

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

著者	Yang Wei, Matsushita Bunkei, Chen Jin, Fukushima Takehiko
journal or publication title	Remote sensing of environment
volume	115
number	5
page range	1247-1259
year	2011-05
権利	(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	<a href="http://hdl.handle.net/2241/113240">http://hdl.handle.net/2241/113240</a>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21

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

Wei Yang<sup>1,2</sup>, Bunkei Matsushita<sup>2\*</sup>, Jin Chen<sup>1</sup>, Takehiko Fukushima<sup>2</sup>

<sup>1</sup> State key laboratory of earth surface processes and resource ecology, Beijing Normal  
University, Beijing, 100875, China

E-mail: yangwei1022@gmail.com, chenjin@ires.cn

<sup>2</sup>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

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

24  
25 Abstract

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

43 inland and coastal waters based on satellite observations.

44

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

46

47

## 48 **1. Introduction**

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

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

64 tripton, colored dissolved organic matter (CDOM) and pure water) in case II waters (Doxaran et al.,  
65 2002; Gin et al., 2002; Goodin et al., 1993).

66 To address the difficulties of monitoring case II waters, researchers have made substantial  
67 efforts to accurately estimate water constituent concentrations (including concentrations of  
68 chlorophyll *a* and tripton as well as absorption of CDOM at 440 nm). These efforts have included  
69 the use of derivative values of reflectance spectra (Goodin et al., 1993), examination of the band  
70 ratio of near infrared reflectance and red reflectance (Han et al., 1994), use of the inherent optical  
71 property (IOP) inversion technique based on a bio-optical model (Brando and Dekker, 2003; Garver  
72 and Siegel, 1997; Santini et al., 2010) or on the use of Hydrolight (Mobley et al., 2005; Van Der  
73 Woerd and Pasterkamp, 2008), and use of the spectral mixture analysis technique to minimize the  
74 interactions of the four OASs (Novo et al., 2006; Oyama et al., 2009; Oyama et al., 2007; Svab et al.,  
75 2005; Tyler et al., 2006).

76 Additionally, several semi-analytical algorithms have been proposed to estimate water  
77 constituent concentrations in case II waters (e.g., Ammenberg et al., 2002; Dall'Olmo et al., 2005;  
78 Doxaran et al., 2002; Gitelson et al., 2008; Morel and Gentili, 2009). These algorithms are generally  
79 composed of two key steps. The first step is to develop an index, which could be the reflectance of a  
80 single band or the arithmetic combination of reflectance from several bands, by analyzing the IOPs  
81 of water constituents. The second step is to empirically establish and calibrate the relationships  
82 between the indices obtained from *in situ* reflectance data or satellite data and water constituent  
83 concentrations. The relationship could be a linear function (Ammenberg et al., 2002; Gitelson et al.,  
84 2008), power function (Kutser et al., 2005), or polynomial function (Dall'Olmo et al., 2003),

85 depending on the regression analysis and the dynamic ranges in the calibration data used. Since these  
86 proposed indices are based on the spectral analysis of IOP for each OAS, they can effectively  
87 minimize the effects of the other OASs on the OAS of interest. Therefore, the algorithms based on  
88 these indices not only promise improved performance for predicting the water constituent  
89 concentration of interest but also are easy to implement using satellite data. However, there still are  
90 two major challenges in the application of the semi-analytical algorithms. First, the models for  
91 estimation of water constituent concentrations depend greatly on the calibration process; thus, a  
92 dataset without a sampling bias is needed to calibrate these models. Second, the proposed indices  
93 were developed based on several specific assumptions, some of which may not be universally  
94 applicable. For example, an important assumption in the three-band index for estimating chlorophyll  
95 *a* concentration is that the absorption and backscattering of suspended solids at the near-infrared  
96 band (750-760 nm) can be neglected compared with the absorption of pure water (Gitelson et al.,  
97 2008). However, this assumption is not applicable in some highly turbid case II waters, such as those  
98 of Lake Taihu and Lake Dianchi in China and Lake Kasumigaura in Japan, and thus resulted in large  
99 errors in chlorophyll *a* concentration estimates in these lakes (Le et al., 2009; Yang et al., 2010).

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

105

## 106 2. Methods

### 107 2.1 Bio-optical model

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

$$110 R_{rs}(\lambda, 0^-) = \frac{f}{Q} \times \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \quad (1)$$

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

$$115 R_{rs}(\lambda) = 0.544 \times \frac{f}{Q} \times \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \quad (2)$$

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

$$119 f = 0.975 - 0.629\mu_0. \quad (3)$$

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

127 The spectral total absorption coefficient,  $a(\lambda)$ , is usually expressed as the sum of the  
128 constituents' absorption coefficients, as follows:

$$129 \quad a(\lambda) = a_w(\lambda) + [\text{Chl} - a]a_{\text{ph}}^*(\lambda) + [\text{TR}]a_{\text{tr}}^*(\lambda) + [\text{CDOM}]a_{\text{CDOM}}^*(\lambda), \quad (4)$$

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

$$135 \quad b_b(\lambda) = b_{b,w}(\lambda) + [\text{Chl} - a]b_{b,\text{ph}}^*(\lambda) + [\text{TR}]b_{b,\text{tr}}^*(\lambda), \quad (5)$$

136 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  
137 backscattering coefficients for phytoplankton and tripton, respectively.

## 138 **2.2 Study areas and *In situ* data**

139 Field investigations were carried out in Lake Dianchi (24°50'N; 102°41'E) and Lake  
140 Kasumigaura (36°00'N; 140°25'E). Lake Dianchi is located in a plateau area of the southwestern  
141 part of China (Fig.1A). It has a surface area of 300 km<sup>2</sup> and is the largest lake in Yunnan Province  
142 and the sixth largest lake in China. The mean depth of the lake is 4.3 m and the maximum depth is  
143 11.3 m. Eutrophication has become more and more serious in the lake in the recent 20 years due to  
144 the large quantities of industrial wastewater and municipal sewage discharged into the lake; algal  
145 blooms occur frequently from April to November each year (Gao et al., 2005). Lake Kasumigaura is  
146 located in the eastern part of Japan's Kanto Plain (Fig.1B). It is the second largest lake of Japan, with  
147 a surface area of 171 km<sup>2</sup> and an average depth of 4 m (maximum depth of 7.3 m). The lake is



148 considered eutrophic, because it has a high load of nutrients, and because of its shallow depth  
149 (Fukushima et al., 1996).

150

151 ----- Please insert Fig. 1 here -----

152

153 Three data collection campaigns were undertaken in 2007 (Oct. 23) and 2009 (Mar. 12 and Jul.  
154 24-31) in Lake Dianchi; and other two campaigns were undertaken in 2006 (Feb. 18) and 2008 (Aug.  
155 07) in Lake Kasumigaura. The spatial distribution of sampling sites is shown in Fig. 1. Sample  
156 collections and reflectance measurements were performed between 10:00 and 14:00 h local time.  
157 Water samples were kept in ice boxes and taken to the laboratory within approximately 0.5 hours  
158 after whole data collections. Chlorophyll *a* was extracted using methanol (100%) at 4°C under dark  
159 conditions for 24 hours. The optical density of the extracted chlorophyll *a* was measured at four  
160 wavelengths (750, 663, 645 and 630 nm) and the concentration was calculated according to  
161 SCOR-UNESCO equations (SCOR-UNESCO, 1966). To obtain the concentration of tripton, the  
162 total suspended solids (TSS) were divided into tripton and phytoplanktonic suspended solids (PSS).  
163 Based on the method of Gons et al. (1992) and organic suspended solids (OSS) data collected from  
164 two lakes, it can be assumed that 1 mg m<sup>-3</sup> chlorophyll *a* concentration is approximately equal to  
165 0.148 g m<sup>-3</sup> TSS in Lake Dianchi and 0.12 g m<sup>-3</sup> TSS in Lake Kasumigaura. Tripton concentrations  
166 were then derived by subtracting PSS from TSS. The absorption of CDOM was measured using a  
167 Shimadzu UV-1700 spectrophotometer with filtered water. *In situ* reflectance spectra were collected  
168 according to Method 1 of Mueller et al. (2000).

169 **2.3 Estimation of SIOPs**

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

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

179 
$$a_{tr}^*(\lambda) = a_{tr}^*(440) \exp[-S_{tr}(\lambda - 440)] \quad (6)$$

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

183 
$$a_{CDOM}^*(\lambda) = a_{CDOM}^*(440) \exp[-S_{CDOM}(\lambda - 440)], \quad (7)$$

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

189 The backscattering coefficients were obtained based on an inversion of the bio-optical model

190 used in this study. According to equations 2 and 5, the backscattering coefficient of particles  
 191 (including phytoplankton and tripton),  $b_{b,p}(\lambda)$ , can be obtained as follows:

$$192 \quad b_{b,p}(\lambda) = \frac{a(\lambda)R_{rs}(\lambda)}{0.544(f/Q) - R_{rs}(\lambda)} - b_{b,w}(\lambda), \quad (8)$$

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

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

$$200 \quad b_{b,tr}^*(\lambda) = b_{b,tr}^*(550) \left( \frac{\lambda}{550} \right)^{-n}, \quad (9)$$

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

205

206 ----- Please insert Fig. 2 here -----

207

208 The measured and modeled  $R_{rs}(\lambda)$  at two additional sites were compared to validate the  
 209 estimated SIOPs of Lake Dianchi (Fig. 3). The two sites had moderate ( $59.29 \text{ mg m}^{-3}$ ) and high

210 (131.79 mg m<sup>-3</sup>) concentrations of chlorophyll *a*. The results show strong agreement between the  
211 measured and modeled  $R_{rs}(\lambda)$ , suggesting that the measured specific absorption coefficients and  
212 retrieved specific backscattering coefficients are determined reasonably well.

213

214 ----- Please insert Fig. 3 here -----

215

## 216 ***2.4 MERIS data collection***

217 Two full-resolution MERIS images (level-1b) covering Lake Dianchi were acquired on Oct. 24,  
218 2007 (orbit 29531, starting time 03:32:09 UTC) and Mar. 13, 2009 (orbit 36774, starting time  
219 03:29:13 UTC), which were each one day after the date of the corresponding field survey (Fig. 1A).

220 Other two full-resolution MERIS images covering Lake Kasumigaura were acquired on Feb. 18,  
221 2006 (orbit 20775, starting time 01:15:29 UTC) and Aug. 7, 2008 (orbit 33652, starting time  
222 00:56:50 UTC), on which the two field campaigns were carried out (Fig. 1B). Images were analyzed  
223 using BEAM 4.0 software (Brockmann Consult, Geesthacht, Germany). The images were  
224 geo-located and masked for land, clouds and invalid reflectance. Atmospheric correction was  
225 performed using the SCAPE-M (Self-Contained Atmospheric Parameters Estimation for MERIS  
226 data) atmospheric processor, which outperformed other previous atmospheric correction algorithms  
227 for turbid inland lakes (Guanter et al., 2010).

228

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

### 230 ***3.1 Performances of the original semi-analytical algorithms***

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

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

$$\begin{aligned}
 & [R_{\text{rs}}^{-1}(\lambda_1) - R_{\text{rs}}^{-1}(\lambda_2)] \times R_{\text{rs}}(\lambda_3) \\
 244 \quad & \propto [a_{\text{Chla}}(\lambda_1) + a_w(\lambda_1) - a_w(\lambda_2)] / a_w(\lambda_3) \quad (10) \\
 & \propto [\text{Chl} - a]
 \end{aligned}$$

245 The remote-sensing reflectance in the near-infrared (NIR) band was selected to estimate the  
 246 concentration of tripton (Ammenberg et al., 2002). Generally, backscattering of tripton is noticeably  
 247 larger than that of phytoplankton; that is,  $b_{\text{b,tr}}(\lambda_{\text{NIR}}) \gg b_{\text{b,ph}}(\lambda_{\text{NIR}})$ . In the NIR region (700-800 nm), the  
 248 total absorption was approximately equal to the absorption by pure water; that is,  $a_{\text{total}}(\lambda_{\text{NIR}})$   
 249  $\approx a_w(\lambda_{\text{NIR}})$ . The absorption of pure water is also much larger than the backscattering of particles

250  $(a_w(\lambda_{\text{NIR}}) \gg b_{b,p}(\lambda_{\text{NIR}}))$  (e.g., Babin and Stramski, 2002; Gitelson et al., 2008; Gons, 1999). Thus, the  
 251 remote-sensing reflectance in the NIR band relating [TR] can be expressed as follows:

$$\begin{aligned}
 R_{\text{rs}}(\lambda_{\text{NIR}}) &= \frac{f}{Q} \times \frac{b_{b,w}(\lambda_{\text{NIR}}) + b_{b,\text{tr}}(\lambda_{\text{NIR}}) + b_{b,\text{ph}}(\lambda_{\text{NIR}})}{a_{\text{total}}(\lambda_{\text{NIR}}) + b_{b,\text{tr}}(\lambda_{\text{NIR}}) + b_{b,\text{ph}}(\lambda_{\text{NIR}}) + b_{b,w}(\lambda_{\text{NIR}})} \\
 252 \quad &\approx \frac{f}{Q} \times \frac{b_{b,w}(\lambda_{\text{NIR}}) + b_{b,\text{tr}}(\lambda_{\text{NIR}})}{a_w(\lambda_{\text{NIR}})} \quad (11) \\
 &\propto [\text{TR}]
 \end{aligned}$$

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

$$255 \quad R_{\text{rs}}(664) / R_{\text{rs}}(550) \propto [\text{CDOM}] \quad (12)$$

256  $R_{\text{rs}}(550)$  is influenced by the absorption of CDOM, as well as scattering by all particulate matters, but  
 257 it is not strongly influenced by the absorption of chlorophyll, while  $R_{\text{rs}}(664)$  is affected by both the  
 258 absorption of chlorophyll and the backscattering by phytoplankton and tripton. Using  $R_{\text{rs}}(664)$  in the  
 259 numerator of the band ratio therefore normalizes for variations in both phytoplankton absorption and  
 260 for the effects of variations in backscattering (Ammenberg et al., 2002).

261 To thoroughly investigate the performances of these semi-analytical algorithms, a reflectance  
 262 spectra dataset was generated using the SIOPs and the bio-optical model described above. The  
 263 average value of  $f/Q$  (0.156) obtained from the *in situ* data in Lake Dianchi was used in the spectra  
 264 generation. The concentrations of chlorophyll *a* ([Chl-*a*]) and tripton ([TR]), as well as the  
 265 absorption of CDOM at 440 nm ([CDOM]) were varied in the ranges of 1-300 ( $\text{mg m}^{-3}$ ), 1-250 ( $\text{g}$   
 266  $\text{m}^{-3}$ ) and 0.1-10 ( $\text{m}^{-1}$ ), respectively. Similar with increments used in Kutser et al. (2006), the  
 267 increment of constituent concentrations varied for different ranges. For [Chl-*a*], an increment of 1  
 268  $\text{mg m}^{-3}$  was used in the range of 1-10  $\text{mg m}^{-3}$ , an increment of 2  $\text{mg m}^{-3}$  was used in the range of

269 10-20 mg m<sup>-3</sup>, an increment of 10 mg m<sup>-3</sup> was used in the range of 20-60 mg m<sup>-3</sup> and an increment of  
 270 20 mg m<sup>-3</sup> was used in the range of 60-300 mg m<sup>-3</sup>. For [TR], the increments of 1 g m<sup>-3</sup>, 5 g m<sup>-3</sup> and  
 271 20 g m<sup>-3</sup> were used in the ranges of 1-10 g m<sup>-3</sup>, 10-50 g m<sup>-3</sup> and 50-250 g m<sup>-3</sup>, respectively. For  
 272 [CDOM], the increments of 0.1 m<sup>-1</sup>, 0.5 m<sup>-1</sup> and 1 m<sup>-1</sup> were used in the ranges of 0.1-1.0 m<sup>-1</sup>, 1.0-5.0  
 273 m<sup>-1</sup> and 5.0-10.0 m<sup>-1</sup>, respectively. Thus, 19,964 (i.e., 31×28×23) sample spectra in total were  
 274 generated to establish the relationships between the indices and water constituent concentrations.  
 275 Since the MERIS sensor has great potential in the remote sensing of case II waters, the generated  
 276 spectra were resampled to MERIS bandwidths (shown in Table 1) through corresponding spectral  
 277 response functions to calculate the semi-analytical indices.

278

279 ----- Please insert Table 1 here -----

280

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

$$283 \quad [\text{Chl} - a] = 223.86 \left[ \frac{R_{rs}(b_{10})}{R_{rs}(b_7)} - \frac{R_{rs}(b_{10})}{R_{rs}(b_9)} \right] + 23.95 \quad (13)$$

$$284 \quad [\text{TR}] = 49909 R_{rs}^2(b_{10}) - 61.38 R_{rs}(b_{10}) + 4.74 \quad (14)$$

$$285 \quad [\text{CDOM}] = 3.03 \left[ \frac{R_{rs}(b_7)}{R_{rs}(b_5)} \right] + 0.35, \quad (15)$$

286 where b<sub>5</sub>, b<sub>7</sub>, b<sub>9</sub> and b<sub>10</sub> denote the bandwidths of MERIS bands 5 (555-565 nm), 7 (660-670 nm), 9  
 287 (703.75-713.75 nm) and 10 (750-757.5 nm), respectively. It is clearly seen that the assumptions in  
 288 the indices development process resulted in larger errors in the estimations of the water constituent  
 289 concentrations, even though noise-free simulation data was used. Especially for [CDOM] estimation,

290 a poor  $R^2$  was given (Fig. 4C). This is because the  $R_{rs}(550)$  will be influenced by chlorophyll  
291 absorption for samples with high [Chl-a].

292

293 ----- Please insert Fig.4 here -----

294

295 To improve the performances of the semi-analytical algorithms, the bio-optical model was  
296 theoretically reanalyzed. From equations (1), (4) and (5), it can be seen that the remote-sensing  
297 reflectance in case II waters depends on concentrations of three constituents: phytoplankton, tripton  
298 and CDOM (absorption and backscattering coefficients of pure water are usually treated as  
299 constants). As in a microcosm experiment (i.e., making an artificial ecosystem in which some  
300 conditions are controlled to simulate behaviors of a simplified natural ecosystem; e.g., Hunter et al.,  
301 2008), we can consider an imaginary case II water, in which only one constituent changes while the  
302 other two constituents are controlled as constants. For example, the imaginary case II water has  
303 variable concentrations of chlorophyll *a* but constant concentrations of tripton and CDOM. The  
304 changes of remote-sensing reflectance in this imaginary case II water should only depend on the  
305 changes of the chlorophyll *a* concentration. In this case, the chlorophyll *a* concentration should also  
306 be accurately predicted by a semi-analytical index calculated from the remote-sensing reflectance.  
307 Fig. 5A shows the relationships between chlorophyll *a* concentration and the three-band index, with  
308 a constant tripton concentration and CDOM absorption of low, moderate and high values,  
309 respectively. As the figure shows, by using quadratic functions, accurate estimation models of  
310 [Chl-a] can be constructed by regression analysis (correlation coefficients equal to 1 and root mean



311 square error (RMSE) around 0). For other combinations of [TR] and [CDOM] not shown in Fig.5A,  
 312 the regressions were all statistically significant with  $R^2 > 0.99$  and  $P\text{-value} \ll 0.0001$ . This  
 313 phenomenon provides a good opportunity to avoid the estimation errors of chlorophyll *a*  
 314 concentration due to the assumptions in the index development process. Similar results for the  
 315 estimation of tripton concentration and CDOM absorption coefficient can also be obtained, as shown  
 316 in Fig. 5B and 5C, respectively.

317 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*]  
 318 and [TR] are relatively high (e.g., cases II and III in Fig. 5C with [Chl-*a*] = 100 mg m<sup>-3</sup>, [TR] = 90 g  
 319 m<sup>-3</sup> and [Chl-*a*] = 300 mg m<sup>-3</sup>, [TR] = 250 g m<sup>-3</sup>, respectively). In these cases, the optical properties  
 320 of the water are dominated by particles (i.e. tripton and phytoplankton). Therefore, variation of  
 321 [CDOM] will not largely change the reflectance and thus is not sensitive to the semi-analytical  
 322 index.

323

324 ----- Please insert Fig.5 here -----

325

### 326 **3.2 Semi-analytical model optimizing and look-up tables**

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

$$330 \quad [\text{Chl} - a] = pX_{\text{ph}}^2 + qX_{\text{ph}} + r \quad (16)$$

$$331 \quad [\text{TR}] = aX_{\text{tr}}^3 + bX_{\text{tr}}^2 + cX_{\text{tr}} + d \quad (17)$$

332 
$$[\text{CDOM}] = mX_{\text{CDOM}}^2 + nX_{\text{CDOM}} + h \tag{18}$$

333 where  $X_{\text{ph}}$ ,  $X_{\text{tr}}$  and  $X_{\text{CDOM}}$  denote the indices  $R_{\text{rs}}(b_{10})/R_{\text{rs}}(b_7) - R_{\text{rs}}(b_{10})/R_{\text{rs}}(b_9)$ ,  $R_{\text{rs}}(b_{10})$   
 334 and  $R_{\text{rs}}(b_7)/R_{\text{rs}}(b_5)$  for chlorophyll *a*, tripton and CDOM, respectively. The regression coefficients  $p$ ,  
 335  $q$  and  $r$  are determined according to [TR] and [CDOM];  $a$ ,  $b$ ,  $c$  and  $d$  are determined according to  
 336 [Chl-*a*] and [CDOM]; and  $m$ ,  $n$  and  $h$  are determined according to [Chl-*a*] and [TR], respectively.  
 337 These regression coefficients can be determined for different concentration combinations of any two  
 338 constituents with fine concentration intervals to represent any case of waters, through simulated  
 339 reflectance spectra based on the bio-optical model. For example, in the case of estimating [Chl-*a*],  
 340 increments of  $1 \text{ g m}^{-3}$  and  $0.1 \text{ m}^{-1}$  were used in the ranges of  $1\text{-}250 \text{ g m}^{-3}$  and  $0.1\text{-}10 \text{ m}^{-1}$  for [TR] and  
 341 [CDOM], respectively; while increments for [Chl-*a*] were the same as those shown in Fig. 5A. Thus,  
 342  $250 \times 100 \times 31$  sample spectra in total were generated to establish the relationships between the  
 343 three-band index and [Chl-*a*]. Correspondingly, increments of  $1 \text{ mg m}^{-3}$  and  $0.1 \text{ m}^{-1}$  for [Chl-*a*] and  
 344 [CDOM] were used for the estimation models of [TR]; increments of  $1 \text{ mg m}^{-3}$  and  $1 \text{ g m}^{-3}$  for  
 345 [Chl-*a*] and [TR] were used for the estimation models of [CDOM]. In the end, three 2-dimensional  
 346 look-up tables (LUTs) were constructed, containing the coefficients of the estimation model for one  
 347 constituent of interest determined by the concentrations of other two constituents. Table 2 shows an  
 348 example of the LUT for estimation models of chlorophyll *a* concentration.

349

350 ----- Please insert Table 2 here -----

351

352 A practical problem is that any information on water constituent concentrations is unknown for

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

357 Step 1 is to calculate the selected semi-analytical indices for the corresponding concentrations  
358 of chlorophyll *a*, tripton and CDOM.

359 Step 2 is to obtain initial estimations of chlorophyll *a* concentration ( $[\text{Chl-a}]_0$  in Fig. 6), tripton  
360 concentration ( $[\text{TR}]_0$  in Fig. 6) and the CDOM absorption coefficient at 440 nm ( $[\text{CDOM}]_0$  in Fig. 6)  
361 using equations 13-15.

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

368 Step 4 is to find a more appropriate estimation model from the LUTs for each water constituent  
369 through the iterative use of the newly obtained concentrations of chlorophyll *a* and tripton and the  
370 CDOM absorption coefficient at 440 nm instead of the previous values. For example,  $[\text{Chl-a}]_1$ ,  $[\text{TR}]_1$   
371 and  $[\text{CDOM}]_1$  were used instead of  $[\text{Chl-a}]_0$ ,  $[\text{TR}]_0$  and  $[\text{CDOM}]_0$  in the second iteration,  
372 respectively (Fig. 6). The new concentrations of chlorophyll *a* and tripton ( $[\text{Chl-a}]_2$  and  $[\text{TR}]_2$  in Fig.  
373 6) and the CDOM absorption coefficient at 440 nm ( $[\text{CDOM}]_2$  in Fig. 6) were then calculated again.

374

375

----- Please insert Fig.6 here -----

376

377

The iteration will be ended when the difference between the current and last outputs is

378

adequately small. In this study, the RMSE of the estimated water constituent concentrations in the

379

$n$ -th and  $(n-1)$ -th iteration was used as the criterion to determine the appropriate number of iterations.

380

The RMSEs become stable after the 10th iteration. Therefore, the estimated water constituent

381

concentrations with 10 iterations were used as the final results in the validation process.

382

## 383 **4. Validation results and discussion**

### 384 **4.1 Validation with simulation data**

385

A validation dataset was generated separately from the dataset used for the calibrations of the

386

estimation models using a different sampling strategy. The concentrations of chlorophyll  $a$  and

387

tripton as well as the absorption coefficient of CDOM at 440 nm were varied with a random

388

distribution in the ranges of 0-300 ( $\text{mg m}^{-3}$ ), 0-250 ( $\text{g m}^{-3}$ ) and 0-10 ( $\text{m}^{-1}$ ), respectively. The 1000

389

simulated reflectance spectra were resampled to the bandwidths of MERIS channels.

390

Three indices, namely the root mean square error (RMSE), mean normalized bias (MNB) and

391

normalized root mean square error (NRMS), were used in accuracy assessment, as suggested by

392

Gitelson et al. (2008). These indices are defined as follows:

393

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (X_{\text{esti},i} - X_{\text{meas},i})^2}{N}} \quad (19)$$

394 
$$\text{MNB} = \text{mean}(\varepsilon_i)\% \quad \text{and} \quad (20)$$

395 
$$\text{NRMS} = \text{stdev}(\varepsilon_i)\% , \quad (21)$$

396 where  $X_{\text{esti},i}$  and  $X_{\text{meas},i}$  are the estimated and measured values, respectively;  $N$  is the number of  
397 samples;  $\varepsilon_i = 100 \times (X_{\text{esti},i} - X_{\text{meas},i}) / X_{\text{meas},i}$  is the percent difference between the estimated and measured  
398 values. The MNB denotes the average bias in the estimation, while the NRMS denotes the relative  
399 random uncertainty of the results. The coefficient of determination ( $R^2$ ) between  $X_{\text{esti},i}$  and  $X_{\text{meas},i}$  is  
400 also calculated.

401 Scatter plots of estimated and true water constituent concentrations are shown in Fig. 7, and the  
402 assessment results are summarized in Table 3. The results show that the SAMO-LUT method works  
403 very well for the simulated spectra with  $\text{MNB} < 1.74\%$ ,  $\text{NRMS} < 8.12\%$  and small RMSE values  
404 close to  $0 \text{ mg m}^{-3}$  (or  $\text{g m}^{-3}$  or  $\text{m}^{-1}$ ). In addition, the coefficients of determination ( $R^2$ ) and slopes and  
405 intercepts of the regression analysis between the true and estimated values were around 1.0, 1.0 and  
406 0, respectively (Fig. 7). These results indicate that the SAMO-LUT method is a reasonable approach.

407

408 ----- Please insert Fig. 7 here -----

409

410 To further demonstrate the performance of the proposed method, the three previously presented  
411 semi-analytical algorithms for estimating [Chl-a], [TR] and [CDOM] (Eqs. 13-15) were also  
412 assessed using the same validation dataset. The results are summarized in Table 3. Estimation results  
413 show noticeable errors for these methods with average bias (MNB) and relative random uncertainty  
414 (NRMS) in the range of 15.26%–139.62% and 328.84%–1160.92%, respectively.

415

416

----- Please insert Table 3 here -----

417

418

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.

425

426

427

428

429

430

431

432

433

434

435

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.

436 **4.2 Validation with in situ data**

437 *In situ* data collected from Lake Dianchi in Jul., 2009 (19 samples) was also used to validate  
438 the SAMO-LUT method. The results are shown in Fig. 8 and Table 3. It can be seen that the  
439 SAMO-LUT method performed well even for these *in situ* collected data. The RMSE, MNB and  
440 NRMS were  $3.37 \text{ mg m}^{-3}$ ,  $-1.58\%$  and  $3.65\%$  for the [Chl-a] estimation;  $1.81 \text{ g m}^{-3}$ ,  $3.83\%$  and  
441  $6.73\%$  for the [TR] estimation, and  $0.21 \text{ m}^{-1}$ ,  $-7.07\%$  and  $31.19\%$  for the [CDOM] estimation,  
442 respectively. The determination coefficients were 0.98 for the [Chl-a] estimation, 0.89 for the [TR]  
443 estimation and 0.78 for the [CDOM] estimation.

444

445 ----- Please insert Fig. 8 here -----

446

447 The results obtained from the *in situ* data showed less accuracy for each water constituent  
448 concentration compared with those obtained from the simulation data. This is probably because of  
449 biases in the optical closure processes. Optical closure refers to the testing of the theoretical  
450 interrelationship between measured inherent optical properties (IOPs) and apparent optical properties  
451 (AOPs) of a water body (Gallegos et al., 2008). Since calibration of the semi-analytical models is  
452 based on simulation data, optical closure plays a crucial role in the successful application of the  
453 SAMO-LUT method. In this study, SIOPs were measured or estimated from four sampling sites in  
454 Lake Dianchi and then were assumed as constants for the whole lake. Results shown in Figs. 3 and 8  
455 indicate this assumption is reasonable in the study areas. In addition, compared with the simulation  
456 data, measurement errors included in the *in situ* data can also result in relatively larger estimation

457 errors.

458       The semi-analytical algorithms were also applied to the *in situ* data. Since the relationships  
459 between semi-analytical indices and water constituent concentrations in the semi-analytical  
460 algorithms strongly depend on the calibration dataset used (e.g., dynamic range and number of data),  
461 the semi-analytical algorithms established from the simulation dataset (i.e. Eqs. 13-15) were applied  
462 to the *in situ* data for fair comparisons rather than using the *in situ* data for model calibration. The  
463 results show that the estimation accuracy for [Chl-a] was slightly decreased (the RMSE, MNB and  
464 NRMS were  $3.60 \text{ mg m}^{-3}$ ,  $-1.77\%$  and  $4.06\%$ , respectively). However, the estimation accuracies for  
465 [TR] and [CDOM] were dramatically reduced compared with those of the SAMO-LUT (Table 3, the  
466 RMSE, MNB and NRMS were  $8.50 \text{ mg m}^{-3}$ ,  $-4.49\%$  and  $35.87\%$  for [TR] and  $1.42 \text{ m}^{-1}$ ,  $388.64\%$   
467 and  $375.26\%$  for [CDOM], respectively). The poor performances of the semi-analytical algorithms  
468 were mainly caused by the invalid assumptions in the semi-analytical indices.

469

#### 470 **4.3 Validation with MERIS data**

471       Two MERIS images for Lake Dianchi and other two MERIS images for Lake Kasumigaura (as  
472 shown in Fig.1) were used to further validate the SAMO-LUT method. Pixels contaminated by  
473 clouds were excluded; and only sampling sites located more than one pixel away from the bank were  
474 remained. Accordingly, there are 21 and 18 sites remained for Lake Kasumigaura in Feb. 2006 and  
475 Aug. 2008, respectively (25 sampling sites in total for both field works). For Lake Dianchi, there are  
476 3 and 5 sites available in Oct. 2007 and Mar. 2009, respectively. To enable comparison with *in situ*  
477 reflectance, the average reflectance of the pixel nearest to the sampling location along with the 8



478 surrounding pixels was computed. These 9 pixels represent a surface area of approximately 0.8 km<sup>2</sup>.  
479 A comparison of *in situ* and atmospherically corrected satellite reflectance for Lake Dianchi is  
480 shown in Fig. 9. The results of atmospheric correction are acceptable with high agreement between  
481 the *in situ* and MERIS-derived remote-sensing reflectance spectra. For Lake Kashumigaura, similar  
482 results were obtained (data not shown).

483

484 ----- Please insert Fig. 9 here -----

485

486 Fig. 10 and Table 3 show the performance of the SAMO-LUT method for the MERIS data. It  
487 should be noted that the CDOM measurements were unavailable for Lake Kasumigaura on Feb.  
488 18, 2006. The MERIS-derived constituent concentrations were in good agreement with *in situ*  
489 measured values with all the points close to the 1:1 line (Fig. 10). The RMSE, MNB and NRMS  
490 were 12.64 mg m<sup>-3</sup>, 7.58% and 16.81% for the [Chl-a] estimation, 4.44 g m<sup>-3</sup>, -2.85% and 23.32%  
491 for the [TR] estimation, and 0.32 m<sup>-1</sup>, 8.74% and 33.62% for the [CDOM] estimation, respectively  
492 (Table 3). The determination coefficients were 0.76 for the [Chl-a] estimation, 0.89 for the [TR]  
493 estimation and 0.26 for the [CDOM] estimation. The low determination coefficient for the  
494 [CDOM] estimation is probably due to errors in the atmospheric corrections for MERIS bands 5  
495 and 7 and its relatively narrow distribution range (0.2-1.5 m<sup>-1</sup>). The semi-analytical algorithms  
496 (Eqs. 13-15) were also used for estimating water constituent concentrations from MERIS data.  
497 The SAMO-LUT method noticeably outperformed these algorithms (Table 3). The [CDOM]  
498 showed the lowest accuracy with MNB and NRMS larger than 110%. The estimations of [TR] and

499 [Chl-a] from semi-analytical algorithms yielded noticeably larger RMSE, MNB and NRMS values  
500 than those from the SAMO-LUT method.

501

502 ----- Please insert Fig. 10 here -----

503

504 It is noted that the estimation accuracy yielded by the SAMO-LUT method for satellite data is  
505 lower than those for both the *in situ* and simulated datasets. The largest error source in the  
506 application of satellite image is the atmospheric correction procedure. Although the results shown in  
507 Fig. 9 suggest that SCAPE-M is a feasible atmospheric correction algorithm, improving the  
508 atmospheric corrections for case II waters is still a challenge due to the complex aerosol composition,  
509 sun glint, adjacency effect and others. Therefore, the three indices used in the SAMO-LUT (Eqs.  
510 10-12) were still contaminated by atmosphere due to the imperfect atmospheric corrections, and then  
511 limited estimation accuracy of the SAMO-LUT. If the atmospheric factor could be more effectively  
512 accounted for, more accurate and reliable monitoring of water quality could be derived from satellite  
513 images.

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

520

## 521 **5. Conclusions**

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

535

536

537 **Acknowledgement:** This research was supported in part by National Natural Science Foundation  
538 of China (Grant No. 40871162), by the Grants-in-Aid for Scientific Research of MEXT from Japan  
539 (No. 20510003, No. 19404012) and also by the Global Environment Research Fund (M-10) of the  
540 Ministry of the Environment, Japan. The authors would like to acknowledge Dr. L. Guanter, who

541 developed the SCAPE-M processor and performed atmospheric correction on the MEIRS images  
542 used in this study. The authors would like to thank the two anonymous reviewers for the useful  
543 comments on the manuscript.

544  
545  
546

#### 547 References:

- 548 Ammenberg, P., Flink, P., Lindell, T., Pierson, D., & Strombeck, N. (2002). Bio-optical modelling  
549 combined with remote sensing to assess water quality. *International Journal of Remote Sensing*, 23,  
550 1621-1638
- 551 Austin, R.W. (1980). Gulf of Mexico, ocean-colour surface-truth measurements. *Bound.-Layer Meteorol.*,  
552 18, 269-285
- 553 Ayres, W., Busia, A., Dinar, A., Hirji, R., Lintner, S., McCalla, A., & Robelus, R. (1996). *Integrated lake  
554 and reservoir management: World bank approach and experience*. Washington, DC.: World Bank
- 555 Babin, M., & Stramski, D. (2002). Light absorption by aquatic particles in the near-infrared spectral  
556 region. *Limnology and Oceanography*, 47, 911-915
- 557 Brando, V.E., & Dekker, A.G. (2003). Satellite hyperspectral remote sensing for estimating estuarine and  
558 coastal water quality. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 1378-1387
- 559 Dall'Olmo, G., & Gitelson, A.A. (2005). Effect of bio-optical parameter variability on the remote  
560 estimation of chlorophyll-a concentration in turbid productive waters: experimental results. *Applied  
561 Optics*, 44, 412-422
- 562 Dall'Olmo, G., Gitelson, A.A., & Rundquist, D.C. (2003). Towards a unified approach for remote  
563 estimation of chlorophyll-a in both terrestrial vegetation and turbid productive waters. *Geophysical  
564 Research Letters*, 30, doi:10.1029/2003GL018065
- 565 Doxaran, D., Froidefond, J.M., & Castaing, P. (2002). A reflectance band ratio used to estimate suspended  
566 matter concentrations in sediment-dominant coastal waters. *International Journal of Remote Sensing*,  
567 23, 5079-5085
- 568 Fukushima, T., Park, J., Imai, A., Matsushige, K. (1996). Dissolved organic carbon in a eutrophic lakes;  
569 dynamics, biodegradability and origin. *Aquatic Sciences*, 58, 139-157.
- 570 Gallegos, C.L., Davies-Colley, R.J., & Gall, M. (2008). Optical closure in lakes with contrasting extremes  
571 of reflectance. *Limnology and Oceanography*, 53, 2021-2034
- 572 Gao, L., Zhou, J.M., Yang, H., & Chen, J. (2005). Phosphorus fractions in sediment profiles and their  
573 potential contributions to eutrophication in Dianchi Lake. *Environmental Geology*, 48, 835-844
- 574 Garver, S., & Siegel, D. (1997). Inherent optical property inversion of ocean color spectra and its  
575 biogeochemical interpretation 1. Time series from the Sargasso Sea. *Journal of Geophysical  
576 Research-Oceans*, 102, 18607-18625
- 577 Giardino, C., Brando, V.E., Dekker, A.G., Strömbeck, N., & Candiani, G. (2007). Assessment of water  
578 quality in Lake Garda (Italy) using Hyperion. *Remote Sensing of Environment*, 109, 183-195

- 579 Gin, K.Y.H., Koh, S.T., & Lin, I.I. (2002). Study of the effect of suspended marine clay on the reflectance  
580 spectra of phytoplankton. *International Journal of Remote Sensing*, 23, 2163–2178
- 581 Gin, K.Y.H., Koh, S.T., & Lin, I.I. (2003). Spectral irradiance profile of suspended marine clay for the  
582 estimation of suspended sediment concentration in tropical waters. *International Journal of Remote*  
583 *Sensing*, 24, 3235–3245
- 584 Gitelson, A.A., Dall'Olmo, G, Moses, W., Rundquist, D.C., Barrow, T., Fisher, T.R., Gurlin, D., & Holz, J.  
585 (2008). A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters:  
586 Validation. *Remote Sensing of Environment*, 112, 3582-3593
- 587 Gons, H.J. (1999). Optical teledetection of chlorophyll a in turbid inland waters. *Environ. Sci. Technol.*,  
588 33, 1127-1132
- 589 Gons, H.J., Burger-Wiersma, T., Otten, J.H., & Rijkeboer, M. (1992). Coupling of phytoplankton and  
590 detritus in a shallow, eutrophic lake (Lake Loosdrecht, The Netherlands). *Hydrobiologia*, 233, 51-59
- 591 Goodin, D.G., Han, L., Fraser, R.N., Rundquist, D.C., Stebbins, W.A., & Schalles, J.F. (1993). Analysis of  
592 suspended solids in water using remotely sensed high resolution derivative spectra.  
593 *Photogrammetric Engineering and Remote Sensing*, 59, 505–510
- 594 Gordon, H.R., Brown, O.B., Jacobs, M.M. (1975). Computed relationship between the inherent and  
595 apparent optical properties of a flat homogeneous ocean. *Applied Optics*, 14, 417-427
- 596 Guanter, L., Ruiz-Verdu, A., Odermatt, D., FGiardino, C., Simis, S., Estelles, V., Heege, T.,  
597 Dominguez-Gomez, J.A., & Moreno, J. (2010). Atmospheric correction of ENVISAT/MERIS data  
598 over inland waters: Validation for European lakes. *Remote Sensing of Environment*, 114, 467-480
- 599 Hale, G.M., & Querry, M.R. (1973). Optical constants of water in the 200 nm to 200 micrometer  
600 wavelength region. *Applied Optics*, 12, 555–563
- 601 Han, L., Rundquist, D.C., Liu, L.L., Fraser, R.N., & Schalles, J.F. (1994). The spectral responses of algal  
602 chlorophyll in water with varying levels of suspended sediment. *International Journal of Remote*  
603 *Sensing*, 15, 3707–3718
- 604 Hunter, P.D., Tyler, A.N., Presing, M., Kovacs, A.W., & Preston, T. (2008). Spectral discrimination of  
605 phytoplankton colour groups: The effects of suspended particulate matter and sensor spectral  
606 resolution. *Remote Sensing of Environment*, 112, 1527-1544
- 607 Khorram, S., Cheshire, H., Geraci, A., & Rosa, G. (1991). Water quality mapping of Augusta Bay, Italy  
608 from Landsat TM data. *International Journal of Remote Sensing*, 12, 803-808
- 609 Kirk, J.T.O. (1994). *Light and Photosynthesis in Aquatic Ecosystem*. Cambridge, U.K.: Cambridge Univ.  
610 Press
- 611 Kutser, T., Metsamaa, L., Strombeck, N., & Vahtmae, E. (2006). Monitoring cyanobacterial blooms by  
612 satellite remote sensing. *Estuarine coastal and shelf science*, 67, 303-312
- 613 Kutser, T., Pierson, D.C., Kallio, K.Y., Reinart, A., & Sobek, S. (2005). Mapping lake CDOM by satellite  
614 remote sensing. *Remote Sensing of Environment*, 94, 535-540
- 615 Le, C., Li, Y., Zha, Y., Sun, D., Huang, C., & Lu, H. (2009). A four-band semi-analytical model for  
616 estimating chlorophyll a in highly turbid lakes: The case of Taihu Lake, China *Remote Sensing of*  
617 *Environment*, 113, 1175-1182
- 618 Liu, Y., Islam, M.A., & Gao, J. (2003). Quantification of shallow water quality parameters by means of  
619 remote sensing. *Progress in Physical Geography*, 27, 24-43
- 620 Mitchell, B.G. (1990). Algorithms for determining the absorption coefficient of aquatic particulates using

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

622 Mobley, C.D., Sundman, L.K., Davis, C.O., Bowles, J.H., Downes, T.V., Leathers, R.A., Montes, M.J.,  
623 Bissett, W.P., Kohler, D.D.R., Reid, R.P., Louchard, E.M., & Gleason, A. (2005). Interpretation of  
624 hyperspectral remote-sensing imagery by spectrum matching and look-up tables. *Applied Optics*, 44,  
625 3576-3592

626 Morel, A. (1974). Optical properties of pure sea water. *Optical Aspects of Oceanography*. London, U.K.:  
627 Academic

628 Morel, A., & Gentili, B. (1993). Diffuse reflectance of oceanic waters. II. Bidirectional aspects. *Applied*  
629 *Optics*, 32, 6864-6879

630 Morel, A., & Gentili, B. (2009). A simple band ratio technique to quantify the colored dissolved and  
631 detrital organic material from ocean color remotely sensed data. *Remote Sensing of Environment*, 113,  
632 998-1011

633 Morel, A., & Prieur, L. (1977). Analysis of variations in ocean color. *Limnology and Oceanography*, 22,  
634 709-722

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

639 Nichol, J.E. (1993). Remote sensing of water quality in the Singapore–Johor–Riau growth triangle.  
640 *Remote Sensing of Environment*, 43, 139-148

641 Novo, E.M.L.M., Barbosa, C.C.F., Freitas, R.M., Shimabukuro, Y.E., Melack, J.M., & Filho, W.P. (2006).  
642 Seasonal change in chlorophyll distributions in Amazon floodplain lakes derived from MODIS  
643 images. *Limnology*, 7, 153-161

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

647 Oyama, Y., Matsushita, B., Fukushima, T., Nagai, T., & Imai, A. (2007). A new algorithm for estimating  
648 chlorophyll-a concentration from multi-spectral satellite data in Case II waters: A simulation based  
649 on a controlled laboratory experiment. *International Journal of Remote Sensing*, 28, 1437-1453

650 Pope, R.M., & Fry, E.S. (1997). Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity  
651 measurements. *Applied Optics*, 36, 8710

652 Santini, F., Alberotanza, L., Cavalli, R.M., & Pignatti, S. (2010). A two-step optimization procedure for  
653 assessing water constituent concentrations by hyperspectral remote sensing techniques: An  
654 application to the highly turbid Venice lagoon waters. *Remote Sensing of Environment*, 114, 887-898

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

657 Svab, E., Tyler, A.N., Perston, T., Presing, M., & Balogh, K.V. (2005). Characterizing the spectral  
658 reflectance of algae in lake waters with high suspended sediment concentrations. *International*  
659 *Journal of Remote Sensing*, 26, 919-928

660 Tyler, A.N., Svab, E., Preston, T., Présing, M., & Kovács, W.A. (2006). Remote sensing of the water  
661 quality of shallow lakes: A mixture modeling approach to quantifying phytoplankton in water  
662 characterized by high-suspended sediment. *International Journal of Remote Sensing*, 27, 1521-1537

663 Van Der Woerd, H.J., & Pasterkamp, R. (2008). HYDROPT: A fast and flexible method to retrieve  
664 chlorophyll-a from multispectral satellite observations of optically complex coastal waters. *Remote*  
665 *Sensing of Environment*, 112, 1795-1807

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

669 Zhang, Y., Liu, M., Qin, B., Woerd, H.J., Li, J., & Li, Y. (2009). Modeling Remote-Sensing Reflectance  
670 and Retrieving Chlorophyll-a Concentration in Extremely Turbid Case-2 Waters (Lake Taihu, China).  
671 *IEEE Transactions on Geoscience and Remote Sensing*, 47, 1937-1948

672  
673  
674

675 List of Tables

676  
677

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

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

678  
679  
680  
681

682 Table 2 Demonstration of a look-up table (LUT) for the [Chl-a]-estimation models, where [TR] and  
683 [CDOM] have units of  $g\ m^{-3}$  and  $m^{-1}$ , respectively

[CDOM] \ [TR]	1	2	...	...	249	250
0.1	( $p_{1,1}; q_{1,1}; r_{1,1}$ )	( $p_{1,2}; q_{1,2}; r_{1,2}$ )	...	...	( $p_{1,249}; q_{1,249}; r_{1,249}$ )	( $p_{1,250}; q_{1,250}; r_{1,250}$ )
0.2	( $p_{2,1}; q_{2,1}; r_{2,1}$ )	( $p_{2,2}; q_{2,2}; r_{2,2}$ )	...	...	( $p_{2,249}; q_{2,249}; r_{2,249}$ )	( $p_{2,250}; q_{2,250}; r_{2,250}$ )
0.3	( $p_{3,1}; q_{3,1}; r_{3,1}$ )	( $p_{3,2}; q_{3,2}; r_{3,2}$ )	...	...	( $p_{3,249}; q_{3,249}; r_{3,249}$ )	( $p_{3,250}; q_{3,250}; r_{3,250}$ )
...	...	...	...	...	...	...
...	...	...	...	...	...	...
9.9	( $p_{99,1}; q_{99,1}; r_{99,1}$ )	( $p_{99,2}; q_{99,2}; r_{99,2}$ )	...	...	( $p_{99,249}; q_{99,249}; r_{99,249}$ )	( $p_{99,250}; q_{99,250}; r_{99,250}$ )
10	( $p_{100,1}; q_{100,1}; r_{100,1}$ )	( $p_{100,2}; q_{100,2}; r_{100,2}$ )	...	...	( $p_{100,249}; q_{100,249}; r_{100,249}$ )	( $p_{100,250}; q_{100,250}; r_{100,250}$ )

684  
685  
686  
687

688

689

690 Table 3. Accuracy assessment of the SAMO-LUT and previous semi-analytical algorithms using  
 691 noise-free simulation data, *in situ* collected data and MERIS data

	<u>Simulation data</u>		<u>In situ data</u>		<u>MERIS data</u>	
	SAMO-LUT	Semi-analytical Algorithms (Eq. 13-15)	SAMO-LUT	Semi-analytical Algorithms (Eq. 13-15)	SAMO-LUT	Semi-analytical Algorithms (Eq. 13-15)
<b>[Chl-a]</b>						
RMSE (mg m <sup>-3</sup> )	0.43	45.83	3.37	3.60	12.64	28.54
MNB (%)	-0.06	15.26	-1.58	-1.77	7.58	23.82
NRMS (%)	2.41	328.84	3.65	4.06	16.81	34.34
R <sup>2</sup>	1.00	0.83	0.98	0.90	0.76	0.31
<b>[TR]</b>						
RMSE (g m <sup>-3</sup> )	0.42	16.07	1.81	8.50	4.44	11.53
MNB (%)	1.74	48.64	3.83	-4.49	-2.85	45.97
NRMS (%)	2.26	359.69	6.73	35.87	23.32	106.26
R <sup>2</sup>	1.00	0.96	0.89	0.14	0.89	0.32
<b>[CDOM]</b>						
RMSE (m <sup>-1</sup> )	0.06	3.36	0.21	1.42	0.32	1.11
MNB (%)	-0.62	139.62	-7.07	388.64	8.74	115.07
NRMS (%)	8.12	1160.92	31.19	375.26	33.62	130.57
R <sup>2</sup>	0.99	0.04	0.78	0.26	0.26	0.02

692

693

694

## 695 List of Figures

696

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

702

703 Fig. 2. Specific inherent optical properties (SIOPs) of Lake Dianchi. (A):  $a_{ph}^*$  [m<sup>2</sup> mg<sup>-1</sup>],  $a_{tr}^*$  [m<sup>2</sup> g<sup>-1</sup>],  
 704 and  $a_{CDOM}^*$  [dimensionless] are specific absorption coefficients of phytoplankton, tripton, and  
 705 colored dissolved organic matter, respectively;  $a_w$  [m<sup>-1</sup>] is the absorption coefficient of pure water.  
 706 (B):  $b_{b,ph}^*$  [m<sup>2</sup> mg<sup>-1</sup>] and  $b_{b,tr}^*$  [m<sup>2</sup> g<sup>-1</sup>] are the specific backscattering coefficients of phytoplankton  
 707 and tripton, respectively; and  $b_{b,w}$  [m<sup>-1</sup>] is the backscattering coefficient of pure water.

708

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

711

712 Fig. 4. Initial estimation models for (A) chlorophyll *a* concentration; (B) tripton concentration; and  
 713 (C) CDOM absorption at 440 nm based on previous semi-analytical indices.

714



715 Fig. 5. Examples of estimation models for (A) chlorophyll *a* concentration when tripton  
716 concentration and CDOM absorption at 440 nm are constants; (B) tripton concentration when  
717 chlorophyll *a* concentration and CDOM absorption at 440 nm are constants; (C) CDOM absorption  
718 at 440 nm when chlorophyll *a* concentration and tripton concentration are constants.

719

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

721

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

724

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

727

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

731

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

734

735

736

737

738

739

740

741

742

743

744

745

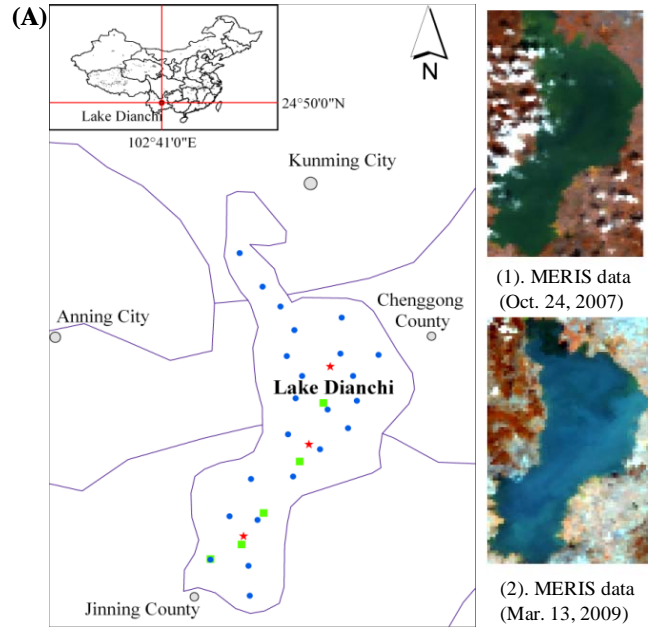
746

747

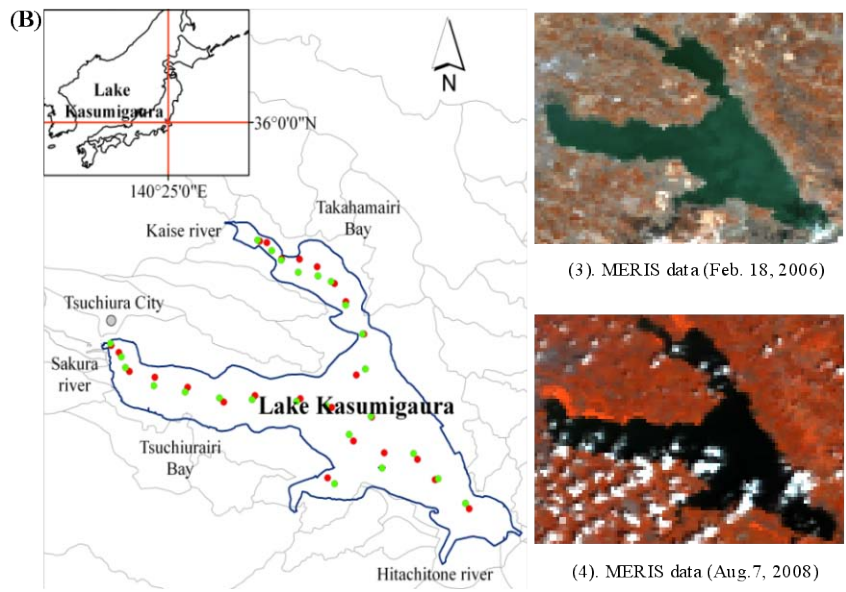
748

749

750



751



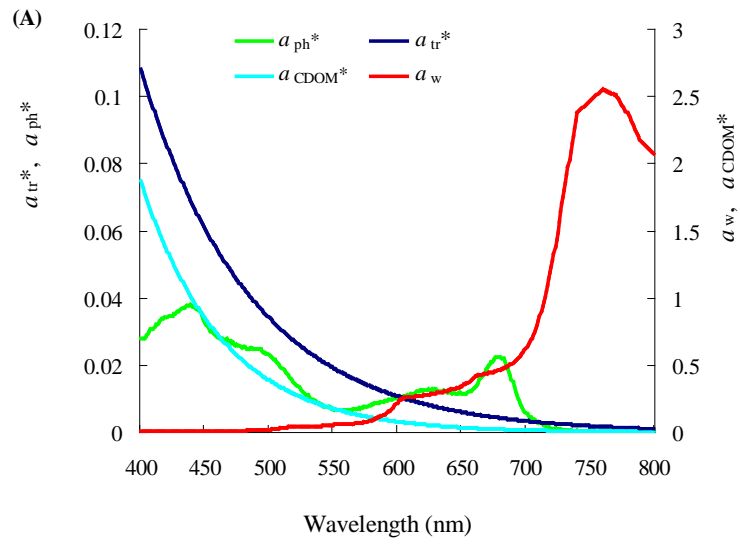
752

753

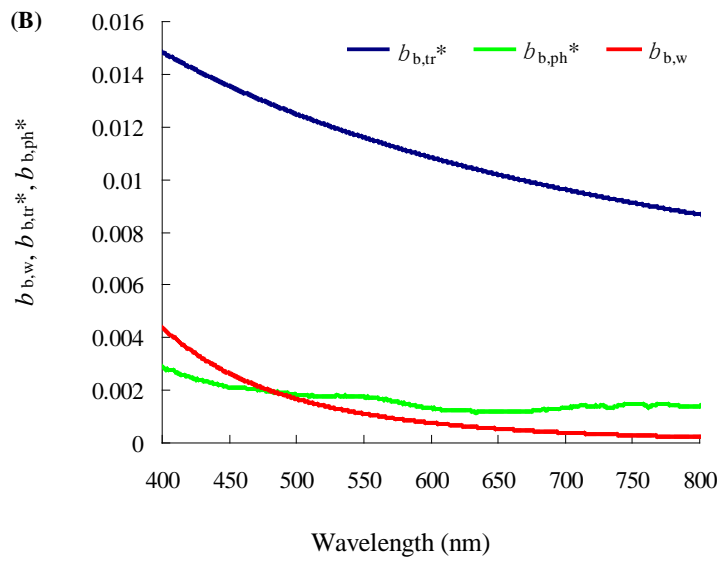
Fig. 1.

754

755



756

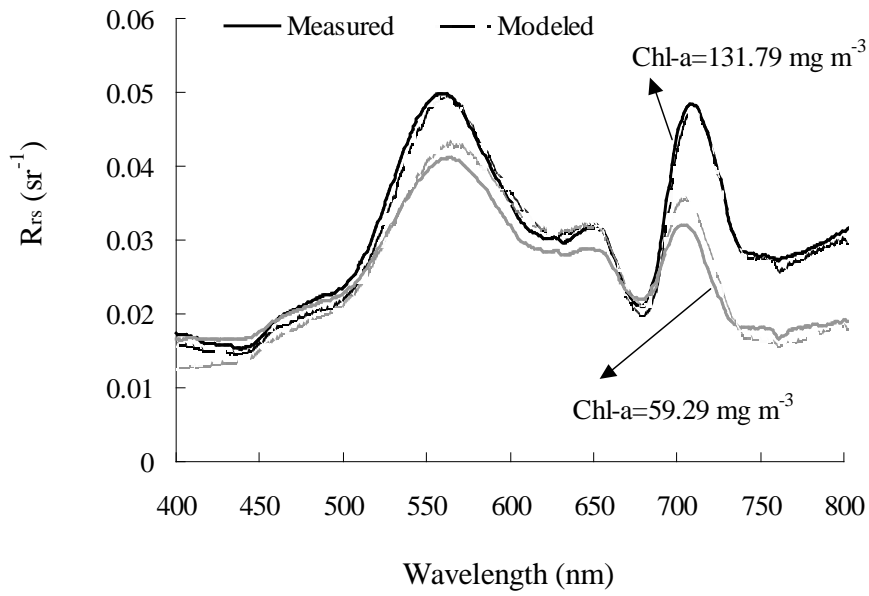


757

758

759

Fig. 2.



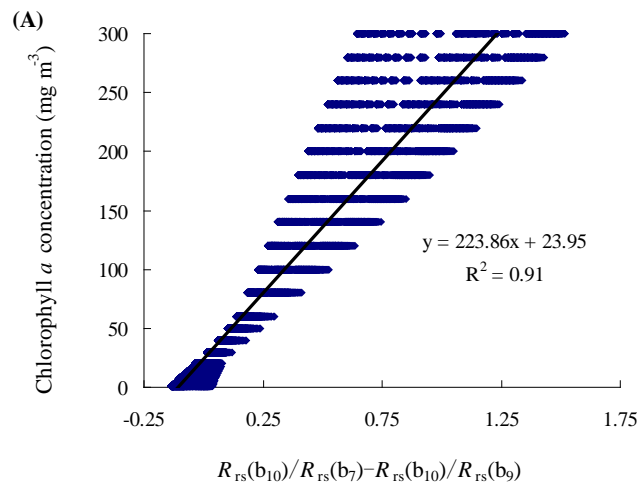
760

761

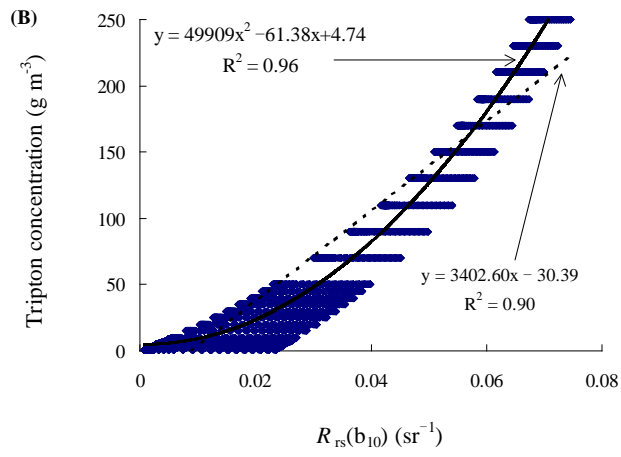
762

763

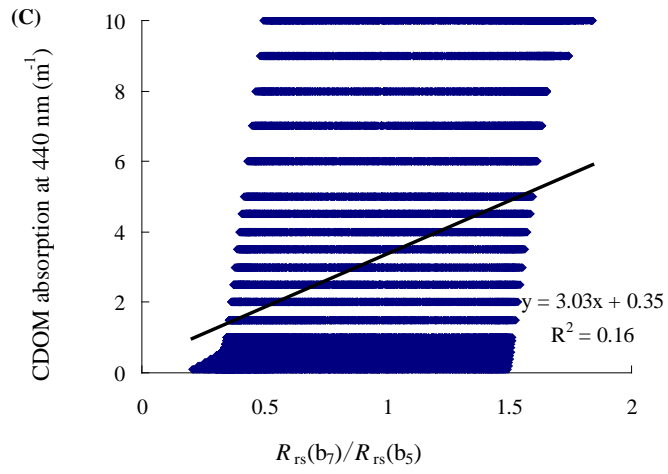
Fig. 3.



764



765



766

767

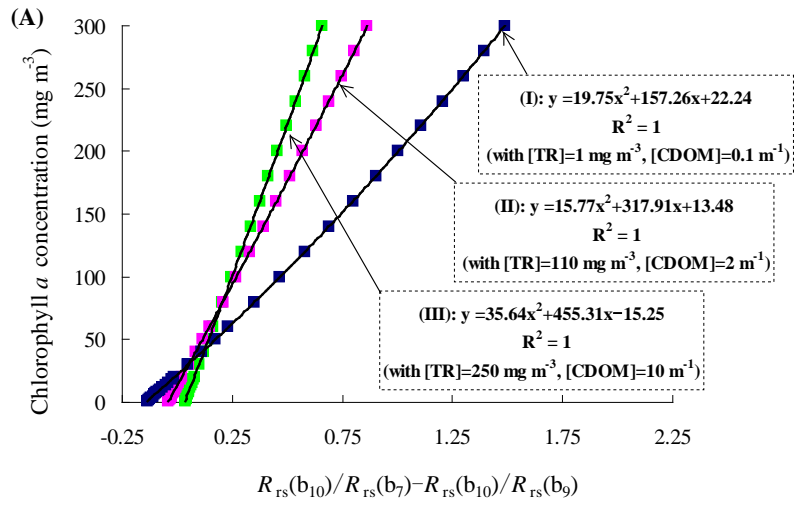
768

769

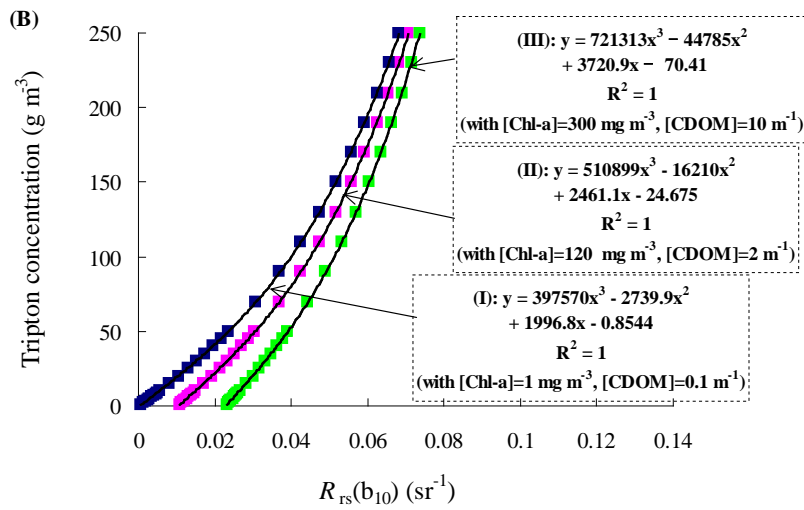
770

771

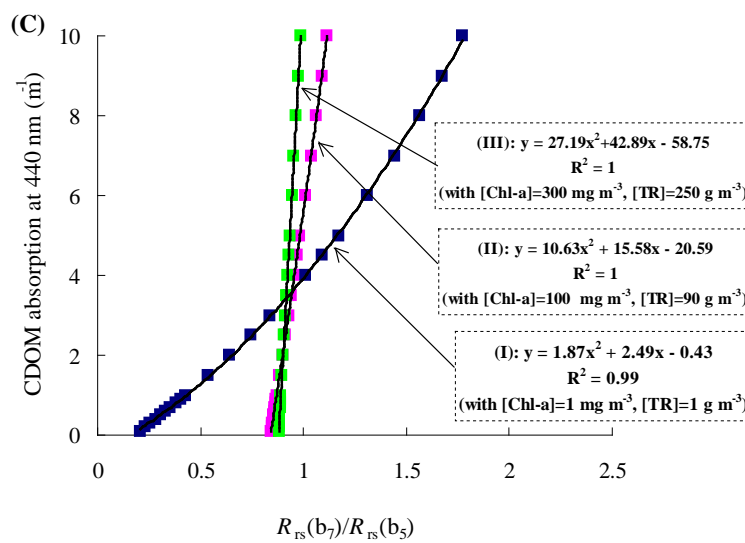
Fig. 4.



772



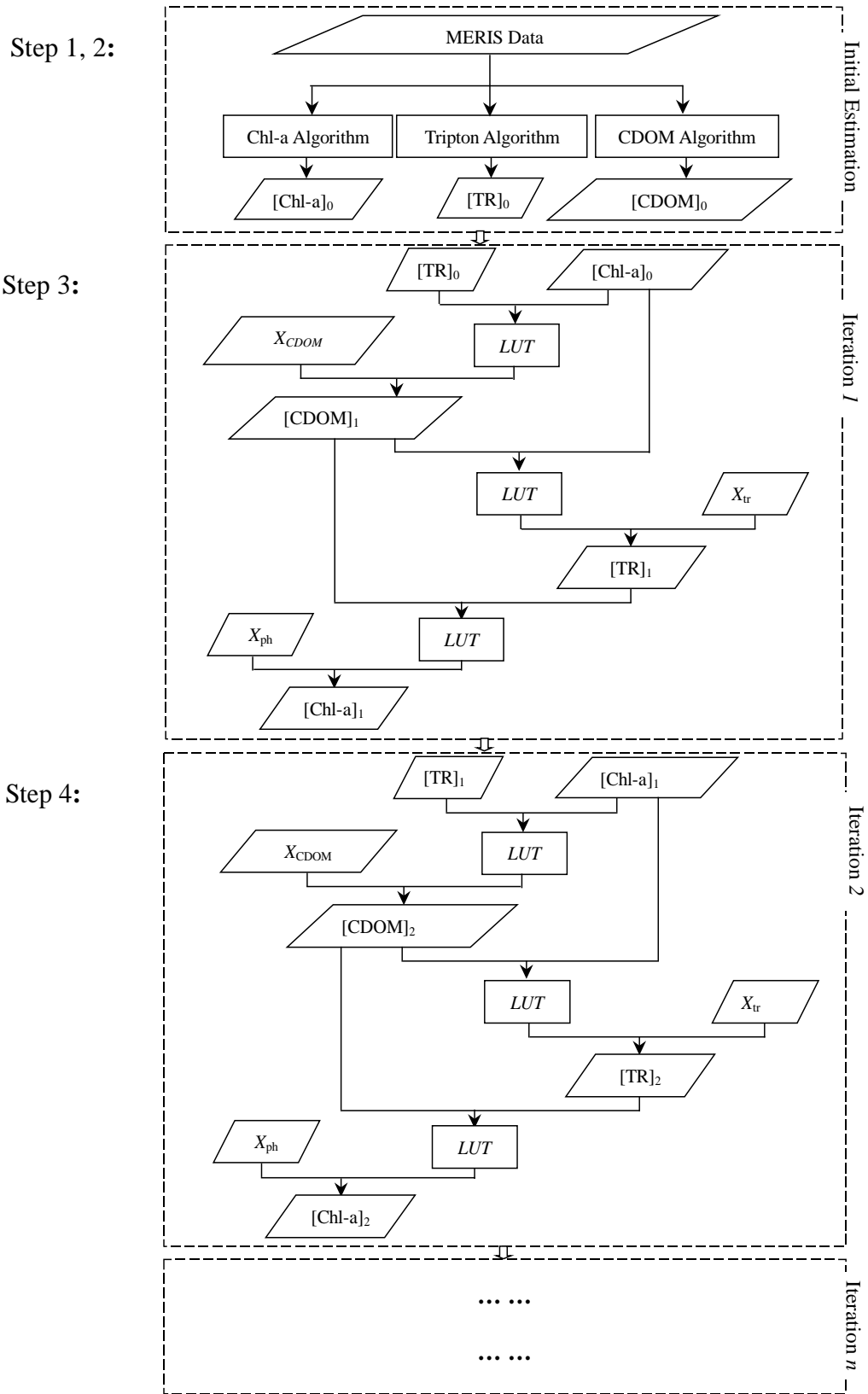
773



774

775

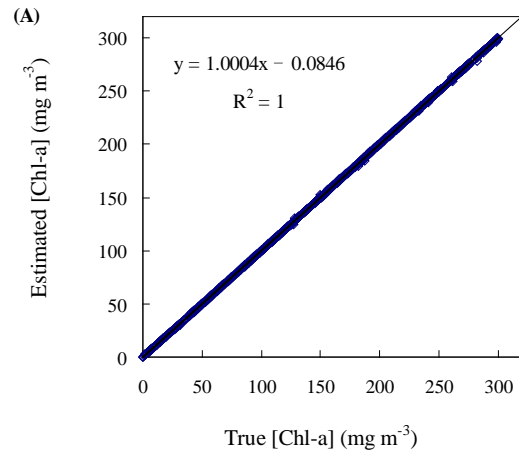
Fig. 5.



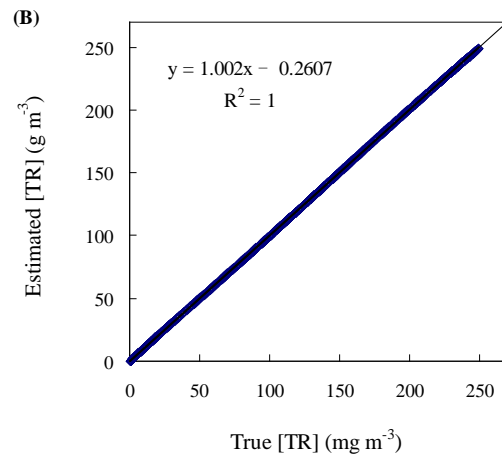
776

777

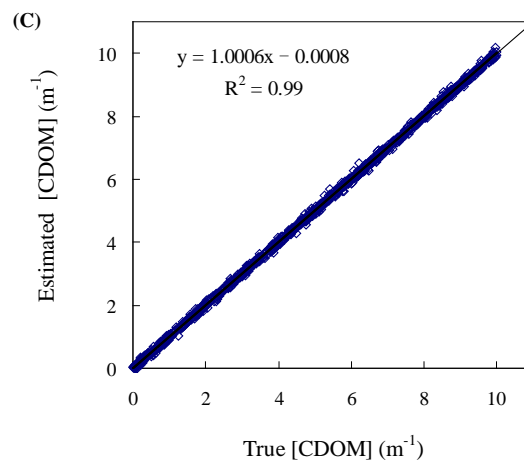
Fig. 6.



778



779



780

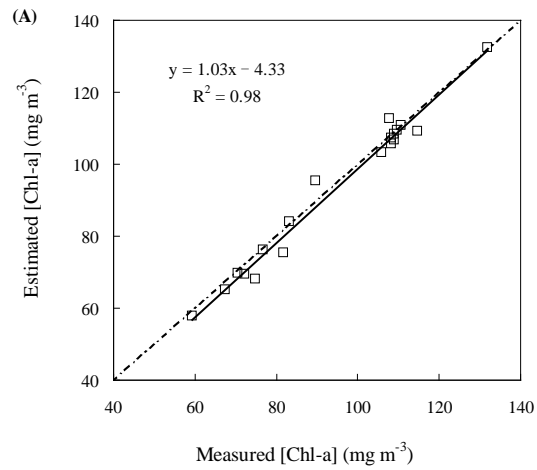
781

782

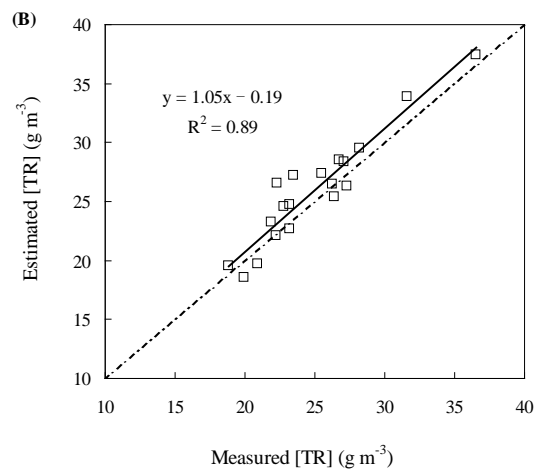
783

Fig. 7.

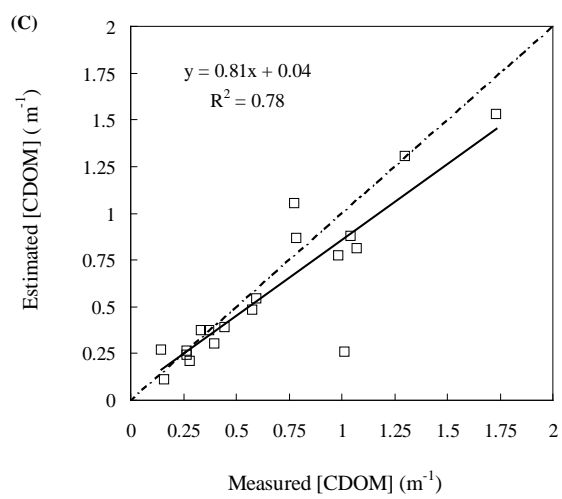




784



785



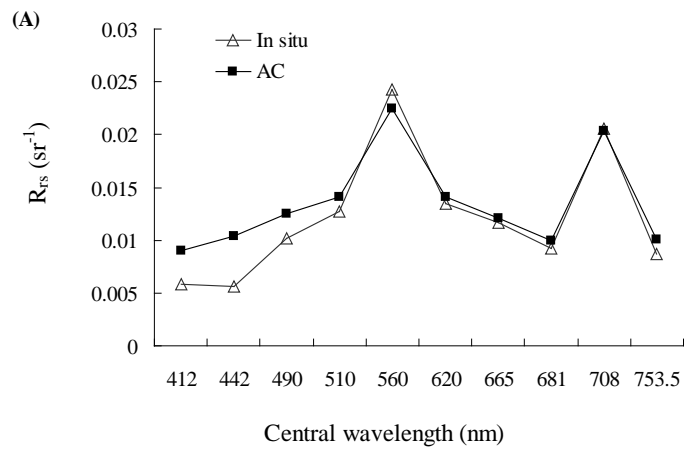
786

787

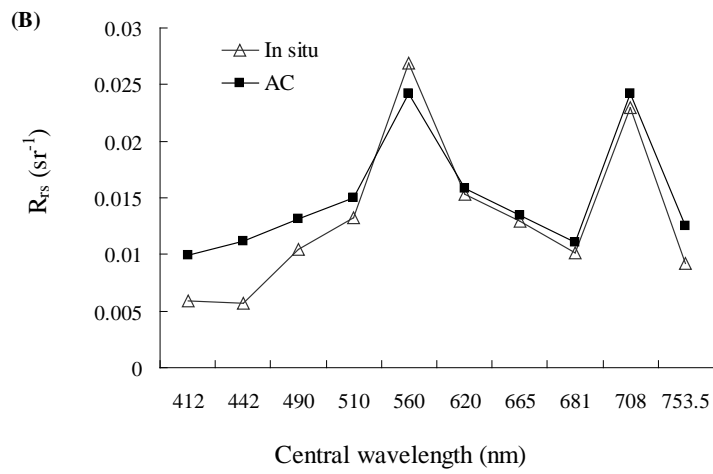
788

Fig. 8.

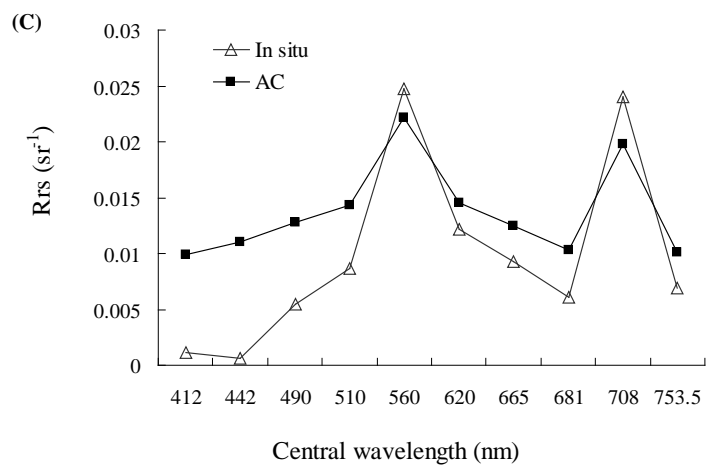
789



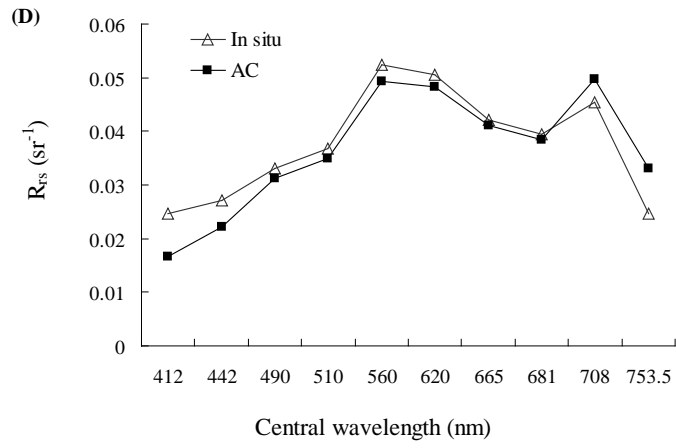
790



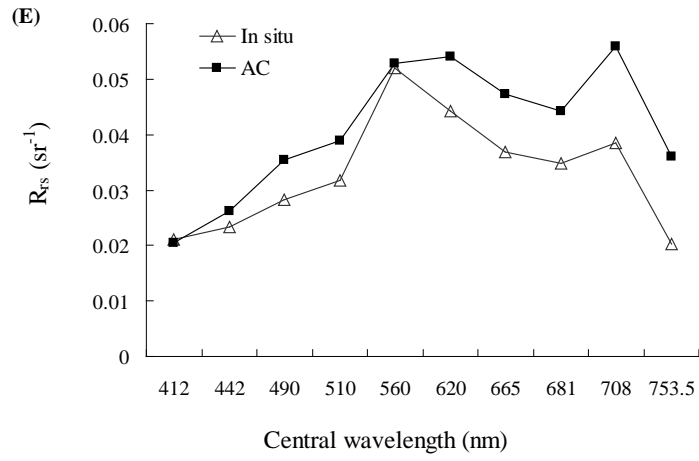
791



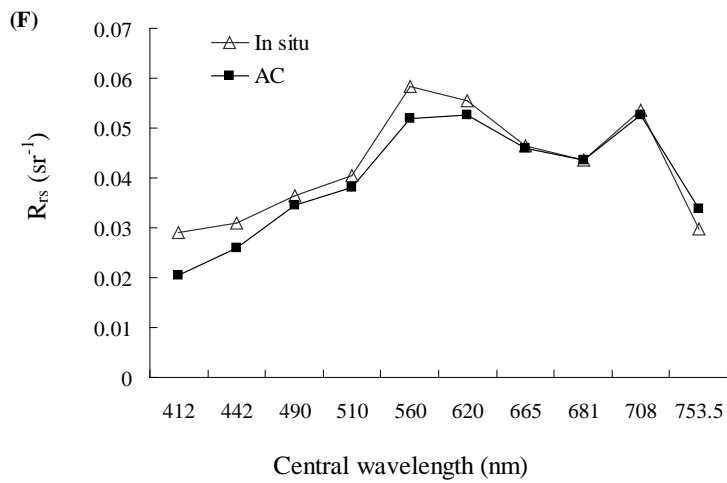
792



793



794

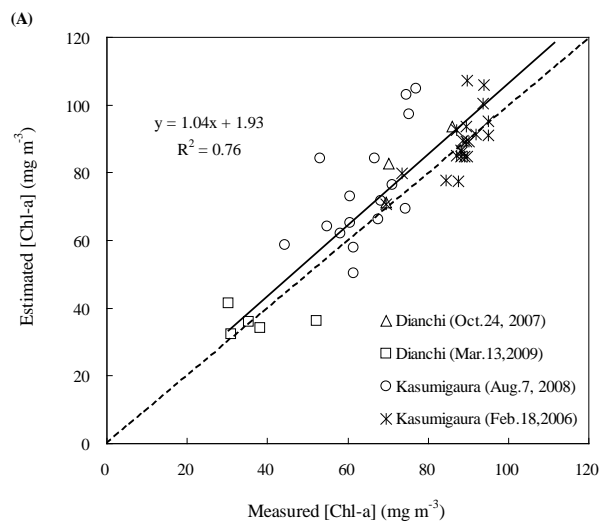


795

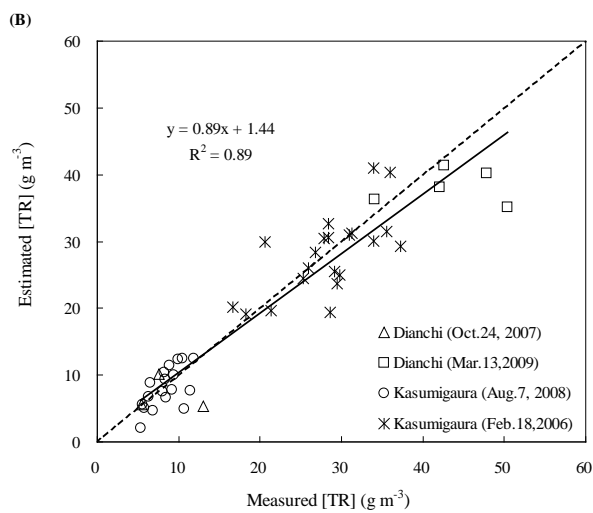
796

797

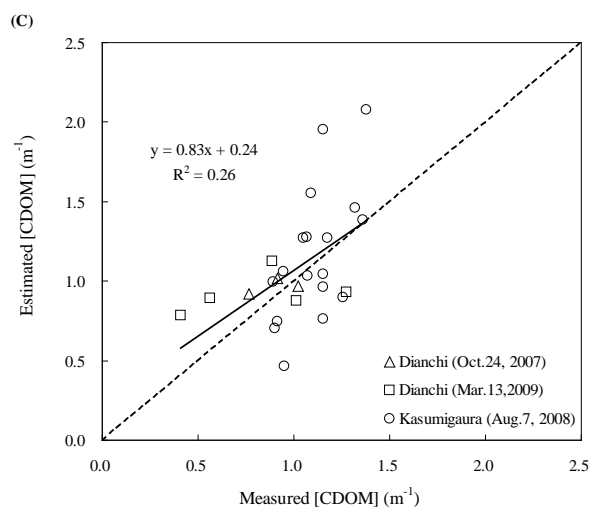
Fig. 9.



798



799



800

801

802

Fig. 10.