Int J Appl Earth Obs Geoinformation 74 (2019) 138–149

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

A soft-classification-based chlorophyll-*a* estimation method using MERIS data in the highly turbid and eutrophic Taihu Lake

Fangfang Zhang[a], Junsheng Li[a,*], Qian Shen[a], Bing Zhang[a,b,*,1], Liqiao Tian[c], Huping Ye[a], Shenglei Wang[a], Zhaoyi Lu[a,b]

[a] Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Keywords:
Chlorophyll-*a* estimation
Optical water types (OWT*s*)
Soft-classification-based method
MEdium Resolution Imaging Spectrometer (MERIS)

ABSTRACT

Soft-classification-based methods for estimating chlorophyll-*a* concentration (C_{chla}) by satellite remote sensing have shown great potential in turbid coastal and inland waters. However, one of the most important water color sensors, the MEdium Resolution Imaging Spectrometer (MERIS), has not been applied to the study of turbid or eutrophic lakes. In this study, we developed a new soft-classification-based C_{chla} estimation method using MERIS data for the highly turbid and eutrophic Taihu Lake. We first developed a decision tree to classify Taihu Lake into three optical water types (OWT*s*) using MERIS reflectance data, which were quasi-synchronous (± 3 h) with *in situ* measured C_{chla} data from 91 sample stations. Secondly, we used MERIS reflectance and *in situ* measured C_{chla} data in each OWT to calibrate the optimal C_{chla} estimation model for each OWT. We then developed a soft-classification-based C_{chla} estimation method, which blends the C_{chla} estimation results in each OWT by a weighted average, where the weight for each MERIS spectra in each OWT is the reciprocal value of the spectral angle distance between the MERIS spectra and the centroid spectra of the OWT. Finally, the soft-classification based C_{chla} estimation algorithm was validated and compared with no-classification and hard-classification-based methods by the leave-one-out cross-validation (LOOCV) method. The soft-classification-based method exhibited the best performance, with a correlation coefficient (R^2), average relative error (ARE), and root-mean-square error (RMSE) of 0.81, 33.8%, and 7.0 μg/L, respectively. Furthermore, the soft-classification-based method displayed smooth values at the edges of OWT boundaries, which resolved the main problem with the hard-classification-based method. The seasonal and annual variations of C_{chla} were computed in Taihu Lake from 2003 to 2011, and agreed with the results of previous studies, further indicating the stability of the algorithm. We therefore propose that the soft-classification-based method can be effectively used in Taihu Lake, and that it has the potential for use in other optically-similar turbid and eutrophic lakes, and using spectrally-similar satellite sensors.

1. Introduction

The optical properties of Case I open ocean waters are relatively simple, and dominated by phytoplankton and their degraded materials. Blue-green band ratio algorithms can be effectively used to estimate chlorophyll-*a* concentration (C_{chla}) (Carder et al., 1999; Gordon and Morel, 2012; O'Reilly et al., 1998). The optical properties of Case II coastal and inland waters are more complex, and dominated by a combination of phytoplankton, non-algal particulate matter, and chromophoric dissolved organic matter (Morel and Prieur, 1977). The absorption coefficient of non-algal particulate matter and chromophoric dissolved organic matter is strong among blue and green wavelengths, which makes it difficult to estimate C_{chla} in Case II waters using blue-green band ratio algorithms (Dall'Olmo and Gitelson, 2005; Lavender et al., 2004; Zimba and Gitelson, 2006). In Case II waters, a model is used for only one water body (or season) for C_{chla} estimation by red to near-infrared wavelengths (Gitelson et al., 2008; Gurlin et al., 2011). However, optical properties might vary substantially between different areas of a water body, and among seasons in the same area. Therefore, a single model cannot fully capture all water areas or seasons

* Corresponding authors at: Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China.
E-mail addresses: lijs@radi.ac.cn (J. Li), zb@radi.ac.cn (B. Zhang).
[1] Part-time professor at University of Chinese Academy of Sciences, Beijing 100049, China.

https://doi.org/10.1016/j.jag.2018.07.018
Received 23 December 2017; Received in revised form 15 July 2018; Accepted 21 July 2018
0303-2434/

of a single body of water, let alone those of multiple water bodies.

To resolve this problem of model applicability, a hard-classification-based method was introduced for C_{chla} estimation (Bao et al., 2015; Zhang et al., 2015). In this method, waters are first classified into several optical water types (OWT*s*), and specific estimation models for each OWT are then developed (Eleveld et al., 2017; Le et al., 2011; Liu et al., 2013; Lubac and Loisel, 2007; Shen et al., 2015). Classification methods mainly include unsupervised classification (Ainsworth and Jones, 1999; de Lucia Lobo et al., 2012; Feng et al., 2005; Molleri et al., 2010; Shi et al., 2013), and empirical threshold segmentation of spectral indexes (Le et al., 2011; Li et al., 2012; Sun et al., 2014; Vantrepotte et al., 2012). Although the hard-classification-based method has achieved better performance than traditional, single models in optically-complex waters, some critical problems remain. First, because the water on either side of an OWT boundary requires a different estimation model, there is discontinuity in the estimation results that does not reflect with natural conditions (Bao et al., 2015; Jackson et al., 2017). Second, the optical properties of some water bodies may lie far from the mean (or OWT centroid), yet these waters are still classified as a particular OWT. Finally, water optical properties may differ within the same OWT; thus, using the same model might generate unreliable results.

To resolve the limitations of the hard-classification-based method, a soft-classification-based method was introduced for C_{chla} estimation (Bao et al., 2015; Jackson et al., 2017; Moore et al., 2001, 2009,2014; Zhang et al., 2015). The soft-classification-based method includes four key points: 1) defining the centroid spectra of each OWT by the fuzzy classification method; 2) calculating the distance weights of each spectrum to every OWT centroid spectra; 3) calibrating optimal C_{chla} estimation models for each OWT; and 4) blending the C_{chla} estimation results of all optimal models using inverse distance weights. In recent years, the soft-classification-based method has received more attention because of the following advantages over the hard-classification-based method (Jackson et al., 2017; Melin et al., 2011; Nazeer and Nichol, 2016): 1) no obvious boundaries between two OWT*s* and continuous distance weights, resulting in continuous estimation of C_{chla}; 2) each spectrum (or pixel) has a different distance weight to compensate for the difference within an OWT; and 3) it does not forcibly classify a specific spectrum among several OWT*s* into a particular OWT, instead using the distance weights to every OWT centroid spectra and the blending of several models, which can reduce the estimation errors of a single model.

Currently, the soft-classification-based method for C_{chla} estimation has shown great potential in turbid coastal waters; however, one of the most important water color sensors, the MEdium Resolution Imaging Spectrometer instrument (MERIS), has not been applied and validated in turbid and eutrophic lakes. In this study, we developed a new soft-classification-based C_{chla} estimation algorithm using MERIS data in the highly turbid and eutrophic Taihu Lake. The goal of this study was to: 1) build a new soft classification framework for C_{chla} estimation in Taihu Lake, including an OWT decision tree based on the spectral index and normalized distance weights based on the spectral angular distance; 2) develop optimal algorithms for each OWT of Taihu Lake; 3) validate the soft-classification-based C_{chla} estimation method; and 4) analyze spatial and temporal variations in C_{chla} for Taihu Lake from 2003 to 2011.

2. Study area and data acquisition

2.1. Study area

Taihu Lake is located in the southern Yangtze River Delta, China (latitude: 30°55′40″N-31°32′58″N; longitude: 119°52′32″E-120°36′10″E; Fig. 1). It is China's third largest freshwater lake, with a total water area of ~2338 km^2, and an average depth of ~1.89 m (Qin, 2004). The optical properties of Taihu Lake are complex, varying

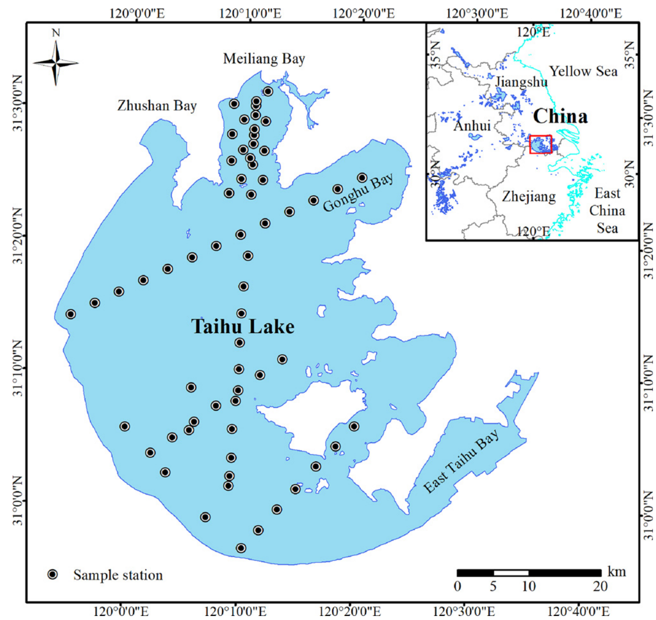

Fig. 1. Location of Taihu Lake, China, and sampling stations of eight cruises, conducted in 2006, 2007, and 2009. Data from a total of 91 sampling stations were obtained. Detailed information for each cruise is listed in Table 1.

substantially by season and location, and even with a period of 24 h at the same location due to bottom sediment resuspension and algal blooms. The water in the northern Meiliang Bay, Zhushan Bay, and the west coast is often characterized by hypereutrophic conditions and frequent outbreaks of blue-green algae. The water in the central area of the lake is usually highly turbid with high non-algal particle (NAP) concentrations, and the water in the southeast is typically clearer than other regions. In the summer and fall, Taihu Lake is typically hypereutrophic, while in the winter and spring, it is generally turbid. Considering the complex and highly variable optical properties of Taihu Lake, it is difficult to develop a single C_{chla} estimation method with high accuracy in all seasons and all areas. Therefore, Taihu Lake presents a suitable study area for testing the robustness of soft-classification-based C_{chla} estimation methods.

2.2. Field data

We conducted eight sampling cruises in Taihu Lake in 2006, 2007, and 2009, concurrent with MERIS image acquisition. Samples were collected from 91 stations (Fig. 1), including 46 stations in January, 14 stations in March, 21 stations in April, 10 stations in August, and 7 stations in October (Table 1). At each station, water samples were

Table 1
Dates, time, and sampling site numbers for the eight cruise surveys. Sampling was concurrent with MERIS image acquisition on eight days in 2006, 2007, and 2009.

Concurrent data		*In-situ* experiment	MERIS image	
Acquisition date	Number of stations	Local time	Greenwich time	Local time
Jan. 9, 2006	15	8:55–13:58	2:41	10:41
Aug. 1, 2006	8	9:04–11:47	2:02	10:02
Oct. 12, 2006	7	10:34–13:50	2:38	10:38
Jan. 7, 2007	19	9:25–13:26	2:04	10:04
Jan. 9, 2007	11	8:55–12:51	2:41	10:41
Apr. 25, 2007	11	8:20–13:12	2:10	10:10
Mar. 15, 2009	14	8:33–13:50	2:24	10:24
Apr. 26, 2009	6	9:17–13:31	2:04	10:04