Estimating constituent concentrations in case II waters from MERIS satellite data by semi-analytical model optimizing and look-up tables

Yang Wei, Matsushita Bunkei, Chen Jin, Fukushima Takehiko

Remote sensing of environment

Volume 115

Number 5

Page range 1247-1259

Year 2011-05

(C) 2011 Elsevier Inc.

NOTICE: this is the author’s version of a work that was accepted for publication in Remote Sensing of Environment. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in PUBLICATION, 115, 5, 2011, DOI:10.1016/j.rse.2011.01.007

URL http://hdl.handle.net/2241/113240
doi: 10.1016/j.rse.2011.01.007
Estimating constituent concentrations in case II waters from MERIS satellite data by semi-analytical model optimizing and look-up tables

Wei Yang¹,², Bunkei Matsushita²*, Jin Chen¹, Takehiko Fukushima²

¹ State key laboratory of earth surface processes and resource ecology, Beijing Normal University, Beijing, 100875, China
E-mail: yangwei1022@gmail.com, chenjin@ires.cn

² Graduate School of Life and Environmental Sciences, University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki, 305-8572, Japan
E-mails: mbunkei@sakura.cc.tsukuba.ac.jp, fukusima@sakura.cc.tsukuba.ac.jp

Submittal date: 2010-9-17

*Corresponding Author
E-mail: mbunkei@sakura.cc.tsukuba.ac.jp
Abstract

Remote estimation of water constituent concentrations in case II waters has been a great challenge, primarily due to the complex interactions among the phytoplankton, tripton, colored dissolved organic matter (CDOM) and pure water. Semi-analytical algorithms for estimating constituent concentrations are effective and easy to implement, but two challenges remain. First, a dataset without a sampling bias is needed to calibrate estimation models; and second, the semi-analytical indices were developed based on several specific assumptions that may not be universally applicable. In this study, a semi-analytical model-optimizing and look-up-table (SAMO-LUT) method was proposed to address these two challenges. The SAMO-LUT method is based on three previous semi-analytical models to estimate chlorophyll $a$, tripton and CDOM. Look-up tables and an iterative searching strategy were used to obtain the most appropriate parameters in the models. Three datasets (i.e., noise-free simulation data, in situ data and MEdium Resolution Imaging Spectrometer (MERIS) satellite data) were collected to validate the performance of the proposed method. The results show that the SAMO-LUT method yields error-free results for the ideal simulation dataset; and is able also to accurately estimate the water constituent concentrations with an average bias (mean normalized bias, MNB) lower than 9% and relative random uncertainty (normalized root mean square error, NRMS) lower than 34% even for in situ and MERIS data. These results demonstrate the potential of the proposed algorithm to accurately monitor...
inland and coastal waters based on satellite observations.

Key words: semi-analytical models; look-up table; bio-optical model; case II water

1. Introduction

Accelerated eutrophication of inland water is becoming a significant environmental issue all over the world (Ayres et al., 1996). The sustainable management of freshwater ecosystems requires the routine monitoring of water quality. However, the spatial and temporal heterogeneity of water bodies coupled often result in inadequate monitoring and characterization of water quality using conventional sampling methods (Khorram et al., 1991; Liu et al., 2003). Consequently a combined approach utilizing the spatial and temporal coverage of remote sensing with conventional water sampling provides a potentially effective solution to monitoring freshwater ecosystems.

From the remote sensing perspective, aquatic environments can be classified as either case I or case II waters (Morel and Prieur, 1977). While case I waters are those dominated by phytoplankton (e.g. open ocean), case II waters contain tripton, dissolved organic matter in addition to phytoplankton. It has been shown that concentrations of total suspended solids (TSS) and organic matter are not necessarily correlated with chlorophyll $a$ concentration in both coastal and inland case II waters (Gin et al., 2003; Nichol, 1993). The use of a remote sensing technique for water quality monitoring in case II waters has been far less successful compared with that in case I waters, due mainly to the complex interactions of the four optically active substances (OASs; i.e., phytoplankton,
tripton, colored dissolved organic matter (CDOM) and pure water) in case II waters (Doxaran et al., 2002; Gin et al., 2002; Goodin et al., 1993).

To address the difficulties of monitoring case II waters, researchers have made substantial efforts to accurately estimate water constituent concentrations (including concentrations of chlorophyll \(a\) and tripton as well as absorption of CDOM at 440 nm). These efforts have included the use of derivative values of reflectance spectra (Goodin et al., 1993), examination of the band ratio of near infrared reflectance and red reflectance (Han et al., 1994), use of the inherent optical property (IOP) inversion technique based on a bio-optical model (Brando and Dekker, 2003; Garver and Siegel, 1997; Santini et al., 2010) or on the use of Hydrolight (Mobley et al., 2005; Van Der Woerd and Pasterkamp, 2008), and use of the spectral mixture analysis technique to minimize the interactions of the four OASs (Novo et al., 2006; Oyama et al., 2009; Oyama et al., 2007; Svab et al., 2005; Tyler et al., 2006).

Additionally, several semi-analytical algorithms have been proposed to estimate water constituent concentrations in case II waters (e.g., Ammenberg et al., 2002; Dall'Olmo et al., 2005; Doxaran et al., 2002; Gitelson et al., 2008; Morel and Gentili, 2009). These algorithms are generally composed of two key steps. The first step is to develop an index, which could be the reflectance of a single band or the arithmetic combination of reflectance from several bands, by analyzing the IOPs of water constituents. The second step is to empirically establish and calibrate the relationships between the indices obtained from \textit{in situ} reflectance data or satellite data and water constituent concentrations. The relationship could be a linear function (Ammenberg et al., 2002; Gitelson et al., 2008), power function (Kutser et al., 2005), or polynomial function (Dall'Olmo et al., 2003),
depending on the regression analysis and the dynamic ranges in the calibration data used. Since these proposed indices are based on the spectral analysis of IOP for each OAS, they can effectively minimize the effects of the other OASs on the OAS of interest. Therefore, the algorithms based on these indices not only promise improved performance for predicting the water constituent concentration of interest but also are easy to implement using satellite data. However, there still are two major challenges in the application of the semi-analytical algorithms. First, the models for estimation of water constituent concentrations depend greatly on the calibration process; thus, a dataset without a sampling bias is needed to calibrate these models. Second, the proposed indices were developed based on several specific assumptions, some of which may not be universally applicable. For example, an important assumption in the three-band index for estimating chlorophyll concentration is that the absorption and backscattering of suspended solids at the near-infrared band (750-760 nm) can be neglected compared with the absorption of pure water (Gitelson et al., 2008). However, this assumption is not applicable in some highly turbid case II waters, such as those of Lake Taihu and Lake Dianchi in China and Lake Kasumigaura in Japan, and thus resulted in large errors in chlorophyll concentration estimates in these lakes (Le et al., 2009; Yang et al., 2010).

Consequently, the main objective of the present study was to propose a novel method by integrating several semi-analytical algorithms with a look-up-table method to address the two challenges described above. To evaluate the performances of the proposed method, three datasets obtained from bio-optical model simulation, field surveys and Medium Resolution Imaging Spectrometer (MERIS) data were used in this study.
2. Methods

2.1 Bio-optical model

According to Gordon et al. (1975), the remote-sensing reflectance just beneath the water surface can be expressed as:

\[ R_{s}(\lambda,0^-) = \frac{f}{Q} \times \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \]  \hspace{1cm} (1)

where \( a(\lambda) \) and \( b_b(\lambda) \) are the spectral total absorption and backscattering coefficients, respectively, \( f \) is the anisotropic factor of the downwelling light field; and \( Q \) is the geometrical factor. Austin (1980) proposed the factor of 0.544 for relating radiance just above the surface to radiance just beneath the surface. Thus, remote-sensing reflectance just above the water surface is determined as follows:

\[ R_{s}(\lambda) = 0.544 \times \frac{f}{Q} \times \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \]  \hspace{1cm} (2)

Kirk (1994) found that \( f \) is a function of the solar elevation angle that was reasonably well expressed as a linear function of \( \mu_0 \), the mean cosine of the zenith angle of the refracted photons as follows:

\[ f = 0.975 - 0.629 \mu_0 \] \hspace{1cm} (3)

The value of \( \mu_0 \) depends on the solar elevation and the proportion of direct and diffuse radiations. It is calculated according to the sampling time, locations and solar zenith angle. \( Q \) is usually expected to range from 3 to 4 (Morel and Gentili, 1993). Gons (1999) proposed an empirical equation of \( Q = 2.38/\mu_0 \) for turbid inland waters under different solar elevation angles. This equation has been successfully applied for modeling remote-sensing reflectance in an extremely turbid case II water (i.e. Lake Taihu, China; Zhang et al., 2009). Since Lake Dianchi is also a turbid lake, the same equation was used in this study.
The spectral total absorption coefficient, $a(\lambda)$, is usually expressed as the sum of the constituents' absorption coefficients, as follows:

$$a(\lambda) = a_w(\lambda) + [\text{Chl-a}] a_{\text{ph}}(\lambda) + [\text{TR}] a_{\text{tr}}(\lambda) + [\text{CDOM}] a_{\text{CDOM}}(\lambda),$$

where $[\text{Chl-a}]$ and $[\text{TR}]$ denote concentrations of chlorophyll $a$ and tripton, respectively; $[\text{CDOM}]$ denotes the absorption of CDOM at 440 nm; $a_w(\lambda)$ is the absorption coefficient of pure water; and $a_{\text{ph}}(\lambda)$, $a_{\text{tr}}(\lambda)$ and $a_{\text{CDOM}}(\lambda)$ are specific absorption coefficients for phytoplankton, tripton and CDOM, respectively. The spectral total backscattering coefficient is expressed as the sum of the backscattering coefficients for each constituent in water except for CDOM, as follows:

$$b_b(\lambda) = b_{b,w}(\lambda) + [\text{Chl-a}] b_{b,\text{ph}}(\lambda) + [\text{TR}] b_{b,\text{tr}}(\lambda),$$

where $b_{b,w}(\lambda)$ is the backscattering coefficient of pure water, and $b_{b,\text{ph}}(\lambda)$ and $b_{b,\text{tr}}(\lambda)$ are the backscattering coefficients for phytoplankton and tripton, respectively.

### 2.2 Study areas and In situ data

Field investigations were carried out in Lake Dianchi (24°50′N; 102°41′E) and Lake Kasumigaura (36°00′N; 140°25′E). Lake Dianchi is located in a plateau area of the southwestern part of China (Fig.1A). It has a surface area of 300 km$^2$ and is the largest lake in Yunnan Province and the sixth largest lake in China. The mean depth of the lake is 4.3 m and the maximum depth is 11.3 m. Eutrophication has become more and more serious in the lake in the recent 20 years due to the large quantities of industrial wastewater and municipal sewage discharged into the lake; algal blooms occur frequently from April to November each year (Gao et al., 2005). Lake Kasumigaura is located in the eastern part of Japan’s Kanto Plain (Fig.1B). It is the second largest lake of Japan, with a surface area of 171 km$^2$ and an average depth of 4 m (maximum depth of 7.3 m). The lake is
considered eutrophic, because it has a high load of nutrients, and because of its shallow depth (Fukushima et al., 1996).

Three data collection campaigns were undertaken in 2007 (Oct. 23) and 2009 (Mar. 12 and Jul. 24-31) in Lake Dianchi; and other two campaigns were undertaken in 2006 (Feb. 18) and 2008 (Aug. 07) in Lake Kasumigaura. The spatial distribution of sampling sites is shown in Fig. 1. Sample collections and reflectance measurements were performed between 10:00 and 14:00 h local time. Water samples were kept in ice boxes and taken to the laboratory within approximately 0.5 hours after whole data collections. Chlorophyll $a$ was extracted using methanol (100%) at 4°C under dark conditions for 24 hours. The optical density of the extracted chlorophyll $a$ was measured at four wavelengths (750, 663, 645 and 630 nm) and the concentration was calculated according to SCOR-UNESCO equations (SCOR-UNESCO, 1966). To obtain the concentration of tripton, the total suspended solids (TSS) were divided into tripton and phytoplanktonic suspended solids (PSS). Based on the method of Gons et al. (1992) and organic suspended solids (OSS) data collected from two lakes, it can be assumed that 1 mg m$^{-3}$ chlorophyll $a$ concentration is approximately equal to 0.148 g m$^{-3}$ TSS in Lake Dianchi and 0.12 g m$^{-3}$ TSS in Lake Kasumigaura. Tripton concentrations were then derived by subtracting PSS from TSS. The absorption of CDOM was measured using a Shimadzu UV-1700 spectrophotometer with filtered water. In situ reflectance spectra were collected according to Method 1 of Mueller et al. (2000).
2.3 Estimation of SIOPs

Four water samples collected from Lake Dianchi under a very clear sky and low wind speed in Jul. 2009 were used to measure/estimate the absorption and backscattering spectra for CDOM, phytoplankton and tripton. The corresponding remote-sensing reflectance spectra were also collected at these four sites to estimate the backscattering spectra for phytoplankton and tripton by the method that will be described below. The absorption coefficients of tripton ($a_{tr}(\lambda)$), CDOM ($a_{CDOM}(\lambda)$) and phytoplankton ($a_{ph}(\lambda)$) were determined according to the quantitative filter technique (QFT) (Mitchell, 1990).

As demonstrated in Giardino et al. (2007), the specific absorption of tripton, $a^*_{tr}(\lambda)$, could be fitted with an exponential wavelength function as follows:

$$a^*_{tr}(\lambda) = a^*_{tr}(440) \exp[-S_{tr}(\lambda - 440)]$$  \hspace{1cm} (6)

where $a^*_{tr}(440)$ is the specific absorption of tripton at 440 nm and $S_{tr}$ is the shape factor of the absorption of tripton. Here, $a^*_{tr}(440)$ and $S_{tr}$ equal 0.0683 and 0.0115, respectively. The specific absorption of CDOM, $a^*_{CDOM}(\lambda)$, can also be described as an exponential function as follows:

$$a^*_{CDOM}(\lambda) = a^*_{CDOM}(440) \exp[-S_{CDOM}(\lambda - 440)],$$  \hspace{1cm} (7)

where $S_{CDOM}$ is the shape factor of the absorption spectra of CDOM. The value of $a^*_{CDOM}(440)$ equals 1 since the absorption spectra of CDOM is normalized by the absorption coefficient at 440 nm; $S_{CDOM}$ equals 0.0157 according to the best-fit results for the in situ collected data of this study. The absorption coefficient of pure water, $a_w(\lambda)$, was taken from Hale and Querry (1973) and Pope and Fry (1997).

The backscattering coefficients were obtained based on an inversion of the bio-optical model...
used in this study. According to equations 2 and 5, the backscattering coefficient of particles (including phytoplankton and tripton), \( b_{b,p}(\lambda) \), can be obtained as follows:

\[
b_{b,p}(\lambda) = \frac{a(\lambda)R_{b}(\lambda)}{0.544(f/Q)-R_{s}(\lambda)} - b_{b,w}(\lambda),
\]

(8)

where \( b_{b,w}(\lambda) \) is the backscattering coefficient of pure water cited from Morel (1974).

The separation of phytoplankton and tripton backscattering is based on the assumption that contributions of backscattering due to phytoplankton cells and tripton are proportional to the ratios of their masses (Brando and Dekker, 2003). As the total backscattering of particles can be derived from Eq. 8, specific backscattering coefficients for these two components can be retrieved. According to Giardino et al. (2007), the specific backscattering coefficient of tripton, \( b_{b,t}(\lambda) \), can also be found from following equation:

\[
b_{b,t}(\lambda) = b_{b,t}(550)\left(\frac{\lambda}{550}\right)^{-n},
\]

(9)

where \( b_{b,t}(550) \) is the specific backscattering coefficient of tripton at 550 nm, and \( n \) is an exponent describing the spectral dependency of tripton backscattering (Giardino et al., 2007). Here, \( b_{b,t}(550) \) and \( n \) equals 0.0116 and 0.7744, respectively. Fig. 2 shows the specific inherent optical properties (SIOPs) for each component collected in Lake Dianchi.

The measured and modeled \( R_{s}(\lambda) \) at two additional sites were compared to validate the estimated SIOPs of Lake Dianchi (Fig. 3). The two sites had moderate (59.29 mg m\(^{-3}\)) and high...
(131.79 mg m\(^{-3}\)) concentrations of chlorophyll \(a\). The results show strong agreement between the measured and modeled \(R_s(\lambda)\), suggesting that the measured specific absorption coefficients and retrieved specific backscattering coefficients are determined reasonably well.

2.4 MERIS data collection

Two full-resolution MERIS images (level-1b) covering Lake Dianchi were acquired on Oct. 24, 2007 (orbit 29531, starting time 03:32:09 UTC) and Mar. 13, 2009 (orbit 36774, starting time 03:29:13 UTC), which were each one day after the date of the corresponding field survey (Fig. 1A). Other two full-resolution MERIS images covering Lake Kasumigaura were acquired on Feb. 18, 2006 (orbit 20775, starting time 01:15:29 UTC) and Aug. 7, 2008 (orbit 33652, starting time 00:56:50 UTC), on which the two field campaigns were carried out (Fig. 1B). Images were analyzed using BEAM 4.0 software (Brockmann Consult, Geesthacht, Germany). The images were geo-located and masked for land, clouds and invalid reflectance. Atmospheric correction was performed using the SCAPE-M (Self-Contained Atmospheric Parameters Estimation for MERIS data) atmospheric processor, which outperformed other previous atmospheric correction algorithms for turbid inland lakes (Guanter et al., 2010).

3. Development of a novel method for retrieving water constituent concentrations

3.1 Performances of the original semi-analytical algorithms
Among a number of semi-analytical indices proposed for case II waters, three indices were
selected for further investigation based on their reasonableness and performance as reported in
previous studies (Ammenberg et al., 2002; Gitelson et al., 2008).

A three-band index proposed by Gitelson et al. (2008) was first selected to estimate chlorophyll
a concentration in case II waters. This index requires the reflectance at three bands as input, i.e.,
$\lambda_1=660-670$ nm, $\lambda_2=700-730$ nm, $\lambda_3=740-760$ nm. The selection of these bands is based on three
assumptions: (1) the chlorophyll $a$ absorption in the first band, $\lambda_1$, should be much larger than that in
the second band, $\lambda_2$; that is, $a_{\text{Chl}}(\lambda_1) >> a_{\text{Chl}}(\lambda_2)$; (2) the tripton and CDOM absorption in the first
band, $\lambda_1$, should be similar to that in the second band, $\lambda_2$; that is, $a_t(\lambda_1) = a_t(\lambda_2)$,
$a_{\text{CDOM}}(\lambda_1) = a_{\text{CDOM}}(\lambda_2)$; (3) the third band, $\lambda_3$, should be minimally affected by the total backscattering
and the absorption of phytoplankton, tripton and CDOM, but should have similar backscattering to
those of the first and second bands; that is, $a_w(\lambda_3) >> a_{\text{ph}}(\lambda_3) + a_t(\lambda_3) + a_{\text{CDOM}}(\lambda_3) + b_b(\lambda_3)$,
$b_b(\lambda_3) \approx b_b(\lambda_2)$. Thus, the three-band index relating [Chl-a] can be expressed as follows:

$$[R_n^{-1}(\lambda_1) - R_n^{-1}(\lambda_2)] \times R_n(\lambda_3)$$

$$\propto [a_{\text{Chl}}(\lambda_1) + a_w(\lambda_1) - a_w(\lambda_2)]/a_w(\lambda_3)$$

$$\propto [\text{Chl-a}]$$

The remote-sensing reflectance in the near-infrared (NIR) band was selected to estimate the
concentration of tripton (Ammenberg et al., 2002). Generally, backscattering of tripton is noticeably
larger than that of phytoplankton; that is, $b_{\text{brf}}(\lambda_{\text{NIR}}) >> b_{\text{bph}}(\lambda_{\text{NIR}})$. In the NIR region (700-800 nm), the
total absorption was approximately equal to the absorption by pure water; that is, $a_{\text{total}}(\lambda_{\text{NIR}})$
$\approx a_w(\lambda_{\text{NIR}})$. The absorption of pure water is also much larger than the backscattering of particles
remote-sensing reflectance in the NIR band relating $[\text{TR}]$ can be expressed as follows:

$$R_n(\lambda_{\text{NIR}}) = \frac{f}{Q} \frac{b_{h,w}(\lambda_{\text{NIR}}) + b_{h,tt}(\lambda_{\text{NIR}})}{a_{\text{wd}}(\lambda_{\text{NIR}}) + b_{h,tt}(\lambda_{\text{NIR}}) + b_{b,tt}(\lambda_{\text{NIR}}) + b_{b,w}(\lambda_{\text{NIR}})}$$

$$\approx \frac{f}{Q} \frac{b_{h,w}(\lambda_{\text{NIR}}) + b_{h,tt}(\lambda_{\text{NIR}})}{a_{\text{w}}(\lambda_{\text{NIR}})}$$

(11)

A band-ratio index proposed by Ammenberg et al. (2002) was selected to estimate the absorption of CDOM at 440 nm. The relationship between the index and $[\text{CDOM}]$ is expressed as follows:

$$R_n(664) / R_n(550) \propto [\text{CDOM}]$$

(12)

$R_n(550)$ is influenced by the absorption of CDOM, as well as scattering by all particulate matters, but it is not strongly influenced by the absorption of chlorophyll, while $R_n(664)$ is affected by both the absorption of chlorophyll and the backscattering by phytoplankton and tripton. Using $R_n(664)$ in the numerator of the band ratio therefore normalizes for variations in both phytoplankton absorption and for the effects of variations in backscattering (Ammenberg et al., 2002).

To thoroughly investigate the performances of these semi-analytical algorithms, a reflectance spectra dataset was generated using the SIOPs and the bio-optical model described above. The average value of $f/Q$ (0.156) obtained from the in situ data in Lake Dianchi was used in the spectra generation. The concentrations of chlorophyll $a$ ([Chl-a]) and tripton ([TR]), as well as the absorption of CDOM at 440 nm ([CDOM]) were varied in the ranges of 1-300 (mg m$^{-3}$), 1-250 (g m$^{-3}$) and 0.1-10 (m$^{-1}$), respectively. Similar with increments used in Kutser et al. (2006), the increment of constituent concentrations varied for different ranges. For [Chl-a], an increment of 1 mg m$^{-3}$ was used in the range of 1-10 mg m$^{-3}$, an increment of 2 mg m$^{-3}$ was used in the range of
10-20 mg m$^{-3}$, an increment of 10 mg m$^{-3}$ was used in the range of 20-60 mg m$^{-3}$ and an increment of 20 mg m$^{-3}$ was used in the range of 60-300 mg m$^{-3}$. For [TR], the increments of 1 g m$^{-3}$, 5 g m$^{-3}$ and 20 g m$^{-3}$ were used in the ranges of 1-10 g m$^{-3}$, 10-50 g m$^{-3}$ and 50-250 g m$^{-3}$, respectively. For [CDOM], the increments of 0.1 m$^{-1}$, 0.5 m$^{-1}$ and 1 m$^{-1}$ were used in the ranges of 0.1-1.0 m$^{-1}$, 1.0-5.0 m$^{-1}$ and 5.0-10.0 m$^{-1}$, respectively. Thus, 19,964 (i.e., $31 \times 28 \times 23$) sample spectra in total were generated to establish the relationships between the indices and water constituent concentrations.

Since the MERIS sensor has great potential in the remote sensing of case II waters, the generated spectra were resampled to MERIS bandwidths (shown in Table 1) through corresponding spectral response functions to calculate the semi-analytical indices.

Through regression analysis between the indices and constituent concentrations, the initial estimation models for [Chl-a], [TR] and [CDOM] could be obtained (Fig. 4):

\[
[\text{Chl-a}] = 223.86 \left[ \frac{R_{\alpha}(b_{10})}{R_{\alpha}(b_{5})} - \frac{R_{\alpha}(b_{10})}{R_{\alpha}(b_{7})} \right] + 23.95 
\]

(13)

\[
[\text{TR}] = 49909 R_{\alpha}^{2}(b_{10}) - 61.38 R_{\alpha}(b_{10}) + 4.74 
\]

(14)

\[
[\text{CDOM}] = 3.03 \left[ \frac{R_{\alpha}(b_{7})}{R_{\alpha}(b_{5})} \right] + 0.35 , 
\]

(15)

where $b_{5}$, $b_{7}$, $b_{9}$ and $b_{10}$ denote the bandwidths of MERIS bands 5 (555-565 nm), 7 (660-670 nm), 9 (703.75-713.75 nm) and 10 (750-757.5 nm), respectively. It is clearly seen that the assumptions in the indices development process resulted in larger errors in the estimations of the water constituent concentrations, even though noise-free simulation data was used. Especially for [CDOM] estimation,
a poor $R^2$ was given (Fig. 4C). This is because the $R_{ch}(550)$ will be influenced by chlorophyll absorption for samples with high [Chl-a].

To improve the performances of the semi-analytical algorithms, the bio-optical model was theoretically reanalyzed. From equations (1), (4) and (5), it can be seen that the remote-sensing reflectance in case II waters depends on concentrations of three constituents: phytoplankton, tripton and CDOM (absorption and backscattering coefficients of pure water are usually treated as constants). As in a microcosm experiment (i.e., making an artificial ecosystem in which some conditions are controlled to simulate behaviors of a simplified natural ecosystem; e.g., Hunter et al., 2008), we can consider an imaginary case II water, in which only one constituent changes while the other two constituents are controlled as constants. For example, the imaginary case II water has variable concentrations of chlorophyll $a$ but constant concentrations of tripton and CDOM. The changes of remote-sensing reflectance in this imaginary case II water should only depend on the changes of the chlorophyll $a$ concentration. In this case, the chlorophyll $a$ concentration should also be accurately predicted by a semi-analytical index calculated from the remote-sensing reflectance. Fig. 5A shows the relationships between chlorophyll $a$ concentration and the three-band index, with a constant tripton concentration and CDOM absorption of low, moderate and high values, respectively. As the figure shows, by using quadratic functions, accurate estimation models of [Chl-a] can be constructed by regression analysis (correlation coefficients equal to 1 and root mean
square error (RMSE) around 0). For other combinations of [TR] and [CDOM] not shown in Fig. 5A, the regressions were all statistically significant with $R^2 > 0.99$ and $P$-value $<< 0.0001$. This phenomenon provides a good opportunity to avoid the estimation errors of chlorophyll $a$ concentration due to the assumptions in the index development process. Similar results for the estimation of tripton concentration and CDOM absorption coefficient can also be obtained, as shown in Fig. 5B and 5C, respectively.

It is noted that the index of $R_{rs}(b_7)/R_{rs}(b_5)$ is not sensitive to variation of [CDOM] when [Chl-a] and [TR] are relatively high (e.g., cases II and III in Fig. 5C with [Chl-a] = 100 mg m$^{-3}$, [TR] = 90 g m$^{-3}$ and [Chl-a] = 300 mg m$^{-3}$, [TR] = 250 g m$^{-3}$, respectively). In these cases, the optical properties of the water are dominated by particles (i.e. tripton and phytoplankton). Therefore, variation of [CDOM] will not largely change the reflectance and thus is not sensitive to the semi-analytical index.

3.2 Semi-analytical model optimizing and look-up tables

As shown in Fig. 5, unlike the case using the original semi-analytical algorithms, polynomial functions are needed to represent the relationships between semi-analytical indices and water constituent concentrations. These relationships can be expressed as follows:

\[ [\text{Chl-a}] = pX_{ph}^2 + qX_{ph} + r \] \hspace{1cm} (16)

\[ [\text{TR}] = aX_u^3 + bX_u^2 + cX_u + d \] \hspace{1cm} (17)
where $X_{ph}$, $X_{tr}$ and $X_{CDOM}$ denote the indices $R_n(b_{10})/R_n(b_7) - R_n(b_{10})/R_n(b_9)$, $R_n(b_{10})$ and $R_n(b_7)/R_n(b_9)$ for chlorophyll $a$, tripton and CDOM, respectively. The regression coefficients $p$, $q$ and $r$ are determined according to [TR] and [CDOM]; $a$, $b$, $c$ and $d$ are determined according to [Chl-a] and [CDOM]; and $m$, $n$ and $h$ are determined according to [Chl-a] and [TR], respectively. These regression coefficients can be determined for different concentration combinations of any two constituents with fine concentration intervals to represent any case of waters, through simulated reflectance spectra based on the bio-optical model. For example, in the case of estimating [Chl-a], increments of 1 g m$^{-3}$ and 0.1 m$^{-1}$ were used in the ranges of 1-250 g m$^{-3}$ and 0.1-10 m$^{-1}$ for [TR] and [CDOM], respectively; while increments for [Chl-a] were the same as those shown in Fig. 5A. Thus, 250×100×31 sample spectra in total were generated to establish the relationships between the three-band index and [Chl-a]. Correspondingly, increments of 1 mg m$^{-3}$ and 0.1 m$^{-1}$ for [Chl-a] and [CDOM] were used for the estimation models of [TR]; increments of 1 mg m$^{-3}$ and 1 g m$^{-3}$ for [Chl-a] and [TR] were used for the estimation models of [CDOM]. In the end, three 2-dimensional look-up tables (LUTs) were constructed, containing the coefficients of the estimation model for one constituent of interest determined by the concentrations of other two constituents. Table 2 shows an example of the LUT for estimation models of chlorophyll $a$ concentration.

A practical problem is that any information on water constituent concentrations is unknown for
a given pixel of satellite data, except for the remote-sensing reflectance. Thus, it is impossible to
determine which estimation model is the most appropriate for an index calculated for this pixel. To
solve this problem, a method based on semi-analytical model optimizing and look-up tables
(SAMO-LUT) is proposed. Fig. 6 shows the major steps of the SAMO-LUT method as follows:

Step 1 is to calculate the selected semi-analytical indices for the corresponding concentrations
of chlorophyll \( a \), tripton and CDOM.

Step 2 is to obtain initial estimations of chlorophyll \( a \) concentration ([Chl-a]\(_0\) in Fig. 6), tripton
concentration ([TR]\(_0\) in Fig. 6) and the CDOM absorption coefficient at 440 nm ([CDOM]\(_0\) in Fig. 6)

Step 3 is to find an estimation model (i.e., regression coefficients of Equations 16-18 saved in
the LUT) from the LUTs for each water constituent using the known initial values of [Chl-a]\(_0\) and
[TR]\(_0\) obtained in step 2, then use the chosen estimation models to replace the previous ones
(Equations 13-15). The new concentrations of chlorophyll \( a \) and tripton ([Chl-a]\(_1\) and [TR]\(_1\) in Fig.
6) and the CDOM absorption coefficient at 440 nm ([CDOM]\(_1\) in Fig. 6) are then recalculated,
respectively.

Step 4 is to find a more appropriate estimation model from the LUTs for each water constituent
through the iterative use of the newly obtained concentrations of chlorophyll \( a \) and tripton and the
CDOM absorption coefficient at 440 nm instead of the previous values. For example, [Chl-a]\(_2\), [TR]\(_2\),
and [CDOM]\(_2\) were used instead of [Chl-a]\(_0\), [TR]\(_0\) and [CDOM]\(_0\) in the second iteration,
respectively (Fig. 6). The new concentrations of chlorophyll \( a \) and tripton ([Chl-a]\(_2\) and [TR]\(_2\) in Fig.
6) and the CDOM absorption coefficient at 440 nm ([CDOM]\(_2\) in Fig. 6) were then calculated again.
The iteration will be ended when the difference between the current and last outputs is adequately small. In this study, the RMSE of the estimated water constituent concentrations in the \( n \)-th and \((n-1)\)-th iteration was used as the criterion to determine the appropriate number of iterations. The RMSEs become stable after the 10th iteration. Therefore, the estimated water constituent concentrations with 10 iterations were used as the final results in the validation process.

4. Validation results and discussion

4.1 Validation with simulation data

A validation dataset was generated separately from the dataset used for the calibrations of the estimation models using a different sampling strategy. The concentrations of chlorophyll \( a \) and tripton as well as the absorption coefficient of CDOM at 440 nm were varied with a random distribution in the ranges of 0-300 (mg m\(^{-3}\)), 0-250 (g m\(^{-3}\)) and 0-10 (m\(^{-1}\)), respectively. The 1000 simulated reflectance spectra were resampled to the bandwidths of MERIS channels.

Three indices, namely the root mean square error (RMSE), mean normalized bias (MNB) and normalized root mean square error (NRMS), were used in accuracy assessment, as suggested by Gitelson et al. (2008). These indices are defined as follows:

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{N}(X_{\text{esti},i} - X_{\text{meas},i})^2}{N}}
\]  

(19)
MNBI = mean(εi)% and
NRMS = stdev(εi)%,

where \(X_{est,i}\) and \(X_{meas,i}\) are the estimated and measured values, respectively; \(N\) is the number of samples; \(ε_i = \frac{100 \times (X_{est,i} - X_{meas,i})}{X_{meas,i}}\) is the percent difference between the estimated and measured values. The MNBI denotes the average bias in the estimation, while the NRMS denotes the relative random uncertainty of the results. The coefficient of determination (R^2) between \(X_{est,i}\) and \(X_{meas,i}\) is also calculated.

Scatter plots of estimated and true water constituent concentrations are shown in Fig. 7, and the assessment results are summarized in Table 3. The results show that the SAMO-LUT method works very well for the simulated spectra with MNBI < 1.74%, NRMS < 8.12% and small RMSE values close to 0 mg m\(^{-3}\) (or g m\(^{-3}\) or m\(^{-1}\)). In addition, the coefficients of determination (R^2) and slopes and intercepts of the regression analysis between the true and estimated values were around 1.0, 1.0 and 0, respectively (Fig. 7). These results indicate that the SAMO-LUT method is a reasonable approach.

To further demonstrate the performance of the proposed method, the three previously presented semi-analytical algorithms for estimating [Chl-a], [TR] and [CDOM] (Eqs. 13-15) were also assessed using the same validation dataset. The results are summarized in Table 3. Estimation results show noticeable errors for these methods with average bias (MNB) and relative random uncertainty (NRMS) in the range of 15.26%–139.62% and 328.84%–1160.92%, respectively.
These results indicate that: (1) the estimation error due to the assumptions in the development process of a semi-analytical index cannot be ignored in many case II waters; and (2) the calibration process strongly depends on the dataset used. For example, if several randomly generated datasets were used to calibrate models for [Chl-a] estimation, different [Chl-a] estimation models will usually be obtained. This is because the assumptions will introduce different effects for different water samples, then resulting in different indices values even for some water samples with the same [Chl-a] values. This is also why some semi-analytical algorithms were site-specific ones.

In contrast, the results obtained from the SAMO-LUT method indicate that the above problems can be solved by using the semi-analytical indices for each special case (i.e., a set of samples with only one water constituent concentration changed; see Fig. 5) and iteratively searching the most appropriate estimation model (optimization) from a prepared LUT for a given case. Since the effects caused by the assumptions in a semi-analytical index are constants for each special case, the estimation model for each special case can be accurately calibrated. These estimation models can be easily constructed using the simulation data rather than using the data collected from the field investigations, because it is difficult or impossible to collect enough data from actual waters to cover all special cases, whereas the simulation data can be generated under any desired environmental conditions. The optimization process allows us to gradually refine the estimation model for each water constituent concentration and finally to obtain the most accurate results.
4.2 Validation with in situ data

In situ data collected from Lake Dianchi in Jul., 2009 (19 samples) was also used to validate the SAMO-LUT method. The results are shown in Fig. 8 and Table 3. It can be seen that the SAMO-LUT method performed well even for these in situ collected data. The RMSE, MNB and NRMS were 3.37 mg m\(^{-3}\), -1.58% and 3.65% for the [Chl-a] estimation; 1.81 g m\(^{-3}\), 3.83% and 6.73% for the [TR] estimation, and 0.21 m\(^{-1}\), -7.07% and 31.19% for the [CDOM] estimation, respectively. The determination coefficients were 0.98 for the [Chl-a] estimation, 0.89 for the [TR] estimation and 0.78 for the [CDOM] estimation.

The results obtained from the in situ data showed less accuracy for each water constituent concentration compared with those obtained from the simulation data. This is probably because of biases in the optical closure processes. Optical closure refers to the testing of the theoretical interrelationship between measured inherent optical properties (IOPs) and apparent optical properties (AOPs) of a water body (Gallegos et al., 2008). Since calibration of the semi-analytical models is based on simulation data, optical closure plays a crucial role in the successful application of the SAMO-LUT method. In this study, SIOPs were measured or estimated from four sampling sites in Lake Dianchi and then were assumed as constants for the whole lake. Results shown in Figs. 3 and 8 indicate this assumption is reasonable in the study areas. In addition, compared with the simulation data, measurement errors included in the in situ data can also result in relatively larger estimation...
errors. The semi-analytical algorithms were also applied to the *in situ* data. Since the relationships between semi-analytical indices and water constituent concentrations in the semi-analytical algorithms strongly depend on the calibration dataset used (e.g., dynamic range and number of data), the semi-analytical algorithms established from the simulation dataset (i.e. Eqs. 13-15) were applied to the *in situ* data for fair comparisons rather than using the *in situ* data for model calibration. The results show that the estimation accuracy for [Chl-a] was slightly decreased (the RMSE, MNB and NRMS were 3.60 mg m\(^{-3}\), -1.77% and 4.06%, respectively). However, the estimation accuracies for [TR] and [CDOM] were dramatically reduced compared with those of the SAMO-LUT (Table 3, the RMSE, MNB and NRMS were 8.50 mg m\(^{-3}\), -4.49% and 35.87% for [TR] and 1.42 m\(^{-1}\), 388.64% and 375.26% for [CDOM], respectively). The poor performances of the semi-analytical algorithms were mainly caused by the invalid assumptions in the semi-analytical indices.

### 4.3 Validation with MERIS data

Two MERIS images for Lake Dianchi and other two MERIS images for Lake Kasumigaura (as shown in Fig.1) were used to further validate the SAMO-LUT method. Pixels contaminated by clouds were excluded; and only sampling sites located more than one pixel away from the bank were remained. Accordingly, there are 21 and 18 sites remained for Lake Kasumigaura in Feb. 2006 and Aug. 2008, respectively (25 sampling sites in total for both field works). For Lake Dianchi, there are 3 and 5 sites available in Oct. 2007 and Mar. 2009, respectively. To enable comparison with *in situ* reflectance, the average reflectance of the pixel nearest to the sampling location along with the 8
surrounding pixels was computed. These 9 pixels represent a surface area of approximately 0.8 km².

A comparison of *in situ* and atmospherically corrected satellite reflectance for Lake Dianchi is shown in Fig. 9. The results of atmospheric correction are acceptable with high agreement between the *in situ* and MERIS-derived remote-sensing reflectance spectra. For Lake Kashumigaura, similar results were obtained (data not shown).

Fig. 10 and Table 3 show the performance of the SAMO-LUT method for the MERIS data. It should be noted that the CDOM measurements were unavailable for Lake Kasumigaura on Feb. 18, 2006. The MERIS-derived constituent concentrations were in good agreement with *in situ* measured values with all the points close to the 1:1 line (Fig. 10). The RMSE, MNB and NRMS were 12.64 mg m⁻³, 7.58% and 16.81% for the [Chl-a] estimation, 4.44 g m⁻³, -2.85% and 23.32% for the [TR] estimation, and 0.32 m⁻¹, 8.74% and 33.62% for the [CDOM] estimation, respectively (Table 3). The determination coefficients were 0.76 for the [Chl-a] estimation, 0.89 for the [TR] estimation and 0.26 for the [CDOM] estimation. The low determination coefficient for the [CDOM] estimation is probably due to errors in the atmospheric corrections for MERIS bands 5 and 7 and its relatively narrow distribution range (0.2-1.5 m⁻¹). The semi-analytical algorithms (Eqs. 13-15) were also used for estimating water constituent concentrations from MERIS data. The SAMO-LUT method noticeably outperformed these algorithms (Table 3). The [CDOM] showed the lowest accuracy with MNB and NRMS larger than 110%. The estimations of [TR] and...
[Chl-a] from semi-analytical algorithms yielded noticeably larger RMSE, MNB and NRMS values than those from the SAMO-LUT method.

It is noted that the estimation accuracy yielded by the SAMO-LUT method for satellite data is lower than those for both the in situ and simulated datasets. The largest error source in the application of satellite image is the atmospheric correction procedure. Although the results shown in Fig. 9 suggest that SCAPE-M is a feasible atmospheric correction algorithm, improving the atmospheric corrections for case II waters is still a challenge due to the complex aerosol composition, sun glint, adjacency effect and others. Therefore, the three indices used in the SAMO-LUT (Eqs. 10-12) were still contaminated by atmosphere due to the imperfect atmospheric corrections, and then limited estimation accuracy of the SAMO-LUT. If the atmospheric factor could be more effectively accounted for, more accurate and reliable monitoring of water quality could be derived from satellite images.

The presented results indicate that the same LUT can be used to estimate water constituent concentrations in two different lakes and periods. This is probably because the SIOPs were similar in the two lakes during the study periods. The SAMO-LUT has the potential to be implemented in other applications, such as case II waters with different SIOPs or use of different satellite sensors. In this case, the look-up tables need to be reconstructed accordingly. This process will be investigated in a future work.
5. Conclusions

In this study, a semi-analytical model optimizing and look-up table (SAMO-LUT) method was proposed to retrieve water constituent concentrations in case II waters. The SAMO-LUT method is based on three previous semi-analytical indices for estimating chlorophyll a, tripton and CDOM. The semi-analytical algorithms were optimized based on the fact that estimation of one constituent of interest can be highly improved when the other two constituents’ concentrations are known in advance. Look-up tables and an iterative searching strategy were applied to obtain the most appropriate parameters in the estimation models. Three datasets (i.e., noise-free simulation data, in situ collected data and MERIS data) were used to validate the performance of the proposed method. The results show that the SAMO-LUT method yields error-free results for the ideal simulation dataset and can also estimate the water constituent concentrations with an average bias (MNB) lower than 9% and a relative random uncertainty (NRMS) lower than 34% even for in situ and MERIS data. In the application of satellite data, the performance of the SAMO-LUT still depends on the accuracy of atmospheric correction.

Acknowledgement: This research was supported in part by National Natural Science Foundation of China (Grant No. 40871162), by the Grants-in-Aid for Scientific Research of MEXT from Japan (No. 20510003, No. 19404012) and also by the Global Environment Research Fund (M-10) of the Ministry of the Environment, Japan. The authors would like to acknowledge Dr. L. Guanter, who
developed the SCAPE-M processor and performed atmospheric correction on the MEIRS images used in this study. The authors would like to thank the two anonymous reviewers for the useful comments on the manuscript.

References:


the quantitative filter technique (QFT). *Proc. SPIE, 1302*, 137


### List of Tables

#### Table 1. Spectral characteristics of the Medium Resolution Imaging Spectrometer (MERIS)

<table>
<thead>
<tr>
<th>Band</th>
<th>Band Centre (nm)</th>
<th>Bandwidth (nm)</th>
<th>Spatial Resolution (m)</th>
<th>Swath Width</th>
<th>Quantization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>412.5</td>
<td>10</td>
<td>300</td>
<td>1150 km</td>
<td>12 bits</td>
</tr>
<tr>
<td>2</td>
<td>442.5</td>
<td>10</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>490</td>
<td>10</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>510</td>
<td>10</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>560</td>
<td>10</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>620</td>
<td>10</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>665</td>
<td>10</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>681.25</td>
<td>7.5</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>708.75</td>
<td>10</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>753.75</td>
<td>7.5</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>761.75</td>
<td>3.75</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>778.75</td>
<td>15</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>865</td>
<td>20</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>885</td>
<td>10</td>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>900</td>
<td>10</td>
<td>300</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Table 2. Demonstration of a look-up table (LUT) for the [Chl-a]-estimation models, where [TR] and [CDOM] have units of g m⁻³ and m⁻¹, respectively

<table>
<thead>
<tr>
<th>[CDOM]</th>
<th>[TR]</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>249</th>
<th>250</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>(p₁,₁; q₁,₁; r₁,₁)</td>
<td>(p₁,₂; q₁,₂; r₁,₂)</td>
<td>...</td>
<td>(p₁,249; q₁,249; r₁,249)</td>
<td>(p₁,250; q₁,250; r₁,250)</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>(p₂,₁; q₂,₁; r₂,₁)</td>
<td>(p₂,₂; q₂,₂; r₂,₂)</td>
<td>...</td>
<td>(p₂,249; q₂,249; r₂,249)</td>
<td>(p₂,250; q₂,250; r₂,250)</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>(p₃,₁; q₃,₁; r₃,₁)</td>
<td>(p₃,₂; q₃,₂; r₃,₂)</td>
<td>...</td>
<td>(p₃,249; q₃,249; r₃,249)</td>
<td>(p₃,250; q₃,250; r₃,250)</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>9.9</td>
<td>(p₉,₁; q₉,₁; r₉,₁)</td>
<td>(p₉,₂; q₉,₂; r₉,₂)</td>
<td>...</td>
<td>(p₉,249; q₉,249; r₉,249)</td>
<td>(p₉,250; q₉,250; r₉,250)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>(p₁₀,₁; q₁₀,₁; r₁₀,₁)</td>
<td>(p₁₀,₂; q₁₀,₂; r₁₀,₂)</td>
<td>...</td>
<td>(p₁₀,249; q₁₀,249; r₁₀,249)</td>
<td>(p₁₀,250; q₁₀,250; r₁₀,250)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Accuracy assessment of the SAMO-LUT and previous semi-analytical algorithms using noise-free simulation data, *in situ* collected data and MERIS data

<table>
<thead>
<tr>
<th></th>
<th>Simulation data</th>
<th>In situ data</th>
<th>MERIS data</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Chl-a]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSE (mg m⁻³)</td>
<td>0.43</td>
<td>45.83</td>
<td>3.37</td>
</tr>
<tr>
<td>MNB (%)</td>
<td>-0.06</td>
<td>15.26</td>
<td>-1.58</td>
</tr>
<tr>
<td>NRMS (%)</td>
<td>2.41</td>
<td>328.84</td>
<td>3.65</td>
</tr>
<tr>
<td>R²</td>
<td>1.00</td>
<td>0.83</td>
<td>0.98</td>
</tr>
<tr>
<td>[TR]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSE (g m⁻³)</td>
<td>0.42</td>
<td>16.07</td>
<td>1.81</td>
</tr>
<tr>
<td>MNB (%)</td>
<td>1.74</td>
<td>48.64</td>
<td>3.83</td>
</tr>
<tr>
<td>NRMS (%)</td>
<td>2.26</td>
<td>359.69</td>
<td>6.73</td>
</tr>
<tr>
<td>R²</td>
<td>1.00</td>
<td>0.96</td>
<td>0.89</td>
</tr>
<tr>
<td>[CDOM]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSE (m⁻¹)</td>
<td>0.06</td>
<td>3.36</td>
<td>0.21</td>
</tr>
<tr>
<td>MNB (%)</td>
<td>-0.62</td>
<td>139.62</td>
<td>-7.07</td>
</tr>
<tr>
<td>NRMS (%)</td>
<td>8.12</td>
<td>1160.92</td>
<td>31.19</td>
</tr>
<tr>
<td>R²</td>
<td>0.99</td>
<td>0.04</td>
<td>0.78</td>
</tr>
</tbody>
</table>

List of Figures

Fig. 1. Distribution of sampling sites in (A) Lake Dianchi, China and (B) Lake Kasumigaura, Japan; and four corresponding MERIS images. (The red stars, green squares and blue circles in (A) denote sampling sites for Lake Dianchi in Oct. 2007, Mar. 2009 and Jul. 2009, respectively; the red circles and green circles in (B) denote sampling sites in Lake Kasumigaura for Feb. 2006 and Aug. 2008, respectively.)

Fig. 2. Specific inherent optical properties (SIOPs) of Lake Dianchi. (A): $a_{ph}^*$ [m² mg⁻¹], $a_w^*$ [m² g⁻¹], and $a_{CDOM}^*$ [dimensionless] are specific absorption coefficients of phytoplankton, tripton, and colored dissolved organic matter, respectively; $a_w$ [m⁻¹] is the absorption coefficient of pure water. (B): $b_{ph}^*$ [m² mg⁻¹] and $b_{w}^*$ [m² g⁻¹] are the specific backscattering coefficients of phytoplankton and tripton, respectively; and $b_{w}$ [m⁻¹] is the backscattering coefficient of pure water.

Fig. 3. Comparison of modeled remote-sensing reflectance spectra just above the water surface and measured spectra.

Fig. 4. Initial estimation models for (A) chlorophyll $a$ concentration; (B) tripton concentration; and (C) CDOM absorption at 440 nm based on previous semi-analytical indices.
Fig. 5. Examples of estimation models for (A) chlorophyll \( a \) concentration when tripton concentration and CDOM absorption at 440 nm are constants; (B) tripton concentration when chlorophyll \( a \) concentration and CDOM absorption at 440 nm are constants; (C) CDOM absorption at 440 nm when chlorophyll \( a \) concentration and tripton concentration are constants.

Fig. 6. Flowchart of the SAMO-LUT method.

Fig. 7. Comparison of true and estimated (A) [Chl-a]; (B) [TR]; and (C) [CDOM] for a noise-free simulation dataset.

Fig. 8. Comparison of measured and estimated (A) [Chl-a]; (B) [TR]; and (C) [CDOM] for an \textit{in situ} collected dataset in Lake Dianchi.

Fig. 9. Comparison between \textit{in situ} collected and atmospheric corrected remote-sensing reflectance spectra in Lake Dianchi. (A)-(C) are for three sampling sites on Oct. 24, 2007, and (D)-(F) are for three sampling sites with lower latitudes on Mar. 13, 2009 (See Fig. 1A).

Fig. 10. Comparison of measured and estimated (A) [Chl-a]; (B) [TR]; and (C) [CDOM] for MERIS data in Lake Dianchi and Lake Kasumigaura during four periods.
Fig. 1.
Fig. 2.
Fig. 3.

\[ y = 223.86x + 23.95 \]

\[ R^2 = 0.91 \]

Chlorophyll \( a \) concentration (mg m\(^{-3}\))

(A)
Fig. 4.
Fig. 5.
Step 1, 2:

MERIS Data

Chl-a Algorithm

Tripton Algorithm

CDOM Algorithm

\([\text{Chl-a}]_0\)

\([\text{TR}]_0\)

\([\text{CDOM}]_0\)

Step 3:

\(X_{\text{CDOM}}\)

LUT

\([\text{CDOM}]_1\)

LUT

\(X_r\)

\([\text{TR}]_1\)

\([\text{Chl-a}]_1\)

LUT

\([\text{CDOM}]_2\)

LUT

\(X_r\)

\([\text{TR}]_2\)

\([\text{Chl-a}]_2\)

... ...

Fig. 6.
Fig. 7.
Fig. 8.
Fig. 9.
Fig. 10.