A novel aggregate cyanobacterial biomass proportion index for estimating cyanobacteria succession in early eutrophic Lake Erhai, China

Shangbo Yang 1, Xiaoling Chen 1 and Jianzhong Lu 1,*

1 State Key Laboratory of Information of Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
* Correspondence: lujzhong@whu.edu.cn; Tel.: (+86)27-68778755

Abstract: Erhai Lake, located on the Yungui plateau in southwest China, has been considered to be in a transition period of ecological process, posing an urgent need for understanding its historical succession of cyanobacteria and further detecting the early signals of cyanobacteria accumulation for developing management strategies in advance. For this reason, an aggregate cyanobacterial biomass proportion index (ACBPI) was introduced as bio-indicator for reflecting increased accumulation of cyanobacteria, through targeting cyanobacteria-associated indexes derived from satellite remote sensing using principal component analysis. Thresholds for ranking the cyanobacteria abundance state were then determined through in situ phytoplankton composition data and the entire ACBPI time series. The results showed that the ACBPI correlated with cyanobacteria biomass proportion with an accuracy level of 66% and cyanobacteria biovolume proportion with coefficient of determination 80%. Dense bloom appeared primarily in northern regions, with 5.5% occurring in 2003, 9.1% in 2006, and 6.7% in 2008. The frequency of moderate bloom in northern lake made up a higher share (14.1±16.0%) across the whole periods, with 6.2±10.7% in central lake and 2.5±4.0% in southern lake. Apparent mitigation of cyanobacterial dominance condition was observed in 2016-2019 in contrast to 2003-2011 with obvious reduction occurring in 2018, probably resulting from series of strict protection initiatives implemented in recent years. However, moderate bloom in northern bays occurred again in 2019, indicating that strict nutrient reduction especially phosphorus pollution should be strengthened under global warming and wind speed decreasing scenario.

Keywords: cyanobacterial biomass proportion; cyanobacterial dominance level; Spatio-temporal variability; remote sensing

PACS: J0101

1. Introduction

The problem of cultural eutrophication in inland waters caused by excessive input of nutrients has attracted wide attention in the world. The eutrophication process is essentially induced by the imbalance of material exchange and the degradation of the structure and function of lake ecosystem. Especially in early eutrophication state, cyanobacteria prone to become the dominant species and sporadically grow into cyanobacterial bloom once appropriate environment conditions are met, through competing with eukaryotic algal (i.e., green algae, diatom) owing to its special characteristic, thus seriously affecting the sustainable development of freshwater ecosystem. Therefore, understanding long-term succession and abnormal growth of cyanobacteria and the main influential factors is key to determine reasonable lake management strategy and then address the pressing issues.

Several pioneering studies have demonstrated cyanobacteria-related ocean color indices as one promising approach for long-term detection and quantification of cyanobacterial bloom such as cyanobacterial dominance [1], cyanobacteria cell count [2], cyanobacterial biomass[3], proportion of cyanobacterial biomass [4], cyanobacterial phycocyanin pigment (PC) [5] and floating algae index...
Cyano-dominant waters show unique sun-induced phycobilipigment fluorescence (SIPF) induced peak near 664nm, which is caused by the effect of phycocyanin absorption at 620nm and sparse absorption near 664nm band [7]. Modified Cyanobacterial Index (CI-cyano), termed as negative spectral shape algorithm SS (665) [8] or computationally equivalent to the negative of SIPF, is capable to capture this discernable feature of cyanobacteria blooms and has been used to identify high biomass cyano-dominant waters [7,9]. Besides, a PC index (PCI) has been demonstrated a reliable index for estimating PC pigment concentrations or identifying cyanobacterial bloom [5]. Additionally, when extremely high cyanobacteria biomass occurs, the FAI introduced in Hu et al. 2009 [6] is generally to characterize the intense blooms of cyanobacteria or floating algae scums (surface accumulations). For this reason, in terms of early eutrophic lake with sporadically occurred cyanobacterial bloom compared to hypertrophic lakes with frequently occurrence of surface scums, PCI and SIPF appears to be a promising index for characterizing succession of cyanobacteria. In addition, Miao et al. 2020 introduced a normalized index of the proportion of cyanobacterial biomass (NIPCB) to successfully map proportion of cyanobacteria biomass (PCB) using semianalytical algorithm based on OLCI data processed with fully atmospheric correction [4]. Miao’s study performed a bio-optical study to find a spectral index NIPCB to exactly account for the proportion of cyanobacteria, isolating it from other interfering effects. However, to our knowledge, fully atmospheric corrected satellite data would lead to decreased data coverage, which is unfavorable for long-term time series research. The Rayleigh-corrected spectral reflectance (Rrc) based PCI and SIPF algorithm makes it nearly immune to most perturbations from the atmosphere, thus leading to much increased usable data coverage as compared with NIPCB algorithm.

As a typical early eutrophic lake, Lake Erhai experience critical transition since 2000s [10]. Pressures facing Lake Erhai involving external pollution, long hydraulic residence time and shortage of inflowing water jointly brings challenges to the eutrophication control of Lake Erhai. Several study has focused on documenting long-term satellite estimated chlorophyll concentration[11,12], however chlorophyll concentration could not capture the specific algal population information. Wang et al. 2018 has explored early warning method for algal cell densities by using two-year monthly chlorophyll fluorescence parameters [13] and Chen et al. 2015 has investigate population dynamics of phytoplankton during cyanobacterial bloom [14]. However, long-term cyanobacteria proportion dynamic has not been explored and documented, which could reveal the phytoplankton composition structure and provide a specific view of seasonal cyanobacteria growth. Moreover, the underlying influential factors for seasonal growth cycle and abnormal growth of cyanobacteria (cyanobacterial bloom) have not been explored in Lake Erhai. The object of this study are as followed: (1) to establish an empirical model for estimating cyanobacteria proportion based on cyanobacteria-associated indices (2) to explore the temporal pattern and environmental drivers of cyanobacteria succession and abnormal growth in Lake Erhai.

2. Results

2.1. Spatial-Temporal trajectory of satellite estimated cyanobacterial biomass proportion
Figure 1. Long-term spatial and temporal trajectory of MERIS derived CBP (2003-2011) and OLCI derived CBP (2016-2019) and its seasonal dynamics

As presented in Figure 1, the spatial pattern of satellite derived CBP in entire lake remained relatively stable with highest CBP level occurring in northern lake, followed by central and southern lake. Obvious temporal trajectory change was observed in summer and winter. Specifically, there is a declining trend in estimated CBP in summer while winter bloom appears to frequent during 2016-2019. The result reveals apparent low CBP in 2016-2019 occurred in 2018, however, the bounce back of CBP occurred in winter 2019 but the CBP level was not as high as the level in winter 2016 and 2017. In addition, the temporal trajectory in spring remain stably low while in autumn season the CBP level generally kept relatively high compared to other seasons in the whole year.

2.2. Seasonal dynamic of cyanobacterial biomass proportion
Figure 2. The monthly changes in biomass of the major groups of phytoplankton and relative cyanophyta biomass in three parts of Erhai Lake during 2015-2016.

The phytoplankton communities of the entire lake are dominated alternately the three groups of algae population such as cyanophyta, bacillariophyta and chlorophyte (together comprising up to more than 80% of the total biomass based on our measurement), with all exhibiting distinct seasonal succession in phytoplankton. As shown in Figure 2, generally cyanophyta exhibits a marked seasonal cycle from June to December with peak occurring in July and October, while chlorophyte exhibits relatively one high peak in December, and bacillariophyta displays two relatively high peaks occurring in June and September. However, in addition to minor difference in peak time (almost one month lag) of each algae population from the three parts of lake, there is marked difference in magnitude of total biomass and relative cyanophyta biomass across the lake.

The average of the amount of total biomass in northern, central and southern lake is 0.63±0.31 mg/L, 0.77±0.32 mg/L and 0.95±0.4 mg/L, while the average of the magnitude of relative cyanophyta biomass is 23.9±11.6%, 23.4±14% and 22.3±12.2%, respectively. The field sampling data that are currently available during 2015-2016 indicate that southern lake exhibited high total biomass but low relative cyanophyta biomass. Fortunately, long–term satellite-estimated PCB can be utilized for spatial and temporal analysis of cyanophyta dynamics.

2.3 Cyanobacterial bloom and its potential drivers

2.3.1. Cyanobacterial bloom level

The spatial and temporal distribution of bloom severity level was clearly distinguished by classification into four levels according to the magnitude in PCEH concentration (Figure 4a). As shown in Figure 4b, cyanobacterial bloom severity condition in northern lake was most severe, followed by the central (similar to condition in entire lake), then the minimum in the southern lake across the study periods. Specifically, dense bloom appeared frequently in northern regions, with 5.5% occurring in 2003, 9.1% in 2006, and 6.7% in 2008. The frequency for moderate bloom made up a relatively significant share (14.1±16.0%) across the whole periods in northern lake, with 6.2±10.7% in central lake and 2.5±4.0% in southern lake, while the frequency of sparse bloom was similar, with 22.2±3.1% among the different parts of lake.
Figure 4. (a) Latitude-Time plot of cyanobacterial bloom type classified by bloom level. (b) Accumulative percentage of cyanobacterial bloom level for the entire lake and each segment.
Figure 5. GAM partial dependency plots demonstrate the effects of deseasonalized influencing factors on deseasonalized PCB. Panel (a) show the modeling result for entire lake in 2003-2011; Panels (b) illustrate the model for entire lake, which incorporates TN, water level and water temperature, respectively; The connected line is the spine, black dots represent the raw data, and gray shadow denotes 95% confidence interval. The yAxis tick label presents estimated smooth term denoted as s (covariate, edf) with edf displaying the effective degrees of freedom for the term. The P-value is expressed as * (P<0.05), ** (P<0.01) and *** (P<0.001).
In period 2003-2011 for summer datasets, the GAM using two predictor variables indicate that 96.8% of the variation in deseasonalized PCB was interpreted with RMSE of 50.23%. The influence of rainfall intensity and N:P ratio on deseasonalized PCB was 54.2% and 33.9%, respectively. In first period for autumn data, the result shows that 56.6% of the variation in deseasonalized PCB was explained (RMSE=37.6%), with 48.4% of the variance in deseasonalized PCB was explained by the sunshine duration variable and 29.3% of the variance in PCB interpreted by surface air pressure. As for winter datasets, the GAM result shows that MDWspd had the highest relative influence (73.0% of the total deviance explained) on deseasonalized PCB with RMSE of 47.6%. In period 2016-2019 for summer and fall datasets, 70.0% of the variation in deseasonalized PCB was explained by water temperature(RMSE=44.32%), and 69.6% of the variation in PCB was interpreted by air temperature (RMSE=18.37%). As for winter datasets, TP interpreted 46.7% of the PCB variation with RMSE of 54.8%. In a word, deseasonalized TN and TP both have significant impact on deseasonalized PCB especially in summer 2003-2011, while deseasonalized TP has obvious influence on PCB in winter 2016-2019. Moreover, other environment factors (i.e, rain intensity, air pressure, water and air temperature, sunshine duration, and MDWspd) have significant modulating effect on abnormal cyanobacterial bloom.

3. Discussion

3.1 Uncertainty of estimated PCB and cyanobacterial bloom

Though there was a four years data gap (2012-2015) in estimated PCB time series during 2003 and 2019, it did not have much influence to find out the substantial variability of seasonal and inter-annual cyanobacteria dynamics during the study periods. In addition, satellite estimated PCB time series were derived from different satellite instrument and there were sensor-associated differences (spectral responses, wavelength, etc.) between OLCI and MERIS. An ideal solution is to develop sensor-specific models; however, it is difficult to obtain field measured PCB concentration data during 2003-2011 and it is almost impossible to find relatively stable target among different images considering potential speckle noise, acquisition time and atmospheric failures. Additionally, there is no statistically significant difference between the spectral bands (required for constructing PCI algorithm) centered at 560nm, 620nm and 665nm from OLCI and MERIS image. Thus, it has little impact on the following trend analysis of the long-term estimated PCB time series.

3.2 Implications for lake nutrient management

The decline trend in satellite estimated PCB in summer (especially suitable temperature conditions) reflect the decrease of nutrient inflowing, probably contributing to the long term effort for series of lake management act. However, winter bloom occurred in 2016 and 2017, probably induced by abnormal TP enhancement. Fortunately, the PCB level showed a sharp decrease in 2018 under strict control of nutrient loadings. In fact, banning garlic planting in the whole watershed in 2018 was implemented as the part of Seven Major Actions to “urgently rescue Erhai Lake”, which is the strictest environmental protection campaign in this region to date. This drop is probably linked with strict implemented policies of agricultural planting structure adjustment. Moreover, the increase of PCB also exhibits an warning to consistently reduce TP pollution especially in winter season to cope with rising air temperature.

4. Materials and Methods

4.1. Study area and In situ datasets
Figure 6. The location of Lake Erhai and its hydrological map. The bathymetry of Lake Erhai is processed using underwater elevation points with 5-m depth intervals acquired from the Bureau of Erhai Protection and Administration of Dali, and the yellow arrow displays the main throughput flow.

Lake Erhai (25°57’-25°36’N, 100°05’-100°17’E, Figure 6), is a typical plateau freshwater lake in the western part of Yunnan Province in Southwest China. This lake is the main water resource for residents with water retention time up to almost 1000 days. The basin is influenced by southwest monsoons and has a subtropical climate with four identical seasons: spring (March-May), summer (June-August), fall (September-November), and winter (December-February). The average annual rainfall is 984 mm, of which approximately 80% occurred from May to October in 2002-2019. The
water supply is mainly from precipitation and inflowing rivers, which are dominated by the Miju River, Luoshi River, and Yongan River in the north of Lake Erhai basin.

Monthly field sampling was measured at the site of eight ecological in-situ observation points throughout 2015 and 2016. After each fieldwork, phytoplankton density and wet weight algal biomass were determined through experiments [15]; The algal samples were firstly fixed with Lugol’s iodine solution (1.5%v/v) and later in the laboratory, a Leica microscope (DM750, Leica) was used to sediment and count the samples until a significant volumetric cell number was achieved. The volume for each species was measured by calculating the cell dimensions and suitable geometric configurations [16]. Assuming cell density equivalent to water, the volumes in cell L$^{-3}$ were transformed into biomass in mg L$^{-1}$. The biomass was acquired at least at the genus level and the absolute biomasses were calculated for the following taxonomic classes: Cyanophyta, Bacillariophyta, Chlorophyta, Chrysophyta, Xanthophyta, Cryptomonad, Pyrrophyta, and Euglenophyta.

Satellite images were collected from the Medium Resolution Instrument Sensor (MERIS) onboard the ENVISAT satellite (2002-2012) and Ocean Land Color Instrument (OLCI) on the Sentinel-3A satellite (2016-now) by ESA (https://www.esa.int/ESA). As an improved successor of the MERIS sensor, the OLCI has the same spectral bands as that, plus six extra bands at 400nm, 673.75nm, 764.37nm, 767.5nm, 940nm, and 1020nm with higher accuracy, greater wavelength, and coverage. Besides, a series of candidate environmental driving factors of cyanobacteria variability were obtained including chemical variables (nutrient concentrations such as total nitrogen (TN) and total phosphorus (TP), mg/L), hydrological variable (lake water level, m), and meteorological variables such as sunshine duration (hours), wind speed (m/s), air pressure (hPa), air temperature (°C) and rainfall intensity (mm/day). In addition, the maximum consecutive days when wind speed less than 3m/s in a month is represented by MDWspd. The chemical and hydrological datasets were acquired from the Bureau of Erhai Protection and Administration of Dali from monthly sampling. Meteorological datasets were obtained from the nearest national meteorological station of Lake Erhai (Dali Station), which were sourced from the National Meteorological Administration of China (http://data.cma.cn)

4.2. Methods

4.2.1. Estimation of Satellite derived PCB based on Bayesian hierarchical linear model

As shown in Figure A1, the scatter plot SIPF versus PCI for all valid pixels, it could be found that there was a significant positive correlation between SIPF and PCI ($R^2=0.52$, MAPE=264%), and relatively high FAI value ranging from 0.024 to 0.066 occurred when PCI is in excess of 0.015 or SIPF in excess of 0.004. Owing to the redundancy existing in SIPF and PCI, principal component analysis (PCA) was firstly performed on these two indices to maximum the variance of the cyanobacteria-related index. The cyanobacteria aggregate index was obtained according to the first principal component of PCA. Then, available matchups (N=28) of in situ sampling data of phytoplankton species composition involving algal abundance and biomasses of all eukaryote and cyanobacteria species from June to November (covering cyanobacteria growing season) in 2016 was used to investigate the relationship between the aggregate index and these algal species composition parameters. The aggregate index was found to correlated well with proportion of cyanobacterial cell counts(Figure 7a), and the proportion of cyanobacterial biomass in total phytoplankton biomass(Figure 7b), respectively. Logistic regression model was applied to obtain the decision boundary for cyanobacteria and eukaryotic algal according to the reference proposed by Zhou et al.2019, that the relative cell counting of the single bloom-dominated groups exceeded 60% [17]. Bayesian hierarchical (Multilevel) linear model was constructed to model the relationship between the aggregate cyanobacterial biomass proportion index (ACBPI) and PCB owing to its constructing hyperpriors on group-level parameter to allow the model sharing the individual properties of PCB among the groups. As shown in Figure 7b, matchups of in situ PCB were divided
based on three regions. The coefficient of determination ($R^2$) for the regression model is 0.73 with root mean square error (RMSE) of 13.58%.

The estimated cyanobacteria proportion in hierarchical linear Model at a given APCBI value (0.003) for north region matchups is 20%, which is termed as threshold of sparse bloom. Liu et al. 2019 suggested when the cyanobacteria proportion more than 50%, cyanobacteria bloom could lead to moderate health hazards according to WHO [18]. For this reason, the 50% cyanobacteria biomass proportion is considered as threshold of moderate bloom. Then the 100% of PCB is termed as dense bloom threshold when cyanobacteria forming surface scums.

Figure 7. (a) Logistic regression model for classifying cyanobacteria and eukaryotic algal; (b) Bayesian hierarchical linear model for estimating PCB based on APCBI.

4.2.2. Time Series decomposition and exploration for abnormal growth influenced by environment factors

Breaks For Additive Seasonal and Trend (BFAST) method is widely applied to decomposes the time series remote sensing data into trend component, seasonal component and residual component, and then detects the mutation in trend component and seasonal component [19]. Seasonal component and deseasonal component of satellite estimated PCB time series were obtained through BFAST approach. Then, generalized additive models (GAMs) was constructed to describe the effects of deseasonal regional environment change on cyanobacteria abnormal growth represented by deseasonal satellite estimated PCB. The GAM was performed using the mgcv package in R studio for estimating the relationship between the response and smooth functions of explanatory variables in an additive form [20].

5. Conclusions

In this study, an empirical bayesian hierarchical linear model was developed to estimate the PCB in Lake Erhai based on aggregate cyanobacterial biomass proportion index using atmospherically Rayleigh-corrected reflectance levels (Rrc) product. The BFAST and GAM method have been used successfully to identify deseasonalized environmental variables influencing abnormal cyanobacteria growth or cyanobacterial bloom in Lake Erhai.

The finding and the approach demonstrated here have significant implications for long-term monitoring of lake environments. Not only will such monitoring provide accurate data for ecological management but they can also be used to evaluate the effectiveness of ecological restoration initiatives. Generally, this study shows additive and synergistic effects of climate change and human activities on PCB on seasonal and interannual timescales, and provides long-term baseline information for future remediation efforts to improve eco-environment of Erhai lake.
Acknowledgments: This work was funded by the National Natural Science Foundation of China (41971402), the Frontier Project of Applied Foundation of Wuhan (2019020701011502), the Natural Science Foundation of Hubei Province (2019CFB736), the ESA-MOST Dragon Program, and the LIESMARS Special Research Funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix B

Figure A1 (a) Comparison of the PCI index versus in situ proportion of cyanobacterial abundance or biomass to the total and the best sigmoid or linear fit and coefficient of determination are also provided, respectively. The range boundary I, II and III of PCI value is 0.0041, 0.0065 and 0.0107, corresponding to estimated PCEH concentration value at 41, 88 and 334 ug/L, respectively. (b) Scatter plot of SIPF and PCI index values for all valid MERIS and OLCI pixels displaying either immersed cyanobacteria or eukaryotes. The color bar denotes FAI index value of pixels flagged as floating cyanobacteria. The dashed vertical line represents specific PCI value (0.006) when SIPF is equivalent to 0 in the linear regression line (c) Histogram of the PCI index values for all the valid pixels (N=1518,057), The dashed vertical line represents the location of the global bloom threshold (0.0041) that is calculated as the mean plus upper quartile standard deviations.

References


© 2020 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).