Article

A Remote Sensing Approach to Estimate Vertical Profile Classes of Phytoplankton in a Eutrophic Lake

Kun Xue [1,2], Yuchao Zhang [1], Hongtao Duan [1], Ronghua Ma [1,*], Steven Loiselle [3] and Minwei Zhang [4]

[1] State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, 73 East Beijing Road, Nanjing 210008, China; E-Mails: klunar@126.com (K.X.); yczhang@niglas.ac.cn (Y.Z.); htduan@niglas.ac.cn (H.D.)

[2] University of Chinese Academy of Sciences, Beijing 100049, China

[3] Dipartimento di Biotecnologie, Chimica e Farmacia, University of Siena, CSGI, Via Aldo Moro 2, Siena 53100, Italy; E-Mail: loiselle@unisi.it

[4] College of Marine Science, University of South Florida, 140 Seventh Avenue, South, St. Petersburg, FL 33701, USA; E-Mail: zhangminwei2004@126.com

* Author to whom correspondence should be addressed; E-Mail: rhma@niglas.ac.cn; Tel.: +86-025-8688-2168; Fax: +86-025-5771-4759.

Abstract: The extension and frequency of algal blooms in surface waters can be monitored using remote sensing techniques, yet knowledge of their vertical distribution is fundamental to determine total phytoplankton biomass and understanding temporal variability of surface conditions and the underwater light field. However, different vertical distribution classes of phytoplankton may occur in complex inland lakes. Identification of the vertical profile classes of phytoplankton becomes the key and first step to estimate its vertical profile. The vertical distribution profile of phytoplankton is based on a weighted integral of reflected light from all depths and is difficult to determine by reflectance data alone. In this study, four Chla vertical profile classes (vertically uniform, Gaussian, exponential and hyperbolic) were found to occur in three *in situ* vertical surveys (28 May, 19–24 July and 10–12 October) in a shallow eutrophic lake, Lake Chaohu. We developed and validated a classification and regression tree (CART) to determine vertical phytoplankton biomass profile classes. This was based on an algal bloom index (Normalized Difference algal Bloom Index, NDBI)

applied to both *in situ* remote sensing reflectance (R_{rs}) and MODIS Rayleigh-corrected reflectance (R_{rc}) data in combination with data of local wind speed. The results show the potential of retrieving Chla vertical profiles information from integrated information sources following a decision tree approach.

Keywords: vertical distribution profile classes; chlorophyll-a; Normalized Difference algal Bloom Index (NDBI); classification and regression tree (CART); remote sensing

1. Introduction

The eutrophication of coastal and inland waters is a major environmental and social-economic problem around the world. The increasing occurrence and intensity of algal blooms have severely affected the security of drinking water and food sources, biodiversity, and economic activities in a large number of ecosystems, such as Lake Taihu in China [1], Baltic Sea [2], Seto-Inland Sea in Japan [3], and Gulf of Mexico [4]. Beginning in the late 1990s, major lakes in eastern China underwent increasing eutrophication and significant algal bloom events. One of these, Lake Chaohu (the fifth largest freshwater lake in China), has experienced recurring harmful algal blooms with significant impacts on the local population [5].

Remote sensing has been widely used to monitor the extent and frequency of algal blooms [6–8] as well as determine key optically active substances, such as chlorophyll-a concentrations (Chla) [9,10], and phycocyanin concentrations (PC) [10,11]. However, these estimates assume that the distribution of phytoplankton is vertically uniform or at least vertically consistent, leading to inaccurate estimates of total phytoplankton biomass across depth. Variability of the phytoplankton vertical distribution can be evidenced by rapid changes in the spatial distribution of surface phytoplankton estimates. For example, using the Geostationary Ocean Color Imager (GOCI), algal blooms in the East China Sea increased 100% in a single day [12]. Algal biomass estimates in shallow lakes do not typically increase at such a rapid pace, but significant vertical movement can occur over short time periods [13]. The variability in the vertical distribution of algal biomass represents a major challenge to the remote determination of lake optical properties, as well as the estimation of total phytoplankton biomass and primary production [14].

The vertical distribution of phytoplankton is mainly governed by meteorological, biological and hydrological parameters [13]. In particular, vertical distributions of phytoplankton in shallow water are directly and indirectly influenced by wind [15]. When the wind speed is higher than 2–4 m/s, phytoplankton is likely to be evenly distributed [13,16–19]. Some cyanobacteria, such as *microcystis aeruginosa*, are capable of regulating their buoyancy, and maintain cyanobacterial cells in suspension, moving vertically through the water in response to changing conditions of temperature [20], light intensity [21] and nutrients [22,23]. However, the extent to which cyanobacteria can control their vertical distribution is also indirectly affected by wind, which controls mixing (temperature), surface reflection (light) and upwelling (nutrients). Generally, weak winds lead to less uniform vertical distributions of algae. In such conditions, cyanobacteria may float to the surface, even accumulating in extreme high concentrations to form scums [24].

In ocean optics, many studies have focused on the vertical distribution of phytoplankton and chlorophyll-a [14,25,26] and its influence on the remote sensing reflectance [27–29]. Gaussian models [30–32] or shifted Gaussian models [14] have been used to represent the vertical profile of chlorophyll-a concentrations in marine environments. It is often assumed that the typical shape of the chlorophyll-a profile is stable in a given region or season [14,33]. However, the vertical distribution of phytoplankton in lakes can change rapidly [34]. Unfortunately, there are few *in situ* datasets of the vertical distribution of phytoplankton in shallow lakes.

Understanding the vertical profile classes of phytoplankton is important to estimate the column-integrated phytoplankton biomass and biogeochemical applications involving primary production [35]. In addition, it is also the key parameters in radiative transfer simulation of the underwater light field.

The present study was directed to: (1) analyze the vertical profiles of algal biomass for a large shallow eutrophic lake, Lake Chaohu; (2) develop an integrated approach to identify the vertical distribution profile class based on remote sensing reflectance data and wind speed; and (3) develop and apply the integrated approach to map the Chla vertical profile classes using MODIS image data (Rayleigh-corrected reflectance, R_{rc}).

2. Study Region

Lake Chaohu is the fifth largest fresh water lake in China, with an area of 770 km^2 (31°25′–31°43′N, 117°17′–117°51′E, Figure 1) and a mean depth of 3.0 m [36]. The maximum Secchi depth was 0.60 m, allowing for optical deep conditions where bottom effects (reflectance from lake bottom) were not important [37,38]. Several rivers (Nanfei, Paihe and Hangbu rivers) flow into the lake, while the Yuxi River flows out and into the Yangtze River [39]. Lake Chaohu has had increasing problems of water pollution and eutrophication over the past three decades [40], jeopardizing its use as a potable water source. Massive algal blooms have been associated to an increase in domestic and industrial sewage from the expanding urban and industrial activities in adjacent areas. Most blooms occur between July and September, and are particularly severe in the west part of the lake, with higher frequency and intensity [41]. The diversity of phytoplankton, including *Chlorophyta*, *Bacillariophyta*, *Cyanophyta*, *Cryptophyta*, and *Euglenuphyta*, varies seasonally. *Cyanophyta* has the highest average annual density (more than 90%) compared to other phytoplankton species [42].

3. Methods

3.1. Field Measurements

Three field surveys were performed in 2013 (28 May, 19–24 July and 10–12 October, Figure 1). Water samples were obtained from 9 depths (surface, 0.1, 0.2, 0.4, 0.7, 1.0, 1.5, 2.0 and 3.0 m) collected in separate 1 liter Niskin bottles using an ad hoc vertical collection device, comprised of a 3.5 m perforated tube (2.5 cm in diameter), a small vacuum pump (about 10 cm in diameter), connective tubes and a scale bar. The depth of water inlet was controlled and determined by the scale bar. The samples of water surface (0 m) were collected directly using a water sampler (organic glass hydrophore). When the wind speed was high, wind-driven waves influenced the collection of samples at precise depths, in particular near the water

surface (surface sampling, 0.1 m and 0.2 m). When these conditions occurred, we sampled twice at each depth and mixed both samples together to create a single sample for each depth. Water samples were stored in the dark at low temperature (4 ℃) before filtration (<6 hours). Following filtration, the samples were frozen and the concentrations of chlorophyll-a (Chla, μg/L), suspended particulate inorganic matter (SPIM, mg/L) and dissolved organic carbon concentration (DOC, μg/L) were measured in the laboratory at the conclusion of the survey (usually 3 days, no more than 5 days). Simultaneously, environmental parameters such as the surface wind speed and cloud conditions were recorded. Wind speed and direction were measured every 10 to 15 minutes (about 7.5 to 10 kilometers between adjacent two sites) in numerous sites across the lake over several days in June 2014 (Figure 1). Water transparency was measured using a 20 cm Secchi disk from the shaded side of the boat.
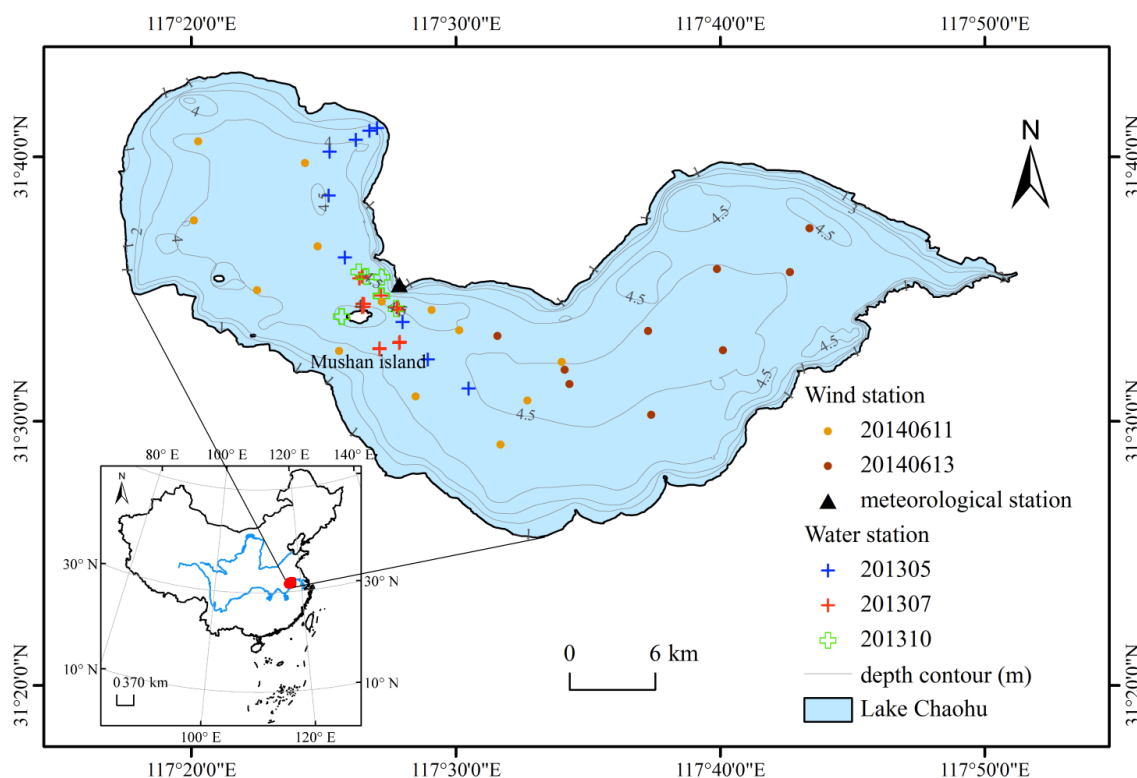

Figure 1. Study region and sampling locations in Lake Chaohu, China. *In situ* measurements of bio-optical parameters made during three cruise surveys in 28 May 2013 (N = 9), 19–24 July 2013 (N = 32) and 10–12 October 2013 (N = 27). The solid circles indicate field measurement sites of wind speed during two cruise surveys on 11 June (N = 13) and 13 June (N = 9) 2013, respectively. The solid black triangle represents the location of meteorological station.

Laboratory analyses: The water samples were filtered with Whatman GF/C glass-fibre filters (pore size of 1.1μm) and pigments were extracted using 90% acetone extraction. The Chla value was calculated using absorbance at 630, 645, 663 and 750 nm measured with a Shimadzu UV-2600 spectrophotometer [38]. For SPIM, 47 mm Whatman GF/F glass fiber filters were pre-combusted at 450 ℃ for 6 h and pre-weighted. Water samples were filtered and filters were dried at 105 ℃ for 4–6 h, suspended particulate inorganic matter (SPIM) was derived gravimetrically by burning organic matter from the filters at 450 ℃ for 6 h

and weighting the filters again [11]. The filtered water was used to determine the concentration of DOC, using a Shimadzu TOC-5000A analyzer [43,44].

Remote sensing reflectance: Following NASA protocols [45], an ASD field spectrometer (FieldSpec Pro Dual VNIR, Analytica Spectra Devices., Inc) was used to measure downwelling radiance and upwelling total radiance above water surface. This instrument has a spectral range of 350 to 1050 nm with two probes and a viewing field of 25 °. Measurements of the total water leaving radiance (L_{sw}), radiance of gray panel (L_p), and sky radiance (L_{sky}) were performed.

Each water spectrum was sampled 90 ° azimuth with respect to the sun and with a nadir viewing angle of 45 °. $L_{sw}(\lambda)$ was measured using the target probe at approximately 0.5 m above the water surface under low cloud (<10%, gathered from the nearest weather station) conditions, while the another probe measured $L_{sky}(\lambda)$. Radiance of a Lambertian gray panel ($L_p(\lambda)$) with reflectance ρ_p was used to determine the incident downwelling irradiance ($E_d(\lambda, 0^+)$) (Equation (1)).

$$E_d(\lambda, 0^+) = L_p(\lambda) \times \pi/\rho_p \tag{1}$$

$L_{sw}(\lambda)$ consisted of the desired water-leaving radiance $L_w(\lambda)$ and skylight reflection ($L_{sky}(\lambda)$), and was corrected by:

$$L_w(\lambda) = L_{sw}(\lambda) - \rho\, L_{sky}(\lambda) \tag{2}$$

The water surface reflectance factor ρ depended on sky conditions, wind speed, and solar zenith angle, and was assumed to be 0.028 [46]. Remote sensing reflectance ($R_{rs}(\lambda)$) was then derived by the ratio of water-leaving radiance $L_w(\lambda)$ to incident downwelling spectral plane irradiance $E_d(\lambda, 0^+)$.

$$R_{rs}(\lambda) = L_w(\lambda)/E_d(0^+, \lambda) \tag{3}$$

3.2. MODIS Satellite Data

Data collected by the Moderate Resolution Imaging Spectroradiometer (MODIS Terra and MODIS Aqua) were used to assess spatial and temporal coverage of the study lake for the three sampling campaigns. The ground resolution of the MODIS data was 250 m for the 645 and 859 nm bands and 500 m for the bands centered at 469, 555, 1240, 1640, and 2130 nm. Level-0 data were obtained from the U.S. NASA Goddard Space Flight Center (GSFC) (http://oceancolor.gsfc.nasa.gov) and converted to calibrated radiance (Level-1B) using SeaDAS (version 7.0). At present, there is no reliable atmospheric correction to produce accurate MODIS Rrs data for Lake Chaohu. A partial atmospheric correction to correct for the gaseous absorption (mainly by ozone) and Rayleigh (molecular) scattering effects was applied to the Level-1B data using routines and look up tables (LUTs) available in SeaDAS [47]. After correction of Rayleigh scattering and gaseous absorption effects, the Rayleigh-corrected reflectance ($R_{rc}(\lambda)$, dimensionless) was determined [48]:

$$R_{rc}(\lambda) = \rho_t(\lambda) - \rho_r(\lambda) = \rho_a(\lambda) + \pi t(\lambda) t_0(\lambda) R_{rs}(\lambda) \tag{4}$$

where ρ_t is the top of atmosphere (TOA) reflectance after adjustment of the gaseous absorption, ρ_r is the reflectance due to Rayleigh scattering, ρ_a is that due to aerosol scattering and aerosol-Rayleigh interactions, t and t_0 are the diffuse transmittance from the lake surface to the satellite and from the sun to the lake surface, respectively. ρ_a, t, and t_0 are functions of aerosol type, aerosol optical thickness, and

solar/viewing geometry. The above formulation assumes negligible contributions from whitecaps and sun glint, and was used to show the relationship of R_{rc} and R_{rs} here. The R_{rc} data were geo-referenced into a cylindrical equidistance (rectangular) projection [49].

Then, after excluding pixels identified as clouds (these pixels had extremely high reflectance). There were 8 cloud free MODIS R_{rc} images coincident with the sampling campaigns on 28 May, 24 July, 10–12 October 2013. The spatial resolution of R_{rc} data was resampled to 250 m. Four MODIS bands of R_{rc} data with the nominal central wavelengths of 555, 645, 859, and 1240 nm were used in bloom indices and true color images.

3.3. Parameterization of Vertical Chla Profiles

To describe the shape of each vertical chlorophyll profile, we wrote a function to curve fit automately using the Curve Fitting Toolbox of MATLAB R2012a software (The Math Works, Inc.). Each Chla vertical profile data was fitted to 5 functions: linear function, quadratic polynomial, Gaussian, exponential, and power function. The sum of squares due to error (SSE), lowest root-mean-square error (RMSE), and coefficient of determination (R^2) of each function were compared. The fitted function with the highest R^2 ($R^2 > 0.8$) and lowest SSE and RMSE was used to identify the vertical distribution class of the Chla vertical profile.

To compare the variabilty of the vertical distribution with respect to a uniform distribution, the coefficient of variation (CV) was used with the standard deviation (SD) and mean:

$$CV = \frac{SD}{mean} \times 100\% \tag{5}$$

3.4. Algal Bloom Indexes

Resampled remote sensing reflectance $R_{rs,MODOS}$ was constructed using the *in situ* measured reflectance (R_{rs}) and the spectral response function of MODIS (B(t)) (Equation (6)).

$$R_{rs,MODIS} = \frac{\int_{t_1}^{t_2} R_{rs}(t) \cdot B(t)\, dt}{\int_{t_1}^{t_2} B(t)\, dt} \tag{6}$$

Several algal bloom indexes were used to distinguish the algal bloom waters from non-bloom waters, Normalized Difference Vegetation Index (NDVI, Equation (7)) [50]; Chlorophyll Spectral Index (CSI, Equation (8)) [51], Floating Algae Index (FAI, Equation (9)) [52], and the Normalized Difference algal Bloom Index (NDBI, Equations (10) and (11)):

$$NDVI = \frac{R_{rs}(748) - R_{rs}(675)}{R_{rs}(748) + R_{rs}(675)} \tag{7}$$

$$CSI = \frac{R_{rs}(700) - R_{rs}(675)}{R_{rs}(700) + R_{rs}(675)} \tag{8}$$

$$\text{FAI} = R_{rc,\text{NIR}} - R'_{rc,\text{NIR}} \, , \ \ R'_{rc,\text{NIR}} = R_{rc,\text{RED}} + \frac{\lambda_{\text{NIR}} - \lambda_{\text{RED}}}{\lambda_{\text{SWIR}} - \lambda_{\text{RED}}} \times (R_{rc,\text{SWIR}} - R_{rc,\text{RED}}) \qquad (9)$$

$$\text{NDBI}_{Rrs} = \frac{R_{rs}(550) - R_{rs}(675)}{R_{rs}(550) + R_{rs}(675)} \qquad (10)$$

$$\text{NDBI}_{Rrc} = \frac{R_{rc,\text{GREEN}} - R_{rc,\text{RED}}}{R_{rc,\text{GREEN}} + R_{rc,\text{RED}}} \qquad (11)$$

Among these algal bloom indexes, NDVI, CSI, and NDBI$_{Rrs}$ were derived by *in situ* R_{rs} data while FAI and NDBI$_{Rrc}$ were applied to MODIS R_{rc} data. In MODIS algorithms, λ_{GREEN}, λ_{RED}, λ_{NIR}, and λ_{SWIR} are associated to wavebands centered at 555 nm, 645 nm, 859 nm, and 1240 nm, respectively.

3.5. Classification and Regression Tree (CART)

Statistical Package for the Social Sciences (SPSS) was used to generate decision trees to classify Chla vertical distribution classes, using the classification and regression tree (CART) approach. CART [53] is a non-parametric procedure for predicting continuous dependent variable with categorical and/or continuous predictor variables where the data are partitioned into nodes on the basis of conditional binary responses to questions [54]. The maximum number of tree levels, minimum number of data sets for the parent nodes and minimum number of data sets for the child nodes was pre-defined. The CART model concluded after one of the stopping rules was satisfied. Stopping rules were set to prevent the model from over-fitting the training data. Accuracy estimates (classification and accuracy), the number of non-terminal nodes, and the number of leaves were used to describe the results. A decision tree to classify the Chla vertical profile classes was developed using the CART algorithm, taking into consideration remote sensing reflectance and wind speed.

Distance to shore provided an indirect influence on Chla vertical distribution, by changing the local wind speed with respect to wind direction and position with respect to the shore or island. The distance (fetch) of this effect will vary with respect to wind speed, wind direction, lake bathymetry and land topography [55]. To overcome this limit, we limited our analysis to those stations with a distance of more than 500 meters (2 MODIS band 1pixels) from the lakeshore. Therefore, the stations whose distance to shore was less than 500 meters were deleted prior to the decision tree development.

Ten-fold cross validation was used to estimate the performance of the decision tree; the data were partitioned into 10 equally (or nearly equally) sized sets. The training and validation were performed 10 times such that 1 segment is held-out for validation, while the remaining 9 are used for training within each iteration [56]. Classification accuracy assessment was accomplished through comparison of the resulting classification classes with *in situ* data through the use of a confusion or error matrix. The number of test sites required for accurate comparisons is discussed in detail below. From the confusion matrix, it was possible to compute a number of metrics that assess the accuracy of the classification, such as overall accuracy, Kappa coefficient, user's accuracy, and producer's accuracy. The traditional F-measure (F1 score), the harmonic mean of precision and recall, of each class was also calculated.

4. Results

4.1. Vertical Characteristics of Optically Active Substances

Chlorophyll-a concentrations (Chla) showed a large range (max/min ratio = 268.78) and variability (SD/mean = 2.62) (Table 1). The CV of Chla vertical profiles ranged from 4% to 239% with an average value of 67%. In the July campaign, the average CV was 20%, while the average CV was 133% in October, when the maximum Chla values at the water surface occurred.

Concentrations of SPIM and DOC had a lower range and variability compared to Chla (Table 1). Mean CVs of their vertical profiles were lower, 28% and 14%, respectively. The nine vertical profiles of DOC were regarded as vertically uniform in the May campaign.

Table 1. Measurements from sampling campaigns in Lake Chaohu in May, July, and October 2013: chlorophyll-a concentration (Chla, $\mu g \cdot L^{-1}$), concentration of suspended particulate inorganic matter (SPIM, $mg \cdot L^{-1}$), Dissolved Organic Carbon (DOC, $\mu g \cdot L^{-1}$) and coefficients of variation (CV, %) from the vertical profiles.

	Water Surface Value					CV of Vertical Profile			
	N	Mean	SD	Min	Max	N	Mean (%)	SD (%)	Min (%)	Max (%)
Chla	64	352.68	922.32	26.00	6988.29	64	67	68	4	239
SPIM	64	31.33	17.47	10.00	88.00	41	28	14	8	64
DOC	64	27.27	23.21	3.23	126.32	9	14	9	6	34

Note: Water surface values of DOC are mixed samples from water surface to 0.4 m, CV = SD/mean $\times 100\%$. The difference of number between vertical samples (SPIM and DOC) and those of water surface samples resulted from that SPIM and DOC profiles were not collected in some stations.

Table 2. Chla vertical profile classes and their fitting functions. R^2 is the coefficient of determination between the raw data and fitted data.

Class	N	Average CV	Chla Vertical Profile Class	Fitted Function	R^2	RMSE
Class 1	27	19.53%	uniform	$f_1(z) = C$	–	9.57
Class 2	9	29.25%	Gaussian	$f_2(z) = C_0 + \dfrac{h}{\sigma\sqrt{2\pi}} \exp[-\dfrac{1}{2}(\dfrac{z}{\sigma})^2]$	0.85	3.36
Class 3	12	97.73%	exponential	$f_3(z) = m_1 \times \exp(m_2 \times z)$	0.91	23.29
Class 4	16	163.60%	hyperbolic	$f_4(z) = n_1 \times z^{n_2}$	0.86	20.15

The Chla vertical profile datasets of the three surveys were processed to determine the most appropriate vertical profile class: vertically uniform (Class 1), Gaussian (Class 2), exponential (Class 3), and hyperbolic (Class 4) (Table 2). Chla profiles of Class 1 (Figure 2a) showed a featureless shape without maximums and with average CV 19.53%, and was classified as vertically uniform (Tables 2 and 3). Class 2 profiles (Figure 2b) exhibited a Gaussian distribution (N = 9) with a maximum Chla value at water surface, where C_0 was the background chlorophyll-a concentration, h was peak height for the Gaussian model and σ represented the peak width [31]. Class 3 profiles (Figure 2c) had an exponential distribution (N = 12) characterized by the scaling and exponent terms m_1 (280.98 ± 146.54) and m_2

(−3.15 ±2.79). Class 4 profiles (Figure 2d) followed a negative power function (N = 16) with the scaling and power terms, n_1 (29.01 ±19.82) and n_2 (−0.71 ±0.26). In Class 3 and 4, the Chla concentration at the water surface was usually higher than 100 μg/L, even up to several thousand μg/L (Table 3).
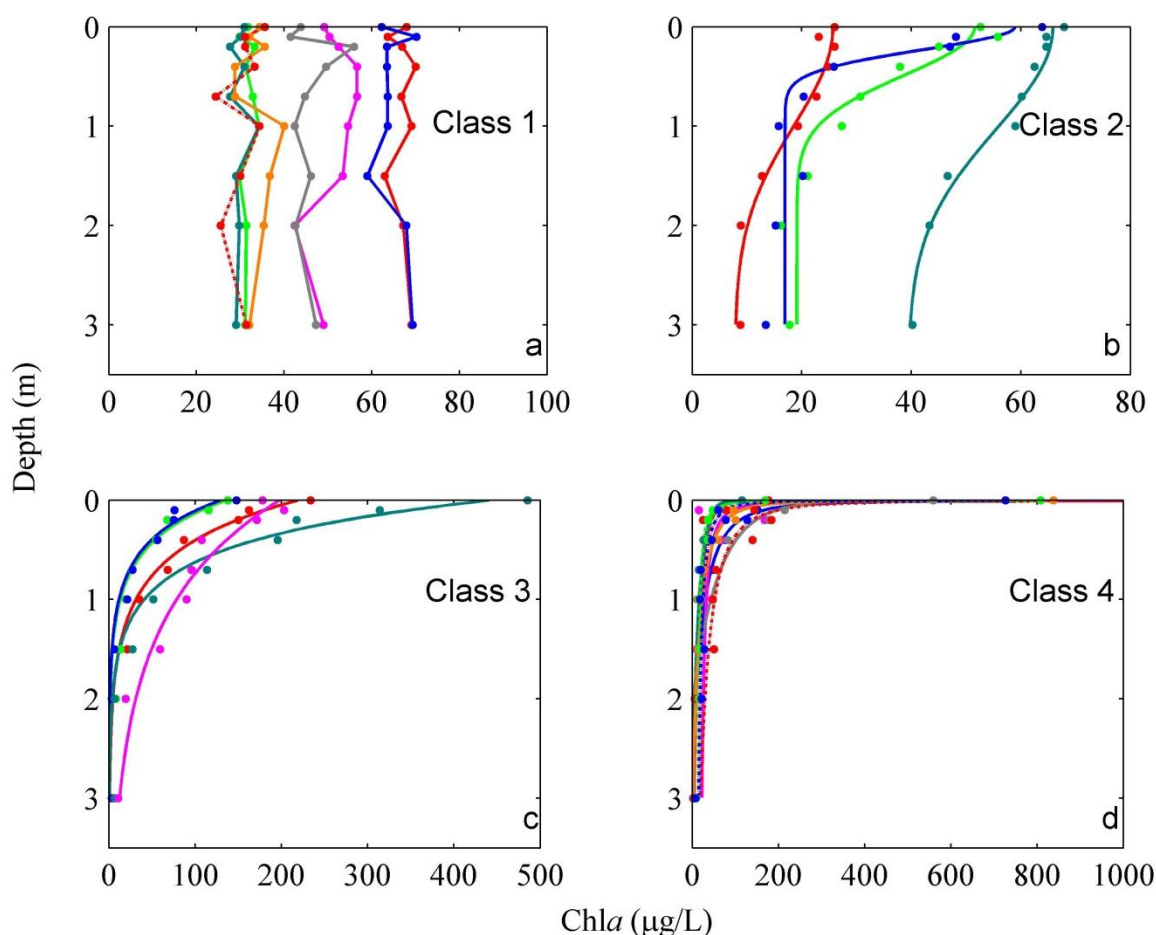

Figure 2. Selected Chlorophyll-a concentration profiles with original data points and fitted curves for the four vertical Chla profile classes: (**a**) Class 1, vertically uniform; (**b**) Class 2, Gaussian distribution; (**c**) Class 3, exponential distribution; and (**d**) Class 4, negative power function distribution.

Table 3. Results of fitted function parameters of chlorophyll-a vertical profile class.

Class	Parameters	Min	Max	Mean	Sd
	C_0	7.76	39.81	22.42	10.85
Class 2	σ	0.02	0.41	0.20	0.15
	h	1.48	75.59	34.08	26.94
Class 3	m_1	129.4	613.30	280.98	146.54
	m_2	−9.67	−0.64	−3.15	2.79
Class 4	n_1	12.46	80.63	29.01	19.82
	n_2	−1.10	−0.28	−0.71	0.26

4.2. Identification of Chla Vertical Profile Classes Using in situ Rrs

4.2.1. R_{rs} Response to Algae Vertical Profile Classes

The spectra curves (400–900 nm) collected in three field surveys were similar in shape and magnitude to the reflectance spectra reported for other highly eutrophic waters (Figure 3) [51,57]. *In situ R_{rs}* in the green region (500–600 nm) was much higher than that in the blue (400–500 nm) and red (600–700 nm) ranges with a maximum at around 550 nm. The variability of spectra in the red spectral range was associated with the combined effect of pure water absorption, scattering by particles, and the absorption and fluorescence of chlorophyll [58]. A minimum reflectance at 675 nm corresponded to the Chla absorption maximum. The peak between 690 and 710 nm resulted from both backscattering and a minimum in absorption by optically active constituents, including pure water. A much larger peak centered at 705 nm has been used as an indicator of dense surface phytoplankton blooms [59]. The dominant effect of absorption in determining the form of this spectrum is indicated by the deep minima at 550 nm and 665 nm, and the lesser minimum at 630 nm, due to chlorophyll a and auxiliary pigments [58]. The reflectance in the near-infrared (NIR) range (700–900 nm) varied widely with a small maximum at 814 nm, due to the high backscattering suspended particles with SPIM > 10 mg/L. Floating plankton like cyanobacteria increases reflectance at wavelengths > 700 nm, in a similar manner to terrestrial plants [58].
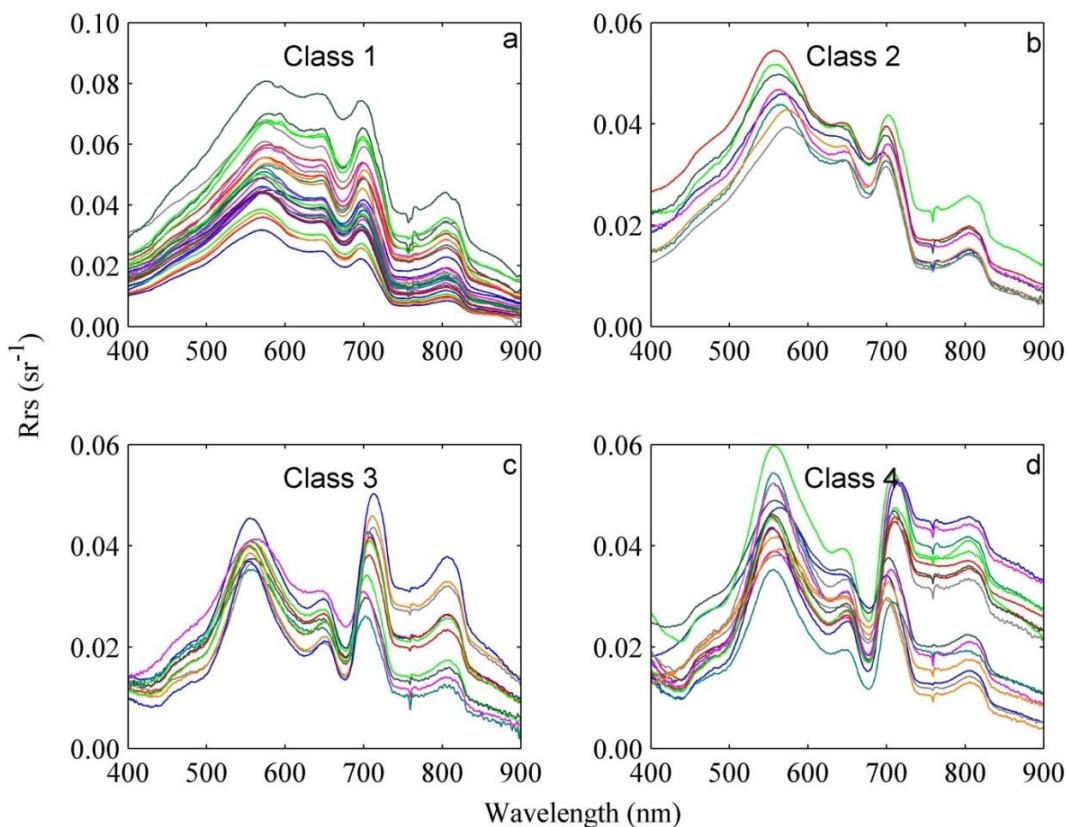

Figure 3. Spectral curves of *Rrs* (λ) from 400 nm to 900 nm of the four vertical Chla profile classes ((**a**) Class 1, vertically uniform; (**b**) Class 2, Gaussian distribution; (**c**) Class 3, exponential distribution; and (**d**) Class 4, negative power function) in Lake Chaohu in May, July and October 2013.

R_{rs} of Class 3 and 4 showed two minima at 440 and 625 nm, which were associated to Chla and phycocyanin absorption, respectively (Figure 3). The R_{rs} in the NIR range of Class 3 and 4 was consistently higher, and was associated to dense surface phytoplankton in bloom conditions [59], and scattering by particulate matter [60]. Based on the above characteristics, spectral differences between Class 1 and 2 and Class 3 and 4 in characterized bands of blue, red and infra-red range indicated a more presence of cyanobacteria and algal blooms in the latter.

Figure 4. (**a**) Outputs of three algal bloom indices; (**b**) the corresponding wind speed (m/s); and (**c**) distance of stations to shore (km). NDVI: Normalized Difference Vegetation Index; CSI: Chlorophyll Spectral Index; NDBI: Normalized Difference algal Bloom Index.

Comparing the algal bloom indexes based on *in situ* R_{rs} (350–1050 nm), $NDBI_{Rrs}$ performed best in distinguishing floating algal scum (lower SD) between Class 1 and 2 and Class 3 and 4 with less variability with respect to NDVI and CSI (Figure 4a). This may have resulted from variations of R_{rs} in the NIR band using in NDBI and CSI, which was caused by high concentrations of suspended matter (Figure 3c,d). There was a general increase of NDBI with the Chla vertical profile classes from Class 1

to Class 4 (Figure 4a). At low NDBI values, Chla vertical profiles were vertically uniform (Class 1) or had a Gaussian distribution (Class 2). At high NDBI, Chla vertical distributions exhibited exponential and negative power function, which were resulted from floating algae gathering near the water surface.

4.2.2. Wind Speed of Different Chla Vertical Profile Classes

The classes of Chla vertical profiles were related to wind speed, with a lower wind speed from Class 1 to Class 4 (Figure 4b). The results confirmed that the vertical distribution of algae followed a uniform distribution (Class 1) at high wind speed, with CV < 30%. At lower wind speeds, there is no obvious separation between Chla vertical profile Class 2 and Class 3, hence the need to use both wind speed and estimated Chla indices to identify profile classes.

Additionally, distance to the shore indirectly influenced Chla vertical profiles. For example, samples collected in the lee side of the Mushan island showed a Gaussian vertical distribution of Chla (Class 2), while Class 1 dominated in the open lake area (Figure 4c). Distance to shore provided an indirect influence on Chla vertical distribution, by changing the local wind speed with respect to wind direction and position with respect to the shore or island. To overcome this limit, we limited our analysis to those stations with a distance of more than 500 meters (2 MODIS pixels at band 1) from the lakeshore in the following sections.

4.2.3. Decision Tree

The vertical profiles decision tree was developed based on data from remote sensing reflectance indices and wind speed (Figure 5). The initial separation was based on Chla concentrations, $NDBI_{Rrs} < 0.25$, where Chla vertical profiles followed either a vertically uniform (Class 1) or Gaussian (Class 2) distribution. When $NDBI_{Rrs}$ was greater than 0.25, profiles were well represented by either an exponential (Class 3) or hyperbolic (Class 4) distribution. Surface wind speeds were then used to further separate vertical profile classes. Class 1 profiles occurred when wind speed was >2.75 m/s, while Class 2 profiles occurred when wind speeds were <2.75 m/s. Class 3 profiles occurred for wind speeds > 1.75 m/s, while Class 4 profiles occurred for wind speeds < 1.75 m/s.
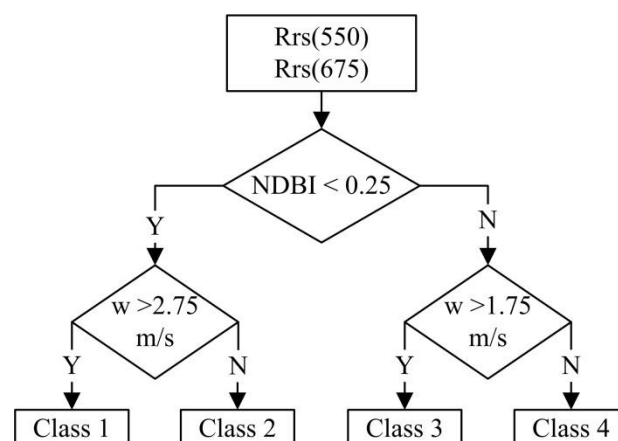

Figure 5. Decision tree (CART) of identifying Chla vertical profile class based on $NDBI_{Rrs}$ and wind speed, where $NDBI_{Rrs}$ is the NDBI index derived from *in situ* measurements, and "w" represents wind speed.

Results for the analyses followed a confusion matrix, where numbers on the diagonals corresponded to correctly classified cases, and off-diagonal entries represent misclassification. The classification accuracy of the decision trees was evaluated (Table 4) using 10-fold cross validation method with overall accuracy 79%, and Kappa coefficient 0.71. The F1-measure of each class is 88%, 47%, 80% and 84%, respectively. Thresholds of the decision tree were evaluated with changing value, and the results showed that the NDBI and wind speed in the decision tree (Figure 5) derived the highest accuracy (Figure 6).
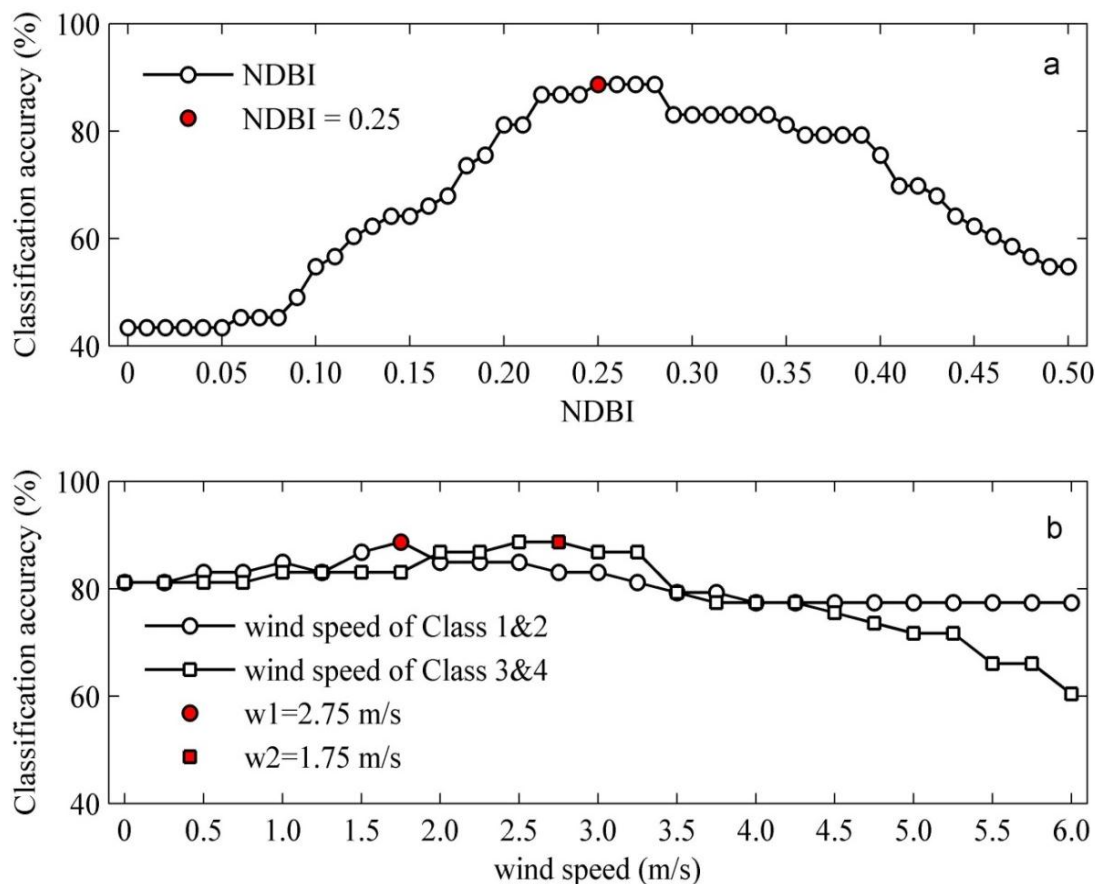

Figure 6. Sensitivity analysis of threshold value of the CART decision tree: (**a**) classification accuracy (%) with changing NDBI; and (**b**) classification accuracy (%) with changing wind speed. The red markers (NDBI = 0.25, w1 = 2.75 m/s, w2 = 1.75 m/s) are the thresholds used in this study with highest classification accuracy.

As the most important environmental force of Chla vertical distribution, wind speed had a direct and rapid influence on Chla vertical profile. Lake Chaohu has few surrounding land masses with high elevations, allowing the assumption that the wind speed across the lake remained relatively uniform over two hours, the association of measurements and satellite overpass (±1 h) was made. After analyzing the wind speed measured in June 2014, we found that the variation of wind speed was small and wind direction was similar throughout the lake. Average wind speed value obtained from *in situ* measurements was 2.8 m/s (2.4–4 m/s) and 0.58 m/s (0–1.4 m/s) on 11 and 13 June, respectively.

Table 4. Confusion matrix for Chla vertical profile classes of Chaohu Lake, 2013. Note that "Overall Acc." is overall accuracy of validation, "Kappa" represents Kappa coefficient.

		Actual					User's
		Class 1	**Class 2**	**Class 3**	**Class 4**	**Totals**	**Accuracy**
	Class 1	19	3	0	0	22	86%
	Class 2	2	4	1	2	9	44%
Predicted	Class 3	0	0	6	1	7	86%
	Class 4	0	1	1	13	15	87%
	Totals	21	8	8	16	53	
Producer's Accuracy		90%	50%	75%	81%	**Overall Acc.**	79%
F1-measure		88%	47%	80%	84%	**Kappa**	0.71

4.3. Identification of Chla Vertical Profile Classes Using MODIS R_{rc} Data

The relationship between R_{rs} of MODIS images and *in situ* R_{rs} requires information on the spectral response of MODIS wavebands and the atmospheric conditions above the study lake. The correlation between *in situ* NDBI_{Rrs} and $\text{NDBI}_{Rrs,\text{MODIS}}$ was high, $R^2 = 0.98$ (N = 67, Figure 7a, Equation (12)). Converting the *in situ* NDBI threshold (0.25) to a MODIS R_{rs} threshold, the value of 0.15 was determined (Equation (12)).

$$\text{NDBI}_{Rrs,\text{MODIS}} = 0.72 \times \text{NDBI}_{Rrs} - 0.03 \tag{12}$$

MODIS R_{rs} was difficult to determine due to complex atmospheric conditions and the lack of available atmospheric data necessary for atmospheric correction. Non-coupling ocean-atmosphere radiative transfer simulation using different aerosol types and optical thicknesses (extracted from SeaDAS LUTs, $\tau_a(555) = 0.1, 0.5, 1.0$) was used to explore the relationships between $\text{NDBI}_{Rrc,\text{MODIS}}$ and $\text{NDBI}_{Rrs,\text{MODIS}}$. For $\text{NDBI}_{Rrs,\text{MODIS}} = 0.15$, the $\text{NDBI}_{Rrc,\text{MODIS}}$ was 0.125, indicating a near independence of aerosol types and thicknesses (Figure 7b). This threshold of 0.125 was used to detect algal blooms in R_{rc} data from MODIS images.

There were eight cloud free MODIS images coincident with the sampling campaigns on 28 May, 24 July, and 10–12 October, 2013. The MODIS $\text{NDBI}_{Rrc,\text{MODIS}}$ threshold value (0.125) with the average wind speed in a two hour window around the satellite data acquisition were used to determine the Chla vertical profile class throughout the lake (Figure 8a–f). Class 1 distributions were observed nearly all parts of the lake on 24 July 2013 with wind speed 5 m/s. The wind speed at 5:30 a.m. GMT had fallen to 1.75 m/s, which produced a reduction in mixing and increased stability of water column. This increased stability would have been essential to enable algae to migrate back to near-surface waters, and the Chla vertical distribution of large part of MODIS images changed from Class 1 to Class 2 (Figure 8d). On 10 October 2013, Class 4 dominated nearly 80% of the lake in the morning, changing to Class 3 with increased wind in the afternoon. The Chla vertical profile class was similar for both Terra and Aqua images taken on the same day in similar wind conditions, and dominant class changed over two consecutive days when wind conditions changed.
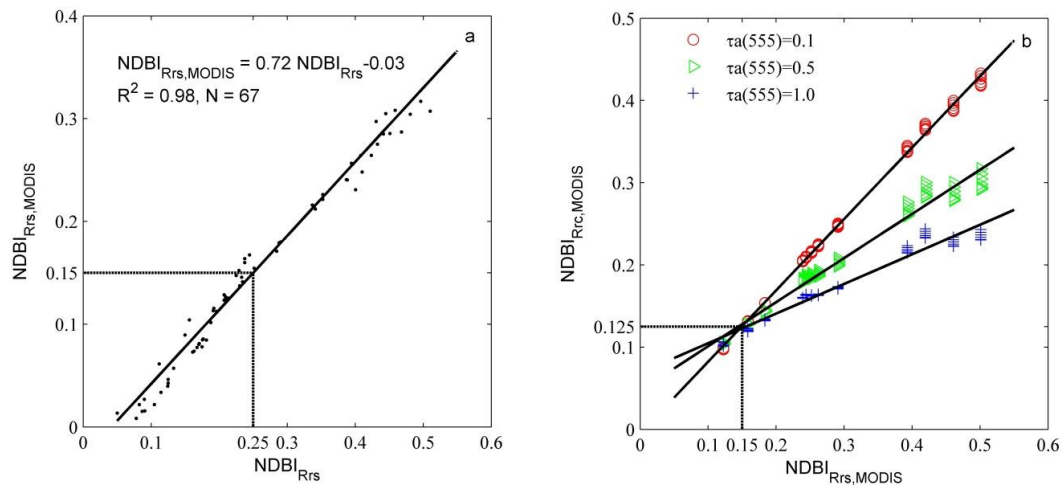
Figure 7. Relationships between (**a**) NDBI$_{Rrs}$ and NDBI$_{Rrs}$,MODIS; and (**b**) NDBI$_{Rrc}$,MODIS and NDBI$_{Rrs}$,MODIS at different atmospheric conditions (seven aerosol types and three optical thickness) through model simulations based on SeaDAS Look up table. $\tau_a(555)$ is the aerosol optical thickness at 555 nm. Threshold value of NDBI$_{Rrs}$,MODIS to detect algal bloom is 0.15, while the corresponding average NDBI$_{Rrc}$ value is 0.125, which is independent of aerosol type and optical thickness. Note that NDBI$_{Rrs}$ is NDBI index derived from *in situ* measurements, NDBI$_{Rrs}$,MODIS is resampled NDBI$_{Rrs}$ according to spectral response function of MODIS, and NDBI$_{Rrc}$,MODIS is NDBI index derived from MODIS images.
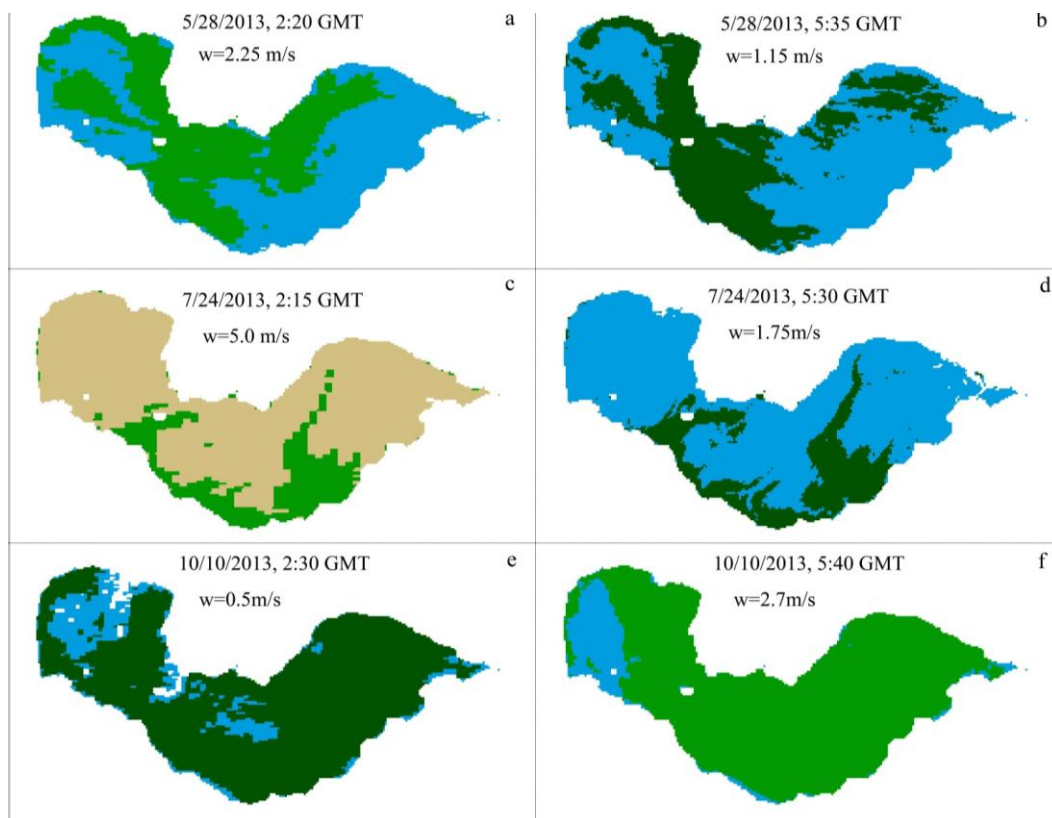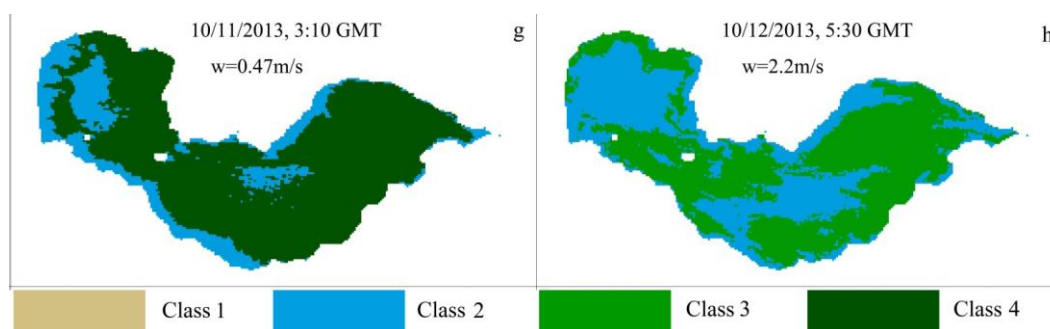

Figure 8. *Cont.*

Figure 8. Chla vertical profile class for Lake Chaohu estimated from MODIS measurements using NDBI$_{Rrc,MODIS}$ (0.125) and measured wind speed on (**a,b**) 28 May 2013, (**c,d**) 24 July 2013, (**e,f**) 10 October 2013, (**g**) 11 October 2013, and (**h**) 12 October 2013.

Previous studies used a narrow time window for coincident *in situ* and satellite data records (*i.e.* no more than ±3 h) [61]. Considering the effects of temporal variability on Chla vertical distribution, nearly concurrent measurements (±1 h) of satellite and *in situ* measurements were used to validate the assigned Chla vertical distribution class. Of the 12 matching pairs, 10 samples had the same Chla vertical profile class as *in situ* measurements (Table 5). For two samples belonged to Class 2 *in situ* measurements, the MODIS estimated Chla vertical class was incorrectly assigned (Class 1), indicating an error related to wind speed. These two stations were located on the lee side of Mushan island, where wind speed value is lower over than that measured in the other areas of the lake.

Table 5. Validation of Chla vertical profile class derived using concurrent *in situ* and MODIS based measurements with different wind speed on 28 May, 24 July, and 11–12 October 2013.

Day	GMT Time	Wind Speed (m/s)	NDBI$_{RRC,MODIS}$	Chla Vertical Class	
				in situ Class	Estimated Class
28 May 2013	1:00	2.25	0.118	2	2
28 May 2013	2:25	2.25	0.159	4	4
28 May 2013	3:25	2.25	0.160	4	4
24 July 2013	1:55	5.00	0.101	2	1
24 July 2013	2:25	5.00	0.101	2	1
24 July 2013	2:55	5.00	0.101	1	1
11 October 2013	2:30	0.43	0.135	4	4
11 October 2013	3:10	0.43	0.135	4	4
11 October 2013	3:50	0.43	0.135	4	4
12 October 2013	1:30	2.20	0.101	2	2
12 October 2013	2:05	2.20	0.101	2	2
12 October 2013	2:40	2.20	0.101	2	2

4.4. Relationships of Structure Parameters to Remotely Sensed Data

Open ocean studies show that it is possible to extract and then parameterize typical profiles as a function of surface Chla concentration (Chla$_s$) [35]. We built the relationships of vertical structure parameters to surface Chla concentration (using remote sensing reflectance) and tried

deriving the vertical profile parameters using remote sensing data or easily obtained optical parameters. Figures 9 and 10 showed the best relationships. For Class 2, C_0 was difficult to retrieve from water surface parameters, but had good correlations with absorption coefficient of CDOM ($a_g(443)$). This relationship required additional information to further explore (Figure 9a,b). Parameters h and σ changed inversely with $R_{rs}(709)$ ($R^2 = 0.56, 0.55$, Figure 9c,d). For Class 3 and 4, the quantitative relationships between vertical structure parameters and R_{rs} did not perform well ($R^2 < 0.5$), while the relationships between m_1, n_1, n_2 and surface Chla concentrations performed relatively well (Figure 10).
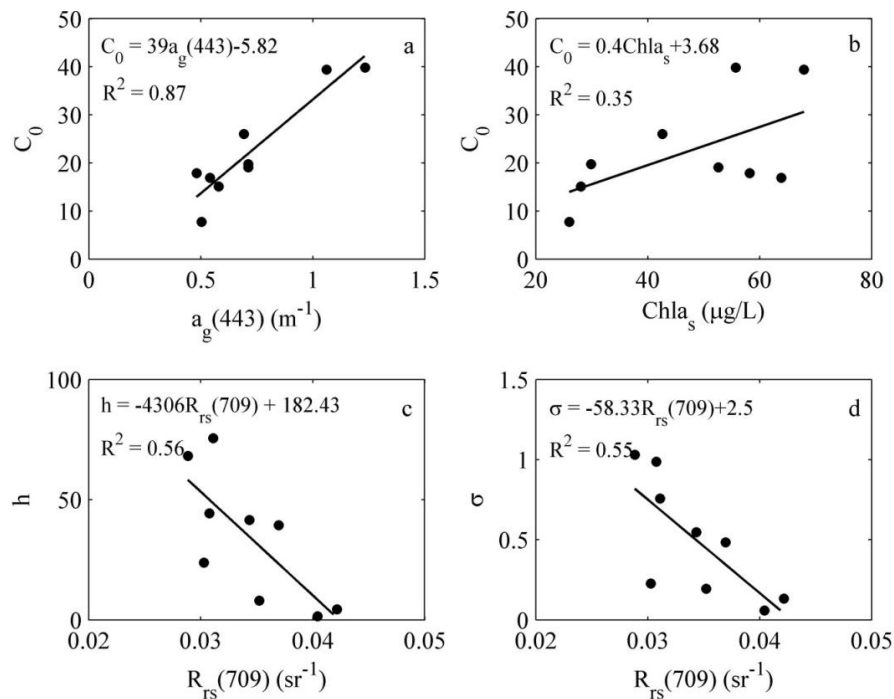

Figure 9. Relationships between structure parameters of Chla vertical Class 2 and surface variables: (**a**) C_0 *versus* $a_g(443)$, (**b**) C_0 *versus* Chla$_s$, (**c**) h *versus* $R_{rs}(709)$, and (**d**) σ *versus* $R_{rs}(709)$. Note that Chla$_s$ is average Chla concentration from water surface to 0.4 m.
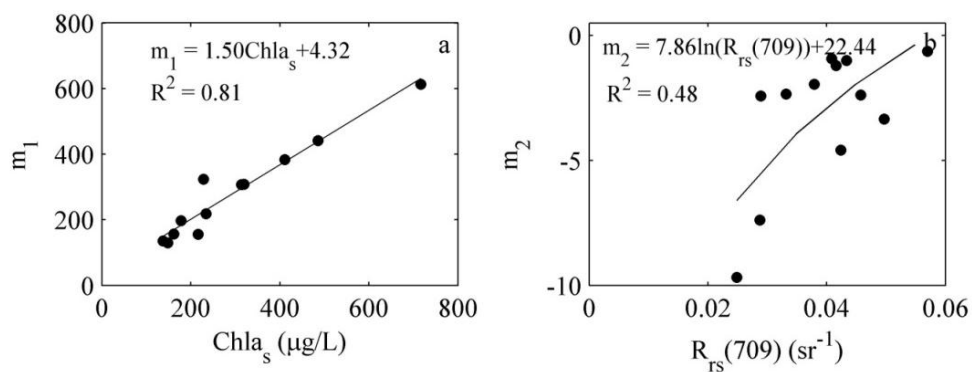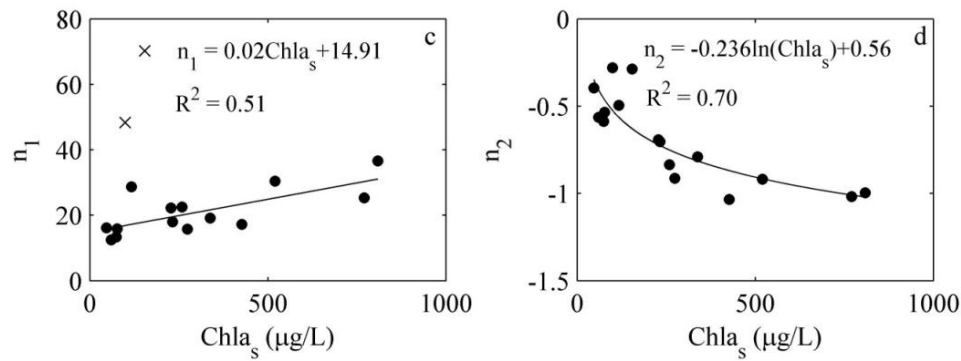

Figure 10. *Cont.*

Figure 10. Relationships between structure parameters of Chla vertical Class 3 and 4 and surface variables: (**a**) m_1 of Class 3 *versus* $R_{rs}(709)$, (**b**) m_2 of Class 3 *versus* $Chla_s$, (**c**) n_1 of Class 4 *versus* $Chla_s$, and (**d**) n_2 of Class 4 *versus* $Chla_s$. Note that $Chla_s$ is average Chla concentration from water surface to 0.4 m.

5. Discussion

Identifying the vertical distribution of phytoplankton is essential to accurately access biomass and evaluate lake trophic conditions. Four Chla vertical profile classes with different structure parameters were observed in Lake Chaohu, and the results indicated that the Chla vertical class may change in short times. As different Chla vertical profile classes may occur at the same time across the lake, understanding the Chla vertical profile classes is necessary to estimate the vertical profile of Chla in complex inland waters. Remote sensing reflectance spectra contain information about the optical properties of components within the effective upwelling depth, below which the optical properties of the water body no longer directly influence the water leaving radiance [62]. Typical of eutrophic shallow lakes, the effective upwelling depth in Lake Chaohu was limited, requiring additional non-optical information to determine Chla vertical profile classes. By combining NDBI and wind speed, a decision tree was developed to estimate the vertical distribution profile of phytoplankton biomass. In the decision tree, optical conditions of the surface water were the first parameter to distinguish the classes, with wind speed as the key parameter in final classification of the Chla vertical profiles. The results suggested that Chla vertical distribution was spatially heterogeneous and also helped to explain why the area and intensity of algal blooms changed over a short time.

Ideally, the R_{rs} of every pixel should initially be derived from the satellite images by atmospheric correction, and then used in the $NDBI_{Rrs}$ algorithm to detect algal blooms. Unfortunately, the assumption of black water at NIR [63] or SWIR [64] bands is invalid in many eutrophic inland lakes, which leads to the difficulty in correcting aerosol effects. There is currently no reliable aerosol correction that can be used to derive R_{rs} accurately in optically complex inland lakes [52]. Moreover, the mixed pixel problem is complicated in MODIS images with 250 m spatial resolution, which results in uncertainty of comparing $NDBI_{Rrs}$ derived from *in situ* measurements to MODIS images directly.

A robust index to detect and quantify floating algae must be relatively stable against changing environmental and observational conditions. The FAI is the difference between Rayleigh-corrected reflectance in the NIR and a baseline formed by the red and SWIR bands, and has been shown to be an effective index to detect floating algae by removing most of the atmospheric effects [52]. Due to the

limitations of spectral range of R_{rs} data measured by ASD field spectrometer (350–1050 nm), FAI of each sample could not be derived by *in situ* R_{rs}. We developed the NDBI based on *in situ* R_{rs} data and MODIS R_{rc} images. The sensitivity of $NDBI_{Rrc,MODIS}$ was compared to FAI estimates of algal boom coverage. The left panels in Figure 11 show the representative RGB images where blooms are present, and the middle and right panels show the corresponding NDBI and FAI distributions, respectively. $NDBI_{Rrc,MODIS}$ and FAI provided similar patterns (Figure 11), especially when blooms occur. However, in high turbid waters (Figure 11j), FAI provided an erroneous bloom coverage, indicating that $NDBI_{Rrc,MODIS}$ is more robust when high suspended inorganic matter is present. For example, in the northwest of the MODIS images on 14 April 2013, areas circled in red with high FAI (Figure 11k) and low NDBI (Figure 11l) appear to be caused by high turbidity.
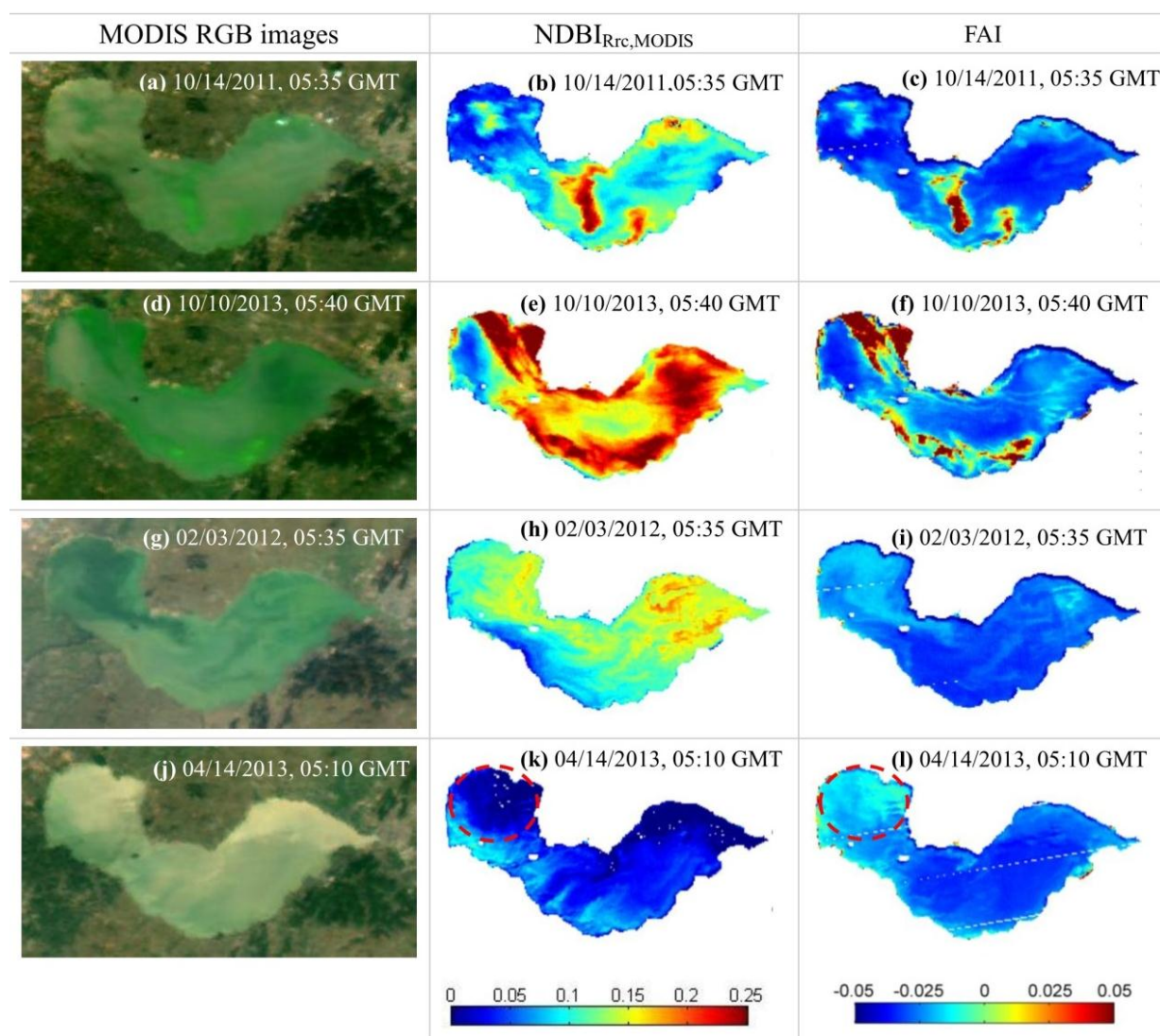

Figure 11. Comparison of FAI and NDBI derived by MODIS images on (**a**–**c**) 14 October 2011, (**d**–**f**) 10 October 2013, (**g**–**i**) 03 February 2012, and (**j**–**l**) 14 April 2013, respectively. The high NDBI and FAI values (dimensionless) in the middle of the lake in (b) and (c) indicate a cyanobacterial bloom. Areas circled in red with high FAI (k) and low NDBI (l) appear to be caused high turbidity. The dashed lines represent MODIS detector error.

An increased spatial resolution of the wind field would provide a more accurate estimation of vertical algal distribution. Satellite based microwave scatterometers, SCATs, have a limited spatial resolution (25 km) with an accuracy of wind speed of ±2 m/s [65]. SARs have higher spatial resolution (sub-kilometer) with all-weather capability and large spatial coverage [66]. However, the accuracy of SAR wind speed at low winds (<5 m/s) is low, as low wind speeds produce little backscatter [65]. Our wind speed thresholds were below this minimum value, thus, these approaches are unlikely to provide better information than shore-based measurements. In-lake meteorological stations were not available in Lake Chaohu at the time of this study. Measurements from a station located at the north shore of Lake Chaohu were the best source of continuous data. A two-hour average wind speed from *in situ* measurement would be the most accurate information to estimate Chla vertical distribution classes in this shallow eutrophic lake.

Vertical mixing in lakes is strongly influenced by wind direction topography and bathymetry. This spatial variability represents a limit to this approach, in particular when applied to other lakes. For lakes where nearshore topography, local bathymetry or internal waves create mixing conditions that are not directly wind related [67,68], an accurate determination of local wind conditions should be determined.

The determination of vertical profile classes based on R_{rs} and wind speed showed accurate results in Lake Chaohu, but was based on a relatively small dataset of 64 profiles. A larger *in situ* dataset would help to better define the functional coefficients for the nonlinear profiles. Radiative transfer models (e.g., Hydrolight) or other modeling approaches (e.g., optimal algorithm, artificial neural network) could be used to further study relationships between upwelling irradiance and vertical conditions, and will be further explored. The capability to associate non-homogeneous Chla profiles to water surface Chla concentration from satellite color sensors would improve the estimates of phytoplankton biomass in large lakes and coastal areas.

6. Conclusion

Identification of vertical profile class of phytoplankton is essential to evaluate the vertical distribution of phytoplankton, algal biomass, and explain the short-term changes of algal bloom conditions. Our dataset, acquired over three surveys (28 May, 19–24 July and 10–12 October) in Lake Chaohu, China, showed vertical profiles of Chla with four different profile classes (vertically uniform, Gaussian, exponential and hyperbolic). Using the combined information provided by the NDBI index and local wind conditions, a decision tree was built to identify the vertical profile classes using both *in situ* R_{rs} data and MODIS R_{rc} images. The threshold value was 0.025 and 0.125 for NDBI$_{Rrs}$ and NDBI$_{Rrc,MODIS}$, respectively. Threshold value of wind speed to distinguish vertically uniform from Gaussian distribution was 2.75 m/s, and to distinguish exponential from hyperbolic distribution was 1.75 m/s. The classification accuracy of the decision tree was evaluated using 10-fold cross validation method with overall accuracy 79%, and Kappa coefficient 0.71. The dominant Chla profile class of each pixel varies with time, and can be determined by the above rules.

The NDBI index uses commonly available green and red wavebands and is therefore easily applied to other satellite data. By using the same approach with Geostationary Ocean Color Imager (GOCI), Visible Infrared Imaging Radiometer (VIIRS) or other satellite data, continuous assessment of the lake's algal biomass and primary production can be made. In broadly similar hydrological (shallow, open lake), trophic (eutrophic) and topographic (non-mountain) conditions, the approach used in this study can be

applied. Application to new lakes requires that similar decision tree can be developed and calibrated using local *in situ* datasets. For lakes where nearshore morphometry or internal waves create mixing conditions that are not directly wind related [67,68], higher resolution data of local wind conditions and hydrology would be required.

Qualitative and quantitative relationships between vertical parameters and surface variables provide new insights, but more *in situ* data or simulated large dataset are required to determine Chla concentrations at different water depths. Further validation is required to test the applicability of the thresholds developed for Lake Chaohu for other satellite sensors and other waterbodies, and to validate the robustness of NDBI in the effects of atmosphere, cloud and mixed pixels.

Acknowledgments

Author Contributions

Kun Xue assembled and analyzed the data. Yuchao Zhang, Hongtao Duan and Ronghua Ma contributed to the data collection and analysis of the data. Minwei Zhang helped the ocean–atmosphere radiative transfer simulation. Kun Xue, Yuchao Zhang and Steven Loiselle developed discussion and revised the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Duan, H.; Ma, R.; Xu, X.; Kong, F.; Zhang, S.; Kong, W.; Hao, J.; Shang, L. Two-decade reconstruction of algal blooms in China's Lake Taihu. *Environ. Sci. Technol.* **2009**, *43*, 3522–3528.
2. Kahru, M.; Savchuk, O.P.; Elmgren, R. Satellite measurements of cyanobacterial bloom frequency in the Baltic Sea: Interannual and spatial variability. *Mar. Ecol. Prog. Ser.* **2007**, *343*, 15–23.
3. Siswanto, E.; Ishizaka, J.; Tripathy, S.C.; Miyamura, K. Detection of harmful algal blooms of Karenia mikimotoi using MODIS measurements: A case study of Seto-Inland Sea, Japan. *Remote Sens. Environ.* **2013**, *129*, 185–196.
4. Tomlinson, M.C.; Stumpf, R.P.; Ransibrahmanakul, V.; Truby, E.W.; Kirkpatrick, G.J.; Pederson, B.A.; Vargo, G.A.; Heil, C.A. Evaluation of the use of SeaWiFS imagery for detecting Karenia brevis harmful algal blooms in the eastern Gulf of Mexico. *Remote Sens. Environ.* **2004**, *91*, 293–303.

5. Jia, X.; Shi, D.; Shi, M.; Li, R.; Song, L.; Fang, H.; Yu, G. Formation of cyanobacterial blooms in Lake Chaohu and the photosynthesis of dominant species hypothesis. *Acta Ecol. Sin.* **2011**, *31*, 2968–2977.

6. Kutser, T.; Metsamaa, L.; Strömbeck, N.; Vahtmäe, E. Monitoring cyanobacterial blooms by satellite remote sensing. *Estuar. Coast. Shelf Sci.* **2006**, *67*, 303–312.

7. Odermatt, D.; Pomati, F.; Pitarch, J.; Carpenter, J.; Kawka, M.; Schaepman, M.; Wüest, A. MERIS observations of phytoplankton blooms in a stratified eutrophic lake. *Remote Sens. Environ.* **2012**, *126*, 232–239.

8. Bresciani, M.; Adamo, M.; De Carolis, G.; Matta, E.; Pasquariello, G.; Vaičiūtė, D.; Giardino, C. Monitoring blooms and surface accumulation of cyanobacteria in the Curonian Lagoon by combining MERIS and ASAR data. *Remote Sens. Environ.* **2014**, *146*, 124–135.

9. Duan, H.; Ma, R.; Zhang, Y.; Loiselle, S.A.; Xu, J.; Zhao, C.; Zhou, L.; Shang, L. A new three-band algorithm for estimating chlorophyll concentrations in turbid inland lakes. *Environ. Res. Lett.* **2010**, 5, doi:10.1088/1748-9326/5/4/044009.

10. Song, K.; Li, L.; Tedesco, L.P.; Li, S.; Duan, H.; Liu, D.; Hall, B.E.; Du, J.; Li, Z.; Shi, K.; *et al.* Remote estimation of chlorophyll-a in turbid inland waters: Three-band model *versus* GA-PLS model. *Remote Sens. Environ.* **2013**, *136*, 342–357.

11. Duan, H.; Ma, R.; Hu, C. Evaluation of remote sensing algorithms for cyanobacterial pigment retrievals during spring bloom formation in several lakes of East China. *Remote Sens. Environ.* **2012**, *126*, 126–135.

12. Lou, X.; Hu, C. Diurnal changes of a harmful algal bloom in the East China Sea: Observations from GOCI. *Remote Sens. Environ.* **2014**, *140*, 562–572.

13. Cao, H.; Kong, F.; Luo, L.; Shi, X.; Yang, Z.; Zhang, X.; Tao, Y. Effects of wind and wind-induced waves on vertical phytoplankton distribution and surface blooms of *Microcystis aeruginosain* Lake Taihu. *J. Freshw. Ecol.* **2006**, *21*, 231–238.

14. Silulwane, N.F.; Richardson, A.J.; Shillington, F.A.; Mitchell-Innes, B.A. Identification and classification of vertical chlorophyll patterns in the Benguela upwelling system and Angola-Benguela front using an artificial neural network. *South Afr. J. Mar. Sci.* **2010**, *23*, 37–51.

15. Ndong, M.; Bird, D.; Nguyen-Quang, T.; de Boutray, M.L.; Zamyadi, A.; Vincon-Leite, B.; Lemaire, B.J.; Prevost, M.; Dorner, S. Estimating the risk of cyanobacterial occurrence using an index integrating meteorological factors: Application to drinking water production. *Water Res.* **2014**, *56*, 98–108.

16. George, D.; Edwards, R. The effect of wind on the distribution of chlorophyll a and crustacean plankton in a shallow eutrophic reservoir. *J. Appl. Ecol.* **1976**, *13*, 667–690.

17. Webster, I.T.; Hutchinson, P.A. Effect of wind on the distribution of phytoplankton cells in lakes revisited. *Limnol. Oceanogr.* **1994**, *39*, 365–373.

18. Kanoshina, I.; Lips, U.; Leppänen, J.-M. The influence of weather conditions (temperature and wind) on cyanobacterial bloom development in the Gulf of Finland (Baltic Sea). *Harmful Algae* **2003**, *2*, 29–41.

19. Huang, C.; Li, Y.; Yang, H.; Sun, D.; Yu, Z.; Zhang, Z.; Chen, X.; Xu, L. Detection of algal bloom and factors influencing its formation in Taihu Lake from 2000 to 2011 by MODIS. *Environ. Earth Sci.* **2013**, *71*, 3705–3714.

20. Tsujimura, S.; Tsukada, H.; Nakahara, H.; Nakajima, T.; Nishino, M. Seasonal variations of Microcystis populations in sediments of Lake Biwa, Japan. *Hydrobiologia* **2000**, *434*, 183–192.

21. Walsby, A.; Booker, M. Changes in buoyancy of a planktonic blue-green alga in response to light intensity. *Br. Phycol. J.* **1980**, *15*, 311–319.

22. Kromkamp, J.; van den Heuvel, A.; Mur, L.R. Formation of gas vesicles in phosphorus-limited cultures of *Microcystis aeruginosa*. *J. Gen. Microbiol.* **1989**, *135*, 1933–1939.

23. Klemer, A.; Feuillade, J.; Feuillade, M. Cyanobacterial blooms: carbon and nitrogen limitation have opposite effects on the buoyancy of Oscillatoria. *Science* **1982**, *215*, 1629–1631.

24. 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. *Limnol. Oceanogr.* **2010**, *55*, 2025–2036.

25. Morel, A.; Berthon, J.-F. Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications. *Limnol. Oceanogr.* **1989**, *34*, 1545–1562.

26. Xiu, P.; Liu, Y.; Tang, J. Variations of ocean colour parameters with nonuniform vertical profiles of chlorophyll concentration. *Int. J. Remote Sens.* **2008**, *29*, 831–849.

27. Gordon, H.R.; Clark, D.K. Remote sensing optical properties of a stratified ocean: An improved interpretation. *Appl. Opt.* **1980**, *19*, 3428–3430.

28. Stramska, M.; Stramski, D. Effects of a nonuniform vertical profile of chlorophyll concentration on remote-sensing reflectance of the ocean. *Appl. Opt.* **2005**, *44*, 1735–1747.

29. Kutser, T.; Metsamaa, L.; Dekker, A.G. Influence of the vertical distribution of *cyanobacteria* in the water column on the remote sensing signal. *Estuar. Coast. Shelf Sci.* **2008**, *78*, 649–654.

30. Sathyendranath, S.; Platt, T.; Caverhill, C.M.; Warnock, R.E.; Lewis, M.R. Remote sensing of oceanic primary production: Computations using a spectral model. *Deep Sea Res. Part A. Oceanogr. Res. Pap.* **1989**, *36*, 431–453.

31. André, J.-M. Ocean color remote-sensing and the subsurface vertical structure of phytoplankton pigments. *Deep Sea Res. Part A. Oceanogr. Res. Pap.* **1992**, *39*, 763–779.

32. Millán-Núñez, R.; Alvarez-Borrego, S.; Trees, C.C. Modeling the vertical distribution of chlorophyll in the California Current System. *J. Geophys. Res.* **1997**, *102*, doi:10.1029/97JC00079.

33. Hidalgo-Gonzalez, R.M.; Alvarez-Borrego, S. Chlorophyll profiles and the water column structure in the Gulf of California. *Oceanol. Acta* **2001**, *24*, 19–28.

34. D'Alimonte, D.; Zibordi, G.; Kajiyama, T.; Berthon, J.-F. Comparison between MERIS and regional high-level products in European seas. *Remote Sens. Environ.* **2014**, *140*, 378–395.

35. Uitz, J.; Claustre, H.; Morel, A.; Hooker, S.B. Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *J. Geophys. Res.* **2006**, *111*, doi:10.1029/2005JC003207.

36. Chen, X.; Yang, X.; Dong, X.; Liu, E. Environmental changes in Chaohu Lake (Southeast, China) since the mid 20th century: The interactive impacts of nutrients, hydrology and climate. *Limnol.—Ecol. Manag. Inland Waters* **2013**, *43*, 10–17.

37. Lodhi, M.A.; Rundquist, D.C. A spectral analysis of bottom-induced variation in the colour of Sand Hills lakes, Nebraska, USA. *Int. J. Remote Sens.* **2001**, *22*, 1665–1682.

38. Ma, R.; Tang, J.; Dai, J. Bio-optical model with optimal parameter suitable for Taihu Lake in water colour remote sensing. *Int. J. Remote Sens.* **2006**, *27*, 4305–4327.

39. Shang, G.P.; Shang, J.C. Spatial and temporal variations of eutrophication in Western Chaohu Lake, China. *Environ. Monit. Assess.* **2007**, *130*, 99–109.

40. Chen, X.; Yang, X.; Dong, X.; Liu, Q. Nutrient dynamics linked to hydrological condition and anthropogenic nutrient loading in Chaohu Lake (Southeast China). *Hydrobiologia* **2010**, *661*, 223–234.

41. Liu, W.; Qiu, R. Water eutrophication in China and the combating strategies. *J. Chem. Technol. Bio. Technol.* **2007**, *82*, 781–786.

42. Yu, T. Phytoplankton Community Structure in Chaohu Lake. Master's Thesis, Anhui University, Hefei, China, 10 June 2010.

43. Chen, Z.; Li, Y.; Pan, J. Distributions of colored dissolved organic matter and dissolved organic carbon in the Pearl River Estuary, China. *Cont. Shelf Res.* **2004**, *24*, 1845–1856.

44. Jiang, G.; Ma, R.; Loiselle, S.A.; Duan, H. Optical approaches to examining the dynamics of dissolved organic carbon in optically complex inland waters. *Environ. Res. Lett.* **2012**, *7*, doi:10.1088/1748-9326/7/3/034014.

45. Mueller, J.L.; McClain, G.S.F.C.R.; Mueller, J.; Bidigare, R.; Trees, C.; Balch, W.; Dore, J.; Drapeau, D.; Karl, D.; Van, L.H. Ocean optics protocols for satellite ocean color sensor validation, revision 5, volume V: Biogeochemical and bio-optical measurements and data analysis protocols. *NASA Tech. Memo.* **2003**, *211621*, 1–36.

46. Mobley, C.D. Estimation of the remote-sensing reflectance from above-surface measurements. *Appl. Opt.* **1999**, *38*, 7442–7455.

47. Qi, L.; Hu, C.; Duan, H.; Barnes, B.; Ma, R. An EOF-Based algorithm to estimate chlorophyll a concentrations in Taihu Lake from MODIS land-band measurements: Implications for near real-time applications and forecasting models. *Remote Sens.* **2014**, *6*, 10694–10715.

48. Qi, L.; Hu, C.; Duan, H.; Cannizzaro, J.; Ma, R. A novel MERIS algorithm to derive cyanobacterial phycocyanin pigment concentrations in a eutrophic lake: Theoretical basis and practical considerations. *Remote Sens. Environ.* **2014**, *154*, 298–317.

49. Feng, L.; Hu, C.; Chen, X.; Tian, L.; Chen, L. Human induced turbidity changes in Poyang Lake between 2000 and 2010: Observations from MODIS. *J. Geophys. Res.* **2012**, doi:10.1029/2011JC007864.

50. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ.* **1974**, *351*, 309.

51. Li, J.; Wu, D.; Wu, Y.; Liu, H.; Shen, Q.; Zhang, H. Identification of algae-bloom and aquatic macrophytes in Lake Taihu from *in situ* measured spectra data. *J. Lake Sci.* **2009**, *21*, 215–222.

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

53. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; CRC Press: Florida, FL, USA, 1984.

54. Razi, M.A.; Athappilly, K. A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert Syst. Appl.* **2005**, *29*, 65–74.

55. Cózar, A.; Gálvez, J.A.; Hull, V.; García, C.M.; Loiselle, S.A. Sediment resuspension by wind in a shallow lake of Esteros del Iberá (Argentina): A model based on turbidimetry. *Ecol. Model.* **2005**, *186*, 63–76.

56. Bengio, Y.; Grandvalet, Y. No unbiased estimator of the variance of k-fold cross-validation. *J. Mach. Learn. Res.* **2003**, *5*, 1089–1105.

57. Shi, K.; Li, Y.; Li, L.; Lu, H. Absorption characteristics of optically complex inland waters: Implications for water optical classification. *J. Geophys. Res.: Biogeosci.* **2013**, *118*, 860–874.

58. Gower, J.F.R.; Doerffer, R.; Borstad, G.A. Interpretation of the 685 nm peak in water-leaving radiance spectra in terms of fluorescence, absorption and scattering, and its observation by MERIS. *Int. J. Remote Sens.* **1999**, *20*, 1771–1786.

59. Gower, J.; King, S.; Wei, Y.; Borstad, G.; Brown, L. Use of the 709 nm band of MERIS to detect intense plankton blooms and other conditions in coastal waters. In Proceedings of the MERIS User Workshop, Frascati, Italy, 10–13 November 2003.

60. Gitelson, A.A.; Dall'Olmo, G.; Moses, W.; Rundquist, D.C.; Barrow, T.; Fisher, T.R.; Gurlin, D.; Holz, J. A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. *Remote Sens. Environ.* **2008**, *112*, 3582–3593.

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

62. Ma, R.; Jiang, G.; Duan, H.; Bracchini, L.; Loiselle, S. Effective upwelling irradiance depths in turbid waters: A spectral analysis of origins and fate. *Opt. Express* **2011**, *19*, 7127–7138.

63. Gordon, H.R. Atmospheric correction of ocean color imagery in the Earth Observing System era. *J. Geophys. Res.: Atmos. (1984–2012)* **1997**, *102*, 17081–17106.

64. Wang, M.; Shi, W. Estimation of ocean contribution at the MODIS near—Infrared wavelengths along the east coast of the US: Two case studies. *Geophys. Res. Lett.* **2005**, doi:10.1029/2005GL022917.

65. Carvalho, D.; Rocha, A.; Gómez-Gesteira, M.; Alvarez, I.; Silva Santos, C. Comparison between CCMP, QuikSCAT and buoy winds along the Iberian Peninsula coast. *Remote Sens. Environ.* **2013**, *137*, 173–183.

66. Lin, H.; Xu, Q.; Zheng, Q. An overview on SAR measurements of sea surface wind. *Prog. Nat. Sci.* **2008**, *18*, 913–919.

67. 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. *Limnol. Oceanogr.* **2008**, *53*, 2391–2406.

68. Loiselle, S.; Cozar, A.; Adgo, E.; Ballatore, T.; Chavula, G.; Descy, J.P.; Harper, D.M.; Kansiime, F.; Kimirei, I.; Langenberg, V. Decadal trends and common dynamics of the bio-optical and thermal characteristics of the African Great Lakes. *PloS One* **2014**, *9*, doi:10.1371/journal.pone.0093656.