nature portfolio

https://doi.org/10.1038/s41561-021-00887-x

Supplementary information

Global mapping reveals increase in lacustrine algal blooms over the past decade

In the format provided by the authors and unedited

Supplementary Information

Global mapping reveals increase in lacustrine algal blooms over the past decade

Xuejiao Hou, Lian Feng*, Yanhui Dai, Chuanmin Hu, Luke Gibson, Jing Tang, Zhongping Lee, Ying Wang, Xiaobin Cai, Junguo Liu, Yi Zheng, Chunmiao Zheng

*Correspondence to: fengl@sustech.edu.cn

1 Supplementary Notes:

1.1 Supplementary Note 1: Rationale of the CIE color space-based algal bloom detection algorithm

Since the 1970s, satellite remote sensing techniques have been effectively used to monitor algal blooms and to examine their distributions, onsets, and durations, among many other properties [1-8]. Particularly, some species of phytoplankton can regulate their buoyancy and form subsurface accumulations and surface scums in calm weather, allowing them to be easily detected by remotely sensed observations [1]. Satellite-delineated bloom extents are generally considered more reliable than those determined from field surveys due to the significant heterogeneity of phytoplankton cells both spatially and temporally [9]. The separation of bloom-containing and bloom-free pixels is straightforward for a limited number of satellite images, as accumulated colonies can substantially modulate satellite signals in terms of both magnitudes and spectral shapes [1,10]. However, the application of a bloom detection algorithm to global Landsat data spanning more than 38 years is not trivial. This is because the satellite signals of lacustrine waters change not only with the concentrations of three optically

sensitive components in water (including chlorophyll, suspended sediments, and colored dissolved organic matter (CDOM)) [11] but also with aquatic vegetation, shallow water bottoms and satellite observation conditions [12-14]. All of these factors can be encountered at sometime or location during global long-term monitoring research.

Various spectral wavelengths or indices have been introduced to detect algal blooms in inland and coastal waters, such as a single red band [15], normalized difference vegetation index (NDVI) [16], maximum chlorophyll index (MCI) [10], maximum peak height (MPH) [17], cyanobacteria index (CI) [18], floating algal index (FAI) [12] and false-colored red-green-blue (FRGB) index [19]. The use of the MCI, MPH and CI requires narrow bands at red to NIR spectral ranges, which are applicable to only a few satellite sensors (such as MERIS (2002-2011) and the Ocean Land Color Instrument (OLCI, 2017-now)). Among the above, the FAI algorithm is currently the most commonly used method for broadband instruments (such as the Moderate Resolution Spectroradiometer (MODIS) and Landsat) [20,21], and FAI values can be calculated using three bands in the red, NIR and SWIR spectral regions. The baseline subtraction designation of the FAI makes it insensitive to changes in satellite/solar geometries, sunglint, and aerosol contamination [12].

A predefined threshold, which is typically determined manually, is required before using the FAI or any other spectral wavelength/index to discriminate phytoplankton blooms within remote sensing images. However, the optimal threshold can vary considerably across space and time according to different lacustrine

environments and satellite observation conditions (Supplementary Fig. S2). Thus, determining the best image-specific thresholds for the 2.91 million images used in our study is practically impossible. Therefore, we developed an automatic algal bloom detection algorithm by introducing the color space created by the Commission Internationale del'éclairage (CIE, or International Commission on Illumination in English) in 1931 [22].

The CIE color space quantitatively defines linkages between the distributions of the visible spectrum and all colors perceptible by human eyes [22]. The CIE color space allows us to numerically specify colors using a two-dimensional CIE xy chromaticity diagram (Extended Data Fig. 1a) instead of the traditional three red-green-blue (RGB) tristimulus values. In contrast, it is often challenging to characterize color similarities within the RGB color space, prohibiting a threshold-based approach to distinguish two colors. The three tristimulus values from the RGB color space can be converted into CIE chromaticity coordinates (x, y) using the following equations [23]:

$$x=X/(X+Y+Z)$$

$$y=Y/(X+Y+Z)$$

$$X=2.7689R+1.7517G+1.1302B$$

$$Y=1.0000R+4.5907G+0.0601B$$

$$Z=0.0000R+0.0565G+5.5943B$$

where R, G and B represent the reflectances at red, green and blue wavelengths, respectively. The white point (1/3, 1/3) in the diagram represents equal energy from all three bands, and any point located along the horseshoe-shaped border corresponds to

monochromatic light (or the spectral locus) (Extended Data Fig. 1a). With decreasing distance to the spectral locus from the white point, more energy is expected from the associated monochromatic light. After transforming the Landsat RGB surface reflectance data into CIE xy chromaticity coordinates (x, y), the colors of water bodies can be easily visualized.

The CIE color space has been used previously by several groups to examine changes in water color through the combination of a Forel-Ule (FUI) color scheme [24,25], while no efforts have been made to use this scheme for the automatic detection of algal blooms. A transformation from the RGB color space to the hue-saturation-value (HSV) space has also been successfully applied to classify water bodies and other land cover types [26]. We conducted a comparison, the results of which demonstrated the poorer separability of bloom pixels in the HSV color space than in the CIE color space (results not shown here).

1.2 Supplementary Note 2: Validations of the algal bloom detection algorithm

Three types of validations were performed to ensure the reliability of the algal blooms discriminated using the CIE-based algorithm: comparisons with visually selected samples from both Landsat images and high resolution Planet Dove images, high spectral solution datasets from MERIS, and a literature review.

First, we evaluated the uncertainty of the algorithm through comparisons between the CIE algorithm-classified results and visually interpreted bloom-containing and bloom-free pixels from the same images. All the validation samples

(41,045 bloom-containing pixels and 47,182 bloom-free pixels) were manually selected from Landsat images covering 22 lakes across different regions and times (Supplementary Table 3a). Although the data from several lakes were also used to develop the algorithm, we employed different images to select the validation samples. The bloom-containing pixels were delineated using the same method for the selection of training samples, and these sampled pixels were used to determine if the algorithm can accurately exclude false negatives (blooms detected as non-blooms); the bloom-free pixels were selected within the same images when no greenish slicks or low FAI values were present, and the samples were used to evaluate whether the algorithm could accurately exclude false positives (non-blooms classified as blooms). The confusion matrix[27] in Extended Data Table 1 shows that the producer and user accuracies of the algorithm for bloom area detection are 98.8% and 92.3%, respectively. Likewise, we used visually interpreted validation samples from 3-m resolution Planet Dove images to conduct quantitative accuracy assessments for bloom areas classified by Landsat-8 OLI images. Since Dove images are not available until December 2015, similar evaluations for preceding Landsat instruments are not possible. We visualized the RGB true color composite of Dove images and the corresponding NDVI images; greenish features with high NDVI values were manually delineated as bloom-containing pixels, whereas pixels with other colors and low NDVI values were determined as bloom-free. We randomly selected 16,564 validation sample points (9,585 for bloom-containing and 6,979 for bloom-free, respectively) from 21 lakes distributed globally (Supplementary Table 3b). To

consider the differences in spatial resolutions between Landsat (30-m) and Dove (3-m), we assured that the neighboring pixels around each sample point in Dove images are dominated by the same class as the sample (bloom-containing or bloom-free). These samples were then compared with the classified results from concurrent Landsat 8 images (i.e., images acquired within the same days as Dove), which resulted in high producer and user accuracies of 98.6% and 94.6% for bloom detection, respectively (Extended Data Table 1). In addition to these quantitative evaluations of the uncertainty levels, we also exported the classification results from the Google Earth Engine for at least 10,000 images across six continents. We exported both the RGB true color image and the associated bloom map to verify whether the bloom patterns are consistent with the features in the RGB image (see examples in Extended Data Fig. 1c and Supplementary Fig. S3). Such a qualitative verification further confirmed the fidelity of our algorithm in discriminating greenish bloom features in Landsat images.

Second, we used high spectral resolution MERIS images to assess the Landsat-detected bloom areas. MERIS has narrow bands (up to 10 nm) configured within the visible to NIR spectral range, making it possible to detect several diagnostic signals from algal blooms [3,18]. One of the pronounced signals is the reflectance peak at 709 nm, caused by high phytoplankton absorption at 675 nm and high water absorption for wavelengths longer than ~720 nm [10]. The MCI, developed via this reflectance peak, can be used to highlight the bloom areas in MERIS images [10]. The MCI is estimated as:

$$MCI = R_{\lambda 2} - [R_{\lambda 1} + (R_{\lambda 3} - R_{\lambda 1}) \times (\lambda_2 - \lambda_1)/(\lambda_3 - \lambda_1)]$$

where λ_1 to λ_3 are the MERIS bands centered at 665, 709, and 754 nm, respectively. A modified version of the MCI (denoted as the MCIT, MCIT= MCI/(1+0.1 ×($R_{\lambda 3}$ - $R_{\lambda 4}$)), λ_4=865 nm) recommended by Qi et al. [28] was used here to reduce some impacts from water turbidity.

MERIS images acquired concurrently with the Landsat data (i.e., with the same acquisition dates) were obtained, and the data cover 28 lakes across different regions and times (Supplementary Fig. 2). For each Landsat image, we first identified the bloom-containing and bloom-free pixels within the examined lake using our CIE-based algorithm. Then, we estimated the MCIT for the corresponding MERIS image and compared the histograms of the MCIT for the Landsat-classified bloom-containing and bloom-free pixels. The MCIT values for the bloom-containing pixels were systematically higher than those for the bloom-free pixels in all 28 examined lakes.

Third, we compared the Landsat-derived bloom patterns to previously reported bloom outbreaks through a review of both published journal papers and gray literature. Bloom occurrences (BO, defined as the ratio between the number of detected algal blooms to the number of valid satellite observations) were estimated using Landsat observations spanning 38 years were carefully examined. For example, in the Great Lakes of North America, we found high BO in western Lake Erie, Saginaw Bay in Lake Huron, and Green Bay in Lake Michigan (Fig. 3); these

lake sections have attracted widespread concern from both the general public and the scientific community [29-31]. Extensive studies have revealed higher risks of cyanobacterial blooms in northern and western Taihu Lake in China (Fig. 3), which agrees well with the high BO we found in these regions [4,32]. In addition, vast areas of cyanobacterial blooms have been detected in northern Lake Winnipeg in Canada using satellite imagery by other groups [33,34]; our results also revealed high BO (>10%) in most of this region (Fig. 3). In total, we carefully inspected more than 100 lakes, including all the lakes we used to select the training and validation samples and many other lakes. The results confirm that the Landsat-delineated blooms are consistent with either peer-reviewed literature or online news reports. We acknowledge that an exhaustive check of the >20,000 lakes in our study is impossible; this is also the reason why the results provided here represent the first comprehensive understanding of global lake blooms.

We acknowledge that ideally real ground truth data should be used to ascertain the reliability of the Landsat-detected algal blooms using the novel CIE-based algorithm. However, *in situ* validations for satellite-delineated phytoplankton blooms are challenging due to the remarkable nonuniformities of algal blooms in both vertical and horizontal directions in the water column [35,36], strongly related to the migration capability of phytoplankton cells [35]. The complex responses of phytoplankton cells/colonies to wind stress, sunlight and other lake environmental properties (such as water depth and temperature) make it difficult to predict their transport trajectory in the water column. Indeed, previous field surveys demonstrated that algal blooms often

show extremely patchy features. For example, concentrations of phytoplankton cell counts or chlorophyll a (a key indicator for phytoplankton biomass) varied by up to an order of magnitude between water samples collected on different sides of a research vessel [9]. Therefore, in situ samples collected from such patchy areas are not suitable for validating Landsat-delineated blooms, which represent mean conditions within 30×30-m^2 boxes. The considerable diurnal changes in bloom features also add complexity to this task [37,38]. Therefore, we believe that the methods used herein represent the best practices for global-scale algal bloom validation.

The bloom detection algorithm in this study was developed by training samples with outbreaks of greenish algal blooms (see Extended Data Fig. 1 and Supplementary Table 2). As such, our algorithm may not be useful to determine algal blooms either without densely aggregated surface/subsurface phytoplankton biomass or appearing non-greenish in color. Indeed, other than the most problematic and widely reported summer blooms, our algorithm is capable of detecting the spring blooms in periodical frozen lakes when greenish bloom slicks are distinct on water surface [39]. It is also noted that although blooms with non-greenish colors exist in marine (often called “red tides”) and saline lacustrine systems [40,41], they have rarely been found in freshwater lakes and the exclusion of such blooms is not expected to strongly impact the global statistics of our study.

Algal blooms are harmful only when they produce toxins or of extremely high biomass [41,42]; and the contents of toxins from phytoplankton blooms can vary substantially in both space and time, even within the same genera [43]. Therefore, the

discrimination between harmful and non-harmful phytoplankton blooms is currently challenging with only satellite observations. Moreover, the determination of the phytoplankton species is also not possible by using their colors, and therefore knowledge of the phytoplankton ecology (i.e., latitudinal and seasonal distributions) is required to interpret what phytoplankton can be expected from the satellite-based algal bloom maps.

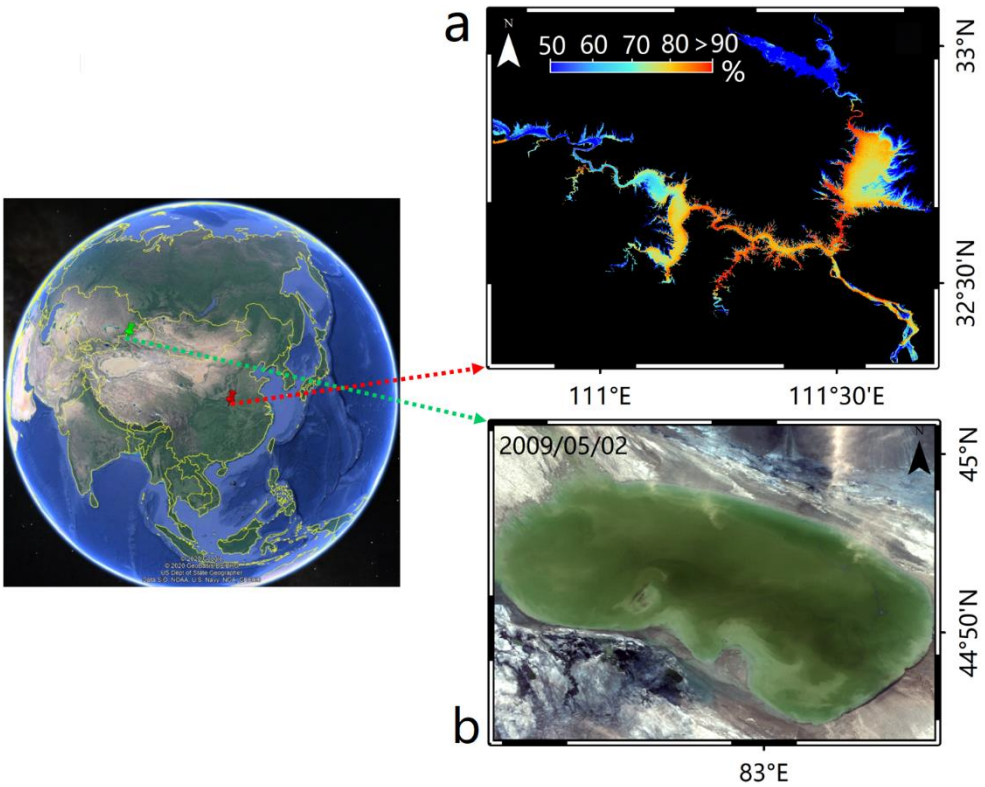
Subsurface accumulation and surface scums of phytoplankton occur under weak wind conditions [18,34], where the bloom signal can be easily detected by satellite remote sensing. However, high wind stress tends to mix the cells/colonies from the surface to deeper waters, making it difficult to be detected with remotely sensed observations [18]. Indeed, such dynamic nature of phytoplankton (e.g., cyanobacteria), as well as their ability to change buoyancy, are also the exact reasons why bloom extents detected from satellite remote sensing are currently considered more reliable than field measurements [1]. More importantly, remote sensing appears to be the only technique that can be used for monitoring algal blooms at continental to global scales. In fact, numerous environmental programs worldwide have incorporated satellite products into the decision-making process, including the US Environmental Protection Agency, NOAA, NASA, the European Environment Agency, Chinese environmental protection agencies, aquaculture industry of southern Africa, etc. [2,5,44,45]. Further, satellite observations are also the only source of historical data for the majority of the global lakes across the planet that were not monitored in the field.

Satellite missions (such as MODIS, the Visible Infrared Imaging Radiometer Suite (VIIRS), Geostationary Ocean Color Imager (GOCI), and OLCI) with more frequent (daily or hourly) observations could potentially address the wind-induced issues by tracking short-term variability of lacustrine algal blooms. However, Landsat datasets provide a much finer spatial resolution, which is particularly important for the detection of bloom features in small lakes (as small as 0.1 km^2). Moreover, the Landsat data archive is longer than any of the other available archives, making it possible to examine the changes in lacustrine environments over the past four decades; in contrast, the other satellites with more frequent observations are not available until circa 2000. In fact, we minimized the impacts of wind and other environmental factors on the global statistics by two procedures when establishing the GBD database: 1) we excluded the areas with insufficient valid observations from the decadal comparisons; and 2) we normalized the number of bloom images over the number of valid Landsat observations for BO to further minimize the associated short-term impacts from wind (see details above). Nevertheless, satellite images with spatial resolutions of tens of meters are currently available from a number of satellite missions, such as the Sentinel series of sensors launched by the European Space Agency and the Gaofen series of satellites launched by China. In particular, daily observations with a resolution of a few meters are being acquired by commercial satellite systems (such as Planet Dove). We expect that the use of multiple satellite missions in the future can provide both high frequency (daily or even hourly) and high

spatial-resolution to help with ongoing monitoring of algal blooms, as well as other environmental monitoring programs.

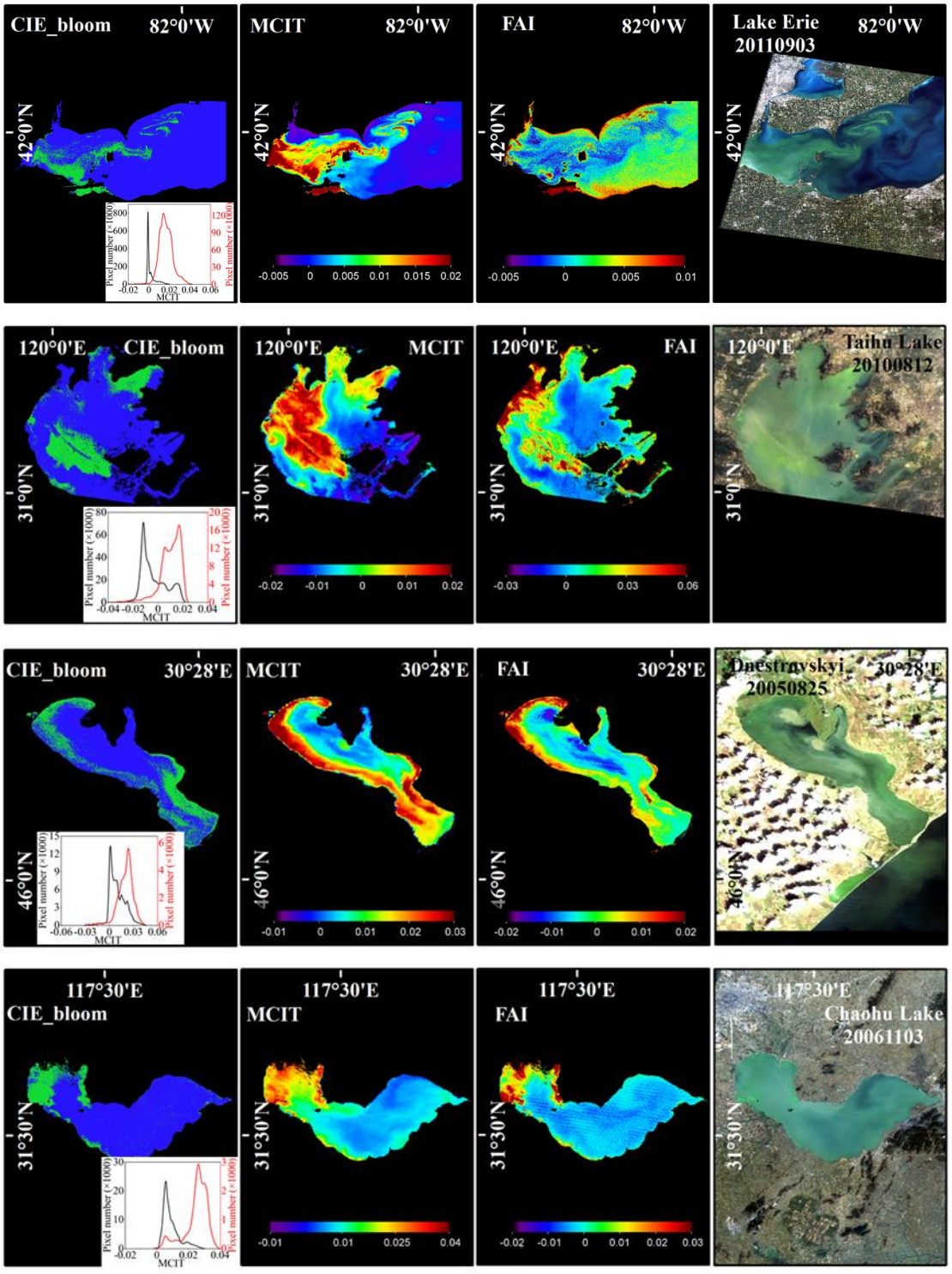
Supplementary Figures

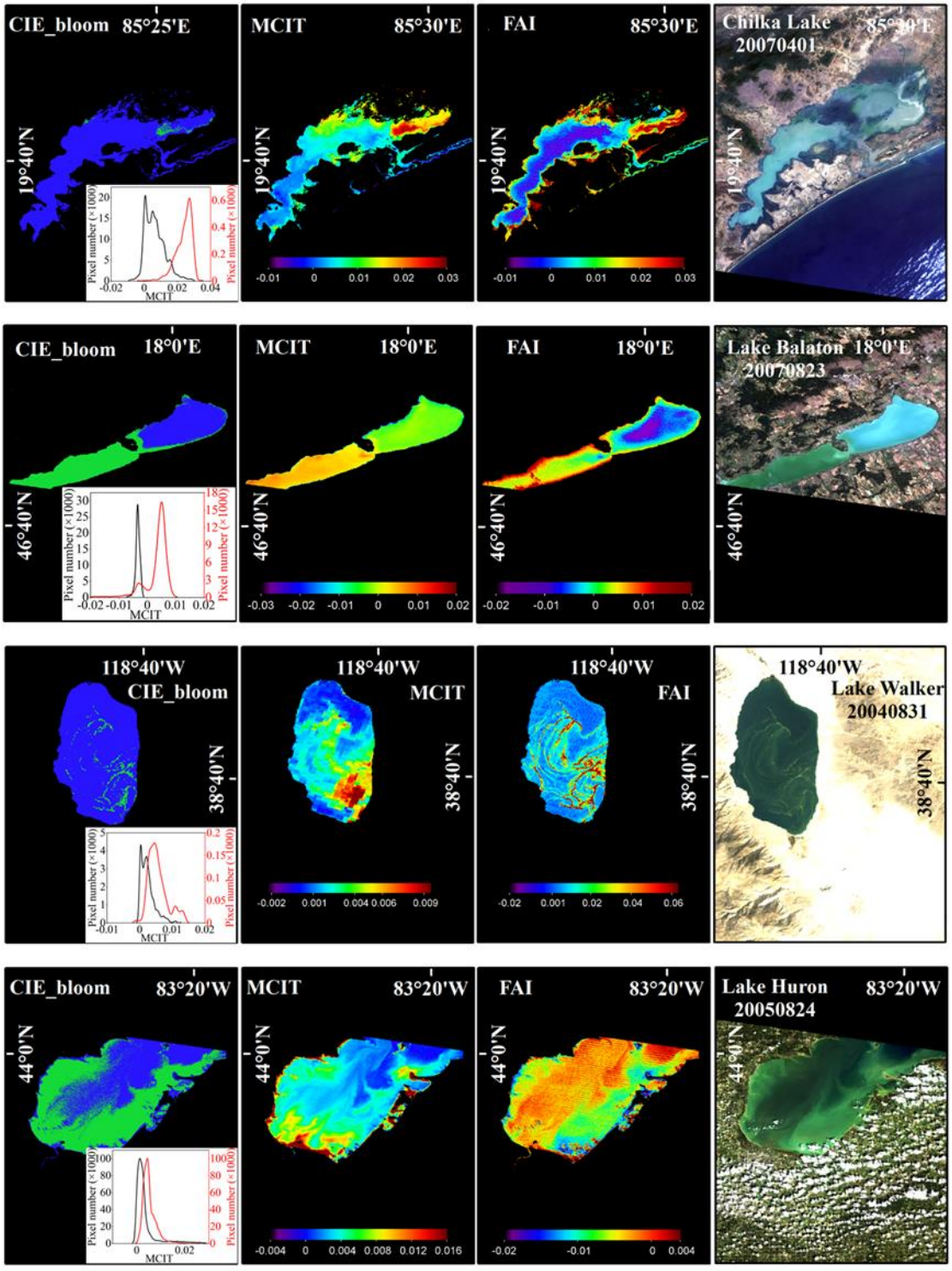
Supplementary Figure S1

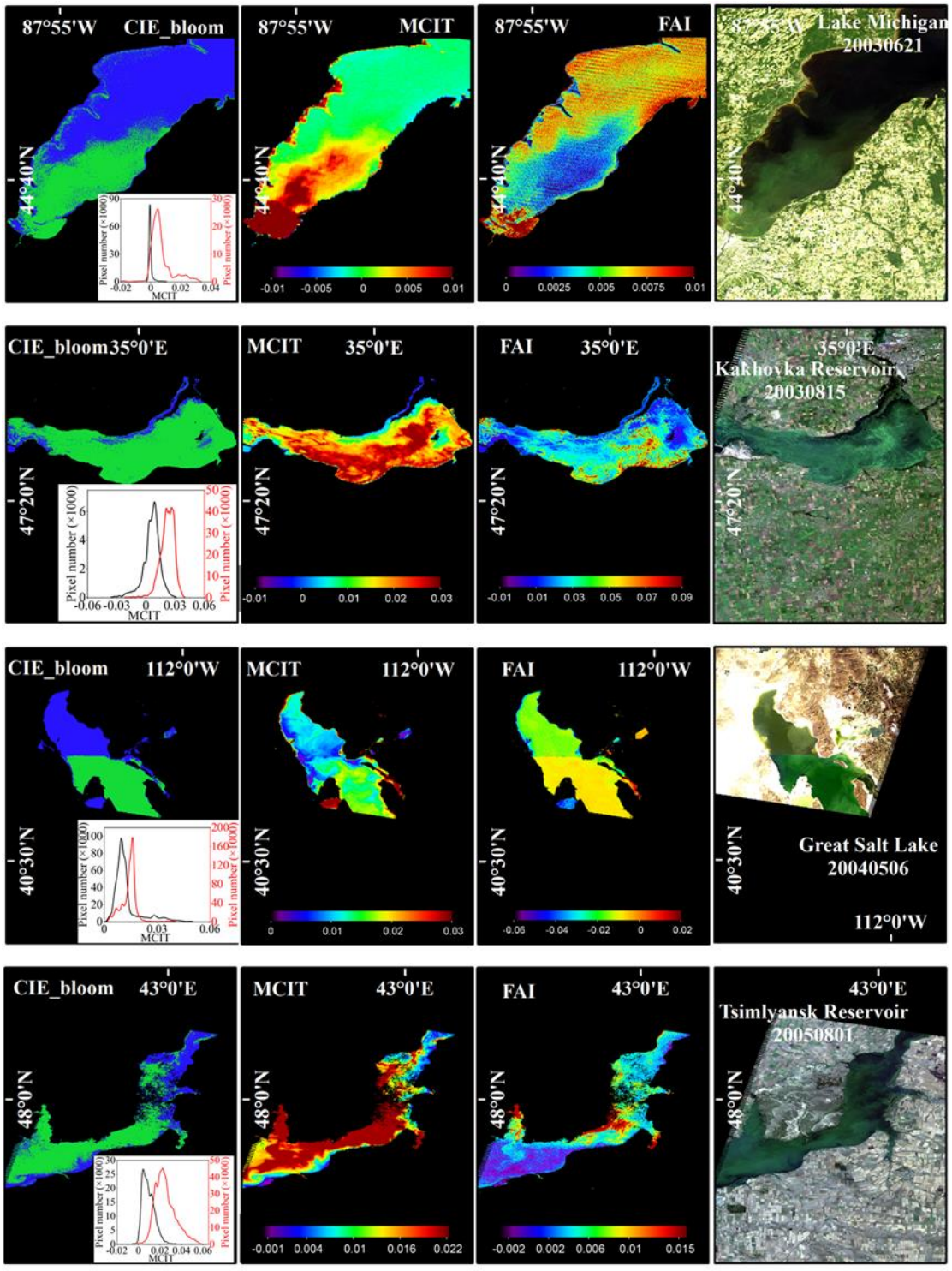


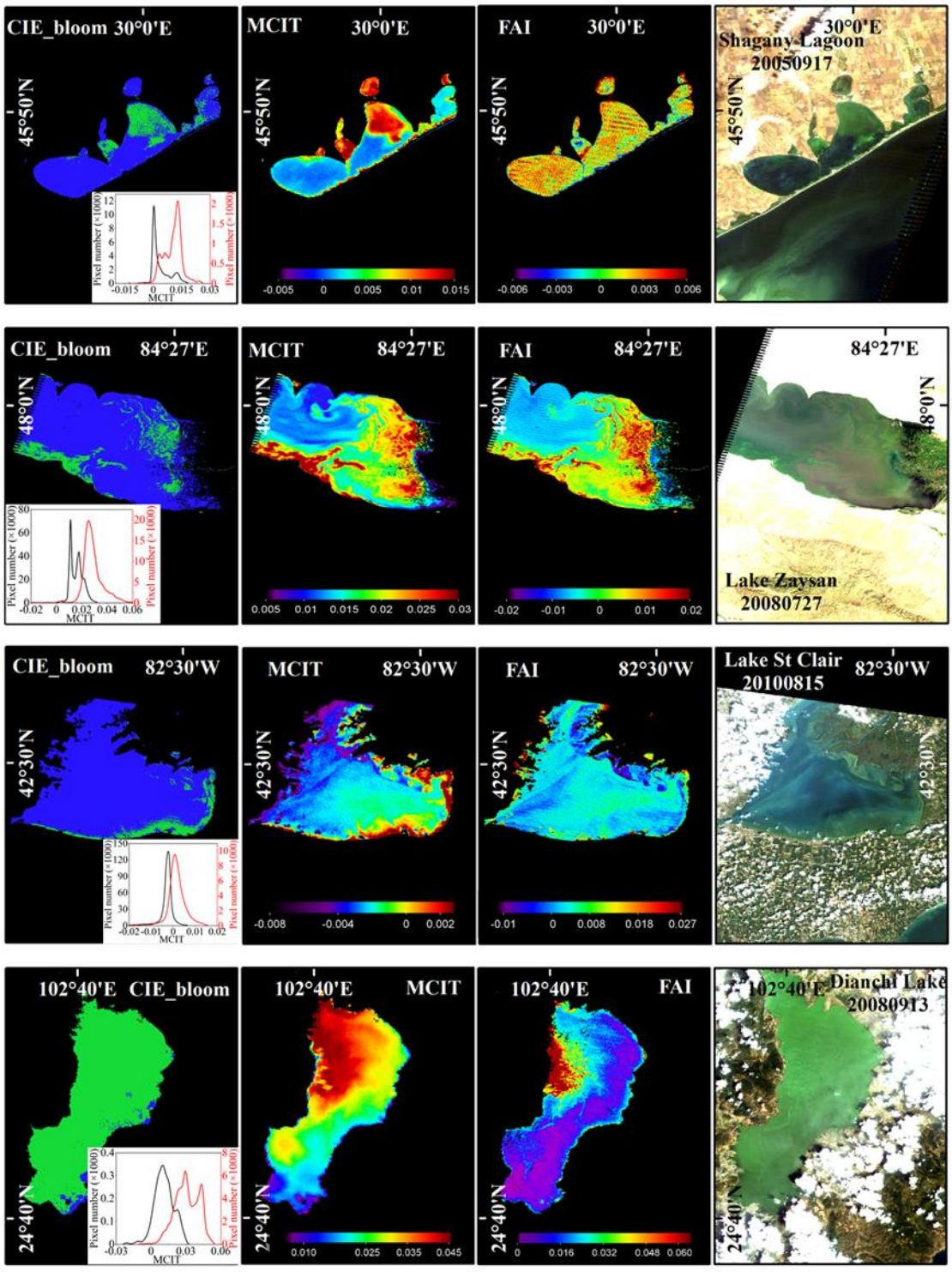
Supplementary Figure S1 | Examples to show two types of disturbances affecting bloom detection. (**a**) Percentage of positive FAI occurrence (from 1982 to 2019) in the Danjiangkou Reservoir in China, where the water has been reported to be of high quality. The signals reflected from adjacent green vegetation (in either a specular or diffuse manner) off the water surface can lead to misclassification with the CIE-based algorithm; (**b**) RGB composite of a clear greenish karst lake (Ebinur Lake in China), which is located in an endorheic basin. The locations of Danjiangkou Reservoir and Ebinur Lake are shown on the left.

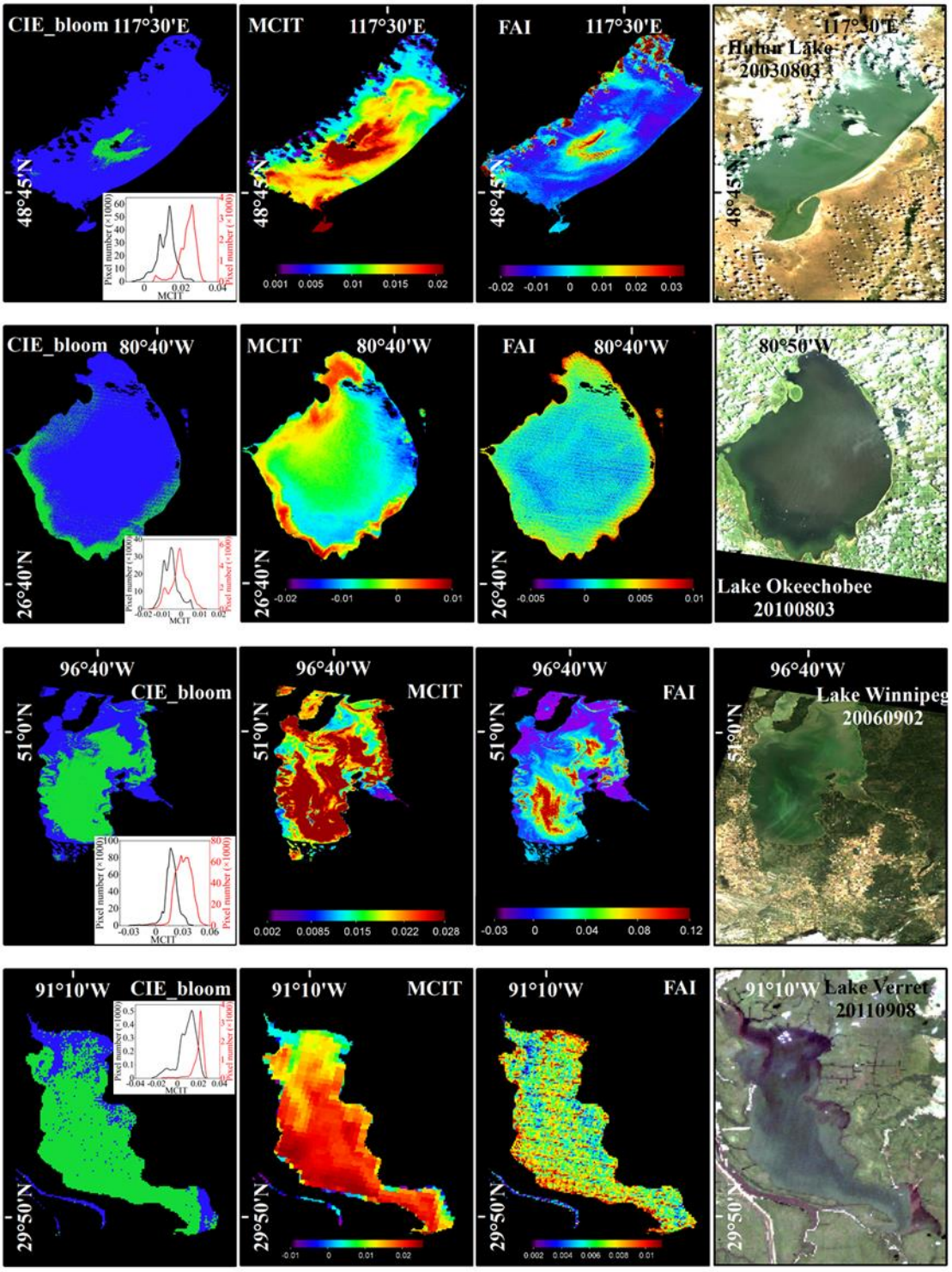
Supplementary Figure S2

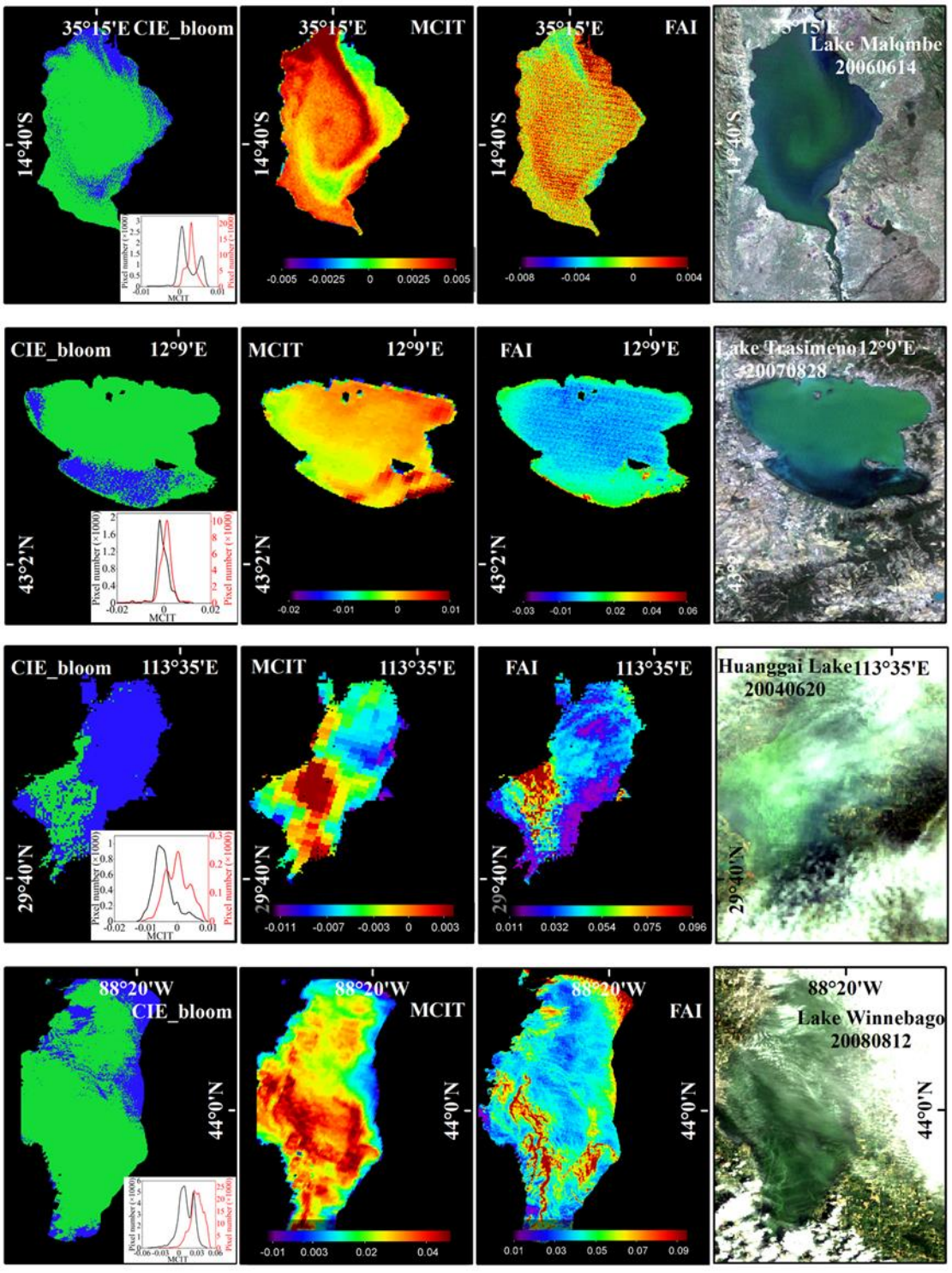


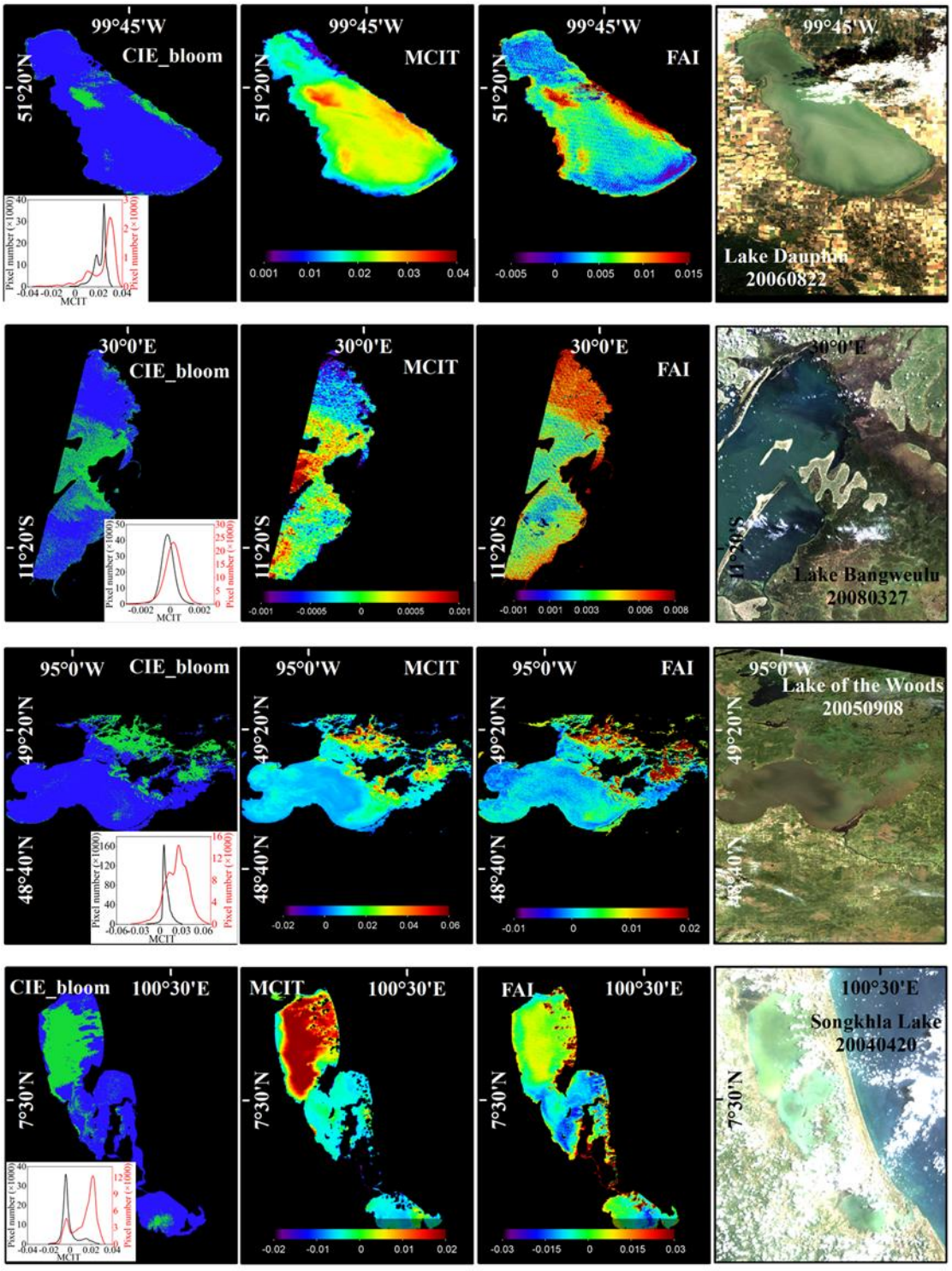

CIE_bloom 85°25'E
MCIT 85°30'E
FAI 85°30'E
Chilka Lake 85°30'E
20070401
19°40'N
CIE_bloom 18°0'E
MCIT 18°0'E
FAI 18°0'E
Lake Balaton 18°0'E
20070823
46°40'N
118°40'W
CIE_bloom
MCIT
FAI
Lake Walker
20040831
38°40'N
CIE_bloom 83°20'W
MCIT 83°20'W
FAI 83°20'W
Lake Huron 83°20'W
20050824
44°0'N
Pixel number (×1000)
MCIT


87°55'W
CIE_bloom
MCIT
FAI
Lake Michigan
20030621
44°0'N
CIE_bloom35°0'E
MCIT 35°0'E
FAI 35°0'E
35°0'E
Kakhovka Reservoir
20030815
47°20'N
CIE_bloom 112°0'W
MCIT 112°0'W
FAI 112°0'W
Great Salt Lake
20040506
40°30'N
112°0'W
CIE_bloom 43°0'E
MCIT 43°0'E
FAI 43°0'E
43°0'E
Tsimlyansk Reservoir
20050801
48°0'N
Pixel number (×1000)
MCIT


CIE_bloom 30°0'E
MCIT 30°0'E
FAI 30°0'E
30°0'E
Shagany-Lagoon
20050917
45°50'N
CIE_bloom 84°27'E
MCIT 84°27'E
FAI 84°27'E
84°27'E
48°0'N
Lake Zaysan
20080727
CIE_bloom 82°30'W
MCIT 82°30'W
FAI 82°30'W
Lake St Clair
20100815
42°30'N
102°40'E CIE_bloom
102°40'E MCIT
102°40'E FAI
102°40'E Dianchi Lake
20080913
24°40'N
Pixel number (×1000)
MCIT

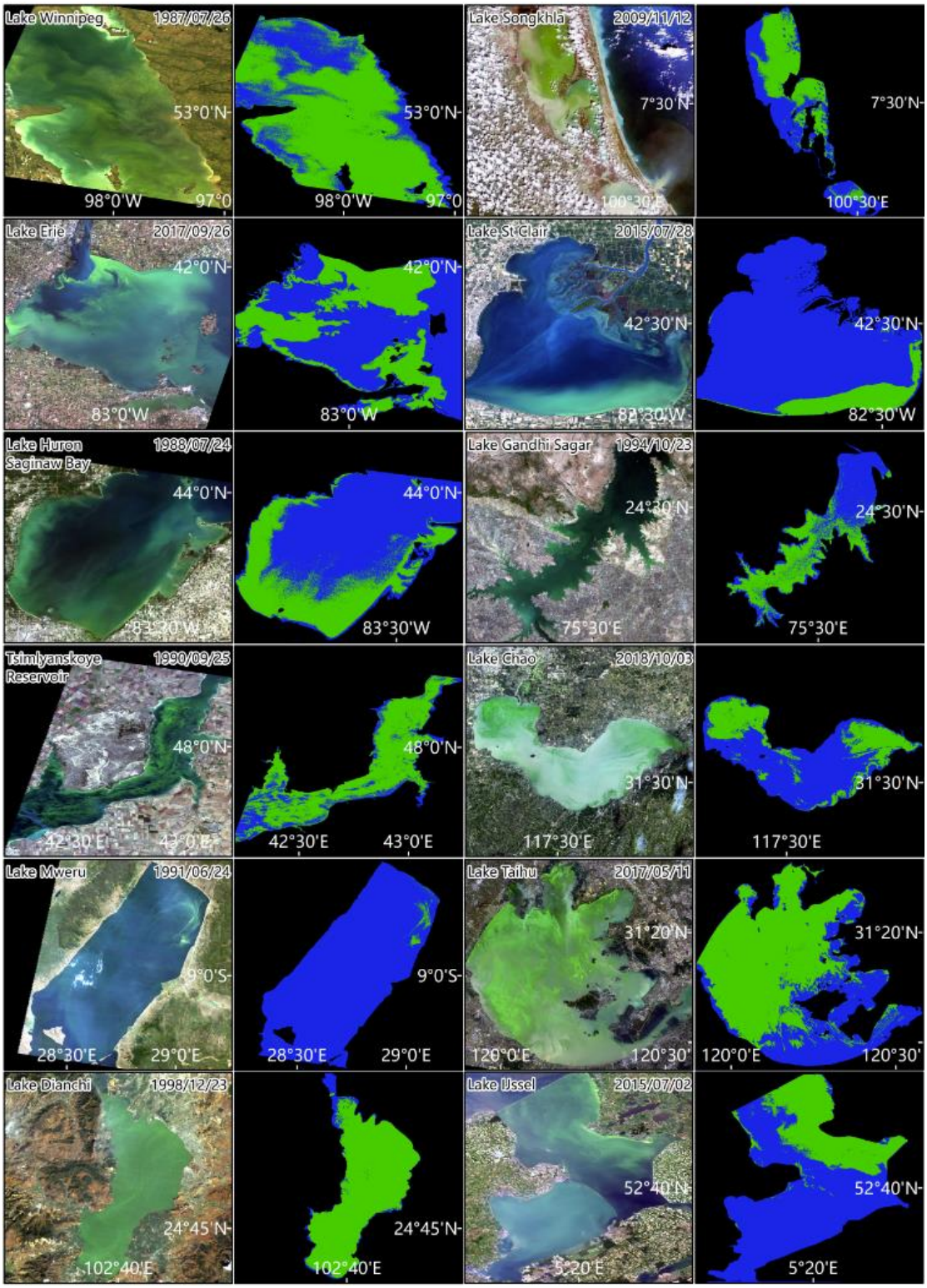

CIE_bloom 117°30'E
MCIT 117°30'E
FAI 117°30'E
117°30'E
Hulun Lake
20030803
48°45'N
Pixel number (×1000)
MCIT
0.001 0.005 0.01 0.015 0.02
-0.02 -0.01 0 0.01 0.02 0.03
CIE_bloom 80°40'W
MCIT 80°40'W
FAI 80°40'W
80°50'W
Lake Okeechobee
20100803
26°40'N
-0.02 -0.01 0 0.01
-0.005 0 0.005 0.01
96°40'W
CIE_bloom
MCIT
FAI
Lake Winnipeg
20060902
51°0'N
0.002 0.0085 0.015 0.022 0.028
-0.03 0 0.04 0.08 0.12
91°10'W
CIE_bloom
MCIT
FAI
Lake Verret
20110908
29°50'N
-0.01 0 0.01 0.02
0.002 0.004 0.006 0.008 0.01


35°15'E CIE_bloom
35°15'E MCIT
35°15'E FAI
35°15'E
Lake Malombe
20060614
14°40'S
-0.005 -0.0025 0 0.0025 0.005
-0.008 -0.004 0 0.004
CIE_bloom 12°9'E
MCIT 12°9'E
FAI 12°9'E
Lake Trasimeno 12°9'E
20070828
43°2'N
-0.02 -0.01 0 0.01
-0.03 -0.01 0.02 0.04 0.06
CIE_bloom 113°35'E
MCIT 113°35'E
FAI 113°35'E
Huanggai Lake 113°35'E
20040620
29°40'N
-0.011 -0.007 -0.003 0 0.003
0.011 0.032 0.054 0.075 0.096
88°20'W
CIE_bloom
MCIT
FAI
Lake Winnebago
20080812
44°0'N
-0.01 0.003 0.02 0.04
0.01 0.03 0.05 0.07 0.09



Supplementary Figure S2 | Validation of the Landsat-delineated bloom areas using MERIS observations in 28 lakes. The RGB composites, FAI and bloom areas (classified using the CIE-based algorithm) were generated using Landsat images, and the MCIT was calculated using concurrent (i.e., on the same acquisition date) MERIS images (see Supplementary Note 2). The MCIT histograms of the bloom-containing (red curves) and bloom-free (black curves) pixels are shown for each image group (left panels).

Supplementary Figure S3



Supplementary Figure S3 | The RGB and corresponding bloom classification maps for 12 selected lakes distributed globally, showing the performance of the algorithm and the severity of algal blooms (represented as green in the right panels).

Supplementary References

1 Kutser, T., Metsamaa, L., Strömbeck, N. & Vahtmäe, E. Monitoring cyanobacterial blooms by satellite remote sensing. *Estuarine, Coastal and Shelf Science* **67**, 303-312 (2006).

2 Stumpf, R. P. *et al.* Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. *Harmful Algae* **54**, 160-173 (2016).

3 Matthews, M. W. Eutrophication and cyanobacterial blooms in South African inland waters: 10years of MERIS observations. *Remote Sensing of Environment* **155**, 161-177, doi:https://doi.org/10.1016/j.rse.2014.08.010 (2014).

4 Hu, C. *et al.* Moderate Resolution Imaging Spectroradiometer (MODIS) observations of cyanobacteria blooms in Taihu Lake, China. *Journal of Geophysical Research: Oceans* **115**, doi:10.1029/2009JC005511 (2010).

5 Mishra, S. *et al.* Measurement of cyanobacterial bloom magnitude using satellite remote sensing. *Scientific Reports* **9**, 1-17 (2019).

6 Coffer, M. M., Schaeffer, B. A., Darling, J. A., Urquhart, E. A. & Salls, W. B. Quantifying national and regional cyanobacterial occurrence in US lakes using satellite remote sensing. *Ecological Indicators* **111**, 105976 (2020).

7 Song, K. *et al.* Climatic versus Anthropogenic Controls of Decadal Trends (1983–2017) in Algal Blooms in Lakes and Reservoirs across China. *Environmental Science & Technology* **55**, 2929-2938 (2021).

8 Binding, C. E., Pizzolato, L. & Zeng, C. EOLakeWatch; delivering a comprehensive suite of remote sensing algal bloom indices for enhanced monitoring of Canadian eutrophic lakes. *Ecological Indicators* **121**, 106999, doi:https://doi.org/10.1016/j.ecolind.2020.106999 (2021).

9 Kutser, T. Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. *Limnology and Oceanography* **49**, 2179-2189 (2004).

10 Gower, J., King, S. & Goncalves, P. Global monitoring of plankton blooms using MERIS MCI. *International Journal of Remote Sensing* **29**, 6209-6216 (2008).

11 Kirk, J. T. *Light and photosynthesis in aquatic ecosystems*. (Cambridge university press, 1994).

12 Hu, C. A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment* **113**, 2118-2129 (2009).

13 Lee, Z., Carder, K. L., Mobley, C. D., Steward, R. G. & Patch, J. S. Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. *Applied optics* **38**, 3831-3843 (1999).

14 Oyama, Y., Matsushita, B. & Fukushima, T. Distinguishing surface cyanobacterial blooms and aquatic macrophytes using Landsat/TM and ETM+ shortwave infrared bands. *Remote Sensing of Environment* **157**, 35-47 (2015).

15 Kahru, M. & Elmgren, R. Multidecadal time series of satellite-detected accumulations of cyanobacteria in the Baltic Sea. *Biogeosciences* **11**, 3619-3633 (2014).

16 Hu, C. & He, M. X. Origin and offshore extent of floating algae in Olympic sailing area. *Eos, Transactions American Geophysical Union* **89**, 302-303 (2008).
17 Matthews, M. W., Bernard, S. & Robertson, L. An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. *Remote Sensing of Environment* **124**, 637-652 (2012).
18 Wynne, T. T., Stumpf, R. P., Tomlinson, M. C. & Dyble, J. Characterizing a cyanobacterial bloom in western Lake Erie using satellite imagery and meteorological data. *Limnology and Oceanography* **55**, 2025-2036 (2010).
19 Qi, L. *et al.* In search of floating algae and other organisms in global oceans and lakes. *Remote Sensing of Environment* **239**, 111659 (2020).
20 Blondeau-Patissier, D., Gower, J. F., Dekker, A. G., Phinn, S. R. & Brando, V. E. A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans. *Progress in Oceanography* **123**, 123-144 (2014).
21 Matthews, M. W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *International Journal of Remote Sensing* **32**, 6855-6899 (2011).
22 Robertson, A. R. The CIE 1976 color‐difference formulae. *Color Research & Application* **2**, 7-11 (1977).
23 Fairman, H. S., Brill, M. H. & Hemmendinger, H. How the CIE 1931 color-matching functions were derived from Wright-Guild data. *Color Research & Application* **22**, 11-23, doi:10.1002/(SICI)1520-6378(199702)22:1<11::AID-COL4>3.0.CO;2-7 (1997).
24 Wang, S. *et al.* Trophic state assessment of global inland waters using a MODIS-derived Forel-Ule index. *Remote Sensing of Environment* **217**, 444-460 (2018).
25 Wernand, M. R., van der Woerd, H. J. & Gieskes, W. W. Trends in ocean colour and chlorophyll concentration from 1889 to 2000, worldwide. *PLOS one* **8**, e63766 (2013).
26 Pekel, J.-F., Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418-422 (2016).
27 Congalton, R. G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment* **37**, 35-46 (1991).
28 Qi, L., Hu, C., Duan, H., Zhang, Y. & Ma, R. Influence of particle composition on remote sensing reflectance and MERIS maximum chlorophyll index algorithm: examples From Taihu Lake and Chaohu Lake. *IEEE Geoscience and Remote Sensing Letters* **12**, 1170-1174 (2015).
29 Michalak, A. M. *et al.* Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proceedings of the National Academy of Sciences* **110**, 6448-6452 (2013).
30 Vanderploeg, H. A. *et al.* Zebra mussel (Dreissena polymorpha) selective

filtration promoted toxic Microcystis blooms in Saginaw Bay (Lake Huron) and Lake Erie. *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 1208-1221 (2001).

31 Bartlett, S. L., Brunner, S. L., Klump, J. V., Houghton, E. M. & Miller, T. R. Spatial analysis of toxic or otherwise bioactive cyanobacterial peptides in Green Bay, Lake Michigan. *Journal of Great Lakes Research* **44**, 924-933 (2018).

32 Guo, L. Doing Battle With the Green Monster of Taihu Lake. *Science* **317**, 1166-1166, doi:10.1126/science.317.5842.1166 (2007).

33 Kling, H., Watson, S., McCullough, G. & Stainton, M. Bloom development and phytoplankton succession in Lake Winnipeg: a comparison of historical records with recent data. *Aquatic ecosystem health & management* **14**, 219-224 (2011).

34 Binding, C., Greenberg, T., McCullough, G., Watson, S. & Page, E. An analysis of satellite-derived chlorophyll and algal bloom indices on Lake Winnipeg. *Journal of Great Lakes Research* **44**, 436-446 (2018).

35 Walsby, A. E., Hayes, P. K., Boje, R. & Stal, L. J. The selective advantage of buoyancy provided by gas vesicles for planktonic cyanobacteria in the Baltic Sea. *New Phytologist* **136**, 407-417 (1997).

36 Kutser, T., Metsamaa, L. & Dekker, A. G. Influence of the vertical distribution of cyanobacteria in the water column on the remote sensing signal. *Estuarine, Coastal and Shelf Science* **78**, 649-654 (2008).

37 Hunter, P., Tyler, A., Willby, N. & Gilvear, D. The spatial dynamics of vertical migration by Microcystis aeruginosa in a eutrophic shallow lake: A case study using high spatial resolution time‐series airborne remote sensing. *Limnology and Oceanography* **53**, 2391-2406 (2008).

38 Qi, L., Hu, C., Visser, P. M. & Ma, R. Diurnal changes of cyanobacteria blooms in Taihu Lake as derived from GOCI observations. *Limnology and Oceanography* **63**, 1711-1726 (2018).

39 Peeters, F., Straile, D., Lorke, A. & Livingstone, D. M. Earlier onset of the spring phytoplankton bloom in lakes of the temperate zone in a warmer climate. *Global Change Biology* **13**, 1898-1909, doi:https://doi.org/10.1111/j.1365-2486.2007.01412.x (2007).

40 Dierssen, H. M., Kudela, R. M., Ryan, J. P. & Zimmerman, R. C. Red and black tides: Quantitative analysis of water-leaving radiance and perceived color for phytoplankton, colored dissolved organic matter, and suspended sediments. *Limnology and Oceanography* **51**, 2646-2659, doi:https://doi.org/10.4319/lo.2006.51.6.2646 (2006).

41 Lopez, C., Jewett, E., Dortch, Q., Walton, B. & Hudnell, H. Scientific assessment of freshwater harmful algal blooms. *United States National Ocean Service*, 65 (2008).

42 Huisman, J. *et al.* Cyanobacterial blooms. *Nature Reviews Microbiology* **16**, 471-483 (2018).

43 Kurmayer, R., Christiansen, G., Fastner, J. & Börner, T. Abundance of active and inactive microcystin genotypes in populations of the toxic cyanobacterium Planktothrix spp. *Environmental microbiology* **6**, 831-841 (2004).

44 Smith, M. E. & Bernard, S. Satellite Ocean Color Based Harmful Algal Bloom Indicators for Aquaculture Decision Support in the Southern Benguela. *Frontiers in Marine Science* **7**, doi:10.3389/fmars.2020.00061 (2020).

45 MA, R.-H. Spatio-temporal distribution of cyanobacteria blooms based on satellite imageries in Lake Taihu, China. *J Lake Sci* **20**, 687-694 (2008).