Using field spectral measurements to estimate chlorophyll-a in waterlogged areas of Huainan, China

Wenming Pei, Suping Yao, Shaochun Dong, Joseph F. Knight, Chong Xu & Yongchun Chen

To cite this article: Wenming Pei, Suping Yao, Shaochun Dong, Joseph F. Knight, Chong Xu & Yongchun Chen (2015) Using field spectral measurements to estimate chlorophyll-a in waterlogged areas of Huainan, China, GIScience & Remote Sensing, 52:6, 660-679, DOI: 10.1080/15481603.2015.1082173

To link to this article: http://dx.doi.org/10.1080/15481603.2015.1082173

Published online: 23 Sep 2015.

Submit your article to this journal

Article views: 5

View related articles

View Crossmark data
Using field spectral measurements to estimate chlorophyll-a in waterlogged areas of Huainan, China

Wenming Pei, Suping Yao*, Shaochun Dong, Joseph F. Knight, Chong Xu and Yongchun Chen

MOE Key Laboratory of Surficial Geochemistry, School of Earth Sciences and Engineering, Nanjing University, 163 Xianlin Ave., Nanjing 210023, China; School of Earth Sciences and Engineering, Nanjing University, 163 Xianlin Ave., Nanjing 210023, China; Department of Forest Resources, University of Minnesota, 1530 Cleveland Ave. N, St. Paul 55108, MN, USA; National Engineering Laboratory of Coal Mine Ecological Environment Protection, Dongshan East Road, Huainan 232033, China

(Received 14 December 2014; accepted 09 August 2015)

Long-term, large-scale underground coal mining activities in Huainan, China, have damaged the stability of the overlying rocks, resulting in large waterlogged areas. The water quality in these areas is compromised by agricultural pollutants and discharge of untreated waste. In this study, chlorophyll-a (Chl-a) concentrations were estimated in two types of waterlogged area, semi-closed (the Panji area [Pj]) and closed (the Guqiao area [Gq]), using information from water samples and in situ hyperspectral data collected in May and November 2013. The absorption coefficients of phytoplankton ($a_{ph}(\lambda)$), colored dissolved organic matter ($a_{CDOM}(\lambda)$), and nonpigmented suspended matter ($a_d(\lambda)$) were determined. Nonpigmented components ($a_d(\lambda)$ and $a_{CDOM}(\lambda)$) had a significant influence on the waterlogged areas, more so in Pj than in Gq because of its connection to the Ni River.

The NIR-red, three-band, and normalized difference chlorophyll index (NDCI) models were calibrated and validated to estimate Chl-a concentrations in Pj and Gq. All three models yielded good estimations of the Chl-a concentrations in waterlogged areas. The three-band model was more accurate for the semi-closed waterlogged area, while the NDCI model was more accurate for the closed waterlogged area. Results indicate that field spectral measurements can be used to accurately estimate Chl-a concentrations in waterlogged areas.

Keywords: Huainan waterlogged area; chlorophyll-a (Chl-a); absorption coefficient; three-band model; normalized difference chlorophyll index

Introduction

Coal, as the main energy resource in China, has played an important role in China’s economic development (Hu, Xu, and Zhao 2012). About 95% of China’s coal comes from underground coal mining (Du et al. 2007). However, extensive long-term underground coal mining compromises the stability of the overlying rocks and causes surface distortion and land subsidence, resulting in the formation of large-scale seasonally or perennially waterlogged areas, especially in areas where groundwater occurs close to the surface (Du et al. 2007; Hu, Xu, and Zhao 2012; Dong et al. 2013). Various pollutants released from the submerged land and discharged from nearby industrial activities converge in the

*Corresponding author. Email: spyao@nju.edu.cn

© 2015 Taylor & Francis
middle of these waterlogged areas, thereby degrading their water quality (Hu, Xu, and Zhao 2012). Over the past 30 years, hundreds of these waterlogged areas have formed in East China, one of the most densely populated areas in China and also a major region for agricultural production (Hu and Wu 2013; Dong et al. 2014). Most of the waterlogged areas are adjacent to villages and have been farmed by the local residents. Therefore, the water quality of the waterlogged areas is closely linked to the sustainable development of the regional economy, the livelihoods of the local residents, and the health of the surrounding environment.

Huainan, a typical industrial mining city, is located in Anhui Province, East China, near the Huaihe River. The coal reserves in the Huainan coal mine account for over 30% of the total coal reserves in East China (Yan et al. 2004). The mining history of Huainan dates back more than 100 years. Due to the long-term exploitation, deep mining depth, and the shallow level of groundwater, the waterlogged areas now cover 40 km² and continue to grow. The water in these areas is used by local residents for irrigation, aquaculture, recreation, and transportation. Previous studies have shown that the waterlogged areas are subject to anthropogenic eutrophication and increased phytoplankton concentrations because of pollutants in runoff from agriculture fields and the discharge of untreated residential and industrial waste (He, Liu, and Gui 2005; Fan et al. 2012). Chlorophyll-a (Chl-a) is an important indicator of water quality, the concentrations of which can be used to evaluate the trophic state of turbid productive waters. Chl-a, found in all algae groups, is an important indicator of the bioproduction of water bodies (Baban 1996; Duan et al. 2008; Li et al. 2010). However, Chl-a monitoring in waterlogged areas has historically been dependent on traditional field sampling methods (Hui 2013), which are both time and cost intensive because of the need to transport and preserve water samples over long distances from the field site to the laboratory. Furthermore, the inappropriate preservation of water samples may introduce errors into the analytical results, thereby reducing the analytical accuracy. Therefore, a more convenient and effective method for estimating Chl-a in waterlogged areas is urgently needed to facilitate efficient water quality monitoring and to ensure the sustainable use of water resources.

Remote sensing techniques are an effective alternative for estimating Chl-a in turbid productive waters, and have the potential to provide all the information required for estimating Chl-a because of their synoptic coverage and the temporal consistency of the data (Navalgund, Jayaraman, and Roy 2007; Duan et al. 2012. D’Sa 2014; Kim et al. 2014). Remote sensing methods have been widely used in water quality monitoring and management to complement traditional field-based sampling methods (Ma and Dai 2005; Gurlin, Gitelson, and Moses 2011; Palmer, Kutser, and Hunter 2015). A number of models, based on remote sensing reflectance from both in situ measurements and images, have been developed to estimate Chl-a concentrations in coastal and inland waters. Although bio-optical models that provide reasonable estimations of Chl-a concentrations in specific waters have been developed (Dall’Olmo, Gitelson, and Rundquist 2003; Zimba and Gitelson 2006), it is difficult to accurately parameterize the specific inherent optical properties of different turbid productive waters (Mishra and Mishra 2012; Song et al. 2012). Therefore, empirical or semi-empirical models remain the most effective and popular method for retrieving Chl-a concentration data (Dekker 1993; Gitelson et al. 2008; Yang et al. 2011; Song et al. 2012). Most of these models have been developed to quantify Chl-a in turbid productive water using wavebands in the red and near-infrared (NIR) range (Dekker 1993; Kallio et al. 2001; Gons, Rijkeboer, and Ruddick 2002; Dall’Olmo and Gitelson 2005; Salas and Henebry 2009; Gilerson et al. 2010). The
NIR/red reflectance ratio is commonly used to estimate Chl-a concentrations in turbid productive waters and have yielded accurate results (Vincent et al. 2004; Gilerson et al. 2010; Yang et al. 2011). A three-band model has also been developed and validated for Chl-a estimation in turbid productive waters (Gitelson, Schalles, and Hladik 2007; Gitelson et al. 2008). This model requires reflectance in three spectral bands from the NIR and red spectral regions: $\lambda_1$, $\lambda_2$, and $\lambda_3$, and is represented by the formula 

$$\frac{R_{rs}(\lambda_1)}{C_1} - \frac{R_{rs}(\lambda_2)}{C_2} \times \frac{R_{rs}(\lambda_3)}{C_3}.$$ 

$\lambda_1$ is highly sensitive to Chl-a absorption from 660 to 690 nm, $\lambda_2$ is minimally sensitive to Chl-a absorption from 690 to 730 nm, and the absorption of nonpigmented suspended matter and colored dissolved organic matter (CDOM) at $\lambda_2$ are very close to those at $\lambda_1$. Absorption of water constituents over the wavelengths 700 to 750 nm has minimal influence on the third band $\lambda_3$. Various studies have demonstrated the efficiency of this model in water bodies with various optical properties (Gitelson, Schalles, and Hladik 2007; Gitelson et al. 2008; Le et al. 2011; Huang et al. 2014).

Mishra and Mishra (2012) recently proposed a novel index, the normalized difference chlorophyll index (NDCI), to estimate Chl-a in optically complex waters. This index is similar to the normalized difference vegetation index (NDVI) for vegetation, which is based on normalization of the difference between the NIR and red bands in vegetation (Mishra, Schaeffer, and Keith 2014). The formula for the NDCI is 

$$\frac{R_{rs}(\lambda_1) - R_{rs}(\lambda_2)}{R_{rs}(\lambda_1) + R_{rs}(\lambda_2)}.$$ 

To avoid the confounding influence of CDOM and TSM on water reflectance spectra at shorter wavelengths, $\lambda_1$ and $\lambda_2$ are closely spaced and the spectral features at the reflectance peak near 700 nm and the absorption peak near 675 nm are selected (Mishra and Mishra 2012). The main advantage of the NDCI is that normalizing by the sum of the reflectance in two bands eliminates uncertainties in estimations of $R_{rs}$, seasonal solar azimuth differences, and atmospheric contributions at these wavelengths (Mishra and Mishra 2012; Mishra, Schaeffer, and Keith 2014).

Satellite images acquired from MODIS, MERIS, and other ocean satellites have been successfully used to estimate Chl-a concentrations in turbid productive waters (Gurlin, Gitelson, and Moses 2011; Le et al. 2011; Sun et al. 2013). These hyperspectral satellite sensors have coarse spatial resolutions (e.g., the resolution of MODIS is 250–500 m, and that of MERIS is 300 m) that make them unsuitable for Chl-a retrieval in small waterlogged areas (e.g., areas smaller than 5 km$^2$). Although airborne hyperspectral remote sensing can provide images with a high spatial resolution that are suitable for retrieving water quality information from small inland water bodies (Kallio et al. 2001; Koponen et al. 2002), the high costs of airborne remote sensing and the lack of available images mean that its application to waterlogged areas in China is impractical. In situ hyperspectral data are therefore most suitable for Chl-a estimation in waterlogged areas of China.

This study examined two representative types of waterlogged area in the Huainan coal fields in order to establish remote sensing models for Chl-a estimation based on in situ hyperspectral data. The aims are to (1) develop an understanding of the characteristics of the absorption coefficients of phytoplankton, nonpigmented suspended matter, and phytoplankton in waterlogged areas; (2) optimize the waveband locations and calibrate the NIR-red, three-band, and NDCI models for accurate Chl-a assessment in waterlogged areas based on field spectral measurements; and (3) compare the performance of the three models with other independent data and analyze the Chl-a concentration data from the waterlogged areas for spatial and temporal variations.
Materials and methods

Study area

In this study, the Guqiao (Gq) and Panji (Pj) waterlogged areas were chosen for analysis as representative of the two main types of waterlogged area in this locality: closed and semi-closed. Both the areas are located in Huainan City, Anhui Province, East China (Figure 1), where there has been extensive underground coal mining for decades (Dong et al. 2013). The two waterlogged areas cover areas of 4.5 and 3.5 km², and have average depths of 4 m and 4.5 m, respectively. The two waterlogged areas are the sites of 2 districts, 5 towns, and nearly 30 villages with a total of ~100,000 inhabitants. Furthermore, two power plants and coal separating plants have been built close to Gq and Pj. The water quality in this area is therefore vulnerable to pollution from various sources, including industrial enterprises, farmland, and local residents. Gq, adjacent to the Gang River, started to form in 2003 in response to exploitation of the Guqiao coal mine. It is a closed waterlogged area and is strongly influenced by seasonal precipitation. Pj started to form in the 1980s in response to exploitation of the Panyi coal mine. It is semi-closed and is connected to the Ni River. Seasonal precipitation has a strong influence on the hydrological characteristics of the Ni River. Sediment loads may be high and bottom sediments may be re-suspended during the rainy season due to increased flow.

![Figure 1. Location of the study area and sampling sites in Gq and Pj. Black filled symbols indicate samples collected in May; gray-filled symbols indicate samples collected in November. For full color versions of the figures in this paper, please see the online version.](image-url)
Field data

Field data were collected at 52 stations (26 in Pj and 26 in Gq) from 21 to 22 May 2013 and at 49 stations (25 in Pj and 24 in Gq) from 16 to 17 November 2013. A standard set of water quality parameters was measured at each station, including water temperature (°C), pH, and water clarity (measured using a standard Secchi disk). Water samples were collected at a depth of 0.5 m with amber bottles and were immediately stored in a cooler containing ice packs. Information about the physicochemical and optical properties of the water samples collected is presented in Tables 1 and 2.

Water samples and remote sensing reflectance spectra were collected at the same time at each station, mostly between 10:00 and 14:00 local time. Field spectra were measured using a portable spectroradiometer (ASD FieldSpec) with a spectral range of 325–1075 nm and a spectral resolution of 3 nm. Water surface reflectance spectral data were measured following Mueller et al. (2003). The radiance spectra of a standard 30% gray reference panel \( L_p \), water \( L_{sw} \) at approximately 1 m above the water surface, and the sky \( L_{sky} \) were measured at each station. To avoid the influence of direct solar radiation and any interference between the boat and the water surface, the optical fiber was positioned at an azimuth angle of 90°–135° to the solar principal plane and at a zenith angle of 30°–40° (Li et al. 2012). After measuring the water radiance, the spectroradiometer was then rotated upwards by 90°–120° to measure the skylight radiance \( L_{sky} \). The view azimuth angle was the same as the angle used for measuring water radiance \( L_{sw} \). To optimize the signal-to-noise ratio and reduce errors in \textit{in situ} measurements, the relevant spectra were measured 10 times at each sampling site. Remote sensing reflectance \( R_{rs}(\lambda) \) was calculated as follows:

\[
R_{rs} = \frac{(L_{sw} - rL_{sky})\rho_p}{\pi L_p}
\]

where \( r \) is the reflectance of skylight at the air–water interface, for which a constant of 0.0226 was used in this study. \( \rho_p \) represents reflectance from the gray reflectance standard.

Laboratory analysis

Water samples on which Chl-a concentrations were determined were filtered with GF/C filters (Whatman). After overnight extraction in a dark freezer with 90% acetone, the samples were centrifuged for 5 min and spectrophotometrically analyzed at 750 and 665 nm. The Chl-a concentration was then calculated following Lorenzen (1967) and Chen, Chen, and Hu (2006). Concentrations of total suspended matter (TSM) were determined gravimetrically (Eaton et al. 2005). Water samples were filtered through pre-combusted GF/C filters (Whatman) to remove suspended organic matter; the filter papers were then dried and weighed (Tables 1 and 2). Water quality parameters for each sampling location were analyzed in triplicate to ensure accurate measurement.

The adsorption spectra of the water samples were determined, and values of the absorption coefficient spectra of CDOM \( [a_{CDOM}(\lambda)] \), total particulate matter \( [a_p(\lambda)] \), phytoplankton \( [a_{ph}(\lambda)] \), and nonpigment suspended matter \( [a_d(\lambda)] \) were calculated using the quantitative filter technique (QFT), employing the method proposed by Zhang et al (2009).

Prior to CDOM absorption measurements, samples were filtered through a 0.2 μm membrane filter into acid-washed and pre-combusted glass bottles. After filtration, CDOM absorbance was measured in a 1 cm quartz cell from 200 to 800 nm using a
Table 1. Descriptive statistics of the water quality parameters measured in Pj.

<table>
<thead>
<tr>
<th></th>
<th>pH</th>
<th>T (°C)</th>
<th>Chl-a (mg m(^{-3}))</th>
<th>SDD (m)</th>
<th>TSM (g m(^{-3}))</th>
<th>(a_p(440)) (m(^{-1}))</th>
<th>(a_p(675)) (m(^{-1}))</th>
<th>(a_d(440)) (m(^{-1}))</th>
<th>(a_{ph}(440)) (m(^{-1}))</th>
<th>(a_{ph}(675)) (m(^{-1}))</th>
<th>(a_{CDOM}(440)) (m(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>May ((N = 26))</td>
<td>Min</td>
<td>7.93</td>
<td>25.6</td>
<td>10.62</td>
<td>0.43</td>
<td>3.4</td>
<td>1.01</td>
<td>0.42</td>
<td>0.57</td>
<td>0.66</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>8.55</td>
<td>27.7</td>
<td>45.06</td>
<td>0.89</td>
<td>16.5</td>
<td>2.76</td>
<td>0.91</td>
<td>1.45</td>
<td>1.81</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>8.32</td>
<td>27.1</td>
<td>26.61</td>
<td>0.65</td>
<td>8.6</td>
<td>1.82</td>
<td>0.74</td>
<td>0.78</td>
<td>1.06</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.21</td>
<td>0.66</td>
<td>9.74</td>
<td>0.13</td>
<td>3.32</td>
<td>0.24</td>
<td>0.16</td>
<td>0.14</td>
<td>0.11</td>
<td>0.09</td>
</tr>
<tr>
<td>November ((N = 25))</td>
<td>Min</td>
<td>7.97</td>
<td>17.7</td>
<td>2.11</td>
<td>0.41</td>
<td>1.24</td>
<td>1.43</td>
<td>0.39</td>
<td>0.69</td>
<td>0.51</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>8.32</td>
<td>19.3</td>
<td>13.89</td>
<td>0.63</td>
<td>8.25</td>
<td>2.54</td>
<td>0.83</td>
<td>1.66</td>
<td>0.94</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>8.09</td>
<td>18.3</td>
<td>6.98</td>
<td>0.49</td>
<td>4.26</td>
<td>1.94</td>
<td>0.60</td>
<td>1.02</td>
<td>0.72</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.08</td>
<td>0.53</td>
<td>3.59</td>
<td>0.06</td>
<td>1.52</td>
<td>0.32</td>
<td>0.11</td>
<td>0.22</td>
<td>0.18</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Note: Chl-a – Chlorophyll-a concentration, SDD – water clarity, TSM – concentration of total suspended matter, \(a_p(440)\) – absorption coefficient spectra of total particulate matter at 440 nm, \(a_p(675)\) – absorption coefficient spectra of total particulate matter at 675 nm, \(a_d(440)\) – absorption coefficient of nonpigment suspended matter at 440 nm, \(a_{ph}(440)\) – absorption coefficient of phytoplankton at 440 nm, \(a_{ph}(675)\) – absorption coefficient of phytoplankton at 675 nm, \(a_{CDOM}(440)\) – absorption coefficient of CDOM at 440 nm, \(N\) – number of samples.
Table 2. Descriptive statistics of the water quality parameters measured in Gq.

<table>
<thead>
<tr>
<th></th>
<th>pH</th>
<th>T (°C)</th>
<th>Chl-a (mg m$^{-3}$)</th>
<th>SDD (m)</th>
<th>TSM (g m$^{-3}$)</th>
<th>$a_p$ (440) (m$^{-1}$)</th>
<th>$a_p$ (675) (m$^{-1}$)</th>
<th>$a_d$ (440) (m$^{-1}$)</th>
<th>$a_{ph}$ (440) (m$^{-1}$)</th>
<th>$a_{ph}$ (675) (m$^{-1}$)</th>
<th>$a_{CDOM}$ (440) (m$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>May</strong> (<em>N</em> = 26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>7.82</td>
<td>26.1</td>
<td>8.49</td>
<td>0.28</td>
<td>1.42</td>
<td>1.62</td>
<td>0.72</td>
<td>0.55</td>
<td>0.78</td>
<td>0.45</td>
<td>0.75</td>
</tr>
<tr>
<td>Max</td>
<td>9.21</td>
<td>27.9</td>
<td>64.45</td>
<td>0.81</td>
<td>12.45</td>
<td>3.54</td>
<td>1.18</td>
<td>1.19</td>
<td>2.55</td>
<td>1.14</td>
<td>1.79</td>
</tr>
<tr>
<td>Mean</td>
<td>8.49</td>
<td>27.3</td>
<td>34.76</td>
<td>0.56</td>
<td>6.02</td>
<td>2.07</td>
<td>1.01</td>
<td>0.72</td>
<td>1.59</td>
<td>0.83</td>
<td>1.22</td>
</tr>
<tr>
<td>SD</td>
<td>0.32</td>
<td>0.61</td>
<td>14.39</td>
<td>0.20</td>
<td>2.28</td>
<td>0.52</td>
<td>0.24</td>
<td>0.32</td>
<td>0.57</td>
<td>0.48</td>
<td>0.27</td>
</tr>
<tr>
<td><strong>November</strong> (<em>N</em> = 24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>8.43</td>
<td>17.6</td>
<td>5.91</td>
<td>0.20</td>
<td>1.31</td>
<td>1.60</td>
<td>0.45</td>
<td>0.63</td>
<td>0.45</td>
<td>0.32</td>
<td>1.28</td>
</tr>
<tr>
<td>Max</td>
<td>8.85</td>
<td>19.9</td>
<td>21.84</td>
<td>0.72</td>
<td>7.62</td>
<td>3.67</td>
<td>0.85</td>
<td>1.61</td>
<td>1.93</td>
<td>0.82</td>
<td>2.16</td>
</tr>
<tr>
<td>Mean</td>
<td>8.55</td>
<td>18.5</td>
<td>13.63</td>
<td>0.41</td>
<td>3.49</td>
<td>2.49</td>
<td>0.67</td>
<td>1.09</td>
<td>0.91</td>
<td>0.51</td>
<td>1.68</td>
</tr>
<tr>
<td>SD</td>
<td>0.11</td>
<td>0.57</td>
<td>5.60</td>
<td>0.17</td>
<td>1.87</td>
<td>0.67</td>
<td>0.37</td>
<td>0.43</td>
<td>0.31</td>
<td>0.30</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Note: Chl-a – Chlorophyll-a concentration, SDD – water clarity, TSM – concentration of total suspended mater, $a_p$ (440) – absorption coefficient spectra of total particulate matter at 440 nm, $a_p$ (675) – absorption coefficient spectra of total particulate matter at 675 nm, $a_d$ (440) – absorption coefficient of nonpigment suspended matter at 440 nm, $a_{ph}$ (440) – absorption coefficient of phytoplankton at 440 nm, $a_{ph}$ (675) – absorption coefficient of phytoplankton at 675 nm, $a_{CDOM}$ (440) – absorption coefficient of CDOM at 440 nm, *N* – number of samples.
spectrophotometer, with Milli-Q water ($A_{MQ}(\lambda)$) as a reference. To avoid scattering and baseline fluctuations, the absorbance data were corrected by subtracting the mean value of the measured absorbance from 700 to 750 nm (Bricaud, Morel, and Prieur 1981; Babin et al. 2003). The value of $a_{CDOM}(\lambda)$ was calculated as follows:

$$a_{CDOM}(\lambda) = 2.303 \left[ A_{CDOM}(\lambda) - A_{MQ}(\lambda) \right] / l$$  \hspace{1cm} (2)$$

where $A_{CDOM}(\lambda)$ is the absorption at wavelength $\lambda$ (nm) and $l$ is the path length (m). The CDOM spectral absorption was fitted to a nonlinear exponential function using a function form (Bricaud, Morel, and Prieur 1981):

$$a_{CDOM}(\lambda) = a_{CDOM}(\lambda_0) e^{-S_{CDOM}(\lambda_0 - \lambda)}$$  \hspace{1cm} (3)$$

where $\lambda_0$ is the reference wavelength and $a_{CDOM}(\lambda_0)$ is the measured absorption at the reference wavelength. In this case, the reference wavelength was 440 nm. $S_{CDOM}$ is the absorption spectral slope coefficient of CDOM.

The absorption of total particulate matter was measured using a spectrophotometer after water samples were filtered through 0.7 $\mu$m Whatman GF/F filters. The absorbance of total particulate matter was recorded at 1 nm intervals from 350 to 800 nm, using a blank filter wetted with Milli-Q water as a reference. The absorption spectra were corrected for path length amplification (Mitchell 1990; Tassan and Ferrari 1995) and baseline offset by subtracting the average absorbance from 700 to 750 nm (Mitchell et al. 2002). The absorbance values were converted to $a_p(\lambda)$ as follows:

$$a_p(\lambda) = \left( 2.303 \left[ A_p(\lambda) \right] \right) A / V$$  \hspace{1cm} (4)$$

where $A_p(\lambda)$ is the absorbance at wavelength $\lambda$, $A$ is the effective filtration area, and $V$ is the filtration volume.

The absorbance of nonpigmented suspended matter was measured after methanol extraction, in the same manner as $A_p(\lambda)$. Values of $A_d(\lambda)$ were converted to $a_d(\lambda)$ using Equation (4). The spectral absorption coefficients of phytoplankton ($a_{ph}(\lambda)$) were obtained by subtracting $a_d(\lambda)$ from $a_p(\lambda)$:

$$a_{ph}(\lambda) = a_p(\lambda) - a_d(\lambda)$$  \hspace{1cm} (5)$$

The spectral absorption by nonpigmented suspended matter was fitted to a nonlinear exponential function using a function form (Bricaud, Morel, and Prieur 1981):

$$a_d(\lambda) = a_d(\lambda_0) e^{-S_d(\lambda_0 - \lambda)}$$  \hspace{1cm} (6)$$

where $\lambda_0$ is a reference wavelength and $a_d(\lambda_0)$ is the absorption measured at the reference wavelength. A reference wavelength of 440 nm was chosen. $S_d$ is the absorption spectral slope coefficient of nonpigmented suspended matter.

**Accuracy assessment**

The data collected in Pj and Gq were randomly divided into two independent subsets: one for model development and the other for model validation. The coefficient of
determination ($R^2$), the mean absolute percent error (MAPE) (7), and the root mean square error (RMSE) (8) were used to evaluate the accuracy of the different models, as follows:

\[
\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - y'_i}{y_i} \right|
\]  

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - y'_i)^2}
\]

where $n$ is the number of samples, $y_i$ is the measured value, and $y'_i$ is the estimated value.

Results

Variation in absorption coefficients

CDOM absorption increased exponentially toward the blue and ultraviolet (UV) wavelengths. The spectral curves and the mean spectral curves for the 400–750 nm range for May and November 2013 are presented in Figure 2d (for Pj) and Figure 3d (for Gq). There was high seasonal and spatial variability in the spectra of $a_{\text{CDOM}}(\lambda)$. The average $a_{\text{CDOM}}(440)$ in November was higher than that in May (Tables 1 and 2). Furthermore, the average $a_{\text{CDOM}}(440)$ in Pj (1.75 ± 0.39 m$^{-1}$ in November, 1.27 ± 0.35 m$^{-1}$ in May) was higher than that in Gq (1.68 ± 0.36 m$^{-1}$ in November, 1.22 ± 0.27 m$^{-1}$ in May) in both May and November.

The spectral slope of CDOM absorption at 440 nm ($S_{\text{CDOM}}$), calculated between 400 and 600 nm, was varied from 0.0107 to 0.0197 nm$^{-1}$ in Pj and from 0.0119 to 0.0191 nm$^{-1}$ in Gq. The average values of $S_{\text{CDOM}}$ in Pj showed little seasonal variation.
and were 0.0158 nm\(^{-1}\) in May and 0.0145 nm\(^{-1}\) in November. The average values of \(S_{CDOM}\) in Gq during May and November were 0.0166 and 0.0143 nm\(^{-1}\), respectively.

The absorption coefficient of the nonpigmented suspended matter decreased as the wavelength increased from 400 to 750 nm in both Pj and Gq (Figures 2b and 3b). Values of \(a_d(440)\) showed considerable variation between May and November 2013 in both areas, and in Pj they varied from 0.57 to 1.45 m\(^{-1}\) in May and from 0.69 to 1.66 m\(^{-1}\) in November; in Gq they varied from 0.55 to 1.19 m\(^{-1}\) in May and from 0.63 to 1.61 m\(^{-1}\) in November. The variation might have resulted from an increase in organic materials from phytoplankton decomposition in November. The average values of \(a_d(440)\) were higher in Pj than in Gq. The spectral slope of \(a_d(\lambda)\) was calculated between 400 and 600 nm. Similar to \(a_d(440)\), \(S_d\) showed considerable temporal variation, and in Pj ranged from 0.01128 to 0.01147 nm\(^{-1}\) in May and from 0.01123 to 0.01174 nm\(^{-1}\) in November; in Gq they varied from 0.01131 to 0.01145 nm\(^{-1}\) in May and from 0.01138 to 0.01169 nm\(^{-1}\) in November. These values of \(S_d\) are influenced by inorganic or mineral-dominated sediments (Bowers, Harker, and Stephan 1996; Babin et al. 2003).

The \(a_{ph}(\lambda)\) spectra obtained during the cruises in Pj and Gq showed two diagnostic absorption peaks: one in the blue wavelength and one in the red wavelength (Figures 2c and 3c). The \(a_{ph}(\lambda)\) value varied markedly in space and time. In May, values of \(a_{ph}(\lambda)\) at 440 and 675 nm ranged from 0.66 to 1.81 m\(^{-1}\) and from 0.32 to 0.75 m\(^{-1}\) in Pj, and from 0.78 to 2.55 m\(^{-1}\) and 0.45 to 1.14 m\(^{-1}\) in Gq, respectively. In November, these values ranged from 0.51 to 0.94 m\(^{-1}\) and from 0.25 to 0.55 m\(^{-1}\) in Pj, and from 0.45 to 1.93 m\(^{-1}\) and from 0.32 to 0.82 m\(^{-1}\) in Gq, respectively. The mean values of \(a_{ph}(440)\) and \(a_{ph}(675)\) were higher in May than in November (Tables 1 and 2). Furthermore, the mean values of \(a_{ph}(\lambda)\) at 440 and 675 nm were higher in Gq (1.21 ± 0.53 in May and 0.71 ± 0.31 in November) than in Pj (0.82 ± 0.1 in May and 0.55 ± 0.14 in November).
Inland waters have multiple bio-optical properties, and are generally dominated by one or more of their inherent optical properties (IOPs), such as phytoplankton, nonpigmented suspended matter, and CDOM (Li et al. 2012). Ternary plots were constructed to illustrate the dominant absorbing constituents in Pj and Gq (Figure 4). These two waterlogged areas were significantly influenced by nonpigmented components (\(a_d(\lambda)\) and \(a_{CDOM}(\lambda)\)), which were important absorbers of light in the blue part of the spectrum; their contributions to total nonwater absorption at 440 nm were more than 50% in May and 60% in November. However, their contributions decreased to 40% at 675 nm because of the exponential decrease in \(a_d(\lambda)\) and \(a_{CDOM}(\lambda)\) with increasing wavelength (Naik et al. 2011). In contrast, the phytoplankton absorption contribution increased as the wavelength increased and generally dominated at 675 nm. Phytoplankton absorption in Pj and Gq exceeded 60% in May and November, and accounted for as much as 80–90% of the total nonwater absorption at some stations in Gq. Further, Spearman correlation analysis \(r_s\) indicated a stronger relationship between \(a_{ph}(675)\) and Chl-a \(r_s = 0.66\) in Pj; \(r_s = 0.69\) in Gq) (Figure 5) than between \(a_{ph}(440)\) and Chl-a \(r_s = 0.47\) in Pj; \(r_s = 0.51\) in Gq). These results highlight the important role of the red spectral region in the development of a model for Chl-a estimation in waters with high CDOM and high absorption by nonpigmented suspended matter.

Figure 4. Ternary plots showing the relative contributions by CDOM \(a_{CDOM}(\lambda)\), phytoplankton \(a_{ph}(\lambda)\) and nonpigmented suspended matter \(a_d(\lambda)\) to total absorption at (a) 440 nm and (b) 675 nm.

Figure 5. Relationships between \(a_{ph}(675)\) and Chl-a in (a) Pj and (b) Gq.
Reflectance spectral characteristics of Chl-a

The reflectance spectra collected in Pj and Gq showed optical conditions typical of turbid productive inland waters (Figure 6). In the blue spectral region, the combined absorption by phytoplankton, nonpigmented suspended matter, and CDOM absorption resulted in a low sensitivity to variations in Chl-a concentrations (Gurlin, Gitelson, and Moses 2011; Zhang et al. 2009). The first reflectance peak emerged around 570 nm, which has previously been attributed to backscattering by particles (Lubac and Loisel 2007). There was a distinct reflectance feature in the vicinity of 675 nm, corresponding to the maximum Chl-a absorption (Bidigare et al. 1990). However, the reflectance in this range was also influenced by absorption and scattering of other constituents, which may have weakened the correlation between $R_{rs}(675)$ and Chl-a (Zimba and Gitelson 2006; Zhang et al. 2009). The second reflectance peak appeared around 700 nm and was caused by the combined effects of reduced Chl-a absorption and pure water (Bricaud et al. 1995; Gitelson et al. 2008; Gilerson et al. 2010). This spectral region included the Chl-a fluorescence emission peak, the magnitude of which varied widely between moderate and high Chl-a concentrations (Gitelson et al. 2008). Furthermore, the position of the reflectance peak near 700 nm was red-shifted as the Chl-a concentration increased in Gq. In the NIR region, Chl-a absorption decreased with increasing wavelength, whereas absorption of pure water increased (Gitelson, Schalles, and Hladik 2007). The high reflectance in this region was caused by interference from other water constituents.

Models for estimating Chl-a concentration

The optimal spectral bands of Chl-a retrieval models may be different in waters with different optical properties (Koponen et al. 2002; Gitelson, Schalles, and Hladik 2007; Gitelson et al. 2009; Zhang et al. 2011). Therefore, it is important to establish specific Chl-a retrieval models for specific water optical characteristics. The NIR-red, three-band, and NDCI models were established to obtain accurate information about variations in the Chl-a concentrations in Pj and Gq, and were calibrated using in situ hyperspectral data. The optimal spectral bands for the NIR-red model were determined using the optimization technique of Gitelson, Schalles, and Hladik (2007). The most sensitive band ratios, $R_{rs}(692)/R_{rs}(675)$ and $R_{rs}(706)/R_{rs}(680)$, were selected to estimate the Chl-a concentrations in Pj and Gq, respectively (Figure 7(a, b)), where $R_{rs}(692)$ and
were close to the average position of the reflectance peak in the NIR region, and $R_{rs}(675)$ and $R_{rs}(680)$ were close to the maximum Chl-a absorption. The NIR-red model, which has been used successfully in many inland water bodies since being proposed by Gitelson, Keydan, and Shishkin (1985), accurately estimated the Chl-a concentrations in this study. Figure 7(a, b) shows a reasonable linear relationship between the NIR-red model and Chl-a in Pj ($R^2 = 0.80$), but that there was a polynomial relationship between this model and Chl-a in Gq ($R^2 = 0.82$).

The optimization approach described earlier for the NIR-red model was used to determine the optimal spectral bands for the three-band model. The initial positions for

![Figure 7. Three models for estimating Chl-a concentrations: (a) NIR-red model in Pj, (b) NIR-red model in Gq, (c) three-band model in Pj, (d) three-band model in Gq, (e) NDCI model in Pj, and (f) NDCI model in Gq.](image)
λ₁ and λ₃ were set at 660 and 750 nm, respectively, to search for the optimal position of λ₂. The optimal position for λ₂ and the initial position for λ₃ were then used to determine the optimal position for λ₁. Subsequently, the optimal positions obtained for λ₁ and λ₂ were selected to determine the optimal position of λ₃. Finally, wavelengths at 665, 698, and 743 nm were selected as the optimal positions for the three-band model in Pj, while wavelengths at 671, 706, and 748 nm were selected as the optimal positions for the three-band model in Gq (Figure 7(c, d)). The performance of the three-band model was superior to that of the NIR-red model. The three-band model showed a polynomial relationship with Chl-a in Pj ($R^2 = 0.87$) and a strong linear relationship with Chl-a in Gq ($R^2 = 0.88$).

The optimal spectral bands for developing the NDCI were selected from the absorption peak near 675 nm and the reflectance peak near 700 nm, as these spectral bands are most sensitive to variations in Chl-a in water (Mishra and Mishra 2012). Application of the optimization technique showed that the optimal spectral bands for developing the NDCI in Pj and Gq were at 675 and 692 nm and at 680 and 706 nm, respectively. The NDCI showed a strong polynomial relationship with Chl-a concentrations in Gq ($R^2 = 0.90$; Figure 7(f)), and a relatively weak polynomial relationship with Chl-a concentrations in Pj ($R^2 = 0.84$; Figure 7(e)).

Model validation

To comprehensively evaluate the performance of the three Chl-a estimation models in Pj and Gq, the independent validation data ($n = 17$ for Pj and $n = 16$ for Gq) were used to validate the models by retrieving the MAPE and RMSE values for Chl-a (Figure 8). The NIR-red model reported an MAPE of 20.05% in Pj and 19.35% in Gq, and an RMSE of 6.13 mg m⁻³ in Pj and 5.52 mg m⁻³ in Gq. The results from the three-band model were better than those from the NIR-red model. The three-band model yielded MAPE values of 11.35% and 11.16% when applied to Pj and Gq, respectively, and RMSE values of 3.87 and 3.28 mg m⁻³ in Pj and Gq, respectively. The NDCI model, with an MAPE value of 10.88% and an RMSE value of 3.19 mg m⁻³, yielded more accurate results for Gq than the three-band model, but gave relatively low-accuracy results for Pj, for which the MAPE was 14.52% and the RMSE was 4.69 mg m⁻³. Of note, the Chl-a concentrations were significantly and systematically overestimated by all three models, possibly because of absorption of nonpigmented suspended matter (Gons et al. 2000).

Discussion

The results of in situ monitoring show that the Chl-a concentrations were higher in May (26.61 ± 9.74 mg m⁻³ in Pj; 34.76 ± 14.39 mg m⁻³ in Gq) than in November, corresponding to the timing of obvious algal blooms in some waters. However, Chl-a concentrations were still relatively high in November (6.98 ± 3.59 mg m⁻³ in Pj; 13.63 ± 5.6 mg m⁻³ in Gq), demonstrating that Pj and Gq were eutrophic. Chl-a concentrations were higher in Gp than in Pj because of nutrient accumulation in closed water, which results in the rapid growth of algae. Similarly, the TSM concentrations were significantly higher in both Pj and Gq in May than in November, as a result of the noticeable sediment re-suspension by considerable rain-generated turbulence in May during the rainy season. TSM concentrations were higher in Pj than in Gq due to the influence of the Ni River. Rivers can transport large quantities of sediment particles, and elevated flows can cause the re-suspension of bottom sediments, especially in the rainy season. Furthermore, Chl-a concentrations and TSM concentrations were positively correlated in Gq, but no
A strong interdependence between Chl-a and TSM in the closed water was found in Pj. This result indicates a strong interdependence between Chl-a and TSM because of the frequent algal blooms that make a large contribution to TSM (Li et al. 2012).

There was considerable variation in the concentrations and spectral absorption coefficients of the three optically active substances in Pj and Gq. The contributions of the nonpigmented components $a_d(440)$ and $a_{CDOM}(440)$ were higher in Pj than in Gq because of the influence of the Ni River. The Ni River transports large quantities of humic substances and dissolved organic matter, which drive much of the variability in the...
nonpigmented suspended matter and CDOM absorption. Water flow and the hydrological properties of the water column have a strong influence on the IOPs through modifications of the package effect and the accumulation of suspended material (Bricaud et al. 2004). In contrast, the values of $a_{ph}(\lambda)$ were higher in Gq than in Pj, which corresponds with the higher Chl-a concentrations in Gq. In this study, in situ hyperspectral data were used to retrieve the concentrations of Chl-a in waterlogged areas in Huainan. All three models gave good estimations of Chl-a concentrations in Pj and Gq. The three-band model and the NDCI model yielded higher $R^2$ values and lower RMSE values than the NIR-red model. Moreover, the accuracy of the NDCI model was highest in Gq. Comparison with an earlier study by Mishra (Mishra and Mishra 2012) shows that the optimal spectral band $\lambda_1$ of the NDCI model shifted from 665 to 675 nm (Pj) and from 665 to 680 nm (Gq) because of the high Chl-a concentrations in the waterlogged areas. Furthermore, the waterlogged areas the optimal spectral bands were influenced not only by Chl-a, but also by TSM and CDOM. In contrast, the accuracy of the three-band model was highest in Pj. The optimal spectral bands of the three-band model for the waterlogged areas were within the ranges recommended by previous studies (Gitelson, Schalles, and Hladik 2007; Gitelson et al. 2008). One of the reasons for the superior performance of the three-band model was the removal of interference from backscattering (Zhang et al. 2009). These results indicate that the NDCI model was more suitable for the relatively low TSM and high Chl-a concentrations in the closed waterlogged area, and that the three-band model was more suitable for the relatively high TSM and low Chl-a concentrations in the semi-closed waterlogged area.

The present results indicate the potential of in situ hyperspectral data for monitoring Chl-a concentrations in waterlogged area. In addition, the model and method discussed in this paper may play an important role in estimating Chl-a concentrations for water quality monitoring and water resource management in other waterlogged areas in China and worldwide. This applies not only to waterlogged areas caused by underground coal mining in Xuzhou (Du et al. 2007), Huaiabei (Fan et al. 2012), and Yanzhou (Hu, Xu, and Zhao 2012), which have similar environmental conditions to those at Huainan, but also to waterlogged areas caused by uneven irrigation and over-irrigation in India (Goyal and Arora 2012; Singh, Kumar, and Mathur 2013), Bangladesh (Adri and Islam 2012) and elsewhere.

Although hyperspectral satellite remote sensing data are not appropriate for small waterlogged areas, due to poor spatial resolution, upcoming satellite sensors with higher spatial and spectral resolutions such as Sentinels, HyspIRI, and GEO-CAPE may be appropriate for such areas. These new missions will not only fill a gap in data provision that has existed since previous missions ended, but will result in a step-change in the use of satellite observations for monitoring the quality of inland water bodies. The combination of hyperspectral satellite data with finer spatial resolution and in situ hyperspectral data will enhance the applicability and robustness of Chl-a retrieval models for waterlogged areas. It will also provide a better understanding of the spatiotemporal distribution of water quality parameters and help to predict future changes in water quality in waterlogged areas.

**Conclusions**

Remote sensing techniques were used to estimate the concentrations of Chl-a in waterlogged areas that developed as a result of long-term underground coal mining. The Gq and Pj
waterlogged areas were chosen as representative examples of the two main types of waterlogged area, closed and semi-closed. There was significant difference between Pj and Gq in terms of the spectral absorption coefficients of the three optically active substances, \(a_{ph}(\lambda), a_d(\lambda), \text{ and } a_{CDOM}(\lambda)\). These two waterlogged areas were significantly influenced by non-pigmented components \(a_d(\lambda)\) and \(a_{CDOM}(\lambda)\). The contributions from nonpigmented components were higher in Pj than in Gq because of the connection to the Ni River.

The NIR-red, three-band, and NDCI models were calibrated and validated for estimating Chl-a concentrations using in situ hyperspectral datasets. All three models gave good estimations of Chl-a concentrations in Pj and Gq, which indicates that hyperspectral remote sensing data, based on field spectral measurements, can be used for estimating Chl-a concentrations in waterlogged areas. Furthermore, the three-band model is more accurate for semiclosed waterlogged areas, while the NDCI model is more accurate for closed waterlogged areas.

**Acknowledgments**

The authors thank Gao Liangmin, Li Shaopengm, and Ye Yuanyuan for their assistance in sample collection and laboratory analysis. The authors are grateful to anonymous reviewers for constructive comments and recommendations that improved an earlier version of the manuscript.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported by the National Key Technology R&D Program [No. 2012BAC10B02]; Huainan Mining Group [No. HNKY-JT-JS-(2013)-004].

**References**


