A hybrid algorithm for estimating the chlorophyll-*a* concentration across different trophic states in Asian inland waters

Bunkei Matsushita[1*], Wei Yang[2], Gongliang Yu[3], Youichi Oyama[1], Kazuya Yoshimura[4], Takehiko Fukushima[1]

[1] Faculty of Life and Environmental Sciences, University of Tsukuba, Tennoudai 1-1-1, Tsukuba, Ibaraki, 305-8572, Japan

[2] Department of Environmental Geochemical Cycle Research, Japan Agency for Marine-Earth Science and Technology, Yokohama 2360001, Japan

[3] Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei 430072, China

[4] Headquarters of Fukushima Partnership Operations, Japan Atomic Energy Agency, Tokyo, 100-8577, Japan

Submittal date: December 17, 2014

*Corresponding Author

Phone: 81-29-853-7190

Fax: 81-29-853-7190

Email: matsushita.bunkei.gn@u.tsukuba.ac.jp

A hybrid algorithm for estimating the chlorophyll-*a* concentration across different trophic states in Asian inland waters

Abstract

The chlorophyll-*a* (Chl-*a*) concentration is one of the most important parameters for evaluating the state of water environments, which often vary markedly across both time and space. Here we propose a hybrid algorithm for retrieving the Chl-*a* values from *in situ* remote sensing data. This hybrid algorithm contains three individual Chl-*a* estimation algorithms that were previously developed for clear waters (a blue-green algorithm), turbid waters (a two-band index-based red–near infrared algorithm), and highly turbid waters (a three-band index-based red–near infrared algorithm). The MCI value (maximum chlorophyll index) was used to switch the three algorithms. To evaluate the performance of the proposed hybrid algorithm, we used the *in situ* remote sensing reflectance and Chl-*a* values collected from five Asian lakes, the trophic status of which varied from oligotrophic to hypertrophic. The results showed that the hybrid algorithm performed well for a wide variety of optical properties, with the NMAE (normalized mean absolute error) of 13.3%. Our results indicate that the proposed hybrid algorithm has the potential for use as an operational tool for monitoring Chl-*a* in waters with widely varying trophic conditions without the requirement of reparameterization.

maximum chlorophyll index, Asian lakes

1. Introduction

The chlorophyll-a (Chl-a) concentration is one of the most important parameters for evaluating water environments, such as the trophic status, clarity, and algal biomass of waters. Routine observations of Chl-a are thus requested by researchers and local governments. Compared to the conventional sampling methods (e.g., sampling from a boat), the retrieval of Chl-a values from remote sensing data can not only save costs and labor but also provide synoptic observations for water areas, and it is thus is considered one of the effective techniques for water-quality monitoring.

Generally, the remote sensing reflectances (R_{rs}) at blue and green wavelengths are used to estimate Chl-a in clear waters (e.g., open ocean waters). For example, the OC4 algorithm (oceanic chlorophyll-a algorithm for SeaWiFS) used the R_{rs} at the wavelengths of 443, 490, 510 and 555 nm to estimate Chl-a for global oceans (O'Reilly et al. 1998). This is because the R_{rs} at these wavelengths can provide a high signal-to-noise ratio (SNR) along with the first absorption peak of Chl-a (at the blue spectra domain) for detecting phytoplankton changes.

In contrast, the R_{rs} at longer wavelengths (e.g., red or near-infrared) cannot provide a sufficient SNR due to the strong absorption from pure waters and the relatively weak backscattering from suspended solids in clear waters. The Chl-a estimation algorithms based on the R_{rs} at the blue and green wavelengths are denoted

hereafter as blue-green algorithms. The blue-green algorithms have often failed in optically complex waters (e.g., Case-2 waters) because the R_{rs} at these wavelengths are affected not only by phytoplankton but also by other constituents including non-algal particles (NAP) and colored dissolved organic matter (CDOM) (Gitelson et al. 2009).

Optically complex waters often have higher turbidity along with a sufficient SNR around the second absorption peak of Chl-*a* (at the red spectra domain), and thus the R_{rs} at the red and near-infrared (NIR) wavelengths became an alternative for estimating Chl-*a* in these types of waters. For example, Dall'Olmo et al. (2003) proposed a three-band index which was designed by using the R_{rs} at the wavelengths around 665, 709, and 754 nm (i.e., $[R_{rs}^{-1}(665) - R_{rs}^{-1}(709)] \times R_{rs}(754)$) to minimize the effects from NAP and CDOM. It was found that the Chl-*a* estimation algorithms based on this index could provide higher accuracy even for optically complex waters (Gitelson et al. 2008; Moses et al. 2009; Gilerson et al. 2010; Yang et al. 2010, 2011). However, Gurlin et al. (2011) reported that the Chl-*a* estimation algorithms based on a two-band index (i.e., $R_{rs}(709)/R_{rs}(665)$) could in some cases perform better than those based on the three-band index. The Chl-*a* estimation algorithms based on the R_{rs} at the red and NIR wavelengths are denoted hereafter as red-NIR algorithms.

Many studies have been conducted to retrieve Chl-*a* values from remote sensing data in waters with different optical properties (Matthews 2011; Odermatt et al. 2012). It has gradually become clear that the blue-green ratios are limited to oligotrophic waters, whereas the red-NIR ratios are only suitable for waters with Chl-*a* concentrations > 10 mg m^{-3} (Odermatt et al. 2012). This finding indicates that a hybrid algorithm that can be

applied to global waters is required for operationally retrieving Chl-*a* values under a wide range of optical properties. Several studies attempted to address this issue. For example, Gomez et al. (2011) proposed two normalized difference indices (i.e., bands 705 and 665 nm, bands 560 and 442 nm) to separate Mediterranean Lakes of European Union into two types, and different Chl-*a* retrieval algorithm was used for each water type. Matthews et al. (2012) developed a maximum peak-height (MPH) algorithm, which uses a baseline subtraction procedure to calculate the height of the dominant peak across the red and NIR wavelengths, for estimating Chl-*a* in inland and coastal waters with a wide range of trophic status. Two Chl-*a* retrieval algorithms were developed for eukaryote-dominant and cyano-dominant waters respectively, and two baseline subtraction indices were used to switch the two algorithms (Matthews et al., 2012). Moore et al. (2014) suggested the use of an optical water type (OWT) framework for selecting and blending two Chl-*a* retrieval algorithms. However, this kind of study is still very limited. In addition, although the proposed Chl-*a* retrieval algorithms have been validated in some inland waters, further validation is necessary to show both the advantages and potential limitations of these algorithms in other waters, especially some Asian inland waters, which have different optical properties but are less included in the previous studies.

Consequently, the objectives of the present study were to (1) further test the performances of recently developed red-NIR algorithms as well as a blue-green algorithm, using the independent data collected from several Asian lakes with a wide range of trophic states; (2) propose a simple remote sensing-based classification method

for selecting the most appropriate Chl-*a* retrieval algorithm; and (3) develop a hybrid method for estimating Chl-*a* concentrations across different trophic states in inland waters.

2. Methods

2.1. Study areas and data collection

The data used in this study were collected from five Asian lakes: Lakes Biwa (Shiga Prefecture, Japan), Suwa (Nagano Prefecture, Japan), Kasumigaura (Ibaraki Prefecture, Japan), Erhai (Yunnan Province, China), and Dianchi (Yunnan Province, China). The basic information about the lakes during the field surveys is summarized in Table 1.

In limnology, the average Chl-*a* concentration will generally be < 1 mg m^{-3} in ultra-oligotrophic waters, < 2.6 mg m^{-3} in oligotrophic waters, between 2.6 mg m^{-3} and 7.2 mg m^{-3} in mesotrophic waters, between 7.2 mg m^{-3} and 20 mg m^{-3} in eutrophic waters, and > 20 mg m^{-3} in hypertrophic waters (Carlson 2007). It can be seen in Table 1 that the five Asian lakes cover trophic categories from oligotrophic to hypertrophic during the periods of our field survey. Lake Biwa is in the oligotrophic category, Lakes Suwa and Erhai are in the eutrophic category, and Lakes Kasumigaura and Dianchi are in the hypertrophic category.

Water samples and the corresponding reflectance spectra were collected from Lake Biwa in a campaign undertaken in October 2011 (10 sites available) and from Lake Suwa in a July 2010 campaign (8 sites) and from Lake Kasumigaura in February

2006, August 2008, and May 2010 (46 sites). The Lake Dianchi data collection was carried out in October 2007, July 2008, and July 2009 (28 sites). In Lake Erhai, there were two field campaigns — one in September 2011 and the other in July 2012; in each field campaign, the investigation was performed at 21 sites (42 sites in total). The distribution of sampling sites in each lake is shown in Figure 1.

Reflectance was measured between 10:00 and 14:00 local time. All measurements were taken over optically deep water with depths larger than Secchi disk depth; floating scum was not found at these sampling sites. The water-leaving radiance ($L_u(\lambda)$), the downward irradiance ($E_d(\lambda)$), and the downward radiance of skylight ($L_{sky}(\lambda)$) were measured at each site following Method 1 described by Muller et al. (2000) using a FieldSpec HandHeld spectroradiometer (Analytical Spectral Devices, Boulder, CO) in the range of 325 to 1075 nm at 1 nm intervals. The above-water $R_{rs}(\lambda)$ was calculated approximately using the following equation (Mobley 1999):

$$R_{rs}(\lambda) = (\frac{L_u(\lambda) - rL_{sky}(\lambda)}{E_d(\lambda)}) \times Cal(\lambda) \times 100, \tag{1}$$

where $Cal(\lambda)$ is the spectral reflectance calibration factor for the gray Spectralon® standard reflectance panel (Labsphere, North Sutton, NH), and r is the reflectance of skylight determined as a function of wind speed. The R_{rs} spectra collected from the five lakes are shown in Figure 2.

Water samples were collected at each site and taken to the laboratory within approx. 0.5–1.5 h after the field investigation. Chlorophyll-a was extracted using

7

methanol (100%) at 4°C for 24 h in the dark. The absorbance of the extracted Chl-*a* was measured at four wavelengths (750, 663, 645 and 630 nm) using a spectrophotometer (UV-2450, Shimadzu, Kyoto, Japan), and the concentrations were calculated according to SCOR-UNESCO equations (SCOR-UNESCO 1966). The concentrations of total suspended solids (TSS) and inorganic suspended solids (ISS) were determined gravimetrically. Samples were filtered through pre-combusted Whatman GF/F filters (at 500°C for 4 h) and then dried at 105°C for 4 h and weighed to obtain the TSS. The filters were re-combusted at 500°C for 4 h and then weighed again to obtain the ISS.

The absorption coefficient of CDOM was measured using a spectrophotometer after the water sample was filtered. The descriptive statistics of the water constituent concentrations for the five lakes are shown in Table 1.

2.2. Basic theory of the two- and three-band indices for retrieving Chl-*a* in optically complex waters

According to Gordon et al. (1975) and Austin (1980), the $R_{rs}(\lambda)$ is determined as follows:

$$R_{rs}(\lambda) \propto f \times \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \tag{2}$$

where *f* is dependent on the geometry of the light field emerging from the body of water, and $a(\lambda)$ and $b_b(\lambda)$ are the spectral total absorption and backscattering coefficients, respectively. The $a(\lambda)$ and $b_b(\lambda)$ can be respectively written as:

$$a(\lambda) = a_{\mathrm{w}}(\lambda) + a_{\mathrm{ph}}(\lambda) + a_{NAP}(\lambda) + a_{\mathrm{CDOM}}(\lambda) \tag{3-1}$$

$$b_{\mathrm{b}}(\lambda) = b_{\mathrm{b,w}}(\lambda) + b_{\mathrm{b,ph}}(\lambda) + b_{\mathrm{b,NAP}}(\lambda) \tag{3-2}$$

where $a_{\mathrm{w}}(\lambda)$ is the absorption coefficient of pure water, and $a_{\mathrm{ph}}(\lambda)$, $a_{\mathrm{NAP}}(\lambda)$ and $a_{\mathrm{CDOM}}(\lambda)$ are the absorption coefficients for phytoplankton, NAP and CDOM, respectively; $b_{\mathrm{b,w}}(\lambda)$, $b_{\mathrm{b,ph}}(\lambda)$, and $b_{\mathrm{b,NAP}}(\lambda)$ are the backscattering coefficients for pure water, phytoplankton, and NAP, respectively.

The two-band index was designed as the ratio of the R_{rs} peak around 709 nm to the R_{rs} trough around 665 nm (Dekker 1993; Gitelson and Kondratyev 1991; Han and Rundquist 1997). It can minimize the effects from NAP and CDOM based on four assumptions: (1) $a_{\mathrm{ph}}(665) + a_{\mathrm{w}}(665) \gg a_{\mathrm{NAP}}(665) + a_{\mathrm{CDOM}}(665)$, (2) $a_{\mathrm{w}}(709) \gg a_{\mathrm{ph}}(709) + a_{\mathrm{NAP}}(709) + a_{\mathrm{CDOM}}(709)$, (3) $b_{\mathrm{b}}(665) \approx b_{\mathrm{b}}(709)$; (4) $a(665) \gg b_{\mathrm{b}}(665)$, and $a(709) \gg b_{\mathrm{b}}(709)$ (Gitelson et al. 2008). Then, the two-band index can be approximated as (Le et al. 2013):

$$R_{\mathrm{rs}}(709)/R_{\mathrm{rs}}(665) \approx [a_{\mathrm{ph}}(665) + a_{\mathrm{w}}(665)]/a_{\mathrm{w}}(709) \tag{4}$$

Since $a_{\mathrm{w}}(\lambda)$ can be considered constant, the two-band index varies only with $a_{\mathrm{ph}}(665)$, which is strongly related to Chl-*a*. Therefore, an algorithm based on the two-band index can provide an accurate estimate for retrieving Chl-*a* values. However, it should be noted that all of the above assumptions except for the third one are valid only in

optically complex waters that have relatively low turbidities (Dall'Olmo et al. 2005; Gitelson et al. 2008).

For optically complex waters with higher turbidities, a three-band index was designed to minimize the effects of NAP and CDOM (Dall'Olmo et al. 2003; Gitelson et al. 2008). First, the R_{rs} around wavelengths 665 nm and 709 nm were inverted and their difference was calculated to minimize the absorption effects of NAP and CDOM (i.e., $R_{rs}^{-1}(665) - R_{rs}^{-1}(709)$). This step is based on two assumptions: (1) $a_{NAP}(665) \approx a_{NAP}(709)$ and $a_{CDOM}(665) \approx a_{CDOM}(709)$, (2) $b_b(665) \approx b_b(709)$. Therefore, $R_{rs}^{-1}(665) - R_{rs}^{-1}(709)$ can be approximated as $[a_{ph}(665) + a_w(665) - a_{ph}(709) - a_w(709)]/b_b$. Then, the R_{rs} at the third wavelength around 754 nm was used to further minimize the effects by b_b. This step is based on another two assumptions: (1) $b_b(665) \approx b_b(709) \approx b_b(754)$, and (2) $a_w(754) \gg a_{ph}(754) + a_{NAP}(754) + a_{CDOM}(754) + b_b(754)$. The three-band index is thus designed as $[R_{rs}^{-1}(665) - R_{rs}^{-1}(709)] \times R_{rs}(754)$, and it can be approximated as (Gitelson et al. 2009; Le et al. 2013):

$$[R_{rs}^{-1}(665) - R_{rs}^{-1}(709)] \times R_{rs}(754) \approx [a_{ph}(665) + a_w(665) - a_{ph}(709) - a_w(709)]/a_w(754)$$

(5)

Similar to the two-band index, an algorithm based on the three-band index can provide an accurate estimate for retrieving Chl-*a* values in more turbid waters. However, the use of a longer wavelength (e.g., $R_{rs}(754)$) will limit the applications of the three-band index to waters with low turbidities, because due to the largely increased absorption by pure

water, not enough signals can be detected at this wavelength. On the other hand, the assumption of $a_w(754) \gg a_{ph}(754) + a_{NAP}(754) + a_{CDOM}(754) + b_b(754)$ will be invalid in more turbid waters (Le et al. 2009; Yang et al. 2010, 2011).

2.3. Development of a hybrid algorithm for operationally estimating Chl-*a*

In light of the calculations described above, it is theoretically known that the two-band algorithms are more appropriate for optically complex waters with lower turbidities whereas the three-band algorithms are more appropriate for optically complex waters with higher turbidities. However, for retrieving Chl-*a* values in clear waters, a blue-green algorithm must be selected to replace the red-NIR algorithm. It is thus necessary to develop a classification method for selecting the appropriate algorithm for different waters.

We used the maximum chlorophyll index (MCI), which is a measure of the reflectance peak at 709 nm in water-leaving reflectance, to select the most appropriate algorithm for each pixel. The MCI is defined as (Gower et al. 2005):

$$MCI = R_{rs}(709) - R_{rs}(665) - [\frac{(709-665)}{(754-665)}(R_{rs}(754) - R_{rs}(665))] \tag{6}$$

Figure 3 shows the relationship between the MCI and Chl-*a* values based on our dataset. It can be seen that the MCI is linearly related to Chl-*a* in the lakes that we studied, with the determination coefficient of 0.76. On the other hand, previous studies reported two important Chl-*a* values: 10 mg m^{-3} and 25 mg m^{-3}, which are turning

points for the performances of several algorithms (Gitelson et al. 2008; Gurlin et al. 2011; Odermatt et al. 2012; Moore et al. 2014). These two Chl-*a* values are similar to the Chl-*a* thresholds used to determine the trophic status of waters (Carlson 2007). According to the relationship described in Figure 3, we found that the MCI values 0.0001 and 0.0016 approximately corresponded to the Chl-*a* values 10 mg m^{-3} and 25 mg m^{-3}, respectively. Accordingly, we propose a hybrid algorithm for operationally estimating Chl-*a* in inland waters: (1) if MCI ≤ 0.0001, then a blue-green algorithm should be selected. (2) If $0.0001 < \text{MCI} \leq 0.0016$, then a two-band index-based red-NIR algorithm should be selected. (3) If the MCI is > 0.0016, then a three-band index-based red-NIR algorithm should be selected.

2.4. Candidate blue-green and red-NIR algorithms for estimating the Chl-*a* concentration

First, we selected one blue-green algorithm, OC4E (version 4 for Medium Resolution Imaging Spectrometer [MERIS]) to retrieve Chl-*a* values in clear waters. The OC4Ev4 algorithm is written as (O'Reilly et al. 2000):

$$Chl - a = 10^x \tag{7}$$

$$x = 0.368 - 2.814R_{4E} + 1.456R_{4E}^2 + 0.768R_{4E}^3 - 1.292R_{4E}^4 \tag{8}$$

where R_{4E} is the maximum band ratio among $R_{rs}(443)/R_{rs}(560)$, $R_{rs}(490)/R_{rs}(560)$, and $R_{rs}(510)/R_{rs}(560)$. The $R_{rs}(560)$ is used as a reference because it is relatively insensitive

to changes in Chl-*a* values due to the weak absorption of chlorophyll at this wavelength; the use of R_{rs} at blue wavelengths is because chlorophyll absorbs light strongly in the 443 to 510 nm range, and the use of the maximum band ratio is because the maximum of the R_{rs} spectra will shift toward longer wavelengths with increasing Chl-*a* (O'Reilly et al. 1998). The coefficients in Eq. (8) were obtained from an empirical regression between R_{4E} and Chl-*a*, using a very large dataset (n=2804, O'Reilly et al. 2000).

We then selected seven red-NIR algorithms to retrieve Chl-*a* values in turbid waters, in which there are three two-band index-based algorithms and four three-band index-based algorithms. The three two-band index-based algorithms include one linear algorithm (Gitelson et al. 2011, referred to hereafter as RN2-Git11; "RN" denotes a red-NIR algorithm, the number "2" denotes a two-band index, "Git" denotes the name of the first author, and the number "11" denotes the year published), one power algorithm (Gilerson et al. 2010, referred to hereafter as RN2-Gil10), and one polynomial algorithm (Gurlin et al. 2011, referred to hereafter as RN2-Gur11). These algorithms are written as follows without reparameterization:

$$Chl - a = 72.66 \times RN2 - 46.535 \tag{9}$$

$$Chl - a = (35.75 \times RN2 - 19.3)^{1.124} \tag{10}$$

$$Chl - a = 25.28 \times RN2^2 + 14.85 \times RN2 - 15.18 \tag{11}$$

where RN2 is the two-band index (i.e., Eq. (4)).

The four three-band index-based algorithms also include one linear algorithm

(Gitelson et al. 2011, referred to hereafter as RN3-Git11; the number "3" denotes a three-band index), one power algorithm (Gilerson et al. 2010, RN3-Gil10), and one polynomial algorithm (Gurlin et al. 2011, RN3-Gur11) as well as the SAMO-LUT (Semi-Analytical Model Optimizing and Look-Up Tables) algorithm (Yang et al. 2011). The RN3-Git11, RN3-Gil10, and RN3-Gur11 algorithms are written as follows without reparameterization:

$$Chl - a = 243.862 \times RN3 + 27.219 \tag{12}$$

$$Chl - a = (113.36 \times RN3 + 16.45)^{1.124} \tag{13}$$

$$Chl - a = 315.50 \times RN3^2 + 215.95 \times RN3 + 25.66 \tag{14}$$

where RN3 is the three-band index (i.e., Eq. (5)).

The details of the SAMO-LUT algorithm are given in the Appendix. The main differences between the SAMO-LUT algorithm and the other three-band index-based algorithms are: (1) the SAMO-LUT is capable of retrieving Chl-*a* values in more turbid waters; (2) it is possible to select the most appropriate Chl-*a* estimation model pixel by pixel rather than using only one Chl-*a* estimation model for all pixels of a lake; and (3) the SAMO-LUT can simultaneously retrieve three water quality parameters (i.e., the concentrations of Chl-*a* and NAP as well as CDOM absorption coefficient at 440 nm; Yang et al. 2011).

2.5. Assessment of the accuracy of the algorithms

We used four indices for our assessment of the accuracy of the tested algorithms: the root mean square error (RMSE), the normalized root mean square error (NRMS), the mean normalized bias (MNB), and the normalized mean absolute error (NMAE) (Gitelson et al. 2008; Gurlin et al. 2011). These indices are defined as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(X_{esti,i} - X_{meas,i}\right)^2}{N-1}} \tag{15}$$

$$\text{NRMS} = \text{stdev}(\varepsilon_i)\% \tag{16}$$

$$\text{MNB} = \text{mean}(\varepsilon_i)\% \tag{17}$$

$$\text{NMAE} = \text{mean}(|\varepsilon_i|)\% \tag{18}$$

where $X_{esti,i}$ and $X_{meas,i}$ are the estimated and measured Chl-*a* values, respectively; N is the number of samples, and $\varepsilon_i = 100 \times (X_{esti,i} - X_{meas,i})/X_{meas,i}$ is the percent difference between the estimated and measured Chl-*a* values. The NRMS denotes the relative random uncertainty of the results. The MNB denotes the average bias in the estimation, and the NMAE denotes the average relative error in the estimation. The correlation between the measured and estimated values was also calculated.

3. Results

3.1. Performances of the selected algorithms for each tested lake

The performances of all selected algorithms in the five Asian lakes are summarized in Table 2. Overall, the blue-green algorithm (OC4Ev4) performed well in the two clearer lakes (Lakes Biwa and Suwa) but failed in the other three turbid lakes (Lakes Kasumigaura, Dianchi, and Erhai). The red-NIR algorithms could reduce the estimation errors in Lakes Kasumigaura, Dianchi, and Lake Erhai (RN2-Gil10), but

they showed larger errors in the other two clearer lakes.

Especially in Lake Biwa, the OC4Ev4 algorithm not only provided an accurate Chl-*a* estimation and a small bias (MNB $= -1.1\%$); it also explained 54% of the Chl-*a* variation in the lake. In contrast, all of the red-NIR algorithms showed large biases (MNB $> 171.6\%$) and low correlations between the estimated and measured Chl-*a* values ($R^2 < 0.14$). In Lake Suwa, all tested algorithms show low correlations between the estimated and measured Chl-*a* values because the coefficient of variation for the measured Chl-*a* is very small (only 5.87%, Table 1) in this small lake (area 13.3 km^2).

In the comparisons of the red-NIR algorithms, it can be seen that (1) for Lake Kasumigaura in Japan and Lake Dianchi in China (both are hypertrophic lakes), the performances of the three-band red-NIR models were better than those of the two-band red-NIR models (except for the RN3-Gur11); and (2) for Lake Erhai, China (a eutrophic lake), the two-band red-NIR models outperformed all three-band red-NIR models.

These findings indicate that (1) no single algorithm could provide acceptable accuracy for all five lakes; (2) different algorithms should be selected for different lakes (even for different sites) to obtain an accurate Chl-*a* estimations.

3.2. Performances of the selected algorithms for each classified group by the MCI

According to the MCI thresholds determined in Section 2.3 above, we divided the entire dataset into three groups: (1) group 1, MCI ≤ 0.0001, 10 sites, all at Lake Biwa; (2) group 2, $0.0001 < $ MCI ≤ 0.0016, 33 sites, eight at Lake Suwa and 25 at Lake Erhai; and (3) group 3, MCI > 0.0016, 91 sites, 46 sites of which are at Lake Kasumigaura, 28

at Lake Dianchi, and 17 at Lake Erhai.

Table 3 provides the statistical description of the Chl-*a* values for each group. The mean Chl-*a* concentration increased from group 1 to group 3, and an overlap of Chl-*a* values can be seen between the second and third groups. Since the dataset for the first group is too small (n=10) and has a narrower range, no Chl-*a* overlaps were found between the first and second groups. To further illustrate the corresponding relationship between the Chl-*a* and MCI values, the statistical descriptions of MCI values for each group divided by two Chl-*a* thresholds (10 and 25 mg m^{-3}) are given in Table 4. It can be seen in the table that the mean MCI values increased from the low- to the high-Chl-*a* groups, and overlaps of the MCI values between the moderate- and high-Chl-*a* groups are observed.

We re-evaluated the accuracy of each algorithm for each group clustered using the two MCI thresholds by comparing the estimated Chl-*a* values with the measured Chl-*a* values (Table 5). The results clearly showed that (1) for the first group, the OC4Ev4 (blue-green algorithm) gave the best accuracy. (2) The two-band red-NIR models were more appropriate for the second group; in which the RN2-Gil10 gave the best accuracy. (3) The three-band red-NIR models were more appropriate for the third group, in which the SAMO-LUT algorithm gave the best accuracy.

The above findings indicate that it is possible to combine the blue-green and the red-NIR algorithms for estimating Chl-*a* in most inland waters, and the MCI can serve as an indicator for selecting the most appropriate algorithm.

3.3. Performance of a hybrid algorithm

Based on the results described in the previous section, we designed a hybrid algorithm (shown in Fig. 4). Chl-*a* was estimated for each site using the most appropriate algorithm selected by the MCI value. Figure 5 shows the performances of the new hybrid algorithm as well as those of three selected algorithms for the entire dataset (n=134). It can be seen that the OC4Ev4 algorithm yielded good accuracy for low Chl-*a* values but performed poorly when the Chl-*a* exceeded 10 mg m^{-3} (large underestimation). The RN2-Gil10 algorithm performed well for moderate Chl-*a* values but poorly for low (overestimation) and high Chl-*a* (underestimation) values, and the SAMO-LUT algorithm showed the best performance for the high Chl-*a* values but larger errors for low and moderate Chl-*a*. In contrast, the Chl-*a* values estimated by the hybrid algorithm were in good agreement with the in situ-measured values, with all of the data points close to the 1:1 line. The RMSE, NMAE, MNB, and NRMS were calculated as 8.63 mg m^{-3}, 13.3%, 1.4%, and 16.5%, respectively. The determination coefficient was calculated as 0.94. These values indicate that the hybrid algorithm can perform better than any single algorithm if it is used for the entire dataset.

4. Discussion

As noted earlier, each of the algorithms selected here was developed based on different assumptions, and thus they are applicable only to the waters where the assumptions were valid. The optical properties of inland waters are very complex. Each lake has unique optical conditions (or water-quality properties) due to the different

impacts of human activities. A single body of water can also present different optical properties both spatially and temporally (Moore et al. 2014). It is thus clear that no single algorithm will be universally applicable and that a hybrid algorithm is necessary if we target on waters with unknown optical properties.

In the present study, we proposed a hybrid algorithm, in which the MCI was used to select from among three types of algorithms (one blue-green algorithm, OC4Ev4; one two-band index-based red-NIR algorithm, RN2Gil10; and one three-band index-based red-NIR algorithm, SAMO-LUT) for different lakes or even for different sites (e.g., in the case of Lake Erhai). Moore et al. (2014) suggested the use of an optical water type (OWT) framework for selecting and blending two Chl-a retrieval algorithms (i.e., OC4v6 and a three-band index-based red-NIR algorithm). They first defined seven water types using an aggregate dataset from Spanish and U.S. lakes as well as SeaBASS coastal marine data, and then they developed a membership function for each OWT and used the produced class memberships as the basis of weighting factors for blending two Chl-a retrievals from different estimation algorithms into a single product.

According to Table 2 in the Moore et al. (2014) paper, we identified OWT 1 with the MCI value < 0.0001, OWTs 2–3 with MCI values between 0.0001 and 0.0016, and OWTs 4–7 with MCI values > 0.0016. This finding indicates that it would become easy for a public user to switch different Chl-a estimation algorithms if the MCI value is used alongside the OWT framework. In addition, it is worthwhile to note that we used a two-band index-based red-NIR algorithm to replace the requirement of weighting factors in Moore et al. (2014). This modification will make the proposed hybrid

algorithm also suitable for waters that are between clear and highly turbid (i.e., moderate-turbid waters; Section 2.2, Fig. 5b).

The two- and three-band indices were specially designed to minimize the effects of NAP and CDOM in Chl-a estimations (Section 2.2) and they are thus less influenced by NAP and CDOM (Gilerson et al. 2010). The two- and three-band indices are also less sensitive to changes in the phytoplankton-specific absorption for higher Chl-a values (Chl-a > 10 mg m^{-3}, Gilerson et al. 2010; MCI > 0.0001 in the present study). For highly turbid waters (MCI > 0.0016), although the coefficients of three-band index-based Chl-a estimation algorithms will be affected by changes in the concentrations of NAP and CDOM, the use of the SAMO-LUT algorithm can address this issue (Yang et al. 2011). These characteristics make the selected RN2-Gil10 and SAMO-LUT algorithms useful for turbid waters without the requirement of reparameterization, and thus the hybrid algorithm has the potential to be an operational tool for effectively monitoring Chl-a from satellites in near-real time if the atmospheric correction can be carried out with an acceptable accuracy.

In the present study, the MCI thresholds were empirically determined according to in situ data collected from five Asian lakes (Section 2.3). Since the inherent optical properties (IOPs) of both NAP and CDOM show less variance in the red–NIR spectral regions (Gitelson et al. 2008), the MCI values are mainly determined by the IOPs of phytoplankton (Fig. 3). Therefore, the MCI values can serve as a good indicator to qualitatively classify waters for selecting the appropriate Chl-a estimation algorithm for a given water pixel. However, since the MCI values also showed weaker positive

correlations with the NAP and CDOM concentrations (R^2=0.44 for NAP, R^2=0.38 for CDOM; data not shown), the use of MCI to quantitatively estimate Chl-a may yield not ignorable errors. In addition, the increases of NAP and CDOM will result in slightly larger MCI values, and thus increase the likelihood of selecting a red–NIR algorithm. This is a positive aspect of the proposed hybrid algorithm.

Although the two MCI thresholds were determined based on a dataset with a wider range of optical properties, this dataset is still limited, especially in light of the lack of water samples with Chl-a between 2 and 10 mg m^{-3} (Fig. 5). In addition, the proposed hybrid algorithm has not yet been applied to satellite images due to the atmospheric correction is still a challenge in turbid inland waters (e.g., Lake Kasumigaura, Japan; Jaelani et al., 2013). Future studies of further validations involving continuous data accumulation and satellite images application will be carried out.

5. Conclusions

We proposed a hybrid algorithm to retrieve the Chl-a concentrations in inland waters with a wide range of optical properties. After testing eight individual Chl-a estimation algorithms in five Asian lakes, we selected the OC4Ev4, the RN2-Gil10, and the SAMO-LUT algorithms for clear, moderate-turbidity and high-turbidity waters, respectively, to compose the hybrid algorithm. The MCI, which can be easily calculated from satellite data, was used to switch the three algorithms. The results show that the hybrid algorithm is superior to all of the single algorithms tested when considering the entire observed range of optical properties, with the RMSE, NMAE, MNB, and NRMS

values of 0.072 in log Chl-*a* units, 13.3%, 1.4%, and 16.5%, respectively. These results indicate that the proposed hybrid algorithm has the potential to be used as an operational tool for monitoring Chl-*a* in waters with widely varying trophic conditions without the requirement of reparameterization.

References

Austin, R. W. (1980). Gulf of Mexico, ocean-colour surface-truth measurements. *Bound.-Layer Meteorol., 18*, 269−285.

Carlson, R.E. (2007). Estimating Trophic State. *LakeLine*, Spring 2007, 25-28.

Dall'Olmo, G., Gitelson, A. A., & Rundquist, D. C. (2003). Towards a unified approach for remote estimation of chlrophyll-a in both terrstrial vegetatioin and turbid productive waters. *Geophysical Research Letters, 30*. doi:10.1029/2003GL018065.

Dall'Olmo, G., Gitelson, A. A., Rundquist, D. C., Leavitt, B., Barrow, T., & Holz, J. C. (2005). Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands. *Remote Sensing of Environment, 96*(2), 176–187.

Dekker, A. G. (1993). Detection of water quality parameters for eutrophic waters by high resolution remote sensing. Ph.D. dissertation (Vrije Universiteit, Amsterdam, The Netherlands).

Gilerson, A., Gitelson, A., Zhou, J., Gulrin, D., Moses, W., Ioannou, I., & Ahmed, S. (2010). Algorithms for remote sensing of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Optics Express, 18(23)*, 24109–24125.

Gitelson, A. A., Dall'Olmo, G., Moses, W., Rundquist, D. C., Barrow, T., Fisher, T. R., et al. (2008). A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. *Remote Sensing of Environment, 112*, 3582−3593.

Gitelson, A., Gurlin, D., Moses, W., & Barrow, T. (2009). A bio-optical algorithm for the remote estimation of the chlorophyll-a concentration in case 2 waters. *Environmental Research Letters, 4*(045003), 5p.

Gitelson, A. A., Gurlin, D., Moses, W. J., & Yacobi, Y. Z. (2011). Remote estimation of chlorophyll-a concentration in inland, estuarine and coastal waters. In Q. Weng (Ed.), *Advances in Environmental Remote Sensing.* CRC Press.

Gitelson, A. A., & Kondratyev, K. Y. (1991). On the mechanism of formation of maximum in the reflectance spectra near 700 nm and its application for remote monitoring of water quality. *Transactions Doklady of the USSR Academy of Sciences: Earth Science Sections, 306*, 1–4.

Gomez, J. A. D., Alonso, C. A., Garcia, A. A. (2011). Remote sensing as a tool for monitoring water quality parameters for Mediterranean Lakes of European Union water framework directive (WFD) and as a system of surveillance of cyanobacterial harmful algae blooms (SCyanoHABs). *Environmental Monitoring and Assessment, 181*, 317-334.

Gordon, H. R., Brown, O. B., & Jacobs, M. M. (1975). Computed relationship between the inherent and apparent optical properties of a flat homogeneous ocean. *Applied Optics, 14*, 417−427.

Gower, J., King, S., Borstad, G., & Brown, L. (2005). Detection of intense plankton blooms using the 709 nm band of the MERIS imaging spectrometer. *International Journal of Remote Sensing, 26*, 2005–2012.

Guanter, L., Ruiz-Verdu, A., Odermatt, D., Giardino, C., Simis, S., Estelles, V., et al. (2010). Atmospheric correction of ENVISAT/MERIS data over inland waters: Validation for European lakes. *Remote Sensing of Environment, 114*, 467-480.

Gurlin, D., Gitelson, A. A., & Moses, W. J. (2011). Remote estimation of chl-a concentration in turbid productive waters — Return to a simple two-band NIR-red model? *Remote Sensing of Environment, 115*(12), 3479–3490.

Han, L. H., & Rundquist, D. C. (1997). Comparison of NIR/RED ratio and first derivative of reflectance in estimating algal-chlorophyll concentration: A case study in a turbid reservoir. *Remote Sensing of Environment, 62*, 253–261.

Jaelani, L. M., Matsushita, B., Yang, W., & Fukushima, T. (2013). Evaluation of four MERIS atmospheric correction algorithms in Lake Kasumigaura, Japan. *International Journal of Remote Sensing, 34(24),* 8967-8985.

Le, C., Hu, C., Cannizzaro, J., English, D., Muller-Karger, F., & Lee, Z. (2013). Evaluation of

chlorophyll-a remote sensing algorithms for an optically complex estuary. *Remote Sensing of Environment,* 129, 75-89.

Le, C., Li, Y., Zha, Y., Sun, D., Huang, C., & Lu, H. (2009). A four-band semi-analytical model for estimating chlorophyll a in highly turbid lakes: The case of Taihu Lake, China. *Remote Sensing of Environment, 113*(6), 1175–1182.

Le, C., Li, Y., Zha, Y., Sun, D., Huang, C., & Zhang, H. (2011). Remote estimation of chlorophyll a in optically complex waters based on optical classification. *Remote Sensing of Environment, 115*, 725–737.

Matthews, M. W. (2011). A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *International Journal of Remote Sensing, 32*(21), 6855-6899.

Matthews, M. W., Bernard, S., & Robertson, L. (2012). 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.

Melin, F., Vantrepotte, V., Clerici, M., D'Alimonte, D., Zibordi, G., Berthon, J. -F., et al. (2011). Multi-sensor satellite time series of optical properties and chlorophyll-a concentration in the Adriatic Sea. *Progress in Oceanography, 91*, 229–244.

Mobley, C. D. (1999). Estimation of the Remote-Sensing Reflectance from Above-Surface Measurements. *Applied Optics, 38*(36), 7442–7455.

Moore, T. S., Dowell, M.D., Bradt, S., & Verdu, A.R. (2014). An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. *Remote Sensing of Environment, 143*, 97–111.

Moses, W., Gitelson, A., Berdnikov, S., & Povazhnyy, V. (2009). Satellite estimation of chlorophyll-a concentration using the red and NIR bands of MERIS — The Azov Sea case study. *IEEE Geoscience and Remote Sensing Letters, 4*(6), 845–849.

Mueller, J.L., Davis, C., Arnone, R., Frouin, R., Carder, K., & Lee, Z.P. (2000). Above-water radiance and remote sensing reflectance measurements and analysis protocols. *Ocean Optics protocols for satellite ocean color sensor validation Revision 2* (pp. 98-107). Greenbelt, Maryland: National Aeronautical and Space Administration.

Odermatt, D., Gitelson, A., Brando, V. E., & Schaepman, M. (2012). Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote Sensing of Environment, 118*, 116-126.

O'Reilly, J., Maritorena, S., Mitchell, B., Siegel, D., Carder, K., Garver, S., et al. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research, 103,* 24,937–24,953.

O'Reilly, J.E., and 24 Coauthors, 2000: SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3. NASA Tech. Memo. 2000-206892, Vol. 11, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, 49 pp.

Oyama, Y., Matsushita, B., Fukushima, T., Matsushige, K., & Imai, A. (2009). Application of spectral decomposition algorithm for mapping water quality in a turbid lake (Lake Kasumigaura, Japan) from Landsat TM data. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*, 73-85.

SCOR-UNESCO (1966). Determination of photosynthetic pigment in seawater. *Monographs on oceanographic methodology* In. Paris, France.

Yang, W., Matsushita, B., Chen, J., Fukushima, T., & Ma, R. (2010). An Enhanced Three-Band Index for Estimating Chlorophyll-a in Turbid Case-II Waters: Case Studies of Lake Kasumigaura, Japan, and Lake Dianchi, China. *IEEE Geoscience and Remote Sensing Letters, 7*, 655−659.

Yang, W., Matsushita, B., Chen, J., & Fukushima, T. (2011). Estimating constituent concentrations in case II waters from MERIS satellite data by semi-analytical model optimizing and look-up tables. *Remote Sensing of Environment, 115*, 1247–1259.

Appendix

The SAMO-LUT algorithm was originally proposed by Yang et al. (2011) for simultaneously retrieving the concentrations of three water constituents: phytoplankton, NAP, and CDOM. The basic idea of the SAMO-LUT involves the use of an imaginary Case-2 water body in which only one constituent changes while the other two are controlled as constants. A comprehensive synthetic dataset (rather than in situ data) was used for the model calibration. In this way, Yang et al. attempted to obtain not only a large number of samples without a sampling bias for model calibration, but also a series of special cases to avoid effects from other constituents and thus to improve the model performance (e.g., a dataset with only various Chl-*a* concentrations while the concentrations of NAP and CDOM were constant).

The main procedures of the SAMO-LUT are summarized as follows (Yu et al. 2014):

Step 1: Generation of the simulation dataset. The R_{rs} spectra were generated based on the specific inherent optical properties (SIOPs) from the target water and a bio-optical model. In the present study, only the SIOPs collected from Lake Dianchi were used due to the lack of complete SIOPs data for the other four lakes. The concentrations of Chl-*a* and NAP (tripton in the original paper), as well as the absorption coefficient of CDOM at 440 nm varied over the wide range of 1–300 mg m^{-3} (31 values), 1–250 g m^{-3} (28 values) and 0.1–10 m^{-1} (23 values). In all, 19,964 sample spectra were generated.

Step 2: Computation of the selected semi-analytical indices. Three semi-analytical

indices were selected for the estimation of Chl-*a*, NAP and CDOM, based on their reasonableness and performance. The selected indices were: a three-band index ([1/R$_{rs}$(665)−1/R$_{rs}$(709)]*R$_{rs}$(754)) for Chl-*a*, remote-sensing reflectance for the band centered at 754 nm, R$_{rs}$(754) for NAP, and the band-ratio R$_{rs}$(560)/R$_{rs}$(665) for CDOM. The synthetic reflectances were resampled to the bandwidths of the Medium Resolution Imaging Spectrometer (MERIS) sensor based on its spectral response function, and then the selected indices were calculated.

Step 3: Construction of the look-up tables (LUTs). We constructed three two-dimensional LUTs containing the coefficients of the estimation model for one constituent of interest, determined by the concentrations of the other two constituents. For instance, for the estimation of Chl-*a*, increments of 1 mg m^{-3} for NAP and those of 0.1 m^{-1} for CDOM were used in the ranges of 1–250 mg m^{-3} and 0.1–10 m^{-1}, respectively, and the regression coefficients corresponding to different combinations of NAP and CDOM were stored in the LUT.

Step 4: Initial estimations of Chl-*a* and NAP. We derived the initial values of Chl-*a* and NAP using two general estimation models obtained through a regression analysis between the simulated reflectance and corresponding Chl-*a* and NAP values. The two general estimation models were:

$$\text{Chl}^-a = 223.86(R_{rs,665}^{-1} - R_{rs,708}^{-1}) \times R_{rs,753} + 23.95 \tag{A-1}$$

$$NAP = 49909R_{rs,753}^2 - 61.38R_{rs,753} + 4.74 \tag{A-2}$$

The calculated initial Chl-*a* and NAP values were then used to estimate the initial CDOM value through a prepared LUT in Step 3.

Step 5: Iteration to select more appropriate model coefficients. The estimation models were improved according to the initial Chl-*a*, NAP, and CDOM values. After that, the refined Chl-*a*, NAP, and CDOM values were obtained by using the improved estimation models.

Step 6: End of the iteration. We found a more appropriate estimation model from the LUTs for each water constituent through the iterative use of the newly obtained Chl-*a*, NAP and CDOM values. The iteration was stopped when the difference between the current and last output was sufficiently small. Generally, the differences become stable after the 10th iteration.

Table 1. Basic information and statistical description of water constituent concentrations for Lakes Erhai and Dianchi (China) and Lakes Biwa, Suwa and Kasumigaura (Japan)

	Japanese Lakes			Chinese Lakes	
	Biwa	Suwa	Kasumigaura (western)	Erhai	Dianchi
Location	35.33°N	36.05°N	36.03°N	25.82°N	24.83°N
	136.17°E	138.08°E	140.40°E	100.18°E	102.72°E
Area (km^2)	670	13.3	171	249	300
Average depth (m)	41.0	4.7	4.0	11.0	4.3
Maximum depth (m)	104	7.2	7.3	21	11.3
Chl-a (mg m^{-3})					
Min	1.8	9.8	36.6	9.7	30.3
Max	2.9	11.4	95.0	36.1	153.9
Mean	2.2	10.7	66.5	19.6	87.7
Stdv	0.4	0.6	19.5	5.9	29.2
CV	19.3	5.9	29.3	30.2	33.2
TSS (g m^{-3})					
Min	0.8	4.8	11.7	3.5	24.5
Max	2.0	6.1	47.9	11.3	55.0
Mean	1.0	5.4	24.5	5.8	37.4
Stdv	0.4	0.4	8.2	1.7	7.8
CV	36.6	8.0	33.7	29.5	20.9
ISS (g m^{-3})					
Min	0.5	1.2	3.1	0.2	0.5
Max	1.5	2.2	37.3	1.6	42.3
Mean	0.6	1.7	16.3	0.8	12.4
Stdv	0.3	0.4	7.3	0.4	11.8
CV	49.2	24.4	44.7	51.5	95.1
CDOM (m^{-1})					
Min	0.15	0.41	0.51	0.33	0.41
Max	0.26	0.48	1.78	0.56	3.98
Mean	0.21	0.44	0.90	0.41	1.25
Stdv	0.03	0.02	0.29	0.06	0.84
CV	15.6	5.4	32.0	13.5	66.9

Chl-a: Chlorophyll-a concentration; TSS: total suspended solids; ISS: inorganic suspended solids;

CDOM: colored dissolved organic matters; Stdv: standard deviation;

CV: coefficient of variation in % and calculated as the ratio of Stdv to Average

Table 2. Performances of all selected Chl-*a* estimation algorithms in the five lakes

	OC4E	RN2-Git11	RN2-Gil10	RN2-Gur11	RN3-Git11	RN3-Gil10	RN3-Gur11	SAMO-LUT
Lake Biwa, Japan (N=10)								
RMSE (mg m^{-3})	0.30	5.96	7.04	7.57	22.58	**	6.60	5.90
NMAE(%)	9.4	191.3	282.4	307.5	938.1	**	266.5	221.8
MNB (%)	-1.1	171.6	282.4	307.5	-938.1	**	-255.5	221.8
NRMS (%)	12.3	245.3	190.1	196.4	299.7	**	130.6	141.5
R^2	0.54	-0.14	-0.14	-0.14	-0.11	**	-0.10	0.07
Lake Suwa, Japan (N=8)								
RMSE (mg m^{-3})	1.73	2.93	3.01	3.69	4.19	2.62	2.76	2.77
NMAE(%)	12.2	22.6	23.9	30.2	27.4	19.1	20.0	19.8
MNB (%)	-1.0	20.9	23.9	30.2	-14.6	-2.3	7.0	8.3
NRMS (%)	15.6	16.3	12.8	13.4	35.1	23.8	24.3	23.9
R^2	0.0033	0.0076	0.0077	0.0078	0.0007	0.0009	0.0011	-0.0004
Lake Kasumigaura, Japan (N=46)								
RMSE (mg m^{-3})	59.60	16.23	22.11	15.46	12.98	14.83	16.98	13.26
NMAE(%)	78.8	17.4	22.2	16.0	19.4	15.3	23.4	16.2
MNB (%)	-78.8	-6.5	-20.2	-5.4	12.8	-8.4	18.9	-3.9
NRMS (%)	13.1	20.3	17.4	19.5	20.5	16.7	20.3	18.4
R^2	-0.50	0.50	0.50	0.49	0.63	0.63	0.60	0.64
Lake Dianchi, China (N=28)								
RMSE (mg m^{-3})	79.14	27.03	35.93	22.75	14.61	10.99	41.26	7.39
NMAE(%)	81.3	23.6	32.1	21.4	14.4	11.5	35.9	6.9
MNB (%)	-81.3	-20.7	32.0	17.1	13.8	6.6	34.9	1.3
NRMS (%)	8.7	15.2	13.2	16.0	15.0	12.2	23.2	11.3
R^2	0.05	0.75	0.75	0.76	0.91	0.91	0.92	0.94
Lake Erhai, China (N=42)								
RMSE (mg m^{-3})	12.28	3.79	3.06	3.34	5.61	4.23	4.37	3.88
NAME (%)	50.3	16.3	11.3	13.9	25.8	18.4	18.8	17.1
MNB (%)	-50.1	9.4	2.1	8.6	-6.0	-12.4	-3.5	-2.8
NRMS (%)	18.4	17.4	14.2	15.0	31.8	21.7	23.4	21.2
R^2	0.01	0.73	0.73	0.74	0.74	0.74	0.74	0.71

**: unavailable results

Table 3. Statistical description of the Chl-*a* concentrations for each group divided by two MCI thresholds

Group	Min. Chl-a	Max. Chl-a	Mean Chl-a	Stdv	CV
1. MCI<=0.0001 (N=10)	1.8	2.9	2.2	0.4	19.3
2. 0.0001<MCI<=0.0016 (N=33)	9.7	36.1	16.6	6.8	40.7
3. MCI>0.0016 (N=91)	15.5	153.9	64.6	31.2	48.3

Chl-a: Chlorophyll-a concentration; stdv: standard deviation; CV: coefficient of variation in % and calculated as the ratio of Stdv to average.

Table 4. Statistical description of MCI values for each group divided by the two thresholds of Chl-*a* concentration

Group	Min. MCI	Max. MCI	Mean MCI	Stdv	CV
1. Chl-a<=10 mg m^{-3}(N=12)	-0.0001	0.0004	0.0000	0.0002	**
2. 10 mg m^{-3}<Chl-a<=25 mg m^{-3} (N=42)	0.0004	0.0025	0.0012	0.0007	58.3
3. Chl-a>25 mg m^{-3} (N=80)	0.0011	0.0215	0.0088	0.0052	59.1

Chl-a: Chlorophyll-a concentration; stdv: standard deviation; CV: coefficient of variation in % and calculated as the ratio of Stdv to average.

Table 5. Performances of all selected Chl-*a* estimation algorithms for different classes by MCI as well as the entire dataset

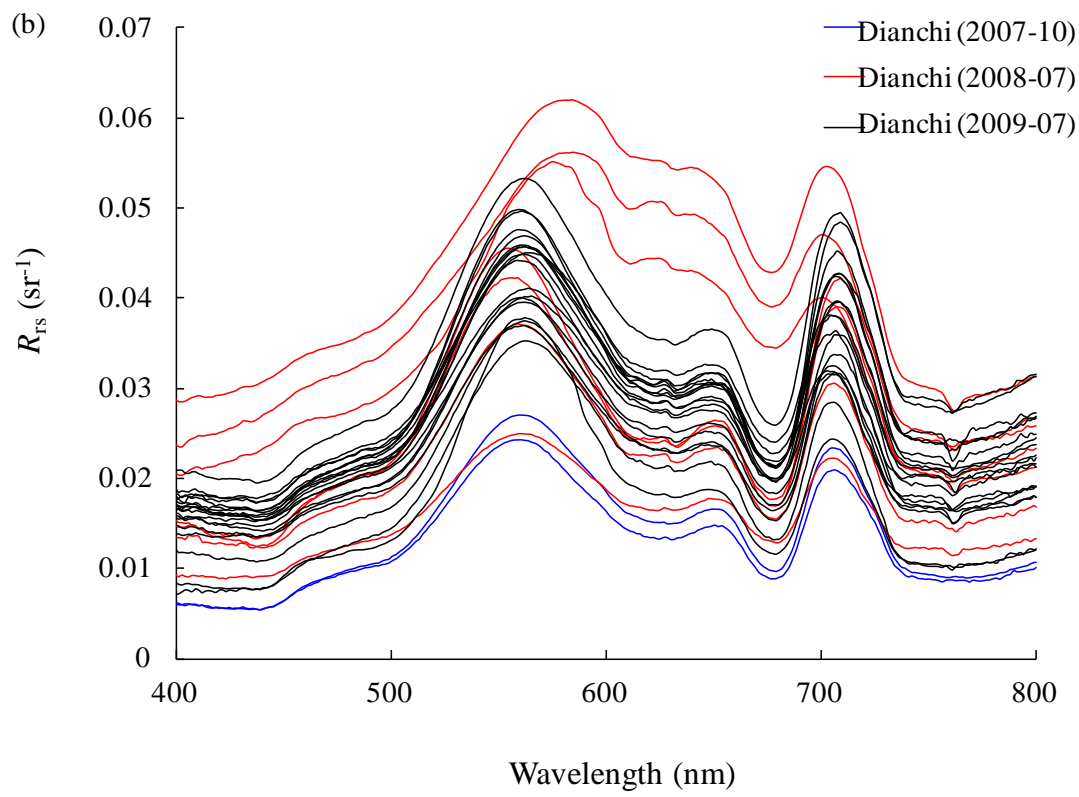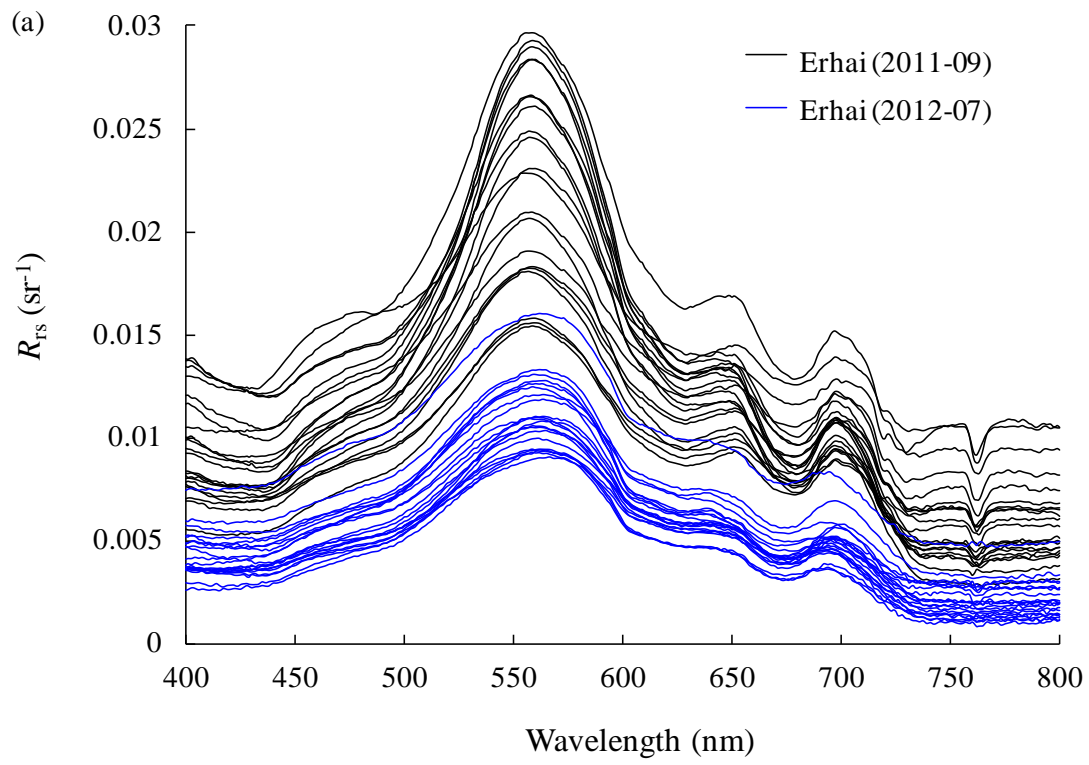|	OC4E	RN2-Git11	RN2-Gil10	RN2-Gur11	RN3-Git11	RN3-Gil10	RN3-Gur11	SAMO-LUT
MCI≤0.0001 (N=10)								
RMSE (mg m^{-3})	0.30	5.96	7.04	7.57	22.58	**	6.66	5.90
NMAE(%)	9.4	191.3	282.4	307.5	938.1	**	266.5	221.8
MNB (%)	-1.1	171.6	282.4	307.5	-938.1	**	-255.5	221.8
NRMS (%)	12.3	245.3	190.1	196.4	299.7	**	130.6	141.5
R^2	0.54	-0.14	-0.14	-0.14	-0.11	**	-0.10	0.07
0.0001<MCI≤0.0016 (N=33)								
RMSE (mg m^{-3})	11.03	3.87	3.48	3.75	6.19	4.74	4.81	4.04
NMAE(%)	43.3	18.8	16.0	19.5	32.5	24.1	23.7	20.1
MNB (%)	-40.5	10.2	7.2	13.4	-15.9	-14.5	-5.6	-4.7
NRMS (%)	27.7	20.2	18.2	19.1	35.9	25.8	27.5	24.4
R^2	-0.03	0.74	0.75	0.75	0.71	0.70	0.70	0.72
MCI>0.0016 (N=91)								
RMSE (mg m^{-3})	60.67	18.78	25.15	16.61	12.26	12.12	25.61	10.29
NMAE(%)	73.4	18.9	22.6	16.8	16.8	12.8	24.8	12.8
MNB (%)	-73.4	7.2	-19.6	-6.3	12.4	-6.9	21.3	-0.4
NRMS (%)	19.2	20.8	19.1	19.3	17.5	14.2	22.2	15.9
R^2	0.01	0.82	0.82	0.82	0.90	0.90	0.87	0.90
Whole dataset (N=134)								
RMSE (mg m^{-3})	50.20	15.64	20.84	13.93	12.06	**	21.27	8.83
NMAE (%)	61.3	31.7	40.4	39.2	89.4	**	42.6	30.2
MNB (%)	-59.9	10.4	9.5	22.0	-65.5	**	-6.0	15.1
NRMS (%)	30.3	81.4	94.6	98.2	261.9	**	82.9	71.7
R^2	0.19	0.89	0.90	0.90	0.93	**	0.93	0.94

**: unavailable results
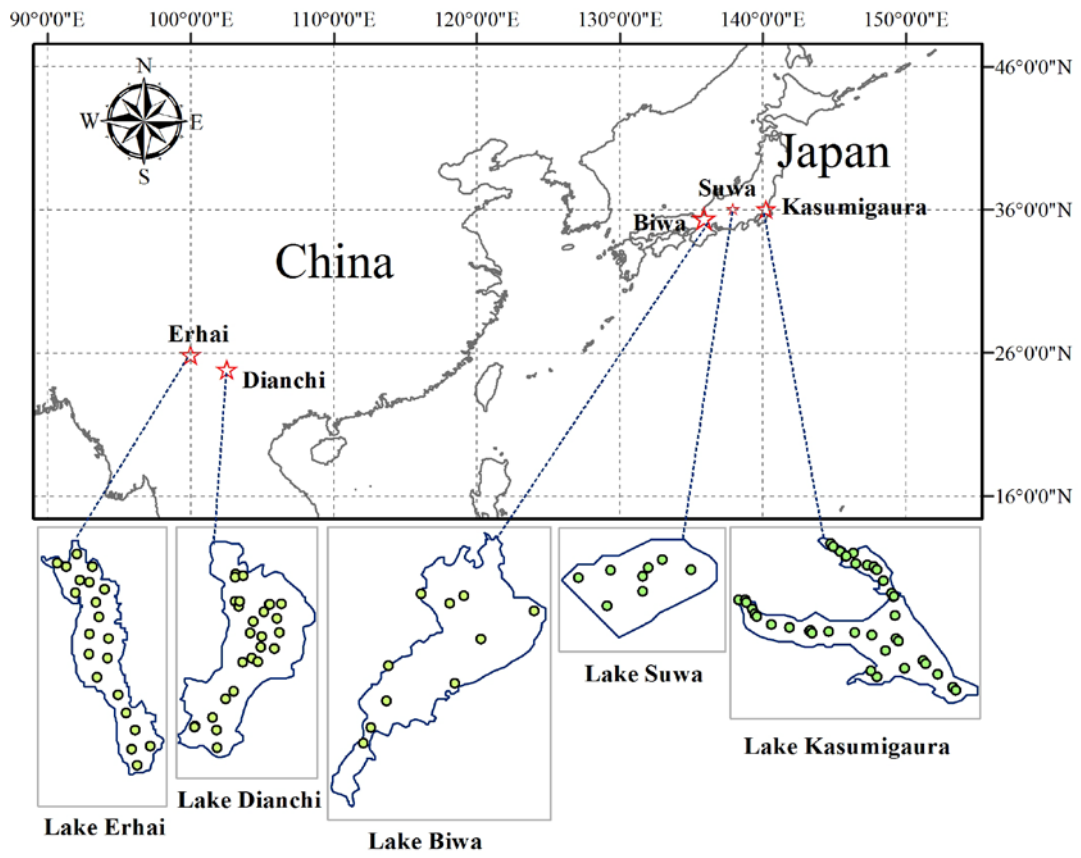
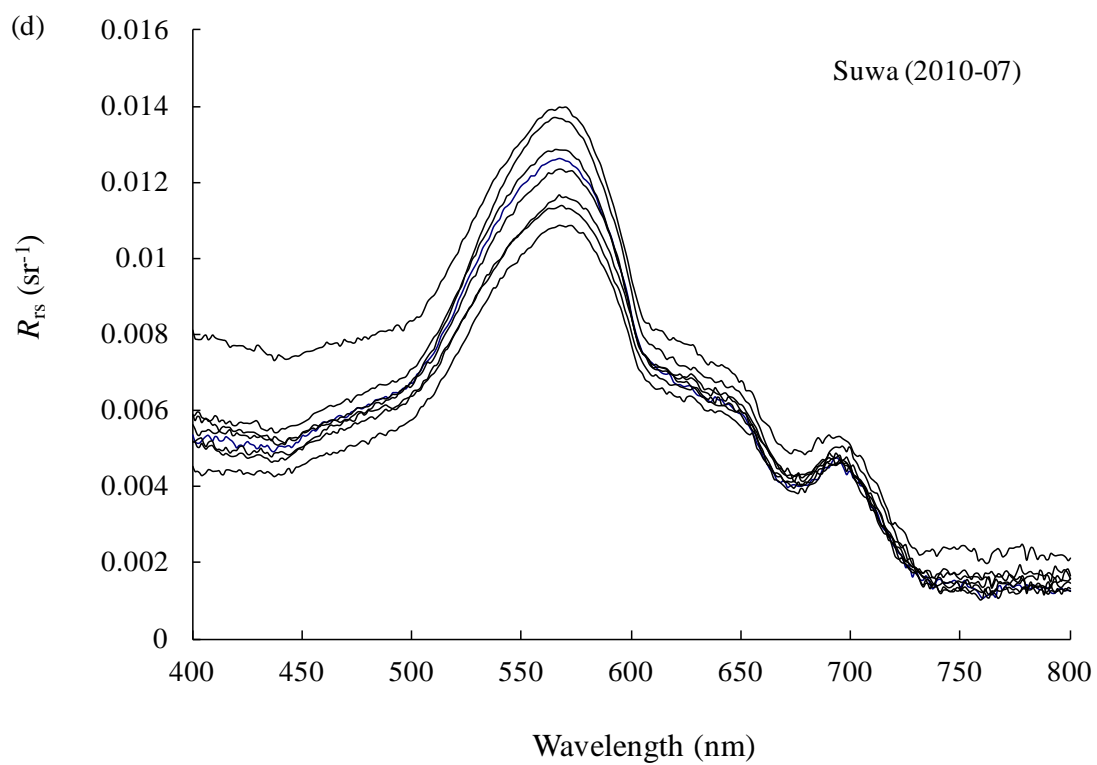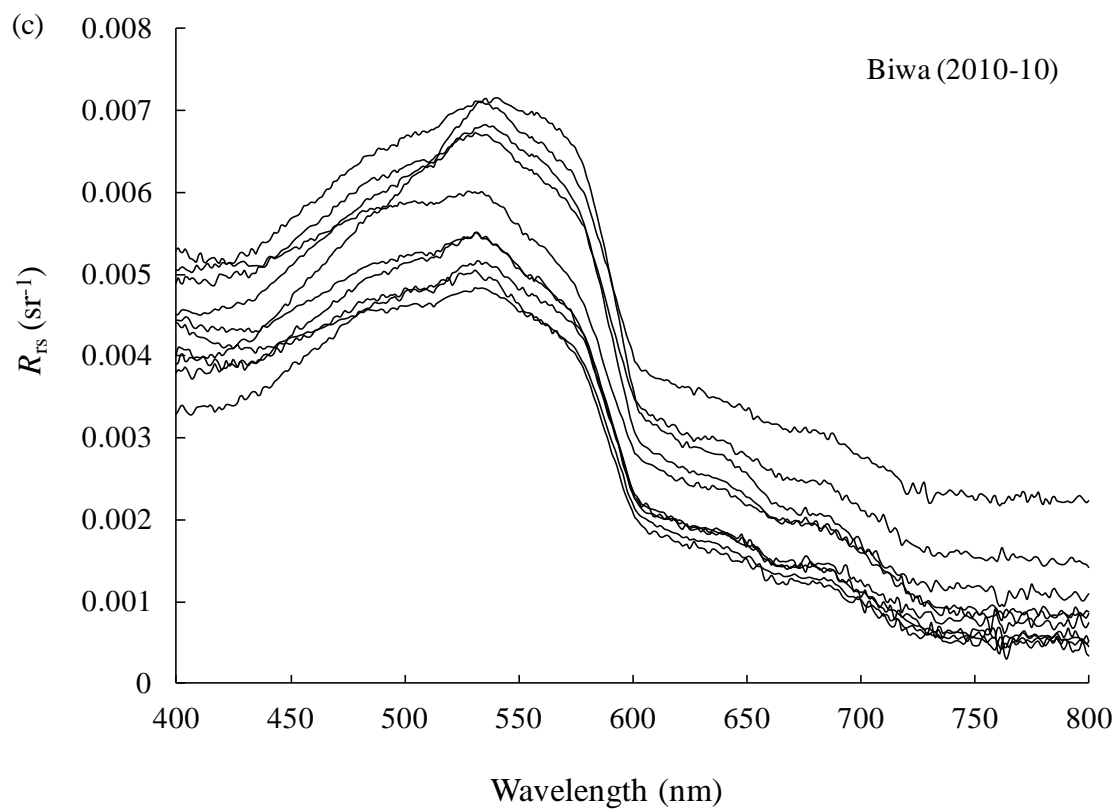**Fig. 1.** Distribution of sampling sites in Lakes Erhai and Dianchi of China and Lakes Biwa, Suwa and Kasumigaura of Japan.

(a)

Erhai (2011-09)
Erhai (2012-07)

(b)

Dianchi (2007-10)
Dianchi (2008-07)
Dianchi (2009-07)
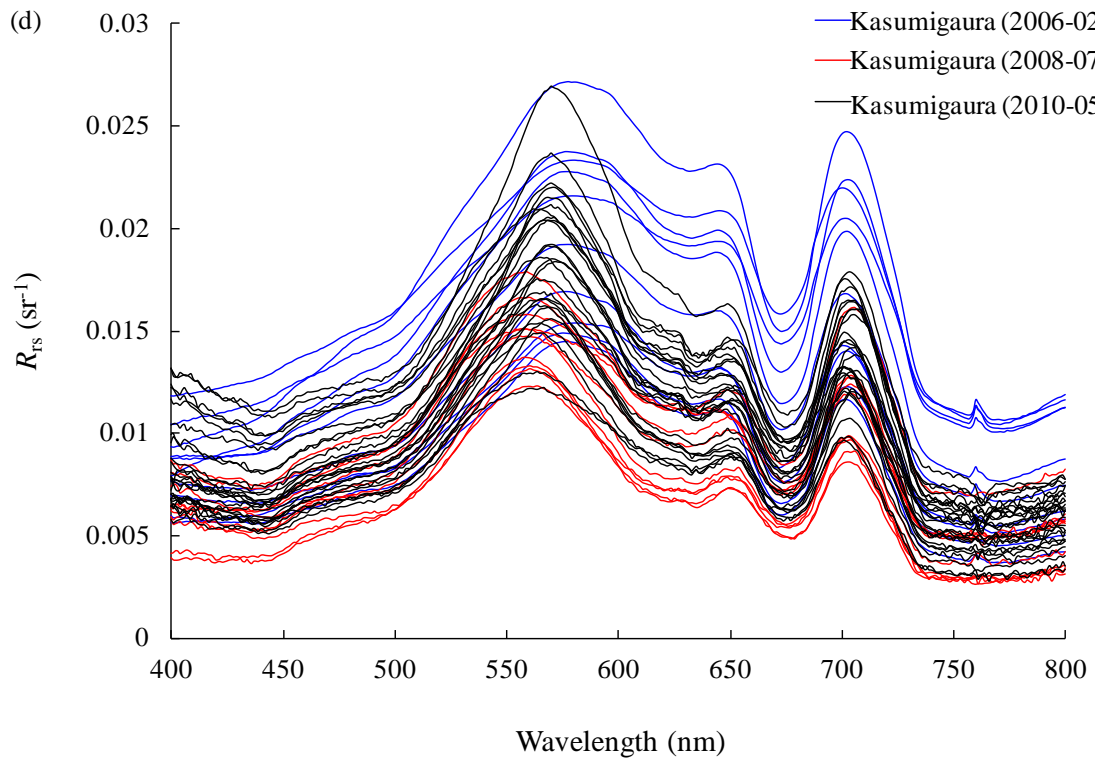
(c) Biwa (2010-10)

(d) Suwa (2010-07)

Fig. 2. Remote-sensing reflectance collected from Lakes Erhai and Dianchi (China) and Lakes Biwa, Suwa and Kasumigaura (Japan).
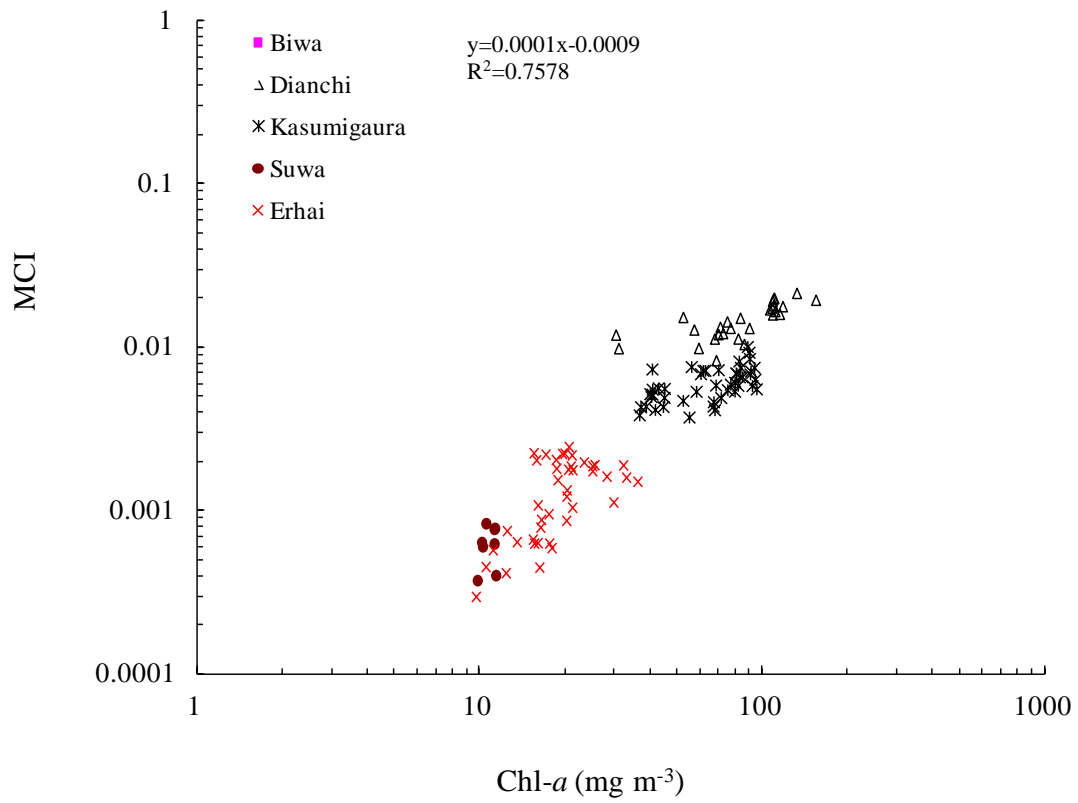
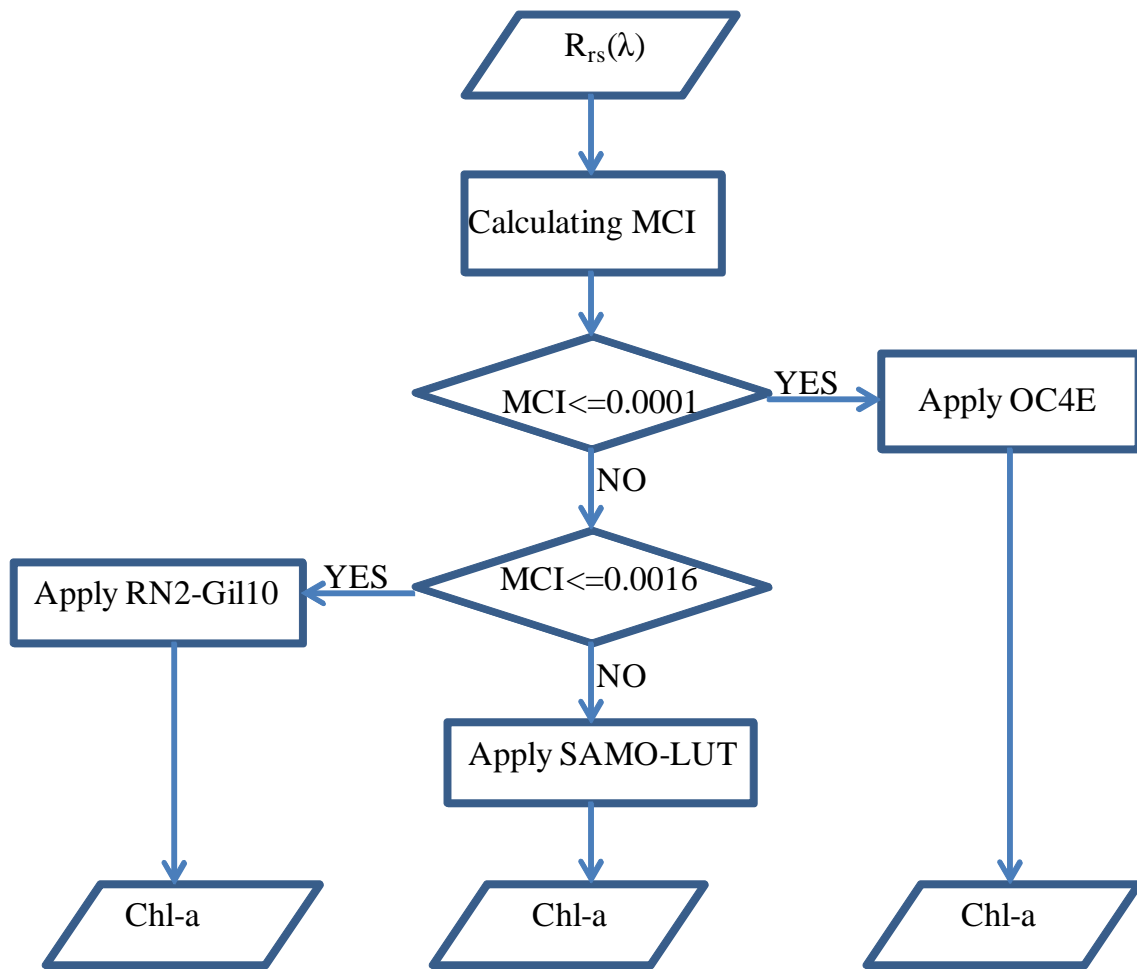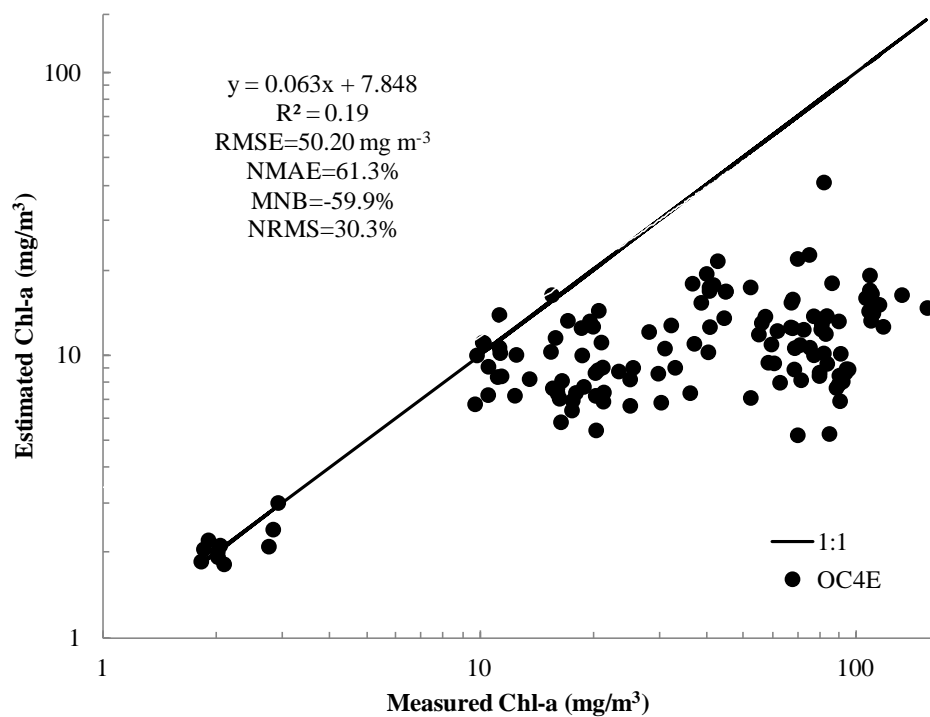**Fig. 3.** Relationship between the MCI and Chl-*a* values based on the data collected from the five lakes.

Fig. 4. Flowchart of the proposed hybrid algorithm.
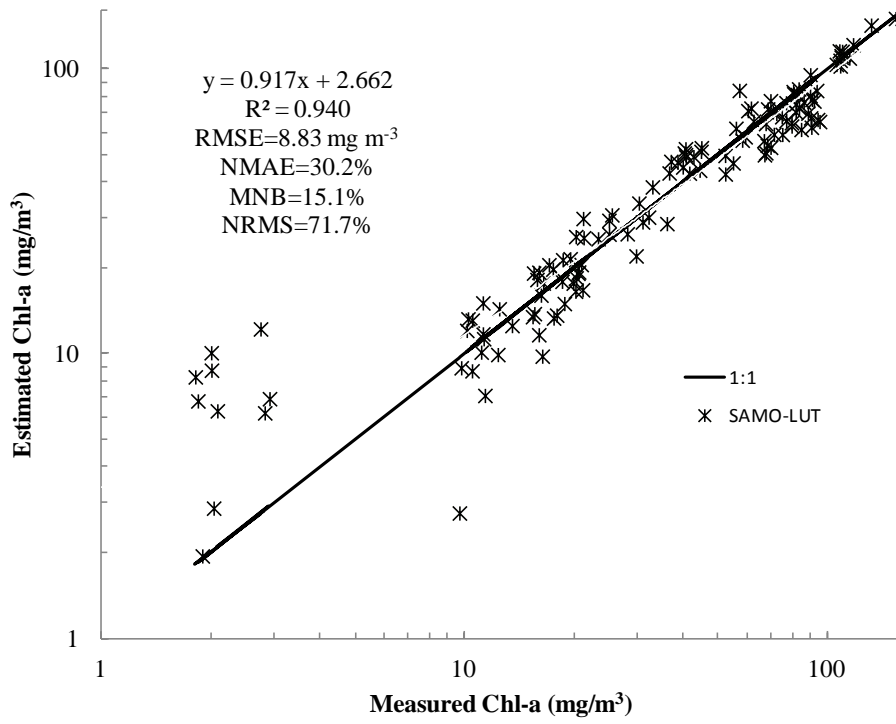
(a)

y = 0.063x + 7.848
$R^2 = 0.19$
RMSE=50.20 mg m^{-3}
NMAE=61.3%
MNB=-59.9%
NRMS=30.3%

——1:1
● OC4E

Estimated Chl-a (mg/m^3)

Measured Chl-a (mg/m^3)

(b)

y = 0.545x + 10.515
$R^2 = 0.898$
RMSE=20.84 mg m^{-3}
NMAE=40.4%
MNB=9.5%
NRMS=94.6%

——1:1
△ RN2-Gil10

Estimated Chl-a (mg/m^3)
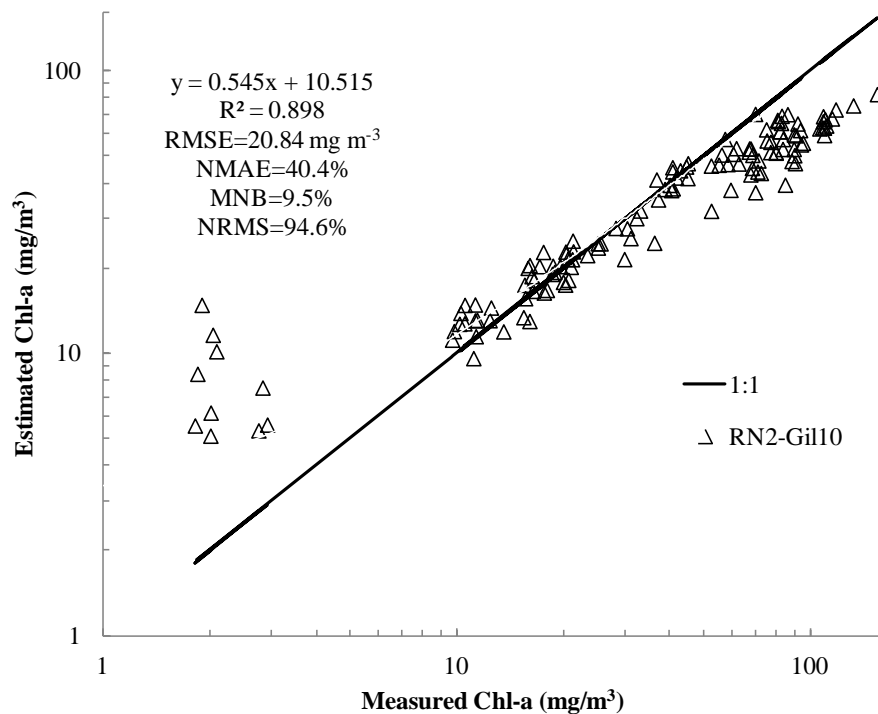
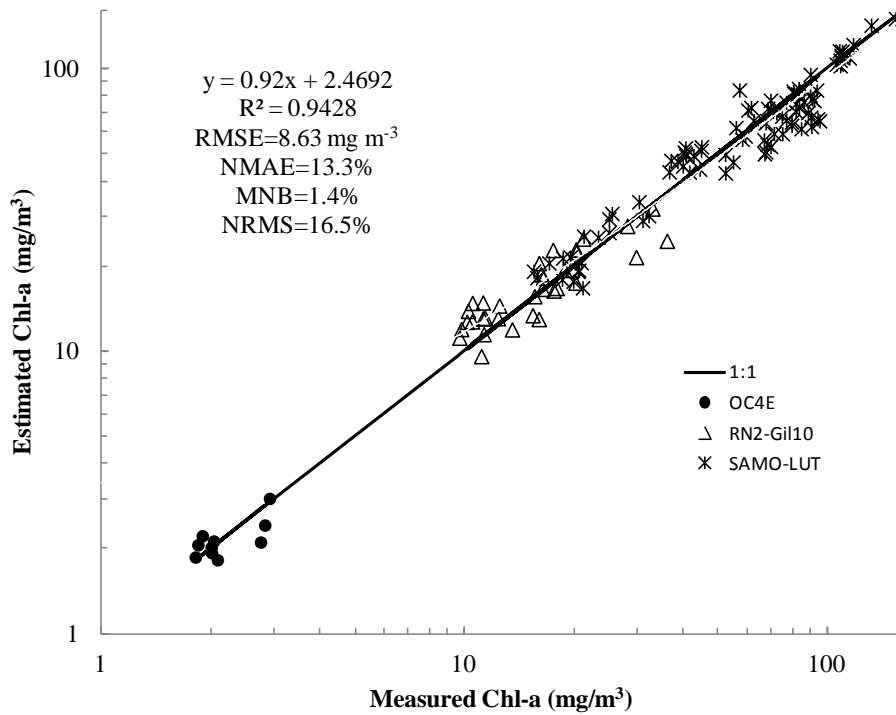**Measured Chl-a (mg/m^3)**

(c)

(d)

Fig. 5. Comparison of the measured and estimated Chl-*a* values for the entire dataset

(n=134). (a) A blue-green algorithm (OC4Ev4), (b) a two-band red-NIR model (RN2Gil10), (c) a three-band red-NIR model (SAMO-LUT), and (d) the proposed hybrid algorithm.