a) determine which pixels are non-observable (i.e., masking cloud pixels and other pixels where algae classification is impossible);

1b) from the remaining pixels, determine which are algae-containing pixels and which are algae-free pixels;

1c) count the number of algae-containing pixels to determine the *U. prolifera* areal coverage in km^2. In addition, the area of water encompassing *U. prolifera* can also be quantified from determining the outer boundaries of the *U. prolifera* pixels.

Basically, Step 1 is an extended version of presence/absence detection.

2) Step 2: quantify *U. prolifera* amount in each algae-containing pixel through pixel unmixing, and then sum up for the entire image. This step is called "objective quantification." The *U. prolifera* amount in each algae-containing pixel can be expressed as % cover (i.e., areal density) within the pixel, equivalent to biomass once a relationship is established between areal density and biomass from laboratory or field experiment (e.g., Hu L et al., 2017; Xiao et al., 2019). Similar to Step 1, the area of water encompassing *U. prolifera* can also be quantified.

In addition to the above two steps applied to individual images, a third step is important as well:

3) Step 3: determine monthly, annual, and inter-annual changing patterns of *U. prolifera* distributions and amounts in order to facilitate studies on their causes and consequences as well as on their modeling and predictions. This is through image composition to remove data gaps due to clouds and other non-observable conditions.

As with any other remote sensing data products (IOCCG, 2019), uncertainty estimates in the derived *U. prolifera* amount are also desired. For brevity, the three steps are focused here but uncertainty estimates are discussed below.

The steps above are not restricted to *U. prolifera* but applicable to all floating macroalgae and microalgae scums. Different studies may have different purposes, and, depending on the specific purpose, not every step above is required and some steps may be simplified. For example, to determine the presence/absence and approximate location of *U. prolifera*, visual inspection of the FRGB images or some other indexes may be sufficient, as long as the colored features can be judged to be *U. prolifera* from either spectral analysis (Qi & Hu, 2021) or from local knowledge of oceanography. In contrast, it is more technically challenging to objectively quantify the *Ulva* amount (in either km^2 coverage or biomass) and to further quantify annual-interannual patterns together with uncertainty estimates. Based on the published literature on the various methodology, below we explain why the above three steps are required for objective mapping of *U. prolifera*, and then lay out the principles that need to be followed either explicitly or implicitly when implementing algorithms or data processing software for the above three steps. These principles include.

1) Step 1, *Ulva* detection through image classification.

This is the fundamental step required for presence/absence detection of any floating matter for obvious reasons, and there is no need to lay out any justifications.

2) Step 2, *Ulva* quantification through pixel unmixing.

This is the step that most published papers missed (Table 1) but is a critical step to objectively quantify *U. prolifera* amount from an individual image unless the pixel resolution is very high (e.g., centimeters). There are two fundamental reasons behind pixel unmixing.

The first is apparent, as most (or all) medium-resolution pixels are not fully covered by *U. prolifera* (Fig. 3), thus requiring pixel unmixing to accurately estimate *U. prolifera* coverage at both pixel level and image level. This is required for nearly all sensors from MODIS (250-m or 1-km resolution) to Landsat (30-m resolution). From statistics of the MODIS 250-m resolution images, 75% of the algae pixels have their *U. prolifera* subpixel coverage <10% (Fig. 7b of Hu L et al., 2017). Even the 2-m resolution pixels of Worldview-2 image do not always contain full algae coverage (Fig. 6 of Hu L et al., 2019). Without pixel unmixing, a pixel containing 0.5% algae is counted the same as a pixel containing 5% or 50% algae, leading to large uncertainties. For the same reason, without pixel unmixing, using images of different resolutions will result in different estimates of *U. prolifera* coverage even for the same region on the same day.

The second is not as apparent but is as important as the first. In classifying algae pixels during Step 1, a lower-bound threshold value is often used, above which the image pixels are classified as algae pixels. In practice, regardless of what method is used (e.g., dynamic selections across different features or different images), there is always a slight error in determining this lower-bound threshold. The slight error will lead to many additional pixels being classified as algae pixels or water pixels, depending on the sign of the error. These additional pixels always have weak signals because they are approaching the lower-bound threshold. With pixel unmixing, their subpixel fractional coverage (α in %) are very small, and therefore will only cause a small error in the total coverage estimates. In contrast, without pixel unmixing, these additional pixels carry the same weight as other algae pixels with stronger signals, leading to disproportional and much higher errors in the total coverage estimates. This is believed to be a major reason why there is a large difference in the total coverage estimates from the same images in different studies (Table 1), as they selected the lower-bound thresholds differently from the same images without using pixel unmixing when calculating the total coverage.

Clearly, pixel unmixing is a critical step in estimating *U. prolifera* coverage from individual images.

3) Step 3, Spatial/temporal patterns through image composition.

This requirement is not as apparent as the above two steps, yet it is a

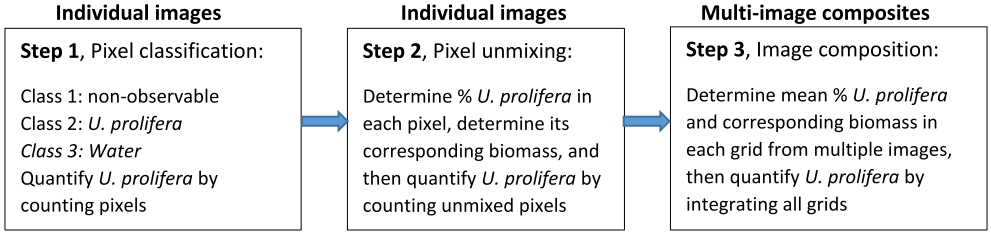


Fig. 2. Conceptual steps to map *U. prolifera* spatial distribution patterns and temporal changes. Most published remote sensing studies only used Step 1, and only two studies used all three steps (Table 1).

Fig. 3. Conceptual illustration of why pixel unmixing is critical in estimating *U. prolifera* coverage. The 1-m resolution image (inset) was collected over Dongjia Bay east of Qingdao, China, with one small region enlarged to show the *U. prolifera* features. The white grid lines illustrate the MODIS 250-m pixel size. There is not a single 250-m pixel that is fully covered by *U. prolifera*, thus requiring pixel unmixing to correctly estimate sub-pixel coverage (α) before total coverage from the image can be derived. The approximate values of α for two grids are illustrated. Statistics of MODIS 250-m resolution images indicates that >99.5% of algae-containing pixels have α < 1.0 (Hu L. et al., 2017).

critical step to derive reliable temporal changing patterns of *U. prolifera* distributions. In the refereed literature, many papers (Table 1) used daily maximum coverage during a year to represent that year. In an ideal situation such as the one illustrated in Fig. 4a, this is a reasonable choice. However, such an ideal situation would never exist, especially for the YS where frequent cloud cover and strong sun glint often make it difficult to obtain valid pixels from satellite images. An example is provided in Fig. S2 of Qi et al. (2017), where the probability of obtaining valid MODIS observations through standard data processing of SeaDAS is shown to be very low from daily images. Fig. 4b shows a hypothetical case where some of the Year 2 images are obscured completely by clouds (i.e., the data gaps) while other images are partially covered by clouds in different regions (annotated by Day X and Day Y). In such a hypothetical

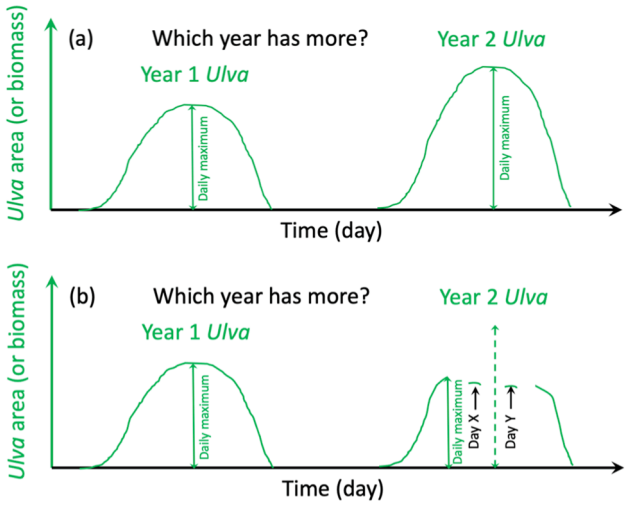


Fig. 4. Conceptual illustration of why extra caution is required when using daily maximum to represent monthly or annual *U. prolifera* amount for (a) hypothetical daily changes of *U. prolifera* in two years, where all daily images are completely cloud free. In this scenario, using a daily maximum to represent the annual mean makes sense and (b) the same hypothetical daily changes, but under a scenario where some of the Year 2 images are obscured completely by clouds (data gaps in the curve) while other images are partially covered by clouds in different regions (Day X and Day Y, see Fig. 5, Scenario 1). In this scenario, the observed daily maximum is an underestimate of the real daily maximum (dashed arrow), and the corresponding timing is also incorrect.

case, the observed daily maximum departs from the real daily maximum, thus cannot be used to represent the mean bloom condition. The same argument can also be applied to monthly or weekly changes within a year: using a single image to represent a month or week may also lead to similar errors as illustrated in Fig. 4b.

The fundamental reason of why it is difficult to use a daily maximum to represent a year or a month comes from the fact that, in a given year, the day of maximum *U. prolifera* may be cloudy (i.e., no valid satellite observation as shown by the dashed arrow in Fig. 4b). The daily maximum obtained from a cloud-free day may be several days or weeks away from the real maximum day, leading to departure of both *U. prolifera* amount and timing of the maximum. Worse than this, such a departure may change from year to year.

One solution to overcome this difficulty is through image composition (Qi et al., 2016; Qi et al., 2022a), where the % cover (or algae density) in each location (pixel or pre-defined grid) is estimated from multiple images rather than from a single snapshot image. This way, as long as clouds move much faster than *U. prolifera* in both space and time and as long as each location (or grid) has enough valid pixels to calculate the average, such a composite would represent the mean condition for that period. This concept is illustrated in Fig. 5 in two possible scenarios. The left side of Scenario 1 represents partial cloud cover in Day X and Day Y of Fig. 4b, where the error is removed in the composite. Scenario 2 shows that the moving *U. prolifera* would not be counted twice because the composite is an average over two valid observations. For comparison, Fig. 5 also shows Scenarios 3 and 4 where the composites would give incorrect answers, yet these scenarios are unlikely because, in reality, clouds move much faster than *U. prolifera* (see Discussion below). Note that in these demonstrations, the composite images are used to represent the mean amount of *U. prolifera* between the two days.

Another possible solution to overcome the same difficulty is through modeling, where the growth rate of *U. prolifera* estimated from daily observations is used to predict the daily maximum (e.g., Yuan et al., 2022). Such predicted daily maximum, however, is subject to uncertainties in the growth rate estimates that are driven by several environmental factors.

3.3. Algorithm conceptual design and recommendations

Based on the above required steps, the following conceptual design is recommended.

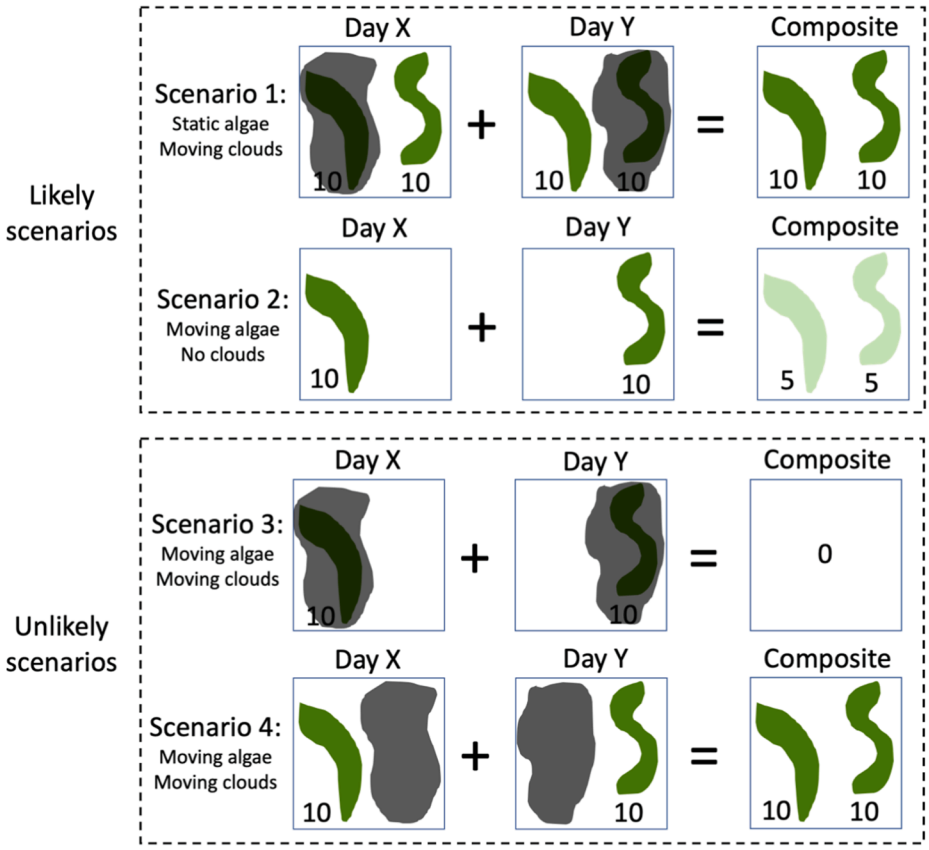


Fig. 5. Illustration of image composition in two likely and two unlikely scenarios, where each square can represent either an image or an image grid. The annotated numbers represent relative algae amount, either blocked by clouds (shaded areas) or cloud free. In Scenario 1, the moving clouds over relatively static algae is removed in the composite image. This scenario represents the reduced algae amount in Day X and Day Y of Fig. 4b, but the composite can "recover" the cloud-obscured algae. In Scenario 2, the moving algae under cloud-free conditions is not double counted in the composite because each location in the composite is an average from two observations. In Scenarios 3 and 4, both algae and clouds move at similar speeds but in opposite directions, so the composite images are incorrect in estimating the algae amount. The last two scenarios are unlikely because, in reality, clouds move much faster than algae.

1) Step 1, Image classification.

For qualitative detection of presence/absence as required by some management agencies and environmental groups, the simplest way is to visually inspect the false-color Red-Green-Blue (FRGB) images (Qi et al., 2020), where land, clouds, sun glint, and *U. prolifera* features can all be identified (e.g., Fig. 6a). Such FRGB images have already been made available in near-real-time for the study region through a Virtual Antenna System (Hu et al., 2014, https://optics.marine.usf.edu/cgi-bin/optics_data?roi=QINGDAO&current=1) using MODIS and OLCI data, and available for global waters through the OCView online portal (https://www.star.nesdis.noaa.gov/socd/mecb/color/ocview/ocview.html; Mikelsons & Wang, 2018) using VIIRS data. However, quantitative classification requires more than FRGB composite images.

Indeed, classifying clouds and other invalid pixels (in simple words, "cloud masking") sounds easy, but in practice can be very difficult because, unlike cloud masking for typical ocean color applications that requires only simple thresholding in a NIR or SWIR band (Wang & Shi, 2006), cloud masking for *U. prolifera* mapping cannot use simple thresholds, otherwise some algae pixels may be classified as clouds because these pixels also have elevated NIR and SWIR signals. While machine learning approaches may be used to classify and mask these pixels, a practical way has been described in Qi et al. (2016) and used in Hu L et al. (2019) to classify these invalid pixels. Briefly, the technique uses a combination of thresholding and spectral shapes in several blue-red-NIR-SWIR bands to mask the invalid pixels while keeping other pixels (including algae pixels and water pixels). As a result, thick clouds, strong sun glint, and extremely shallow waters are all masked as invalid pixels. The reason of not masking thin clouds is because well-designed algorithms can "see through" thin clouds (e.g., up to $R_{rc}(859) \sim 0.05$, nearly doubling the SeaDAS default threshold of 0.027) to classify algae pixels from water pixels. An example of such a cloud masking result is presented in Fig. 6b.

After masking clouds and other invalid pixels, the remaining pixels (i.e., valid pixels) need to be classified into two classes: algae pixels and algae-free (i.e., water) pixels. There are different ways for this classification, including machine learning (e.g., Qiu et al., 2018; Wan et al., 2021; Gao et al., 2022) or through the use of certain indexes such as NDVI, NDAI, FAI, AFAI, EVI, VB-FAH, or DVI (Table 1). The latter require a lower-bound threshold (U_L) to differentiate algae from water pixels, and such a threshold is not trivial to determine because it may depend on observing geometry, aerosol optical thickness, and water type (clear or turbid), and therefore may vary within an image and across different images collected even by the same sensor. In practice, it can be determined dynamically through a moving window (e.g., Garcia et al., 2013) or through the construction and subtraction of a "water background" image (e.g., Hu L et al., 2019). Some published papers simply used visual interpretation of companion medium-resolution and high-resolution images to determine the threshold manually for different locations within the same image and for different images, thus could lead to large subjectivity and uncertainties.

Regardless of the index used, in order to minimize possible errors in determining the lower-bound threshold, a reliable index should be relatively stable under different observing conditions. Through numerical simulations, image data comparison, and sensitivity analysis, Hu (2009) showed that the linear index, FAI, is more tolerant than non-linear indexes (NDVI, EVI) to perturbations by thick aerosols, thin clouds, cloud-adjacent straylight, moderate sun glint, and variable observing geometry. Therefore, FAI and its alternative form, AFAI, may serve as candidate indexes, from which lower-bound threshold is determined to classify algae and water pixels (Fig. 6b). Mathematically, FAI or AFAI is a measure of NIR reflectance against a linear baseline formed by two neighboring bands:

$$FAI(or AFAI) = R_{\lambda_2} - \left[R_{\lambda_1} + \frac{\lambda_2 - \lambda_1}{\lambda_3 - \lambda_1} \times \left(R_{\lambda_3} - R_{\lambda_1}\right)\right] \quad (1)$$

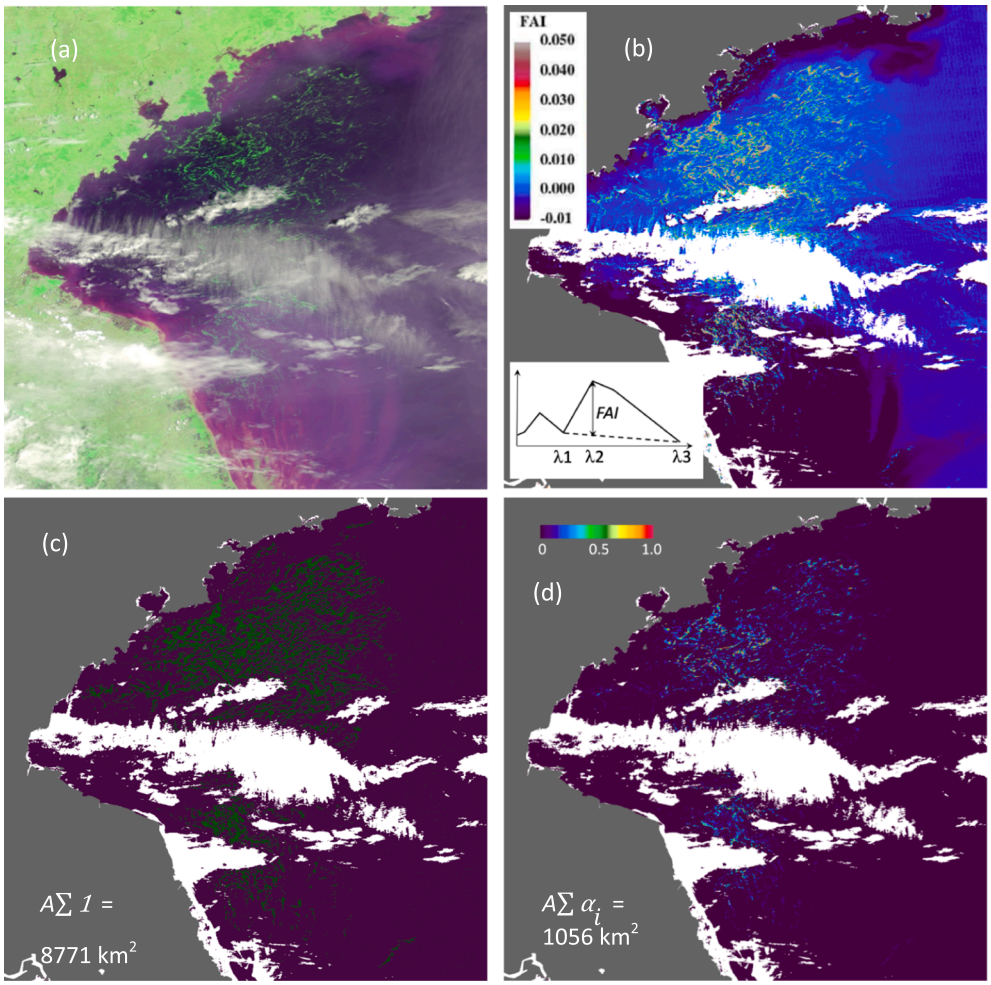


Fig. 6. Demonstration of the first two steps in mapping *U. prolifera*. (a) False-color Red-Green-Blue MODIS image on 19 June 2021 showing greenish image features, representing *U. prolifera* rafts. This type of image is sufficient for presence/absence detection as well as for delineating water area encompassing *U. prolifera*. (b) Cloud masking (white color) on the corresponding floating algae index (*FAI*) image based on the procedure of Qi et al. (2016). The inset figure illustrates the concept of FAI. (c) Classified algae pixels (green) and water pixels (purple) using the procedure described in Hu L et al. (2019). The results in (b) and (c) represent Step 1 described in this text. (d) Same as in (c) but after pixel unmixing using the procedure and threshold values described in Hu L et al. (2019). This represents Step 2 described in this text. The total *U. prolifera* coverage before and after pixel unmixing are annotated in (c) and (d), respectively, where "*A*" stands for the pixel size (in this case, 0.0625 km^2). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

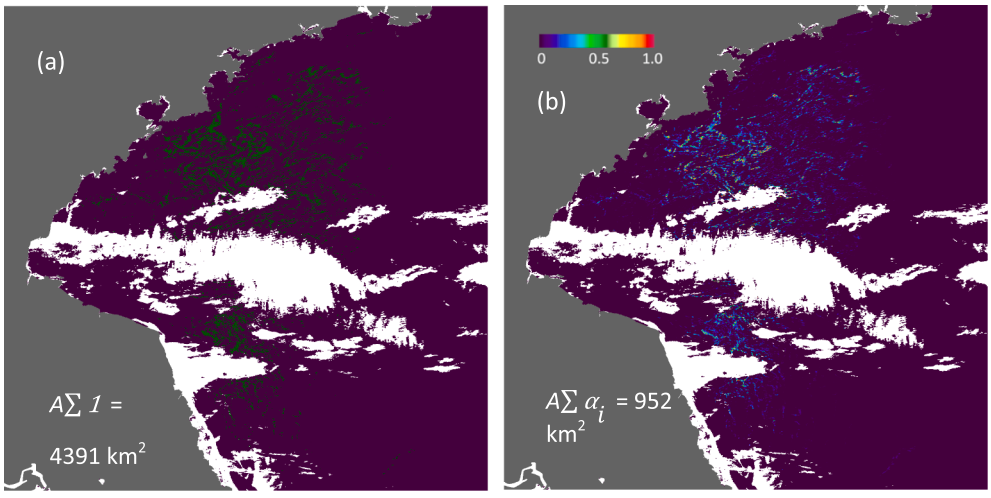


Fig. 7. Sensitivity of estimated *U. prolifera* coverage to errors in the selected lower-bound threshold before (a) and after (b) pixel unmixing. In both cases, the lower-bound threshold to differentiate algae pixels from water pixels was raised by 5% of the dynamic range of (upper-bound – lower-bound), with the original results in Fig. 6c & 6d being used as the reference (i.e., "truth"). In (a), the increased lower-bound resulted in almost a 50% reduction in the estimated *U. prolifera* coverage (from the original 8771 km^2 in Fig. 6c to the current 4391 km^2). In (b), the same increased lower-bound only resulted in a 10% reduction in the estimated coverage (from the original 1056 km^2 in Fig. 6d to 952 km^2). If the lower-bound is raised by 1%, the reduction before and after pixel unmixing is 6.1% (538 km^2) and 0.4% (4 km^2), respectively.

where the first term is the NIR reflectance at λ_2, and the second term is the linear baseline formed between λ_1 and λ_3 (see illustration in the inset of Fig. 6b). For MODIS *FAI*, λ_1, λ_2, and λ_3 are 645 nm, 859 nm, and 1240 nm, respectively. For MODIS *AFAI*, λ_1, λ_2, and λ_3 are 667 nm, 748 nm, and 869 nm, respectively. For other sensors, depending on the band availability, these indexes can be defined accordingly.

The fundamental reason of why *FAI* or *AFAI* is better than a non-linear index in terms of tolerance to variable observing conditions is because of its linear design. Many perturbation factors (e.g., aerosols, whitecaps, straylight, thin clouds, moderate sun glint) are spectrally linear or near linear, and therefore can be minimized through the linear subtraction in Eq. (1). For this reason, when the number of spectral bands is insufficient to calculate *FAI* or *AFAI*, a linear index of DVI may be used, yet its tolerance to perturbations is not as robust because it requires those perturbations to be spectrally flat instead of spectrally

linear.

After applying the lower-bound threshold to the background-referenced *FAI* image of Fig. 6b, two classes are obtained in Fig. 6c: algae pixels (green) and water pixels (purple). Knowing the pixel size of $A = 250\ \text{m} \times 250\ \text{m}$ (0.0625 km^2), the total *U. prolifera* coverage is estimated to be the number of algae pixels multiplied by A, 8,771 km^2. Note, however, that this areal estimate is simply a count of algae pixels without pixel unmixing, and therefore cannot be used objectively.

2) Step 2, Pixel unmixing.

Because the mixing of optical signals within a pixel is linear (Eq. (2)), a linear index such as *FAI* or *AFAI* can facilitate pixel unmixing using linear equations (Eq. (3)). Specifically, in each spectral band, we have

$$R_T = aR_U + (1-a)R_W, \tag{2}$$

where R_T is the reflectance of the target pixel (a mixed pixel, see illustration in Fig. 3), R_U is the reflectance of *U. prolifera* at full (i.e., 100%) coverage that represents the algae endmember, R_W is the reflectance of algae-free water that represents the water endmember, and α is the sub-pixel fraction of *U. prolifera* (in several earlier publications this is termed as χ). For simplicity, the wavelength dependence of each reflectance term is omitted here. In practice, R_U can be determined from laboratory or field measurement (e.g., Hu L et al., 2017) while R_W can be determined from individual images. As argued below, small errors in selecting R_U will propagate proportionally to the final estimates of *U. prolifera* coverage or biomass, but these small errors are not important because they would not change either the spatial distribution patterns or the temporal changing patterns of *U. prolifera*. From Eqs. (1) and (2), it is straightforward to derive

$$FAI_T = aFAI_U + (1-a)FAI_W \tag{3}$$

In contrast, the non-linear indexes such as NDVI would not lead to such a linear combination:

$$NDVI_T \neq aNDVI_U + (1-a)NDVI_W \tag{4}$$

Such a contrast between linear indexes and non-linear indexes has been demonstrated in Fig. 9 of Hu L et al. (2019). Therefore, in addition to the reasons outlined in Step 1, for the sake of pixel unmixing, it is also easier to use a linear index than a non-linear index, although the use of NDVI or other non-linear index can also achieve the same purpose with additional effort to follow a non-linear mixing line. Then, after applying the upper-bound threshold of *FAI* value (e.g., $FAI_U = 0.2$) and water endmember value (e.g., $FAI_W = 0.0$), α in each algae pixel can be derived using Eq. (3) and the pixel's FAI_T value, and this step is therefore called linear unmixing. From the originally classified image of Fig. 6c, after linear unmixing, the resulting image is shown in Fig. 6d. Integration of α from all algae pixels, after accounting for the pixel size of $A = 0.0625$ km^2, leads to a total *U. prolifera* coverage of 1,056 km^2. This is much lower than the original estimate of 8,771 km^2 but is an objective estimate that does not depend on sensor resolution (see Table 3 of Hu L et al., 2019). For this particular case, the average sub-pixel algae fraction is 1056/8771 = 0.12 (or 12% of a MODIS 250-m pixel). Furthermore, with a calibration constant of biomass per area of *U. prolifera* determined from laboratory experiment (2 kg m^{-2}, Hu L et al., 2017), the areal coverage of 1,056 km^2 can be easily converted to total wet biomass of *U. prolifera*.

As argued in Section 3.2 on why pixel unmixing is critical, Fig. 7 further demonstrates the sensitivity of the estimated total *U. prolifera* coverage to errors in the selected lower-bound threshold before and after pixel unmixing. If the lower-bound was raised by only 5% of the entire range of (upper-bound – lower-bound), without unmixing, the estimated *U. prolifera* coverage would be reduced by ~ 50% (Fig. 7a). This means that about half of the algae pixels have subpixel algae coverage of < 5%. In contrast, for the same change of the lower-bound, after unmixing, the reduction in the estimated total coverage was only 10%. Likewise, for a 1% increase in the lower-bound threshold, the reduction in the estimated total coverage before and after pixel unmixing was 6.1% and 0.4%, respectively. Clearly, pixel unmixing significantly lowered the sensitivity of *U. prolifera* estimates to errors in the selected lower-bound threshold.

3) Step 3, Image composition.

The argument of why image composition is required to have accurate representation of *U. prolifera* coverage in a given period is laid out in Section 3.2. Here, Fig. 6d also serves as an example of why image composition is important for mapping *U. prolifera*. Even if the estimated *U. prolifera* coverage in Fig. 6d were to be error free, it still could not be used to accurately represent that day, let alone representing the corresponding month or year. This is simply because a large and unknown portion of *U. prolifera* is obscured by clouds in Fig. 6d. As those clouds-covered pixels will likely be cloud free in another day, this problem can be addressed through image composition.

In the image composition, for a certain period (e.g., month), all individual images are aggregated together. Then, for each pre-defined grid, the mean areal density (i.e., fractional cover) is calculated as

$$\bar{\alpha} = \frac{1}{N_t}\sum_{i=1}^{N_t}\alpha_i \tag{5}$$

where N_t is the number of valid pixels within the grid from all images and i being the pixel number, with $\bar{\alpha}$ ranging between 0.0 and 1.0 (or 0% –100%). This is exactly the same way as in the NASA monthly "binning" practice to determine monthly mean chlorophyll-*a* concentration in a given grid, except that pixel-level Chl is replaced by pixel-level α.

As an example, Fig. 8 shows monthly $\bar{\alpha}$ distributions in each 4-km grid for the months of June 2019, July 2019, June 2021, and July 2021, respectively. Except for the extremely shallow Subei Bank, nearly all open-water areas have valid data (typically > 80 valid 1-km observations used to calculate the mean in each 4-km grid), and these represent the mean distributions of *U. prolifera* abundance during those months.

Integration of $\bar{\alpha}$ over all grids, after accounting for the grid size, yielded the total areal coverage of *U. prolifera* (km^2) in the study region. Likewise, $\bar{\alpha}$ in each grid can be converted to biomass using the predetermined calibration coefficient after accounting for the grid size, and integration of biomass over all grids would result in total biomass during the month. Note that such derived total *U. prolifera* coverage or biomass represents the mean conditions during the month, thus can be much lower than the daily maximum value.

3.4. Time series of U. prolifera amount

Using the above steps and MODIS data collected between 2008 and 2021, the derived annual mean *U. prolifera* biomass values are presented in Fig. 9, together with daily maximum *U. prolifera* coverage (after pixel unmixing) for each calendar year. Here, the annual mean was calculated from the two months of June and July only, as other months (before June and after July) showed no or small amount of *U. prolifera*. For both estimates, uncertainties are presented as vertical or horizontal bars (see Discussion below).

Although the general inter-annual patterns are similar between the two datasets (Fig. 9a), they are not exactly parallel to each other, and some significant data scatter is still observed when one is plotted against the other (Fig. 9b), suggesting inevitable uncertainties if daily maximum were to be used to represent the mean bloom situation in a year. For the same reason, it is also difficult to use daily maximum to represent the mean bloom situation for the month of June because of the data scatter found between the two datasets (Fig. 9c).

These results have significant implications when interpreting inter-

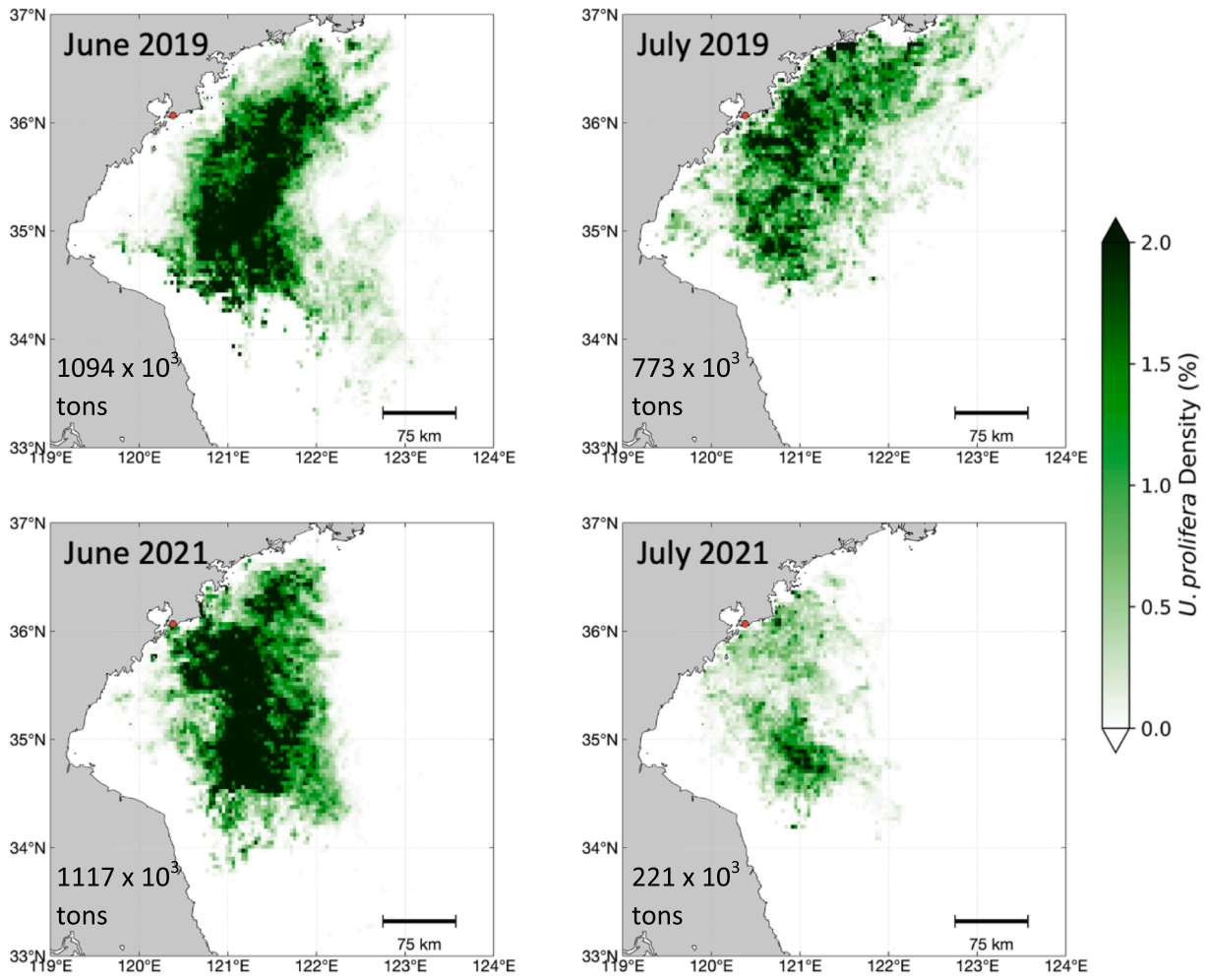


Fig. 8. Monthly mean surface coverage density (in % of each 4-km grid) of *U. prolifera* derived from MODIS/Terra and MODIS/Aqua observations in (a) June 2019, (b) July 2019, (c) June 2021, and (d) July 2021, after image composition (Step 3 of this text) of individual images after classification and pixel unmixing (e.g., Fig. 6d). The mean *U. prolifera* biomass is annotated in each panel, representing the mean coverage during the month. Figures are adapted from Qi et al. (2022a).

annual changes or determining which years had more *U. prolifera* than other years.

First, the inter-annual patterns in daily maximum *U. prolifera* coverage are dramatically different from those reported in earlier studies. For example, for the period of 2008 – 2018, several studies reported the maximum year of 2009 (Zhang Y. et al., 2019; Zhang H et al., 2019), while this study shows that the year of 2009 had much lower daily maximum than those in 2008, 2015, and 2016 for the same period. This is believed to be due to inconsistent thresholds, inconsistent indexes, or inconsistent inter-sensor calibration together with the lack of pixel unmixing used in earlier studies. Likewise, for the entire period of 2008–2021, several studies reported significantly higher (2–5 folds) daily maximum in 2021 than in 2019 (Li et al., 2022; Zheng et al., 2022), which appear to be due to lack of pixel unmixing, selection of an incorrect day to represent daily maximum, and lack of image composition. In contrast to these earlier studies, this study shows that daily maximum in 2021 is only slightly higher (3%) than in 2019 (1589 km^2 versus 1537 km^2, Fig. 9a).

Second, because the time-series of daily maximum is not exactly parallel to the time-series of annual mean, different answers may be obtained when determining which years had more *U. prolifera* than other years. An example is given by comparing 2019 and 2021. In contrast to the similar daily maxima between the two years, 2021 showed much lower annual mean than 2019 (666 versus 934 kilo tons, Fig. 9a), well beyond the estimated uncertainties. This is because although June 2021 showed similar *U. prolifera* amount as in June 2019 (Fig. 8a & 8c), July 2021 showed much lower *U. prolifera* than July 2019 (Fig. 8b & 8d) because the *U. prolifera* season ended earlier in July 2021 . Clearly, even if a daily maximum were to be the true maximal daily value during a year (this is certainly questionable due to variable clouds), it is difficult to use a daily maximum to represent the mean condition during a month or a year. For example, while the daily maximum in the four major bloom years (2015, 2016, 2019, and 2021) changed substantially, the mean biomass in June was rather stable among the four years (Fig. 8c). In this regard, although daily maximum is a simple and useful index to determine the maximum amount *observable* by satellites in any single day during a month or a year, its interpretation requires caution. In the end, if the inter-annual changing patterns were incorrect to begin with, using environmental factors such as light availability, nutrient availability, temperature, winds, and/or circulations, to explain the changing patterns would only lead to incorrect interpretations.

Finally, although not in the scope of work of this review, because it is relevant to the interpretations of inter-annual changing patterns, it is worth mentioning that all these *U. prolifera* estimates are those observed from satellite observations *after* mitigation efforts by human beings (e.g., control of release of seed populations during the harvest season of the seaweed aquaculture, physical removal at sea). Because these mitigation efforts may vary in different years for various reasons and because the exact impacts of these mitigation efforts on the observed *U. prolifera* amount are largely unknown, caution is required when using environmental factors alone to interpret annual or short-term changes.

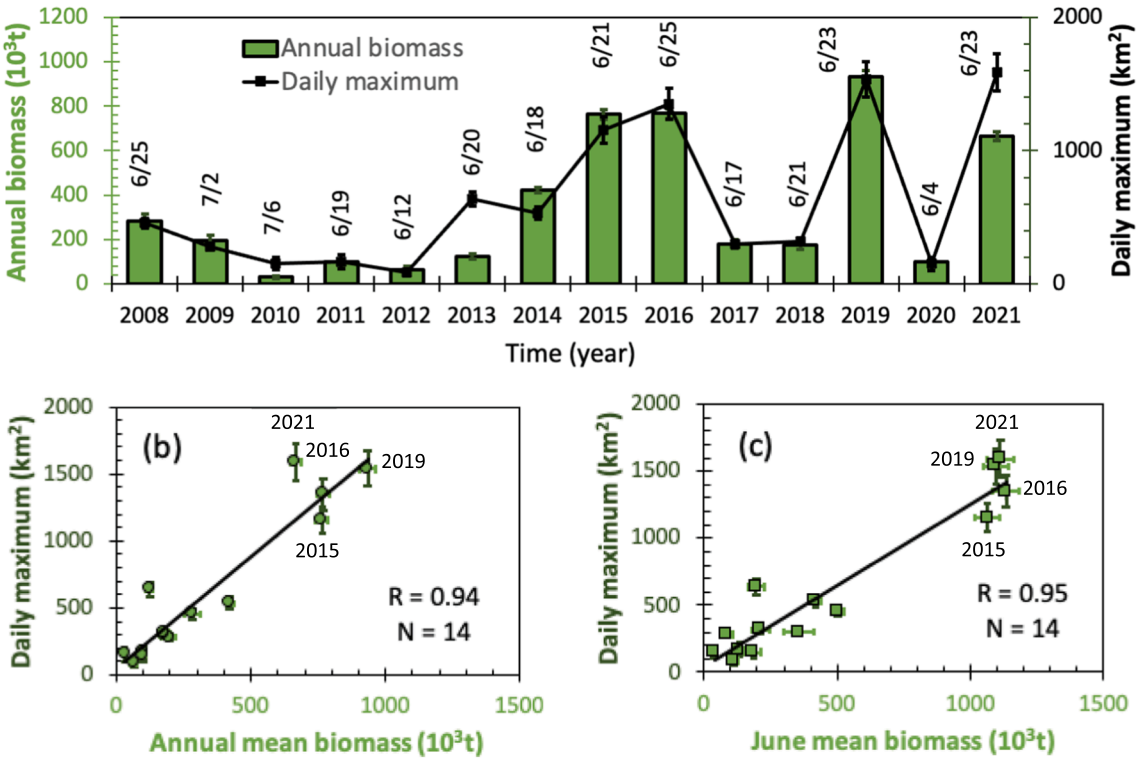


Fig. 9. (a) Time series of *U. prolifera* daily maximum coverage (2008 – 2016 from Hu L et al., 2019, with the 2017 – 2021 estimates derived in this study using the same method) and annual mean biomass (Qi et al., 2022a) between 2008 and 2021, based on MODIS estimates using the approach described here (classification, pixel unmixing, and image composition). The dates (month/day) corresponding to the daily maxima are annotated. The vertical bars represent root-mean-square uncertainties as estimated from paired Sentinel-2 and MODIS images (see text for more explanations). (b) Scatter plot between the two datasets showing high correlation but with some data scatter. (c) Same scatter plot but the mean biomass during June is used here.

4. Discussion

4.1. Validation and uncertainties

Similar to any other remote sensing data product, the *U. prolifera* maps derived from the above approach, or from any other approach, are not error free. The question is how to validate the maps and how to estimate uncertainties.

There are standard protocols to use field measurements to evaluate satellite estimates of bio-optical parameters such as surface reflectance or chlorophyll-*a* concentrations (e.g., Barnes et al., 2019) and there are also proposed protocols to estimate uncertainties in the derived ocean color data products (IOCCG, 2019). However, because of the strong heterogeneity in *U. prolifera* distributions and because every pixel is only partially covered by *U. prolifera* (Fig. 2), such protocols cannot be followed directly. Likewise, it is very difficult to collect *all U. prolifera* within a large region and compare to co-located image pixels. These difficulties are not limited to *U. prolifera* but are fundamental to all floating matters including oil slicks.

Currently, one way to evaluate the *U. prolifera* maps is through the use of high-resolution images, such as those from the Landsat sensors (30 m), Sentinel-2 sensors (10 m), Worldview sensors (2 m), or even airborne sensors (centimeters). The assumption is that these sensors, after proper pixel unmixing, can provide independent estimates of *U. prolifera* to evaluate the MODIS-based estimates. While some limited case studies can be found in Hu L et al. (2019) and Lu et al. (2019), Qi et al., (in prep) recently used all available Sentinel-2 data between 2016 and 2021 to perform a systematic evaluation of MODIS-derived *U. prolifera* amounts. From all concurrent and co-located image pairs, the root-mean-square difference (RMSD in %) between the MODIS estimates and the fitting line was used to represent uncertainties in the MODIS estimates. Such estimated RMSD for individual images decreased with increasing *U. prolifera* amount, with RMSD of ~9% for *U. prolifera* coverage of >200 km^2. For monthly or annual composites, the RMSD further decreased because the number of valid daily observations in each 1-km grid is typically >5 within a month (or >80 in each 4-km grid). Such estimated uncertainties in both daily estimates and annual (or monthly) means are presented in Fig. 9.

For practical reasons, because different sensors and different indexes or approaches may be used for different purposes, based on the laboratory and field-measured reflectance and biomass density (Hu L et al., 2017), we provide the upper-bound thresholds of *U. prolifera* in Table 2a. These represent index values corresponding to the assumed full coverage (i.e., $\alpha = 1.0$ or 100%, or 2 kg m^{-2} biomass density) of *U. prolifera* within a pixel, or the FAI_U endmember value in Eq. (3). The lower-bound thresholds that represent the α value approaching 0.0 (i.e., the FAI_W endmember value in Eq. (3)), however, depend on water type, as shown in Qi and Hu (2021). Nevertheless, for relatively clear waters (most open waters in the YS), the lower-bound thresholds are provided in Table 2b. For moderately turbid waters, the lower-bound thresholds are lower by a certain percentage of the full range (upper-bound – lower-bound) (Table 2c). These relative changes also suggest that among the various indexes, the lower-bound thresholds of *FAI* and *AFAI* are the least sensitive to changes in water type.

All the above arguments as well as the associated uncertainty estimates are for self-consistency evaluations, i.e., to assure that the relative changing patterns from one time to another or from one location to another are correct. The absolute values in either *U. prolifera* coverage or biomass may be off because they depend on the selections of both the upper-bound and lower-bound thresholds and depend on the calibration constant to convert coverage to biomass. For example, due to lack of *in situ* data to determine the thresholds, Hu et al. (2010) used a much lower upper-bound threshold than shown in Table 2a, resulting in much higher *U. prolifera* coverage than reported here. However, as long as the

Table 2

a) Upper-bound threshold values of *U. prolifera* derived from field measured reflectance corresponding to 2 kg wet biomass m^{-2} ($\alpha = 1.0$ or 100% within a pixel.) For FAI, the values represent FAI_U in Eq. (3).

Upper threshold	FAI	AFAI	NDVI	EVI	DVI	VB_FAH
MODIS	0.192	0.154	0.711	0.332	0.188	0.159
MERIS/OLCI	N/A	0.143	0.809	0.384	0.201	0.168
VIIRS	0.189	0.159	0.833	0.393	0.204	0.170
GOCI	N/A	0.149	0.785	0.372	0.197	0.165
Landsat 4 5 7	0.217	N/A	0.798	0.392	0.213	0.182
Landsat 8 9	0.201	N/A	0.772	0.360	0.195	0.162
MSI	0.206	0.160	0.809	0.386	0.201	0.167
HJ	N/A	N/A	0.798	0.391	0.213	0.182
GF	N/A	N/A	0.798	0.392	0.213	0.181
HY/CZI	N/A	N/A	0.809	0.384	0.201	0.168
DOVE	N/A	N/A	0.809	0.384	0.201	0.169

b) Lower-bound threshold values of *U. prolifera* derived from field measured reflectance corresponding to 0 kg wet biomass m^{-2} ($\alpha = 0.0$ or 0% within a pixel). For FAI, the values represent FAI_W in Eq. (3). These were derived from the field measured water reflectance of relatively clear waters in the YS (i.e., "clear" water in Qi and Hu, 2021) and they may change slightly for more turbid waters (see (c)). They are provided as reference only, as their values will depend on the water and imaging conditions and therefore need to be adjusted.

Lower threshold (reference only)	FAI	AFAI	NDVI	EVI	DVI	VB_FAH
MODIS	−0.002	−0.002	−0.688	−0.013	−0.004	−0.011
MERIS/OLCI	N/A	−0.002	−0.547	−0.009	−0.003	−0.010
VIIRS	−0.003	−0.002	−0.527	−0.008	−0.003	−0.010
GOCI	N/A	−0.0002	−0.574	−0.010	−0.003	−0.010
Landsat 4 5 7	−0.002	N/A	−0.466	−0.009	−0.003	−0.009
Landsat 8 9	−0.003	N/A	−0.632	−0.011	−0.004	−0.011
MSI	−0.002	−0.002	−0.547	−0.009	−0.003	−0.010
HJ	N/A	N/A	−0.466	−0.009	−0.003	−0.009
GF	N/A	N/A	−0.466	−0.009	−0.003	−0.010
HY/CZI	N/A	N/A	−0.547	−0.009	−0.003	−0.010
DOVE	N/A	N/A	−0.547	−0.009	−0.003	−0.010

c) Relative difference between the new lower-bound threshold values determined from moderately turbid waters (see Qi and Hu, 2021 for reflectance spectra) and those determined from clear waters in (b), referenced against the full range of upper-bound minus lower-bound. For FAI, these represent ΔFAI_W referenced against ($FAI_U - FAI_W$) in (a) and (b). Clearly, the lower-bound thresholds for FAI and AFAI show the lowest sensitivity to water type changes (i.e., values in the FAI and AFAI columns are closer to 0% than in other columns).

Relative change	FAI	AFAI	NDVI	EVI	DVI	VB_FAH
MODIS	−10%	−9%	−14%	−25%	−17%	−21%
MERIS/OLCI	N/A	−9%	−24%	−21%	−14%	−19%
VIIRS	−13%	−8%	−25%	−20%	−13%	−18%
GOCI	N/A	−9%	−22%	−22%	−15%	−19%
Landsat 4 5 7	−10%	N/A	−25%	−20%	−13%	−17%
Landsat 8 9	−12%	N/A	−18%	−23%	−16%	−21%
MSI	−10%	−9%	−24%	−21%	−14%	−19%
HJ	N/A	N/A	−25%	−19%	−13%	−17%
GF	N/A	N/A	−25%	−20%	−13%	−17%
HY/CZI	N/A	N/A	−24%	−21%	−14%	−19%
DOVE	N/A	N/A	−24%	−21%	−14%	−19%

potential bias in the absolute values, if any, is systematic, the spatial patterns and relative temporal changes are still valid. For example, using a different index and different thresholds, Yuan et al. (2022) showed much lower *U. prolifera* coverage in all years than that shown in Fig. 9a, yet their temporal changing patterns (e.g., 2019 versus 2021) are rather similar. Indeed, it is the relative temporal changes that have been used in the attempt to use environmental factors to explain such changes (e.g., Qi et al., 2022a). In this regard, the absolute values in the upper-bound thresholds are less important than using a time-independent consistent value as long as the upper-bound thresholds are not overly low (otherwise there may be too many "saturated" pixels).

Finally, all the above arguments are based on the assumptions of linear mixing (Eqs. (2) & (3)) as a result of horizontal mixing between *U. prolifera* and water. Above a certain threshold (see Fig. 6 of Hu L et al., 2017), the mixing can be non-linear as *U. prolifera* starts to aggregate in the vertical direction (i.e., becoming thicker), thus making the assumptions invalid and causing uncertainties in the estimated *U. prolifera* amount. While in theory this is certainly a problem, in practice the uncertainties caused by non-linear mixing are expected to be small. The reason is that in an open ocean without boundary, floating materials (including *U. prolifera*) tend to dissipate more horizontally than vertically. Only in some special cases, such as nearshore waters where land serves as a boundary, convergence ocean fronts, or in the middle of large *U. prolifera* mats, may *U. proliferra* aggregate vertically to cause underestimates in the *U. prolifera* amount. More importantly, such underestimates are expected to be systematic rather than random for inter-annual comparisons, thus are unlikely to change the long-term *U. prolifera* patterns. Likewise, strong winds may dissipate *U. prolifera* vertically, causing underestimates in the *U. prolifera* amount (Fig. 5 of Qi et al., 2016). These perturbation factors are also expected to be systematic in different years, thus would not change the *U. prolifera* patterns.

4.2. Daily maximum or image composition

Depending on the needs, both daily maximum and image composition may be used to describe the *U. prolifer* blooms, as they have different meanings. An illustration of their differences as well as their implications are presented and discussed in Fig. 9 above. While daily maximum shows the maximal *observable* algae amount at a specific time during a year, an image composite represents the mean algae amount during a period, with the latter amount often being much lower than the former. In practice, daily maximum is a simple and straightforward way to indicate the severity of the *U. prolifer* bloom in a certain year, therefore has often been used by the management agencies to help implement mitigation strategy. However, due to frequent cloud cover, the image used to estimate the daily maximum may be captured at a time days or weeks apart from the time of the true maximum (Fig. 4b), and such a departure may vary from year to year. For example, the dates corresponding to daily maximum varied between June 4 (in 2020) to July 6 (in 2010) (Fig. 9a). Therefore, interpretation of daily maximum requires extra caution, especially when multi-year sequence is used to interpret the environmental forcing factors or environmental impacts.

The use of image composition is expected to minimize the impacts of cloud cover, and therefore can provide more objective estimates to facilitate time-series studies of long-term changes and to investigate reasons behind such changes. Indeed, such a method to fill data gaps is not new but has been used for decades in generating global maps of chlorophyll-*a* concentration and sea surface temperature. In the SeaDAS software, this is called "binning", where all valid pixels within a certain period (week, month, season) for a given grid (either 4 km or 9 km) were pulled from individual images and used to calculate the average. Image composition is different from a simple average of all daily *U. prolifera* coverage in the same period where the former requires averaging pixels in each grid after accounting for non-valid pixels and sub-pixel coverage.

Of course, the use of image composites may still be problematic *if* clouds are static in time or *if* clouds follow the movement of *U. prolifera*. For the hypothetical case shown in Fig. 4b, if clouds were in the same locations during the partially cloudy days (Day X and Day Y), an image composite would still underestimate the total *U. prolifera* coverage in that period, thus facing the same difficulty as with daily maximum. Likewise, if clouds follow the movement of *U. prolifera* in either the same or the opposite directions (Scenarios 3 and 4 in Fig. 5), the image composites can either underestimate (Scenario 3) or overestimate (Scenario 4) the *U. prolifera* amount. In practice, Scenarios 3 and 4 are unlikely because most clouds move much faster than *U. prolifera* in both space and time (Fig. 5, Scenario 1) and thus can be removed in the image composites. Indeed, typical wind speeds (which drive the movement of clouds) are >2 m s^{-1}, much higher than typical surface current speeds

(which drive the drift speed of *U. prolifera*) of <20–30 cm s^{-1} in the Yellow Sea during June and July 2008 (Qiao et al., 2011). Caution is still required, however, to verify whether a certain portion of the composite period (e.g., the first week in a month) has complete and persistent cloud cover. In such cases, the image composite is biased towards the remaining portion of the composite period. Inspection of the daily sequence of MODIS images suggested that such a scenario is unlikely, as clouds are relatively random in space and time (Qi et al., in prep). In a typical monthly composite, at least 80 valid 1-km MODIS pixels are used in calculating the average *U. prolifera* amount in any given 4-km grid, and such valid pixels are rather random in time.

The use of image composites can also reduce relative uncertainties in the estimated *U. prolifera* amount because many observations are used to calculate the mean. This is why the uncertainties in the annual means or monthly means of *U. prolifera* amounts are lower than those in the daily maxima in Fig. 9.

4.3. Which sensor to use

In the published literature (Table 1), many medium-resolution and high-resolution sensors have been used to map *U. prolifera*, as no single sensor can fit all purposes. For long-term, synoptic patterns, medium-resolution sensors with frequent revisits (e.g., MODIS, VIIRS, OLCI, GOCI) are appropriate. Of these, MODIS sensors provide the longest temporal coverage (2000 – present), thus being the often-used sensors. For more detailed views in a local region, especially during the bloom initiation phase where *U. prolifera* mats are rather small, high-resolution sensors (e.g., OLI, MSI, CZI, GaoFen) are appropriate. In any case, pixel unmixing is required in order to obtain accurate estimates of *U. prolifera* amount. This is true even for the high-resolution sensors because many algae-containing pixels still have partial algae cover.

Because many sensors are currently available, one question is whether they can be combined to remove cloud cover to improve image composition or estimates of daily maximum. In theory this is certainly possible, but caution is required to assure cross-sensor consistency, especially when combining high-resolution and medium-resolution observations. A "calibration" equation obtained from just several pairs of images may not be applicable under all observing scenarios, especially when considering the variable amounts of aerosols and sun glint. On the other hand, SAR sensors provide all-weather observations and thus can complement optical sensors when most images are obscured by clouds. While several case studies demonstrate the feasibility of using SAR to detect *U. prolifera* patches (Cui et al., 2012; Gao et al., 2022; Geng et al., 2020; Shen et al., 2014; Yu et al., 2021), a comprehensive evaluation of Sentinel-1 SAR measurements indicate that the SAR images do not appear to be able to capture small *Ulva* patches (Qi et al., 2022b). Merging SAR and optical observations for quantitative use still requires more work.

4.4. Coverage, biomass, or water area

In previous studies, *U. prolifera* blooms (green tides) have been characterized in three ways: coverage (in km^2), biomass (in kilo tons), and water area (in km^2) encompassing *U. prolifera*, each being used for its own purpose.

Both coverage and biomass are used to characterize the *U. prolifera* amount, and in this regard, they may be used interchangeably although the use of biomass is more applicable in terms of mitigation effort (i.e., the physical removal of *U. prolifera* is measured in kilo tons rather than in area). Indeed, after pixel unmixing, the former can be easily converted to the latter using a calibration constant of biomass per area for *full* (100%) algae coverage within a pixel, often determined from laboratory or field measurements (e.g., Hu L et al., 2017; Xiao et al., 2019). Any errors in the calibration constant, presumably small, would lead to the same errors in the estimated biomass. However, as long as these small errors are systematic in both space and time, they would have no or negligible impacts on time-series studies. In contrast, large uncertainties may be encountered if the calibration constant is no longer a constant but changes with *U. prolifera* thickness, which will violate all linear-mixing assumptions used in Eqs. (2) & (3).

The use of water area encompassing *U. prolifera* is another useful way to describe the size of *U. prolifera* blooms. This is because the water area, and especially through its distribution, is an effective measure of where *U. prolifera* can be found and where biological and ecological consequences may be quantified. In practice, the water area is often delineated through the outer boundary of where *U. prolifera* is found in satellite imagery, using either manual drawing or more objective ways to define the polygon. Depending on how *U. prolifera is* distributed in space and what is the lowest detection limit of the selected satellite sensors, there can be different answers, leading to large degrees of uncertainties. Nevertheless, even though this parameter is less objective than *U. prolifera* coverage or biomass, it is still a useful parameter to describe the spatial extent of the *U. prolifera* bloom.

4.5. Frequently asked questions

Since the first report of the *U. prolifera* bloom in 2008, blooms occurred every year in almost the same period of May – July in the YS, raising many questions on the bloom origin, bloom size, driving factors, long-term patterns, impacts on the ocean and coastal environments, among others. While it is impossible to provide answers to all these questions, for the nature of this revisit on remote sensing methodology, we attempt to address questions that are frequently asked by both researchers and the general public, with specific focus on remote sensing methodology and results. Below we list such frequently asked questions and our answers.

Q: Some small *Ulva* patches may be missed in remote sensing images, so what is the lower detection limit?

A: This depends on the sensor's SNRs and pixel size. Higher SNRs lead to lower detection limits. For MODIS 250-m resolution bands, the lower detection limit is about 1% of a pixel size, or 625 m^2. For MODIS 1-km resolution bands, the lower detection limit is about 0.2% of a pixel size, or 2000 m^2. The lower detection limit of OLCI and MSI is 0.5% and 2% of their pixel size, corresponding to 450 m^2 and 2 m^2, respectively. Note that the lower detection limit is different from the lower-bound threshold.

Q: What atmospheric correction is required? Can top-of-atmosphere data be used?

A: Pixel-wise atmospheric correction over *U. prolifera* pixels often fails because the enhanced NIR and SWIR reflectance of *U. prolifera* makes the atmospheric correction assumptions invalid. Non-pixel-wise atmospheric correction based on dark objects in an image is a good choice. The analysis in this work is based on a partial atmospheric correction (Rayleigh correction) that only requires the observing geometry, surface pressure, and wind speed – none of these depends on the pixel's NIR or SWIR signals. The reason to perform a Rayleigh correction is to remove most of the non-linear reflectance shape before a linear index is applied. However, even without any correction, TOA signals in either radiance, reflectance, or even digital numbers can still be used, but extra caution is required to assure cross-image consistency and cross-sensor consistency.

Q: *Sargassum horneri* macroalgae can often be found in the East China Sea but occasionally it can also be found in the Yellow Sea, so how do we know what the satellite detected are really *Ulva*?

A: *Sargassum* and *Ulva* have different spectral shapes: one is brown rich and the other is green rich although both have enhanced NIR and SWIR reflectance. The difference in their spectral shapes can be used to differentiate between the two as long as their subpixel proportions are above a threshold – typically a few percent of a pixel. Below this limit, other information can be used to separate the two, for example, large amount of *Sargassum* is mostly found in February – May with June being a tailing month, but large amount of *Ulva* is mostly found in May – July

with August being a tailing month. When large amounts of both are found at the same time, in satellite imagery they rarely overlap in space. One exception is on the Subei Bank where both *Sargassum* and *Ulva* have been found on the aquaculture rafts and nearby waters, but these amounts are rather small. Some mixture of *Ulva* and *Sargassum* in the open Yellow Sea has been reported from boat observations, but the mixture is typically dominated by one and the amounts are also relatively small. In the future, improved algorithms are required to determine the proportions of *Ulva* and *Sargassum* if they coexist in a pixel.

Q: Can I estimate total *Ulva* amount from every image in a month, and calculate their average to represent that month?

A: No. This is because most images are obscured partially or entirely by clouds, and the *Ulva* amount estimated from a daily image may not be an accurate representation for that day. Such a calculated average is also sensitive to the number of images used in the calculation. Image composition is the recommended way.

Q: Without pixel unmixing, the *Ulva* area is overestimated. But if it is always overestimated in the same way, perhaps the spatial and temporal patterns are still correct?

A: While the spatial patterns may still be correct, the temporal patterns may be not. This is because the overestimates may be different from day to day because of the morphological changes in *U. prolifera* due to winds, circulation, ocean fronts, among other factors. For example, suppose *U. prolifera* area is 100 km^2 in Day 1 and 120 km^2 in Day 2 if unmixing is performed. Without unmixing, the estimates may be 1000 km^2 (a factor of 10) and 960 km^2 (a factor of 8) for the two days. In this case, the increasing pattern from Day 1 to Day 2 changes to a decreasing pattern. Using multiple sensors without pixel unmixing can make this effect much worse.

Q: Can *Ulva* thickness be estimated from remote sensing images?

A: It is difficult because the reflectance signal is a non-linear function of thickness, and, once above a certain thickness threshold the signal is saturated (i.e., no longer a function of thickness). To date, there has been no attempt to measure thickness in the field or in the laboratory to relate to remote sensing signals, although some of these measurements may implicitly contain thickness information to a certain degree. In the future, perhaps high-resolution (a few meters or even sub-meter) sensors may be used together with field measurements to estimate thickness.

Q: To classify *Ulva* pixels, which way is preferred, using an index or through machine learning?

A: They each have their own pros and cons. Index based classification has a physical meaning and therefore is easy to understand, but the selection of the lower-bound threshold requires caution. Machine learning may get around this problem once properly trained using large quantity of data, but if the training does not include all potential scenarios some unexpected errors may occur. After classification, pixel unmixing will require an index or a reference band in both ways.

Q: To date, which year has the largest green tide?

A: This depends on how to define the "size" of a green tide and how each year is described, after considering uncertainties in the size estimates. If a daily maximum is used to represent the annual condition, then 2021 may have the largest green tide because the daily maximum in 2021 is higher than in 2019, but the small difference is overwhelmed by uncertainties in each daily estimate. If annual mean biomass is used to represent the mean annual condition, then 2019 has the largest green tide, as the difference between 2019 and 2021 exceeds the uncertainties in the annual mean estimates. Such a mean value implicitly includes the information on the duration of the green tide in each year.

Q: Some studies used the water area encompassing *Ulva* as an index for bloom severity. Does that make sense?

A: Yes, and that is another useful way to describe the green tide condition, although it tends to be more subjective than using the *Ulva* coverage or biomass. Extra caution is required to minimize the subjectivity due to pixel resolution and due to the variable methods to draw boundaries around the *Ulva* pixels. The distribution and the water area encompassing *Ulva* provide important information for impact assessments.

Q: Can remote sensing be used to determine the starting and ending days of *Ulva* blooms?

A: It is possible but difficult. The reason is that if some initial *Ulva* mats form in Day 1, due to clouds and other factors, satellites may not be able to detect it until Day 7. In this case, an error of 6 days is inevitable if Day 7 is regarded as the bloom starting day. Same can be said for the bloom ending day. Using fine-resolution images may help because there may be small cloud-free "windows" around clouds. In any case, extra caution is required when using satellite data to infer starting and ending days of *Ulva* blooms.

Q: Can remote sensing be used to estimate daily growth rate of *Ulva*?

A: It is possible when cloud-free images are available in several consecutive days during the growing phase. Pixel unmixing is required to estimate *Ulva* coverage or biomass in each day using a consistent method in order to quantify the relative temporal changes.

Q: According to the satellite observations, what is the history of green tides in the Yellow Sea?

A: High-resolution satellite data going back to the 1980s showed the first appearance of *Ulva* in the Yellow Sea in 1999, but large green tides did not start until 2007, one year before 2008 when green tides in this region became well known. After 2008, the annual patterns fluctuated without statistically significant trends, possibly due to variable mitigation efforts in different years.

5. Concluding remarks

Multi-sensor satellite remote sensing has been used extensively to map *Ulva prolifera* macroalgae blooms (green tides) in the Yellow Sea, yet dramatic differences have been found in the reported *U. prolifera* areal coverage (or biomass) from individual images and in the reported inter-annual patterns from multiple images even from the same data source, a result of different methodology and sensor selection. Based on physical principles, this review presented a conceptual diagram on the algorithm design for various purposes. Specifically:

1) To examine spatial distribution patterns of *U. prolifera* presence/absence and to determine the water areas encompassing *U. prolifera*, color imagery after proper color stretching is sufficient for visual inspection and approximate delineation of the green tide boundary.
2) To quantify *U. prolifera* amount beyond simple presence/absence detection from individual images, pixel unmixing is a critical step because many or most image pixels are only partially covered by *U. prolifera*, and pixel unmixing is the most feasible way to obtain objective results that are independent of sensor resolution and *U. prolifera* morphology.
3) To understand monthly or annual patterns and their potential environmental drivers, image composition is recommended, while some alternative approaches (e.g., applying a growth model to predict the maximum coverage) may also be attempted. This is because variable cloud cover and other factors make it difficult to use a single day to represent a month or a year, but the much faster changing of clouds (than *U. prolifera*) makes it possible and meaningful to compose multiple images into a single image to estimate the mean bloom size.
4) The daily maximum estimates from individual years, on the other hand, provide useful information to help mitigation effort in a practical sense although their interpretation requires caution due to the above effect.
5) For these requirements, linear indexes are preferred over non-linear indexes when determining which algorithms to use.
6) To study short-term and long-term changes including annual fluctuation patterns, it is important to assure algorithm/product consistency, while uncertainty estimates in the data products are always desired.

Finally, the purpose of this review is to demonstrate the possible

problems in mapping *U. prolifera* green tides and to propose a conceptual diagram towards addressing these problems. The use of MODIS and *FAI* to show the individual steps in the diagram is for demonstration only, and by no means should they be interpreted as the *only* data or method to address the problems. In contrast, once the conceptual diagram is followed, either explicitly or implicitly, different data sources, indexes, or mapping methods can still be used.

In this regard, with more and more satellite sensors being available and new satellite missions being planned, we hope these general guidelines can be useful for the remote sensing community when developing new algorithms or approaches in order to reduce data product discrepancy and to generate robust time-series data to facilitate both interdisciplinary research and implementation of mitigation strategy.

CRediT authorship contribution statement

Chuanmin Hu: Conceptualization, Funding acquisition, Methodology, Data curation, Formal analysis, Investigation, Project administration, Resources, Software, Visualization, Writing - original draft, Writing - review & editing. **Lin Qi:** Conceptualization, Methodology, Data curation, Formal analysis, Funding acquisition, Investigation, Validation, Visualization, Writing - original draft, Writing - review & editing. **Lianbo Hu:** Methodology, Data curation, Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Tingwei Cui:** Methodology, Data curation, Funding acquisition, Visualization, Writing - original draft, Writing - review & editing. **Qianguo Xing:** Methodology, Funding acquisition, Writing - original draft, Writing - review & editing. **Mingxia He:** Writing - original draft, Writing - review & editing. **Ning Wang:** Writing - original draft, Writing - review & editing. **Yanfang Xiao:** Writing - original draft, Writing - review & editing. **Deyong Sun:** Funding acquisition, Writing - original draft, Writing - review & editing. **Yingcheng Lu:** Funding acquisition, Writing - original draft, Writing - review & editing. **Chao Yuan:** Writing - original draft, Writing - review & editing. **Mengquan Wu:** Writing - original draft, Writing - review & editing. **Changying Wang:** Writing - original draft, Writing - review & editing. **Yanlong Chen:** Writing - original draft, Writing - review & editing. **Haipeng Xu:** Writing - original draft, Writing - review & editing. **Li'e Sun:** Writing - original draft, Writing - review & editing. **Maohua Guo:** Writing – original draft, Writing – review & editing. **Menghua Wang:** Supervision, Funding acquisition, Writing - review & 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.

Acknowledgements

This work was in part supported by the JPSS/NOAA ocean color cal/val program (ST13301CQ0050/1332KP22FNEED004), the National Natural Science Foundation of China (42076180, 42176179, 42076188), the Fundamental Research Funds for the Central Universities (22lglj01), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA19090140), among others. We thank the U.S. NASA for providing satellite data, and thank the two anonymous reviewers for providing detailed and constructive comments. 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.

References

An, D., Yu, D., Zheng, X., Zhou, Y., Meng, L., Xing, Q., 2021. Monitoring the Dissipation of the Floating Green Macroalgae Blooms in the Yellow Sea (2007–2020) on the Basis of Satellite Remote Sensing. Remote Sens. (Basel) 13 (19), 3811. https://www.mdpi.com/2072-4292/13/19/3811.

Barnes, B.B., Cannizzaro, J.P., English, D., Hu, C., 2019. Validation of VIIRS and MODIS reflectance data in coastal and oceanic waters: An assessment of methods. Remote Sens. Environ. 220, 110–123.

Cao, Y., Wu, Y., Fang, Z., Cui, X., Liang, J., Song, X., 2019. Spatiotemporal Patterns and Morphological Characteristics of Ulva prolifera Distribution in the Yellow Sea, China in 2016–2018. Remote Sens. (Basel) 11 (4), 445. https://www.mdpi.com/2072-4292/11/4/445.

Cui, T., Li, F., Wei, Y., Yang, X., Xiao, Y., Chen, X., Liu, R., Ma, Y., Zhang, J., 2020. Super-resolution optical mapping of floating macroalgae from geostationary orbit. Appl. Opt. 59 (10), C70–C77. https://doi.org/10.1364/AO.382081.

Cui, T., Liang, X., Gong, J., Tong, C., Xiao, Y., Liu, R., Zhang, X., Zhang, J., 2018. Assessing and refining the satellite-derived massive green macro-algal coverage in the Yellow Sea with high resolution images. ISPRS J. Photogramm. Remote Sens. 144, 315–324.

Cui, T., Zhang, J., Sun, L., Jia, Y., Zhao, W., Wang, Z., Meng, J., 2012. Satellite monitoring of massive green macroalgae bloom (GMB): imaging ability comparison of multi-source data and drifting velocity estimation. Int. J. Remote Sens. 33 (17), 5513–5527.

Gao, L., Li, X., Kong, F., Yu, R., Guo, Y., Ren, Y., 2022. AlgaeNet: A Deep-Learning Framework to Detect Floating Green Algae From Optical and SAR Imagery. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 2782–2796. https://doi.org/10.1109/JSTARS.2022.3162387.

Geng, X., Li, P., Yang, J., Shi, L., Li, X.-m., Zhao, J., 2020. Ulva prolifera detection with dual-polarization GF-3 SAR data. IOP Conference Series: Earth and Environmental Science, 502(1), 01202, doi: 10.1088/1755-1315/502/1/012026.

Garcia, R.A., Fearns, P., Keesing, J.K., Liu, D., 2013. Quantification of floating macroalgae blooms using the scaled algae index. J. Geophys. Res. Oceans 118 (1), 26–42. https://doi.org/10.1029/2012JC008292.

Harun-Al-Rashid, A., Yang, C.-S., 2018. Improved Detection of Tiny Macroalgae Patches in Korea Bay and Gyeonggi Bay by Modification of Floating Algae Index. Remote Sens. (Basel) 10 (9), 1478.

He, M.-X., Liu, J., Yu, F., Li, D., Hu, C., 2011. Monitoring green tides in Chinese marginal seas. In: Morales, J., Stuart, V., Platt, T., Sathyendranath, S. (Eds.), Handbook of Satellite Remote Sensing Image Interpretation: Applications for Marine Living Resources Conservation and Management. EU PRESPO and IOCCG, Dartmouth, Canada, pp. 111–124.

Hu, C., 2009. A novel ocean color index to detect floating algae in the global oceans. Remote Sens. Environ. 113 (10), 2118–2129.

Hu, C., 2021. Remote detection of marine debris using satellite observations in the visible and near infrared spectral range: Challenges and potentials. Remote Sens. Environ. 259, 112414 https://doi.org/10.1016/j.rse.2021.112414.

Hu, C., Barnes, B. B., Murch, B., Carlson, P., 2014. Satellite-based virtual buoy system to monitor coastal water quality. Optical Engineering, 53(5), 051402-051402.

Hu, C., He, M.-X., 2008. Origin and offshore extent of floating algae in Olympic sailing area. Eos 89 (33), 302–303.

Hu, C., Li, D., Chen, C., Ge, J., Muller-Karger, F.E., Liu, J., Yu, F., He, M.-X., 2010. On the recurrent Ulva prolifera blooms in the Yellow Sea and East China Sea. J. Geophys. Res.: Oceans (1978–2012) 115 (C5), C05017.

Hu, L., Hu, C., He, M.-X., 2017. Remote estimation of biomass of Ulva prolifera macroalgae in the Yellow Sea. Remote Sens. Environ. 192, 217–227. https://doi.org/10.1016/j.rse.2017.01.037.

Hu, L., Zeng, K., Hu, C., He, M.-X., 2019. On the remote estimation of Ulva prolifera areal coverage and biomass. Remote Sens. Environ. 223, 194–207. https://doi.org/10.1016/j.rse.2019.01.014.

Hu, S., Yang, H., Zhang, J., Chen, C., He, P., 2014b. Small-scale early aggregation of green tide macroalgae observed on the Subei Bank, Yellow Sea. Marine Pollution Bulletin 81 (1), 166–173.

Huo, Y., Zhang, J., Chen, L., Hu, M., Yu, K., Chen, Q., He, Q., He, P., 2013. Green algae blooms caused by Ulva prolifera in the southern Yellow Sea: identification of the original bloom location and evaluation of biological processes occurring during the early northward floating period. Limnol. Oceanogr. 58 (6), 2206–2218.

IOCCG, 2019. Uncertainties in Ocean Colour Remote Sensing. Mélin F. (ed.), IOCCG Report Series, No. 18, International Ocean Colour Coordinating Group, Dartmouth, Canada. https://doi.org/10.25607/OBP-696.

Jin, S., Liu, Y., Sun, C., Wei, X., Li, H., Han, Z., 2018. A study of the environmental factors influencing the growth phases of Ulva prolifera in the southern Yellow Sea, China. Mar. Pollut. Bull. 135, 1016–1025. https://doi.org/10.1016/j.marpolbul.2018.08.035.

Keesing, J.K., Liu, D., Fearns, P., Garcia, R., 2011. Inter-and intra-annual patterns of Ulva prolifera green tides in the Yellow Sea during 2007–2009, their origin and relationship to the expansion of coastal seaweed aquaculture in China. Mar. Pollut. Bull. 62 (6), 1169–1182. https://doi.org/10.1016/j.marpolbul.2011.03.040.

Kim, K., Shin, J., Kim, K.Y., Ryu, J.-H., 2019. Long-Term Trend of Green and Golden Tides in the Eastern Yellow Sea. J. Coast. Res. 90 (SI), 317–323. https://doi.org/10.2112/si90-040.1.

Lee, J.H., Pang, I.C., Moon, I.J., Ryu, J.H., 2011. On physical factors that controlled the massive green tide occurrence along the southern coast of the Shandong Peninsula in 2008: A numerical study using a particle-tracking experiment. J. Geophys. Res. Oceans 116 (C12). https://doi.org/10.1029/2011JC007512.

Li, D., Gao, Z., Wang, Z., 2022. Analysis of the reasons for the outbreak of Yellow Sea green tide in 2021 based on long-term multi-source data. Mar. Environ. Res. 105649.
Li, D., Gao, Z., Xu, F., 2021. Research on the dissipation of green tide and its influencing factors in the Yellow Sea based on Google Earth Engine. Mar. Pollut. Bull. 172, 112801 https://doi.org/10.1016/j.marpolbul.2021.112801.
Li, L., Zheng, X., Wei, Z., Zou, J., Xing, Q., 2018. A Spectral-Mixing Model for Estimating Sub-Pixel Coverage of Sea-Surface Floating Macroalgae. Atmos. Ocean 56 (4), 296–302. https://doi.org/10.1080/07055900.2018.1509834.
Liu, D., Keesing, J.K., Dong, Z., Zhen, Y., Di, B., Shi, Y., Fearns, P., Shi, P., 2010. Recurrence of the world's largest green-tide in 2009 in Yellow Sea, China: Porphyra yezoensis aquaculture rafts confirmed as nursery for macroalgal blooms. Mar. Pollut. Bull. 60 (9), 1423–1432.
Liu, D., Keesing, J.K., He, P., Wang, Z., Shi, Y., Wang, Y., 2013. The world's largest macroalgal bloom in the Yellow Sea, China: Formation and implications. Estuar. Coast. Shelf Sci. 129, 2–10. https://doi.org/10.1016/j.ecss.2013.05.021.
Liu, D., Keesing, J.K., Xing, Q., Shi, P., 2009. World's largest macroalgal bloom caused by expansion of seaweed aquaculture in China. Mar. Pollut. Bull. 58 (6), 888–895.
Liu, X., Li, Y., Wang, Z., Zhang, Q., Cai, X., 2015. Cruise observation of Ulva prolifera bloom in the southern Yellow Sea, China. Estuar. Coast. Shelf Sci. 163, 17–22.
Liu, X., Wang, Z., Zhang, X., 2016. A review of the green tides in the Yellow Sea, China. Mar. Environ. Res. 119, 189–196.
Lu, T., Lu, Y., Hu, L., Jiao, J., Zhang, M., Liu, Y., 2019. Uncertainty in the optical remote estimation of the biomass of Ulva prolifera macroalgae using MODIS imagery in the Yellow Sea. Opt. Express 27 (13), 18620–18627. https://doi.org/10.1364/OE.27.018620.
Ma, Y., Wong, K., Tsou, J.Y., Zhang, Y., 2022. Investigating Spatial Distribution of Green-Tide in the Yellow Sea in 2021 Using Combined Optical and SAR Images. J. Marine Sci. Eng. 10 (2), 127. https://www.mdpi.com/2077-1312/10/2/127.
Mikelsons, M., Wang, M., 2018. Interactive online maps make satellite ocean data accessible. Eos 99. https://doi.org/10.1029/2018EO096563.
Qi, L., Hu, C., 2021. To what extent can *Ulva* and *Sargassum* be detected and separated in satellite imagery? Harmful Algae 103, 102001. https://doi.org/10.1016/j.hal.2021.102001.
Qi, L., Hu, C., Barnes, B. B., Lapointe, B. E., Chen, Y., Xie, Y., Wang, M., 2022a. Climate and anthropogenic controls of seaweed expansions in the East China Sea and Yellow Sea. Geophysical Research Letters, 49(19), e2022GL098185. https://doi.org/https://doi.org/10.1029/2022GL098185.
Qi, L., Hu, C., Mikelsons, K., Wang, M., Lance, V., Sun, S., Barnes, B.B., Zhao, J., Van der Zande, D., 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.
Qi, L., Hu, C., Wang, M., Shang, S., Wilson, C., 2017. Floating algae blooms in the East China Sea. Geophysical Research Letters, 44(22), 11,501-511,509. https://doi.org/10.1002/2017GL075525.
Qi, L., Hu, C., Xing, Q., Shang, S., 2016. Long-term trend of *Ulva prolifera* blooms in the western Yellow Sea. Harmful Algae 58, 35–44. https://doi.org/10.1016/j.hal.2016.07.004.
Qi, L., Wang, M., Hu, C., Holt, B., 2022b. On the capacity of Sentinel-1 synthetic aperture radar in detecting floating macroalgae and other floating matters. Remote Sens. Environ. 280, 113188 https://doi.org/10.1016/j.rse.2022.113188.
Qi, L., et al. (in prep). Uncertainties in the MODIS-derived *Ulva prolifera* amounts in the Yellow Sea: Systematic evaluations using Sentinel-2 observations.
Qiao, F., Wang, G., Lü, X., et al., 2011. Drift characteristics of green macroalgae in the Yellow Sea in 2008 and 2010. Chinese Sci. Bull. 2011 (56), 2236–2242. https://doi.org/10.1007/s11434-011-4551-7.
Qiu, Z., Li, Z., Bilal, M., Wang, S., Sun, D., Chen, Y., 2018. Automatic method to monitor floating macroalgae blooms based on multilayer perceptron: case study of Yellow Sea using GOCI images. Opt. Express 26 (21), 26810–26829.
Shen, H., Perrie, W., Liu, Q., He, Y., 2014. Detection of macroalgae blooms by complex SAR imagery. Mar. Pollut. Bull. 78 (1–2), 190–195.
Shi, W., Wang, M., 2009. Green macroalgae blooms in the Yellow Sea during the spring and summer of 2008. J. Geophys. Res. Oceans 114 (C12). https://doi.org/10.1029/2009JC005513.
Son, Y.B., Choi, B.-J., Kim, Y.H., Park, Y.-G., 2015. Tracing floating green algae blooms in the Yellow Sea and the East China Sea using GOCI satellite data and Lagrangian transport simulations. Remote Sens. Environ. 156, 21–33. https://doi.org/10.1016/j.rse.2014.09.024.
Son, Y.B., Min, J.-E., Ryu, J.-H., 2012. Detecting massive green algae (Ulva prolifera) blooms in the Yellow Sea and East China Sea using geostationary ocean color imager (GOCI) data. Ocean Sci. J. 47 (3), 359–375.
Sun, D., Chen, Y., Wang, S., Zhang, H., Qiu, Z., Mao, Z., He, Y., 2021. Using Landsat 8 OLI data to differentiate Sargassum and Ulva prolifera blooms in the South Yellow Sea. Int. J. Appl. Earth Obs. Geoinf. 98, 102302.
Sun, X., Wu, M., Xing, Q., Song, X., Zhao, D., Han, Q., Zhang, G., 2018. Spatio-temporal patterns of Ulva prolifera blooms and the corresponding influence on chlorophyll-a concentration in the Southern Yellow Sea, China. Sci. Total Environ. 640–641, 807–820. https://doi.org/10.1016/j.scitotenv.2018.05.378.
Wan, X., Wan, J., Xu, M., Liu, S., Sheng, H., Chen, Y., Zhang, X., 2021. Enteromorpha coverage information extraction by 1D-CNN and Bi-LSTM networks considering sample balance from GOCI images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 9306–9317.
Wang, C., Chu, J., Tan, M., Shao, F., Sui, Y., Li, S., 2017. An automatic detection of green tide using multi-windows with their adaptive threshold from Landsat TM/ETM plus image. Acta Oceanol. Sin. 36 (11), 106–114.
Wang, M., Shi, W., 2006. Cloud Masking for Ocean Color Data Processing in the Coastal Regions. IEEE Trans. Geosci. Remote Sens. 44 (11), 3196–13105. https://doi.org/10.1109/TGRS.2006.876293.
Wang, X., Wang, L., Chen, L., Zhang, F., Chen, K., Zhang, Z., Zou, Y., Zhao, L., 2022. AlgaeMask: An Instance Segmentation Network for Floating Algae Detection. J. Marine Sci. Eng. 10 (8), 1099. https://www.mdpi.com/2077-1312/10/8/1099.
Wang, X., Xing, Q., An, D., Meng, L., Zheng, X., Jiang, B., Liu, H., 2021. Effects of Spatial Resolution on the Satellite Observation of Floating Macroalgae Blooms. Water 13 (13), 1761. https://www.mdpi.com/2073-4441/13/13/1761.
Wang, Z., Xiao, J., Fan, S., Li, Y., Liu, X., Liu, D., 2015. Who made the world's largest green tide in China?—an integrated study on the initiation and early development of the green tide in Yellow Sea. Limnol. Oceanogr. 60 (4), 1105–1117. https://doi.org/10.1002/lno.10083.
Xiao, Y., Zhang, J., Cui, T., 2017. High-precision extraction of nearshore green tides using satellite remote sensing data of the Yellow Sea, China. Int. J. Remote Sensing 38 (6), 1626–1641. https://doi.org/10.1080/01431161.2017.1286056.
Xiao, Y., Zhang, J., Cui, T., Gong, J., Liu, R., Chen, X., Liang, X., 2019. Remote sensing estimation of the biomass of floating Ulva prolifera and analysis of the main factors driving the interannual variability of the biomass in the Yellow Sea. Mar. Pollut. Bull. 140, 330–340.
Xing, Q., An, D., Zheng, X., Wei, Z., Wang, X., Li, L., Tian, L., Chen, J., 2019. Monitoring seaweed aquaculture in the Yellow Sea with multiple sensors for managing the disaster of macroalgal blooms. Remote Sens. Environ. 231, 111279.
Xing, Q., Hu, C., 2016. Mapping macroalgal blooms in the Yellow Sea and East China Sea using HJ-1 and Landsat data: Application of a virtual baseline reflectance height technique. Remote Sens. Environ. 178, 113–126. https://doi.org/10.1016/j.rse.2016.02.065.
Xing, Q., Tosi, L., Braga, F., et al., 2015. Interpreting the progressive eutrophication behind the world's largest macroalgal blooms with water quality and ocean color data. Nat. Hazards 78, 7–21. https://doi.org/10.1007/s11069-015-1694-x.
Xing, Q., Wu, L., Tian, L., Cui, T., Li, L., Kong, F., Gao, X., Wu, M., 2018. Remote sensing of early-stage green tide in the Yellow Sea for floating-macroalgae collecting campaign. Mar. Pollut. Bull. 133, 150–156. https://doi.org/10.1016/j.marpolbul.2018.05.035.
Xu, F., Gao, Z., Shang, W., Jiang, X., Zheng, X., Ning, J., Song, D., 2017. Validation of MODIS-based monitoring for a green tide in the Yellow Sea with the aid of unmanned aerial vehicle. J. Appl. Remote Sens. 11 (1), 012007.
Xu, Q., Zhang, H., Cheng, Y., 2016. Multi-sensor monitoring of Ulva prolifera blooms in the Yellow Sea using different methods. Front. Earth Sci. 10 (2), 378–388.
Xu, Q., Zhang, H., Ju, L., Chen, M., 2014. Interannual variability of Ulva prolifera blooms in the Yellow Sea. Int. J. Remote Sens. 35 (11–12), 4099–4113. https://doi.org/10.1080/01431161.2014.916052.
Yu, H., Wang, C., Li, J., Sui, Y., 2021. Automatic Extraction of Green Tide From GF-3 SAR Images Based on Feature Selection and Deep Learning. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 10598–10613. https://doi.org/10.1109/JSTARS.2021.3118374.
Yuan, C., Xiao, J., Zhang, X., Zhou, J., Wang, Z., 2022. A new assessment of the algal biomass of green tide in the Yellow Sea. Mar. Pollut. Bull. 174, 113253 https://doi.org/10.1016/j.marpolbul.2021.113253.
Zhang, B., Guo, J., Li, Z., Cheng, Y., Zhao, Y., Boota, M.W., Zhang, Y., Feng, L., 2022a. Identifying the spatio-temporal variations of Ulva prolifera disasters in all life cycle. J. Water Clim. Change 13 (2), 629–644.
Zhang, G., Wu, M., Wei, J., He, Y., Niu, L., Li, H., Xu, G., 2021a. Adaptive Threshold Model in Google Earth Engine: A Case Study of Ulva prolifera Extraction in the South Yellow Sea, China. Remote Sensing 13 (16), 3240. https://www.mdpi.com/2072-4292/13/16/3240.
Zhang, G., Wu, M., Zhang, A., Xing, Q., Zhou, M., Zhao, D., Song, X., Yu, Z., 2020a. Influence of sea surface temperature on outbreak of Ulva prolifera in the Southern Yellow Sea, China. Chinese Geographical Science 30 (4), 631–642.
Zhang, G., Wu, M., Zhou, M., Zhao, L., 2022b. The seasonal dissipation of Ulva prolifera and its effects on environmental factors: based on remote sensing images and field monitoring data. Geocarto Int. 37 (3), 860–878.
Zhang, H., Qiu, Z., Devred, E., Sun, D., Wang, S., He, Y., Yu, Y., 2019a. A simple and effective method for monitoring floating green macroalgae blooms: a case study in the Yellow Sea. Opt. Express 27 (4), 4528–4548. https://doi.org/10.1364/OE.27.004528.
Zhang, H., Wang, G., Zhang, C., Su, R., Shi, X., Wang, X., 2020b. Characterization of the development stages and roles of nutrients and other environmental factors in green tides in the Southern Yellow Sea, China. *Harmful Algae* 98, 101893.
Zhang, H., Yuan, Y., Xu, Y., Shen, X., Sun, D., Qiu, Z., Wang, S., He, Y., 2021b. Remote sensing method for detecting green tide using HJ-CCD top-of-atmosphere reflectance. Int. J. Appl. Earth Obs. Geoinf. 102, 102371 https://doi.org/10.1016/j.jag.2021.102371.
Zhang, J., Shi, J., Gao, S., Huo, Y., Cui, J., Shen, H., Liu, G., He, P., 2019b. Annual patterns of macroalgal blooms in the Yellow Sea during 2007–2017. PLoS One 14 (1), e0210460.
Zhang, J., Zhao, P., Huo, Y., Yu, K., He, P., 2017. The fast expansion of Pyropia aquaculture in "Sansha" regions should be mainly responsible for the Ulva blooms in Yellow Sea. Estuar. Coast. Shelf Sci. 189, 58–65.

Zhang, J., Huo, Y., Zhang, Z., Yu, K., He, Q., Zhang, L., Yang, L., Xu, R., He, P., 2013. Variations of morphology and photosynthetic performances of Ulva prolifera during the whole green tide blooming process in the Yellow Sea. Mar. Environ. Res. 92, 35–42.

Zhang, Y., He, P., Li, H., Li, G., Liu, J., Jiao, F., Zhang, J., Huo, Y., Shi, X., Su, R., Ye, N., Liu, D., Yu, R., Wang, Z., Zhou, M., Jiao, N., 2019c. Ulva prolifera green-tide outbreaks and their environmental impact in the Yellow Sea, China. National Sci. Rev. 6 (4), 825–838. https://doi.org/10.1093/nsr/nwz026.

Zheng, H., Liu, Z., Chen, B., Xu, H., 2020. Quantitative Ulva prolifera bloom monitoring based on multi-source satellite ocean color remote sensing data. Appl. Ecol. Environ. Res 18, 4897–4913.

Zheng, L., Wu, M., Cui, Y., Tian, L., Yang, P., Zhao, L., Xue, M., Liu, J., 2022. What causes the great green tide disaster in the South Yellow Sea of China in 2021? Ecol. Ind. 140, 108988.