A novel remote sensing algorithm to quantify phycocyanin in cyanobacterial algal blooms

S Mishra\(^1\) and D R Mishra\(^2\)

\(^1\)Dow Agrosciences LLC, Indianapolis, IN 46268, USA
\(^2\)Department of Geography, University of Georgia, Athens, GA 30602, USA

E-mail: sachi.msstate@gmail.com

Received 23 July 2014, revised 3 October 2014
Accepted for publication 3 October 2014
Published 3 November 2014

Abstract

We present a novel three-band algorithm (PC\(_3\)) to retrieve phycocyanin (PC) pigment concentration in cyanobacteria laden inland waters. The water sample and remote sensing reflectance data used for PC\(_3\) calibration and validation were acquired from highly turbid productive catfish aquaculture ponds. Since the characteristic PC absorption feature at 620 nm is contaminated with residual chlorophyll-\(a\) (Chl-\(a\)) absorption, we propose a coefficient (\(\psi\)) for isolating the PC absorption component at 620 nm. Results show that inclusion of the model coefficient relating Chl-\(a\) absorption at 620 nm–665 nm enables PC\(_3\) to compensate for the confounding effect of Chl-\(a\) at the PC absorption band and considerably increases the accuracy of the PC prediction algorithm. In the current dataset, PC\(_3\) produced the lowest mean relative error of prediction among all PC algorithms considered in this research. Moreover, PC\(_3\) eliminates the nonlinear sensitivity issue of PC algorithms particularly at high PC range (>100 \(\mu g\) L\(^{-1}\)). Therefore, introduction of PC\(_3\) will have an immediate positive impact on studies monitoring inland and coastal cyanobacterial harmful algal blooms.

Keywords: cyanobacteria, phycocyanin, hyperspectral remote sensing, water quality, algal blooms

1. Introduction

Cyanobacterial harmful algal blooms (CyanoHABs) are becoming a frequent phenomenon in inland lakes and reservoirs, estuaries, and coastal waters all over the world due to various factors including nutrient loading from agricultural and urban runoff, increased frequency of drought and associated reduction in water quantity, and ecological regime shifts triggered by external events. Increase in nutrient loading has been linked with severe CyanoHABs in numerous water bodies around the world including Lake Taihu, China (Hu \textit{et al} 2010), Patos Lagoon, Brazil (Yunes \textit{et al} 1996), Gulf of Finland (Bianchi \textit{et al} 2000), Lake Erie, USA (Vincent \textit{et al} 2004), Lake Pontchartrain, USA (Mishra and Mishra 2010) eventually leading to anoxia and harmful impacts on aquatic life and human health. CyanoHABs have been broadly recognized as a human and animal health problem because of the wide varieties of toxins (cyanotoxins) associated with them. Some cyanotoxins are strong neurotoxins (anatoxin-\(a\) and saxitoxins), others are primarily hepatotoxins (microcystins, nodularin, and cylindrospermopsin). The hazard to human health caused by cyanotoxins can be estimated from toxicological knowledge in combination with information on their occurrence (World Health Organization 1999). Quantitative surveys on CyanoHABs occurrence are lacking, and therefore, incidence of cyanotoxin exposure through drinking water or during recreational water activities is largely unknown. A recent study reports that climate change may even worsen this environmental problem as rise in global temperature could affect the physiological and molecular changes in...
cyanobacteria and consequently increasing hepatotoxin production (El-Shehawy et al. 2012). Therefore, effective and accurate monitoring solutions should be developed for early warning systems or analyzing the triggering mechanism of a Cyanobacteria event.

For remote sensing applications, researchers have extensively used the accessory photopigment, phycocyanin (PC), as the marker pigment for estimating the presence of cyanobacteria (Dekker 1993, Simis et al. 2005, Mishra et al. 2009, 2015). PC absorption feature at ~620 nm in the remote sensing reflectance, \(R_{\alpha}(\lambda)\), spectrum has been extensively used to develop empirical and semi/semi-analytical algorithms (QAA) to detect PC concentration in inland water bodies (Dekker 1993, Simis et al. 2005, Hunter et al. 2010, Mishra et al. 2013, 2014). Over the years, five broad types of algorithms have been proposed to quantify PC from remote sensing reflectance data based on its absorption feature (a trough in reflectance) at 620 nm. They include (1) reflectance band ratio empirical algorithms (Schalles and Yacobi 2000, Vincent et al. 2004, Mishra et al. 2009); (2) a semi-analytical baseline subtraction algorithm (Dekker 1993); (3) a nested semi-analytical band ratio algorithm (Simis et al. 2005); (4) a three band semi-analytical algorithm (Hunter et al. 2010); and (5) a QAA (Mishra et al. 2013, 2014). Existing algorithms (Dekker 1993, Simis et al. 2005, 2007, Mishra et al. 2013) have reported the effect of accessory photopigments on the PC prediction accuracy. Dekker (1993) documented that other than Chl-a, accessory photopigments such as Chl-b, Chl-c, and allophycocyanin-b also absorb light in the photosynthetically active region of the electromagnetic spectrum with the \textit{in vitro} absorption peaks at 643–645 nm, 628 nm, and 618 nm respectively which overlap the PC absorption peak. Therefore, presence of accessory pigment would lead to overestimation of PC concentrations. Similarly, Simis et al. (2007) reported that accessory pigments such as Chl-b and fucoxanthin could significantly increase absorption at 620 nm consequently overestimating PC estimations especially at low PC concentrations (PC <50 \(\mu\)g L\(^{-1}\)). To address this issue, semi/semi-QAAs have focused on isolating PC absorption components by resolving absorption and particulate back-scattering from non-PC optically active constituents in the water (Dekker 1993, Simis et al. 2005, Mishra et al. 2013, 2014).

Our previously proposed QAA (Mishra et al. 2013) showed an improved performance compared to the nested semi-analytical algorithm (Simis et al. 2005). QAA produced a median relative error of 22.08%, whereas, the nested semi-analytical algorithm produced a model error of 45.96% in the same dataset (Mishra et al. 2013). Most of the model uncertainty in case of QAA was believed to be resulting from the inversion step where the \(R_{\alpha}(\lambda)\) signal is inverted to total absorbance signal using bio-optical models and radiative transfer theory. Because of the physical nature of QAA algorithm, errors associated with the bio-optical inversion can be attributed to the total absorption and particulate back-scattering coefficient modeling error at the reference wavelength, \(\alpha_{\lambda}(\lambda_{0})\) (see Lee et al. 2002). Through a systematic error analysis, (Lee et al. 2010) reported that, for coastal waters, errors from the modeled total absorption at the reference wavelength will have larger effect, as compared to particle-backscattering coefficient, \(b_{\phi}(\lambda)\), modeling errors on the total absorption at any other wavelengths, \(\alpha_{\lambda}(\lambda)\). It was also found that for strongly absorbing waters, when \(a_{\phi}(440)\) approaches 0.5 m\(^{-1}\), uncertainty in retrieving \(\alpha_{\phi}(440)\) goes up to \(\pm 37\%\) (Lee et al. 2010). In case of turbid productive waters similar to our current study site, bio-optical inversion accuracy of total absorption coefficient decreases and eventually reduces the PC estimation accuracy. There could be even higher uncertainties in PC pigment absorption, \(\alpha_{\phi}(\lambda)\), and absorption by colored detrital matter, \(\alpha_{C} DM(\lambda)\), estimations because of the physical nature of the decomposition of \(\alpha_{\phi}(\lambda)\) signal into their individual components (refer to Mishra et al. 2013 for a detailed description of the algorithm and modeling uncertainty). Therefore, a simplified yet accurate approach is still needed where the model uncertainty from bio-optical inversion could be avoided while addressing the noise from the Chl-a absorption at the PC absorption maximum (620 nm).

The proposed algorithm is based on a three-band algorithm conceptualized by Gitelson et al. (2003) to estimate Chl-a concentration in higher plant leaves. The algorithm was designed to isolate the absorption coefficient of pigment of interest i.e., Chl-a by using reflectance measurements at three specific bands as below.

\[
\text{Pigment} \propto \left( R_{\alpha}^{-1}(\lambda_1) - R_{\alpha}^{-1}(\lambda_2) \right) R_{\alpha}(\lambda_3),
\]

where, \(R_{\alpha}^{-1}(\lambda_1)\) is the inverse reflectance at a wavelength region that is maximally sensitive to the absorption by the pigment of interest (PC); \(R_{\alpha}^{-1}(\lambda_2)\) is the inverse reflectance measured at a wavelength that is minimally sensitive to the pigment of interest; and \(R_{\alpha}(\lambda_3)\) is the reflectance measured at a wavelength region where pigment absorption is negligible, and is used to compensate for the residual backscattering of the medium. Thus, subtraction of \(R_{\alpha}^{-1}(\lambda_2)\) from \(R_{\alpha}^{-1}(\lambda_1)\) isolates the signal from the pigment of interest. Similar approaches have been successfully used to estimate phytoplanktonic Chl-a concentration in natural waters (Dall’Olmo et al. 2003, Gilerson et al. 2010, Moses et al. 2012). Hunter et al. (2010) developed a three-band algorithm to quantify PC concentration by setting the \(\lambda_1, \lambda_2,\) and \(\lambda_3\) at 620 nm, 600 nm, and 725 nm. However, during cyanobacterial bloom conditions, PC absorption \(\alpha_{PC}\) at 600 nm and 620 nm are often very close (figure 1(a)). For example, in our dataset, \(\alpha_{PC}\) (620)/\(\alpha_{PC}\) (600) varied between 0.94 and 1.12 with an average value of 1.02. Therefore, it is not advisable to use 600 nm as \(\lambda_2\) because \(\alpha_{PC}(600)\) is almost equally sensitive to PC concentration as \(\alpha_{PC}(620)\). Because of the fact that Chl-a has significant absorption in the spectral region where PC has its absorption maximum (620 nm), developing a three-band algorithm for PC quantification is not straightforward. It is evident from existing literature that the removal of Chl-a contribution at 620 nm would be critical for the successful isolation of PC contribution (Simis et al. 2005, Mishra et al. 2009, 2013, Ogashawara et al. 2013).

In this research, we developed a conceptual three-band PC algorithm (\(PC_{\lambda}\)) using field radiometric and pigment
dataset collected from highly turbid and productive aquaculture ponds dominated by cyanobacteria. We proposed a coefficient ($\psi$) in order to remove the confounding effects of Chl-a absorption and to isolate the PC absorption component at 620 nm in order to enhance PC predictive accuracy. The specific objective of this research was to calibrate and validate the conceptual algorithm using \textit{in situ} dataset with very high Chl-a and PC content to demonstrate the sensitivity of the model at extremely high PC concentrations level.

2. Data and methods

\textit{In situ} data comprising of water samples for pigment extraction and $R_{\text{rs}}(\lambda)$ were collected from 15 aquaculture ponds at the Delta Research Extension Center located near Stoneville, MS, USA during 13–16 July 2010 and 28–29 April 2011.

2.1. Pigment measurements

Water samples for Chl-a and Chl-b analysis were simultaneously collected in 1L Niskin bottles and immediately filtered onto GF/F filters (Whatman, 0.7 $\mu$m pore size) under low vacuum (<5 inch of Mercury). Samples were extracted in triplicates using 90% acetone and concentrations were measured using high performance liquid chromatography following the Environmental Protection Agency method 447 (Arar 1997). Water samples for PC analysis were filtered immediately after collection through a 0.2 $\mu$m nucleopore membrane filters (Milipore) under low vacuum. PC pigments were extracted in 50 mM phosphate buffer and measured using spectrophotometric method. A detailed description of the pigment extraction method has been published in Mishra \textit{et al} (2013).

2.2. Spectral measurements

\textit{In situ} radiometric data were collected within a three-hour window centering the local solar noon. Sky condition was clear during the field sampling days. A dual sensor system with two inter-calibrated Ocean Optics (Ocean Optics, Dunedin, FL, USA) spectroradiometers were used to collect $R_{\text{rs}}$ data in the range of 400–900 nm with a sampling interval of 0.3 nm as given in Mishra \textit{et al} (2013). Radiometer 1, equipped with a 25° field-of-view optical fiber measured the upwelling radiance just below the air water interface; whereas, radiometer 2, equipped with an optical fiber and cosine diffuser (yielding a hemispherical field of view) acquired above surface downwelling irradiance. To match their transfer functions, inter-calibration of the radiometers was accomplished by measuring the upwelling radiance of a white Spectralon reflectance standard (Labsphere, North Sutton, NH) simultaneously with incident irradiance. The two radiometers were inter-calibrated immediately before and after measurements at each field site. After the data acquisition, $R_{\text{rs}}$ was calculated as below:

$$R_{\text{rs}}(\lambda) = \frac{t}{n^2} \frac{L_0(0^-, \lambda)}{E_0(0^+, \lambda)},$$  \hspace{1cm} (2)

where, $t$ is the transmittance at the air–water interface (0.98); $n$ is the refractive index of water (1.34); $L_0(\lambda)$ and $E_0(\lambda)$ are the upwelling radiance measured below the air–water interface and downwelling radiance measured above-surface. For each station, six consecutive scans were recorded and further averaged to calculate a representative $R_{\text{rs}}(\lambda)$ spectrum (figure 1(b)).

2.3. Conceptual algorithm development

Key pigment spectral features generally observed in reflectance data from a cyanobacteria dominated turbid productive
system have been used in the algorithm development and are presented in figure 1(b). Reflectance trough near 620 nm and 665 nm appears due to strong absorption by PC and Chl-a pigments respectively (Schalles and Yacobi 2000, Dall’Olmo et al 2003, Mishra et al 2009). The spectral peak around 650 nm appears because of prominent absorption by chlorophylls and PC on both sides of the peak and PC fluorescence emission at 640 nm (Schalles and Yacobi 2000, Mishra et al 2009). Generally, the reflectance trough (i.e., the absorption feature) observed at 620 nm is assumed to be entirely due to PC absorption in empirical band ratio models (Schalles and Yacobi 2000, Ogashawara et al 2013). However, we have found that the PC absorption feature at 620 nm is contaminated by Chl-a absorption and that interference is quite significant (Simis et al 2005, 2007, Mishra et al 2009). It is quite evident in figure 1(a) that the difference in phytoplankton absorption (≈Chl-a absorption) and PC absorption is quite significant at 620 nm highlighting the residual Chl-a absorption at the PC absorbing wavelength region. Therefore, elimination of residual Chl-a absorption signal is required in order to estimate PC concentration accurately. In this research, we systematically develop the PC3 algorithm in two steps:

2.3.1. The optimal PC3 model. For the optimal PC3 algorithm development, we used the three-band algorithm framework proposed by Gitelson et al (2003) and adapted by Hunter et al (2010). In order to design the optimal PC3 algorithm, we carried out best band analysis to select \( \lambda_1 \), \( \lambda_2 \), and \( \lambda_3 \) at three successive steps. At step 1, we took the initial value of \( \lambda_1 \) and \( \lambda_3 \) as 620 nm and 778 nm. Initially, we picked 620 nm for \( \lambda_1 \) as it is considered as the PC absorption maxima and hence highly sensitive to PC concentration. Similarly, based on the three-band architecture of the PC3, \( \lambda_3 \) is required to be at a wavelength region where pigment absorption is negligible so that it can be used to compensate for the residual backscattering of the medium. We selected 778 nm as \( \lambda_3 \) because it matched ESA’s medium resolution imaging spectrometer (MERIS) band architecture (MERIS band 12). Although out of commission, MERIS has a massive archive of data for spatio-temporal studies of cyanobacterial blooms. Since the ultimate goal was to make the algorithm readily applicable to satellite data, whenever possible even during optimization, initial bands selection for PC3 was performed keeping MERIS or MERIS-like sensors in consideration. The three-band algorithm was calculated iteratively by systematically selecting the value of \( \lambda_2 \) from 400–800 nm at 1 nm interval. Standard error was calculated for each iteration by least-square linear regression between the model, \( \left( R_{rs}^{-1}(620) - R_{rs}^{-1}(\lambda_2) \right) \ast R_{rs}(778) \) and PC concentration. Optimized \( \lambda_2 \) was selected where the model produced minimum standard error of the estimates. At step 2, initial value of \( \lambda_1 \) (620 nm) and optimal value of \( \lambda_2 \) from the step 1 was used to estimate the optimal \( \lambda_3 \). Similarly at step 3, two optimal values of \( \lambda_2 \) and \( \lambda_3 \) were used to estimate the optimal value of \( \lambda_1 \). From three steps, optimal band selection procedure returned 629 nm, 659 nm, and 724 nm as \( \lambda_1 \), \( \lambda_2 \), and \( \lambda_3 \). Figure 2 shows the standards errors normalized to their range for \( \lambda_1 \), \( \lambda_2 \), and \( \lambda_3 \), PC3 in the optimal form can be written as:

\[
PC \propto \left( R_{rs}^{-1}(629) - R_{rs}^{-1}(659) \right) \ast R_{rs}(724).
\]

2.3.2. PC3 for MERIS-like sensors. Considering the direct applicability of the conceptual PC3 algorithm to satellite sensors, we also selected \( \lambda_1 \), \( \lambda_2 \), and \( \lambda_3 \) at 620 nm, 665 nm, and 778 nm matching the MERIS band configuration. We used 620 nm spectral region as \( \lambda_1 \) because it is highly sensitive to PC absorption. Unlike Hunter et al (Hunter et al 2010), we used 665 nm as \( \lambda_2 \) spectral region that is minimally sensitive to PC absorption. Instead of 676 nm (the true Chl-a absorption maxima), we used 665 nm band (MERIS band 7) primarily because of two reasons, (a) 676 nm band could be contaminated by Chl-a fluorescence in hyper-eutrophic systems, and (b) 665 nm matches the MERIS-like satellite sensor’s band architecture; so that the developed algorithm could be easily applied to satellite data. In addition, MERIS does not have a 676 nm band; therefore, we had two choices to capture the Chl-a absorption for PC3, 655 nm and 681 nm band. 681 nm band was discarded mostly because of the fluorescence related interference discussed above. For \( \lambda_3 \), we used the 778 spectral region, where absorption by any photo pigment is negligible. PC3 for
MERIS-like sensors can be written as:

\[ \text{PC} \propto \left( R_{r\alpha}^{-1}(620) - R_{r\alpha}^{-1}(665) \right) \cdot R_{r\alpha}(778). \] (4)

However, it should be noted that 665 nm spectral region is highly sensitive to Chl-\(\alpha\) absorption, and therefore, subtraction of \(R_{r\alpha}(665)\) from \(R_{r\alpha}(620)\) to isolate PC absorption will underestimate the PC component. In order to address this issue, we introduced a coefficient, \(\psi\), which estimates the equivalent of Chl-\(\alpha\) contribution at 620 nm from Chl-\(\alpha\) absorption at 665 nm \((a_{\text{chl}}(665))\) information. \(\psi\) was estimated as a ratio of Chl-\(\alpha\) pigment absorption at 665–620 nm as:

\[ \psi = \frac{a_{\text{chl}}(665)}{a_{\text{chl}}(620)}. \] (5)

The model coefficient \(\psi\) relates Chl-\(\alpha\) absorption at 620 nm and 665 nm. It was shown that \(\psi\) is highly correlated with \textit{in vitro} Chl-\(\alpha\) concentration \((\psi = 2.268 \times (\text{Chl-}\alpha)^{0.006}, \ R^2 = 0.96, \ p < 0.0001)\) (Mishra et al 2013). Acknowledging the challenge of \textit{in vivo} Chl-\(\alpha\) absorption measurement, \(a_{\text{chl}}(\lambda)\), (Mishra et al 2013) retrieved \(\psi_{1}\), (in this research \(\psi\)), values by optimization. For each sample, \(\psi\) was solved using \textit{in vitro} measured PC absorption at 620 nm and 665 nm, \([a_{\text{PC}}(620)\) and \(a_{\text{PC}}(665)]\), \(a_{\text{PC}}(665): a_{\text{PC}}(620)\) ratio, \(\psi_2\), and filter-pad measured phytoplankton pigment absorption, \(a_{\phi}(620)\) and \(a_{\phi}(665)\) as below (Mishra et al 2013).

\[ a_{\text{PC}}(620) = \frac{\psi \cdot a_{\phi}(620) - a_{\phi}(665)}{\psi - \psi_2}. \] (6)

where, \(\psi\) is the Chl-\(\alpha\) pigment absorption ratio from 665 nm to 620 nm, \(a_{\text{chl}}(665)/a_{\text{chl}}(620)\), and \(\psi_2\) is the PC pigment absorption ratio from 665 nm to 620 nm, \(a_{\text{PC}}(665)/a_{\text{PC}}(620)\).

The detailed method explaining the pigment absorption and retrieval of \(\psi\) is available in (Mishra et al 2013).

For remote sensing applications, \(\psi\) was empirically modeled from a reflectance band ratio, \(R_{r\alpha}(560)/R_{r\alpha}(665)\) as provided in Mishra et al (2013) (figure 3). The specific band ratio was selected because \(\psi\) is highly correlated with Chl-\(\alpha\) concentration, which is well correlated with the band ratio \(R_{r\alpha}(560)/R_{r\alpha}(665)\). It should be noted that \(R_{r\alpha}(560)\) is sensitive to the presence of cyanophycoerythrin bearing algae in the water body. However in this study, cyanophycoerythrin will have negligible effect on modeled \(\psi\) due to its minimal concentration as compared to abundant presence of chloropysllous pigments and PC. In other water bodies, if significant cyanophycoerythrin is present than a different reference wavelength needs to be used for modeling \(\psi\).

Finally, we modify PC3 algorithm in equation (4) by incorporating \(\psi\) to remove the confounding effect of Chl-\(\alpha\) at 620 nm and to isolate the PC contribution. PC3 in the final form can be written as:

\[ \text{PC} \propto \left( R_{r\alpha}^{-1}(620) - (\psi R_{r\alpha}(665))^{-1} \right) \cdot R_{r\alpha}(778), \] (7)

where, \(R_{r\alpha}^{-1}(620)\) represents the proxy of absorption component by PC and Chl-\(\alpha\) at 620 nm, and \((\psi R_{r\alpha}(665))^{-1}\) represents the proxy of absorption by Chl-\(\alpha\) at 620 nm making the right side of the equation (5) proportional to PC absorption at 620 nm. PC3 assumes that \(b_2\) within the spectral range of 620–778 nm is spectrally flat and difference in absorption by colored dissolved organic matter (CDOM) and detrital matter at 620 and 665 nm is negligible. Note that the inclusion of \(\psi\) in the model is not a complete analytical treatment for retrieval of \(a_{\text{PC}}(620)\). However, assuming \(R_{665}\) to be a proxy of \([a_{\text{chl}}(665)/b_2]\), we consider that inclusion of \(\psi\) would enable PC3 to compensate for Chl-\(\alpha\) contribution at 620 nm.

3. Results and discussion

Summary of the measured pigment concentrations is presented in table 1. PC concentration varied widely from 68.13 to 3032.47 \(\mu g\ \text{L}^{-1}\) with an average of 418.76 \(\mu g\ \text{L}^{-1}\). Chl-\(\alpha\) concentration varied within 59.4–1376.6 \(\mu g\ \text{L}^{-1}\) with an average of 302.06 \(\mu g\ \text{L}^{-1}\). PC:Chl-\(\alpha\) ratio varied within 0.3–3.29 (mean = 1.27), indicating cyanobacterial dominance in the phytoplankton community structure (table 1). These values are in similar range compared to the previously reported values from Spain and Netherlands (Simis et al 2005, Ruiz-Verdu et al 2008). Strong dependency was found between Chl-\(\alpha\) and PC concentration in the entire dataset \((r = 0.91)\).
3.1. The optimal PC3 model

To explore the performance of PC3 using optimal bands, those were selected from the best band analysis, we used all available data points and analyzed it successively in steps. At step 1, \( R_{39}^{-1}(629) \) was used as an independent explanatory variable to model PC concentration. Results from the linear least-square regression analysis showed that the model explained 86% of the variance in the data \((p<0.0001)\). At step 2, subtraction of \( R_{39}^{-1}(659) \) from \( R_{39}^{-1}(629) \) was believed to compensate for the differences in absorption and backscattering by non-pigment water constituents (tripton and CDOM). Regression analysis showed that \( \left[ R_{39}^{-1}(629) - R_{39}^{-1}(659) \right] \) explained 89% variance in the PC concentration \((p<0.0001)\). Finally at step 3, multiplication of \( \left[ R_{39}^{-1}(629) - R_{39}^{-1}(659) \right] \) by \( R_{39}(724) \) compensated for the backscattering signals and significantly improved the linear relationship between PC and PC3. The optimal PC3 model in the form of \( PC = 476.33*PC3 + 96.16 \) showed strong predictive ability \((R^2=0.95, \text{standard error of estimate} = 150.38 \mu g L^{-1})\).

3.2. PC3 model for MERIS-like sensors

As discussed in the previous section, PC3 was also designed using the channels available in MERIS-like sensors. This analysis was performed in order to make the model applicable to MERIS data or any other sensor with a band at 620 nm. PC3 in the form of \( \left[ R_{39}^{-1}(620) - R_{39}^{-1}(665) \right]*R_{39}(778) \) explained 88% of variance in the PC concentration \((p<0.0001)\). Results from the regression analysis showed that the model was not sensitive to PC concentration below 300 \( \mu g L^{-1} \) (figure 4(a)). Low sensitivity in this range can be attributed to the low PC:Chl-a ratio \( (\text{mean} = 1.97) \) in samples with PC concentration greater than 300 \( \mu g L^{-1} \). Lower PC:Chl-a ratio indicates that the samples in the lower range are dominated by Chl-a.

We addressed this issue in PC3 by using model coefficient \( \psi \) (equation (5)). After incorporating \( \psi \) in the PC3, relationship between PC concentration and PC3 substantially increased \((R^2=0.98, \ p<0.0001)\) (figure 4(b)). The model produced a standard error of 83.7 \( \mu g L^{-1} \), which is almost one third of the error from the model before the incorporation of \( \psi \). Results from the regression analysis produced evidence that the model coefficient \( \psi \) is able to address the issue of interference of Chl-a absorption at the PC absorption band (620 nm) and enables accurate isolation of PC component.

In order to calibrate and validate the PC3 model, we randomly divided the dataset into two groups. 70% of dataset was used for model calibration and the rest was used for model validation. In the calibration dataset, the model in the form of PC \( (\mu g L^{-1}) = 480.92*PC3 + 123.23 \) accounted for 98% of the variance in PC concentration \((p<0.0001)\) (figure 5(a)). For PC3 accuracy assessment, predicted PC values were compared with analytically measured PC concentrations. Results showed that the model produced a mean relative error, \[ \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\text{predicted PC} - \text{measured PC}}{\text{measured PC}} \right) \times 100 \] of 30.71%. The slope of the best-fit line between predicted and measured PC was 0.97 \((R^2=0.99)\) (figure 5(b)). Overall, after correcting for the confounding effects of Chl-a at 620 nm PC3 performed significantly better than all PC models considered in this research (table 2). Mean relative error of PC3 was \( \sim 12\% \) lower than Hunter et al (2010) and \( \sim 4\% \) lower than the quasi-analytical PC algorithm by Mishra et al (2013). Similarly, improvement in accuracy was \( \sim 52\% \) higher as compared to the semi-analytical algorithm (Simis et al 2005) in the current dataset (table 2). We also assessed the performance of PC3 at the lower range of PC concentration \((<300 \mu g L^{-1})\) through a systematic validation analysis and comparison with existing PC algorithms considered in this research. Results showed that PC3 performed most accurately compared to other existing algorithms when validated at the low range (below 300 \( \mu g L^{-1} \)). Overall, PC3, Mishra et al (2013), and Hunter et al (2010) showed high prediction accuracy even in the lower range by producing 38.18%, 40.26%, and 53.67% of MRE respectively (table 2).

It is apparent from the validation analysis that accurate PC prediction for waters with low to moderate PC concentrations.
concentration is still an issue. We believe that the mediocre performance of PC3 and other algorithms at low to moderate range is mainly due to interference from other optically active constituents in the water primarily Chl-a. Since relative PC absorption is low for water bodies with low to moderate PC concentrations, Chl-a interference plays a dominant role and is often difficult to isolate from PC absorption at 620 nm. This is due to the fact that at low PC:Chl-a scenarios, confounding signal from Chl-a absorption at 620 overpowers the weak PC signal thereby decreasing the signal to noise ratio and increasing the uncertainty of PC absorption component retrieval. Although PC3 performs relatively well as compared to other algorithms, the model error is considered to be high (38.18%) for the low to medium PC range. This is probably because the accuracy of the empirical model we used to derive $\psi$ (figure 3) decreases for low to moderate PC concentrations. As $\psi$ is the critical parameter that addresses Chl-a interference, any error with $\psi$ retrieval would be carried over to the final PC prediction stage. In our future research, we will parameterize $\psi$ for the low PC concentration scenarios in order to improve the prediction accuracy of PC3 for low to moderate PC concentration. We believe that inclusion of data points in the 0–70 $\mu$g L$^{-1}$ PC range would improve the model prediction accuracy making it widely applicable.

Table 2. Model validation results showing mean relative error, MRE (%). Slope and $R^2$ from least-square regression between measured and modeled PC concentrations were also provided. Range of PC concentrations originally used for calibration and validation of PC algorithms considered in this research are provided as subscripts in the first column. Numbers in side square brackets are PC ranges ($\mu$g L$^{-1}$).

| PC models                  | Full-range model | Model for PC < 300 ($\mu$g L$^{-1}$) |
|----------------------------|------------------|--------------------------------------|
|                            | MRE (%)  | $R^2$  | Slope | Intercept | MRE (%)  | $R^2$  | Slope | Intercept |
| $R_{rs}(620)/R_{rs}(710)$  | [68.13–3032.47] | 123.10 | 1.00  | 1.31      | 191.94   | 160.44 | 0.91  | 2.01      | -271.29  |
| $R_{rs}(620)/R_{rs}(725)$  | [68.13–3032.47] | 127.90 | 0.99  | 1.25      | -171.76  | 163.23 | 0.60  | 0.97      | -165.12  |
| Dekker (1993)[17–110 ($\mu$g L$^{-1}$)] | 204.05 | 0.64  | 0.45  | 309.65    | 260.71  | 0.01  | 1.05      | 215.20   |
| Schalles and Yacobi (2000)[0–530] | 111.44 | 0.89  | 0.73  | 97.03     | 141.60  | 0.45  | 3.16      | -174.24  |
| Simis et al (2005)$^a$ [0.8–79.8] | 83.38  | 0.18  | -0.04 | 102.90    | 79.10   | 0.47  | -1.04     | 221.32   |
| Hunter et al (2010)[0–93.7] | 42.49  | 0.99  | 1.06  | 9.27      | 53.67   | 0.00  | 0.04      | 138.80   |
| Mishra et al (2013)$^b$ [68.13–3032.47] | 37.88  | 0.98  | 1.56  | -78.48    | 40.26   | 0.48  | 1.14      | 6.99     |
| Ogashawara et al (2013)$^b$ [4.92–831.35] | 71.58  | 0.98  | 1.18  | -89.39    | 91.12   | 0.24  | 1.38      | -101.01  |
| This study[68.13–3032.47] | 30.71  | 0.99  | 0.97  | 29.08     | 38.18   | 0.65  | 1.01      | 37.28    |

$^a$ Model coefficients were locally parameterized and measured $a^*_{PC(620)}$ was used in the validation analysis (Mishra et al 2013).

$^b$ Based on the error analysis reported by (Mishra et al 2013) where the same dataset was used for model calibration and validation.

Reflectance band ratio models often show lower sensitivity to pigment of interest because of saturation in absorption at high pigment concentration. Figure 6 explains this phenomena and shows the nonlinear sensitivity of a band ratio, $R_{rs}(620)/R_{rs}(708)$, to PC concentration. The nonlinearity is most likely due to the saturation in pigment absorption at high concentration range and it is an inherent difficulty...
associated with band-ratio algorithms. To its advantage, PC₃ showed a linear trend within the entire PC range even at very high PC concentration unlike its band ratio counterparts (figure 4(a)), even when compared to PC₃ before the introduction of ψ (equation (4)). So far semi-analytical and QAAs (Simis et al 2005, Mishra et al 2013) needed a reflectance band at 709 nm, which is only available in limited sensors, to estimate PC. However, in this research we have proposed an innovative method to address the overlapping Chl-α absorption issue at 620 nm by using reflectance data measured at 620 nm, 665 nm, and at a longer NIR band, 778 nm. Therefore, application of PC₃ can be extended to any sensor even without the 709 nm band.

4. Conclusion

In this research, we have developed a three-band PC algorithm, PC₃, to accurately retrieve PC concentration in inland water bodies. Using best band selection analysis, we selected three optimal bands centered at 629 nm, 659 nm, and 724 nm for λ₁, λ₂, and λ₃ respectively. The optimal PC₃ model showed strong PC predictive ability in the current dataset (STE = 150.38 μg L⁻¹). However, for practical applicability of the algorithm, PC₃ was also designed using the channels available in MERIS-like sensors i.e., 620 nm, 665 nm and 778 nm as λ₁, λ₂, and λ₃ respectively. PC₃ for MERIS-like sensors produced relatively lower prediction accuracy (STE = 232.41 μg L⁻¹) than the optimal PC₃ algorithm. We concluded that the high standard error observed in both variants of the PC₃ algorithm was mainly due to the interference of Chl-α absorption at PC absorption maximum (620 nm). To address the lack of sensitivity of PC₃ at low to moderate PC range (<300 μg L⁻¹), we introduced a model coefficient, ψ, to remove the effects of Chl-α absorption at 620 nm. In the final form, PC₃ for MERIS-like sensors produced highest PC prediction accuracy (MRE = 30.71%) compared to all algorithms considered in this research. PC₃ model prediction accuracy somewhat decreased when validated for the low to moderate PC range, although it produced the minimum MRE among all algorithms.

It is evident from the results that: (a) PC₃ can be applied to retrieve cyanobacterial PC concentration in inland waters with considerably higher accuracy than existing algorithms; (b) inclusion of ψ in the model compensates for the confounding effect of Chl-α absorption at the PC absorption band, i.e. 620 nm; and (c) unlike reflectance band ratios, PC₃ addresses the nonlinearity issue at higher PC range. Development of remote sensing algorithms for quantifying cyanobacterial PC is a challenging task primarily because specific absorption coefficient of PC and Chl-α varies significantly with change in phytoplankton species composition. In addition, Chl-α absorption at the PC absorbance maxima causes significant interference that adds to the model complexities (Dekker 1993, Simis et al 2005, Mishra et al 2013). Therefore, modeling errors from all PC algorithms considered in this research were quite high (∼30–204%). Results from this analysis supports the previous findings that PC algorithms perform poorly when cyanobacteria population in the phytoplankton community is relatively smaller as documented in Hunter et al (2010), Ruiz-Verdu et al (2008), and Mishra et al (2013). Despite the modeling complexities, the newly developed PC₃ algorithm for MERIS-like sensors performed best among all PC algorithms considered in this research. It should also be noted that PC₃ does not require the 709 nm band, which is used in the semi-analytical algorithm (Simis et al 2005) and QAA (Mishra et al 2013) to retrieve PC absorption coefficients. As PC₃ does not require the 709 nm band, we anticipate that it can be widely applicable to other sensors, i.e., sensors without the 709 nm band and hence another advantage over many existing PC algorithms.

In this research, model calibration and validation analysis was performed on a small dataset with extreme values of PC concentration. It should be noted that the model parameter, ψ, plays a critical role in the model configuration and separation of Chl-α interference. Therefore, in future the empirical model that retrieve ψ from the reflectance band ratio should be considered for recalibration specifically for low PC concentrations for robust performance of PC₃. It should also be noted that PC₃ has been calibrated and validated using a dataset with PC concentration varying from 68.13 to 3032.47 μg L⁻¹. Hence the algorithm should be applied as is to datasets within this PC range. Our future research effort will use a comprehensive dataset including data particularly in the lower range (0–70 μg L⁻¹) of PC concentration to recalibrate and validate the algorithm for wider applicability. A large dataset with a wide range of PC concentration will also help to validate PC₃, and systematically compare its performance with other existing PC algorithms. Findings from this research can be leveraged to any current and future satellite or airborne sensor with the specific bands to monitor cyanobacterial algal blooms. Therefore, introduction of PC₃ will have an immediate positive impact on studies monitoring inland and coastal CyanohABs.

Acknowledgement

The authors are grateful to Chris Downs and Margaret E Dennis for their help with field data collection efforts and microscopic identification of algal community. Authors also thank two anonymous reviewers for their constructive comments and suggestions that significantly improved the presentation of the manuscript. This work was partially funded by the University of Georgia Provost’s Summer grant.

References

Arar E J 1997 In vitro determination of chlorophylls a, b, c₁ + c₂ and pheopigments in marine and freshwater algae by visible spectrophotometry USEPA Method 446.0 pp 1–26
Bianchi T S, Engelhaupt E, Westman P, Andren T, Rolff C and Elmgren R 2000 Cyanobacterial blooms in the Baltic sea: natural or human- induced? Limnol. Oceanogr. 453 716–26
Dall’Olmo G, Gitelson A and Rundquist D C 2003 Towards a unified approach for remote estimation of chlorophyll-a in both terrestrial vegetation and turbid productive waters Geophys. Res. Lett. 30 1938

Dekker A G 1993 Detection of the optical water quality parameters for eutrophic waters by high resolution remote sensing PhD Dissertation (Amsterdam) Free University

El-Shehawy R, Gorokhova E, Fernández-Piñas F and Del Campo F F 2012 Global warming and hepatotoxin production by cyanobacteria: what can we learn from experiments? Water Res. 46 1420–9

Hunter P D, Tyler A N, Carvalho L, Codd G A and Maberly S C 2010 Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes Remote Sens. Environ. 114 2705–18

Gilerson A A, Gitelson A A, Zhou J, Gurlin D, Moses W J, Ioannou I and Ahmed S A 2010 Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands Opt. Express 18 24109–25

Gitelson A A, Gritz Y and Merzlyak M N 2003 Relationship between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves J. Plant Physiol. 160 271–82

Hu C, Lee Z P, Ma R, Yu K and Li D 2010 Moderate resolution imaging spectroradiometer (MODIS) observation of cyanobacteria blooms in Taihu Lake, China J. Geophys. Res. 115 C04002

Lee Z P, Carder K L and Arnone R A 2002 Deriving inherent optical properties from water color: a multi-band quasi-analytical algorithm for optically deep waters Appl. Opt. 41 5755–72

Lee Z P, Arnone R, Hu C, Werdel J and Labac B 2010 Uncertainties of optical parameters and their propagation in an analytical ocean color inversion algorithm Appl. Opt. 49 369–81

Mishra S, Mishra D R and Schluchter W M 2009 A novel algorithm for predicting PC concentrations in cyanobacteria: a proximal hyperspectral remote sensing approach Remote Sens. 1 758–75

Mishra D R and Mishra S 2010 Plume and bloom: effect of the Mississippi River diversion on the water quality of Lake Pontchartrain Geocarto Int. 25 555–68

Mishra S, Mishra D R, Lee Z P and Tucker C G 2013 Quantifying cyanobacteria phycocyanin concentration in turbid productive waters: a quasi-analytical approach Remote Sens. Environ. 133 141–51

Mishra D R, Schaeffer B A and Keith D 2014 Performance evaluation of normalized difference chlorophyll index in northern Gulf of Mexico estuaries using the hyperspectral imager for the Coastal Ocean Gisci. Remote Sens. 51 175–98

Moses W J, Gitelson A A, Perk R L, Gurlin D, Rundquist D C, Leavitt B C, Barrow T M and Brakhage P 2012 Estimation of chlorophyll-a concentration in turbid productive waters using airborne hyperspectral data Water Res. 46 993–1004

Ogashawara I, Mishra D R, Mishra S, Curtarelli M P and Stech J L 2013 A performance review of reflectance based algorithms for predicting phycocyanin concentrations in inland waters Remote Sens. 5 4774–98

Pope R and Fry E 1997 Absorption spectrum (380–700 nm) of pure waters: II. Integrating cavity measurements Appl. Opt. 36 8710–23

Ruiz-Verdu A, Simis S G H, de Hoyos C, Gons H J and Pena-Martinez R 2008 An evaluation of algorithms for the remote sensing of cyanobacterial biomass Remote Sens. Environ. 112 3996–4008

Schalles J F and Yacobi Y Z 2000 Remote detection and seasonal patterns of phycocyanin, carotenoid and chlorophyll pigments in eutrophic waters Arch. Hydrobiol. Spec. Issues Adv. Limnol. 55 153–68

Simis S, Ruiz-Verdu A, Domínguez-Gómez J, Peña-Martínez R, Peters S and Gons H 2007 Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass Remote Sens. Environ. 106 414–427

Simis S, Peters S and Gons H 2005 Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water Limnol. Oceanogr. 50 237–45

Smith R C and Baker K S 1981 Optical properties of the clearest natural waters (200–800 nm) Appl. Opt. 20 177–84

Yunes J S, Salomon P S, Mattheisen A, Beattie K A, Raggett S L and Codd G A 1996 Toxic blooms of cyanobacteria in the Patos Lagoon Estuary, Southern Brazil J. Aqua. Ecosyst. Health 5 223–9

Vincent R K, Qin X M, McKay R M L, Minzer J, Czajkowski K, Savino J and Bridgeman T 2004 Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie Remote Sens. Environ. 89 381–92

World Health Organization 1999 Toxic Cyanobacteria in Water: a Guide to their Public Health Consequences, Monitoring, and Management (London, UK: F & FN Spon)