Characterization and variations of dissolved organic matter in the Lake Taihu area of China

Yong Qiu, Hanchang Shi, He Jing, Rui Liu, Qiang Cai, Minoru Takemura and Satoshi Haraguchi

ABSTRACT

Lake Taihu in China is a eutrophicated lake surrounded by industrial and urbanized zones, thus its water quality often suffers from organic and nutrient contaminants. In this paper, a 1 year water quality survey was conducted around the lake and statistical analysis tools were used to characterize the variations of organic pollutants. Analysis of variance (ANOVA), cluster analysis and principal component analysis (PCA) confirm the seasonal and spatial variations of surface water quality in Lake Taihu. Surface water quality is better during the wet season and worse downstream during the dry season. The dissolved organic matter was further analyzed using a parallel factor analysis (PARAFAC) model with three-dimensional excitation-emission fluorescence matrices. Four components were extracted from the fluorescence data, namely, two autochthonous biodegradation products (C1: amino acids, C4: protein-like materials) and two humic-like substances (C2: from microbial processing, C3: terrestrial). C1 and C4 were dominant in the chromophoric dissolved organic matter (CDOM) fluorophores; this result is similar to those of other inland water bodies in China. The CDOM fluorophores showed similar seasonal and spatial variations with common water quality indices, with the exception of the seasonal responses of C2 in winter. Bivariance correlations between the organic and nutrient concentrations and the fluorescence intensities of the CDOM fluorophores imply possible common sources of the different contaminants. This paper exemplifies advanced statistical methods as a useful tool in understanding the behavior of contaminants in inland fresh water systems.

Key words | dissolved organic matter, Lake Taihu area, parallel factor analysis, statistical data analysis

INTRODUCTION

Lake Taihu is the third largest freshwater lake in East China and eutrophicated by excessive nitrogen (N) and phosphorous (P) brought about by rapid industrialization and urbanization. As a typical shallow lake surrounded by developing areas, Lake Taihu was seriously contaminated in the past and its water quality has continuously deteriorated since the 1980s (Jin et al. 2008). It also suffers from annual algal blooms, which once resulted in a serious drinking water shortage affecting the surrounding municipal cities. The last decade showed a very slow improvement of the water quality in Lake Taihu even though a number of measures have been taken to control lake pollution (Moss 2007), including the construction of additional sewage treatment plants, regulating industrial and domestic discharges, and reforming industrial parks to comply with environmental standards. Therefore, with the goal of restoring lake water quality and the safety of the water supply treatment, the characteristics of the organic components and nutrients in the lake water, as well as their interactions with microorganisms, must be determined and understood.

Chromophoric dissolved organic matter (CDOM) is a component of the total dissolved organic matter (DOM)
that absorbs light in the ultraviolet and visible range of the electromagnetic spectrum (Singh et al. 2010). Changes in the CDOM absorption and fluorescence can reflect the variations in CDOM composition resulting from physical, biological and chemical processes in the water system (Coble 1996). Therefore, fluorescence spectroscopic analysis can be used to characterize the properties of organic matter in water to better understand it (Wu et al. 2006).

Using the pioneering parallel factor analysis (PARAFAC) model, three-dimensional excitation emission matrices (3DEEM) can be interpreted by the N-way partial least square (N-PLS) methods, which can decompose complex mixture of fluorophores into several independent components (Stedmon et al. 2003; Stedmon & Bro 2008). The number of components in PARAFAC can be optimized and validated to facilitate a reasonable data interpretation (Engelen et al. 2009). This method has been applied in the characterization of organics in lake water (Bricaud et al. 1981), wastewater treatment effluent (Lee & Ahn 2007), tryptophan-like fluorophores in surface water (Baker et al. 2005) and fulvic acid and humic acid in river water (Ahmad et al. 2002).

Analysis of 3DEEM spectra by PARAFAC model to characterize CDOM in Lake Taihu has been adapted to optimize the water quality monitoring program (Wang et al. 2007), to understand the correlation between CDOM and microbial activities (Zhang et al. 2009) and to identify the sources and fate of CDOM fluorophores (Yao et al. 2011). However, the seasonal variations of the identified CDOM fluorophores in Lake Taihu and their correlations to bulk water quality indices (WQIs) have not been investigated. Therefore, a 1 year survey was conducted in the Lake Taihu area, and sophisticated statistical analysis tools, including 3DEEM PARAFAC analysis, were used to interpret the dataset and determine the seasonal and spatial variations and sources of the organic components as well as the correlation between them.

MATERIALS AND METHODS

Sampling and sites

Periodic, monthly surveys of surface water quality were conducted from March 2008 to February 2009 to determine the seasonal and spatial variations of water quality in the Lake Taihu area. The monthly average precipitation data in the lake area from May to October 2008, which is identified as the wet season, ranged from 93 to 289 mm, and those of the other months (dry season) ranged from 31 to 85 mm.

Figure 1 shows the sampling sites and water flow directions. At each sampling time, water samples were collected on the bank, from Site 1 to Site 7, using a manufactured deep water sampler to extract water samples at approximately 300 mm below the water surface. The samples were collected into 600 mL acid prewashed plastic bottles, which were then placed in a thermostat box at 4 °C prior to transport back to the laboratory, where the samples were stored in the refrigerator overnight at 4 °C. The WQIs of the samples were then determined the following day.

In the seven sampling sites, Site 1 represents the water in the Jinghang Canal receiving ground runoffs and municipal wastewater treatment effluents. Site 2 is the major effluent river of Lake Taihu, the Taipu River, which heads toward the Shanghai district. Site 3 is on the estuary of River Tiaoxi to Lake Taihu and currently acts as a drinking water source. Sites 4 and 5 are located in the industrialized urban areas around Lake Taihu, at which industrial discharges into the lake should be considered. Sites 6 and 7 are located at historic drinking water sources that have been recently abandoned because of water quality deterioration. The flow directions of the lake water are shown by arrows in Figure 1.

Figure 1 | Map and sampling sites in the Lake Taihu area.
Methods of water quality analysis

WQIs including pH, conductivity, turbidity, dissolved oxygen (DO) and temperature were analyzed by a portable multi-parameter water quality analyzer (Horiba U-10, Horiba Corp., Japan). Chemical analysis including chemical oxygen demand (CODcr), ammonium (NH3-N), ortho- and total phosphorus (PO4-P and TP), nitrate (NO3-N), nitrite (NO2-N) and total nitrogen (TN) were conducted using a Hach portable spectrometer (DR2500, Hach Company, USA). The range of Hach reagent is 0–40, 0.02–2.5, 0.1–1.1, 0.1–1.1, 0–25, 0–0.3 and 0–25 ppm for CODcr, NH3-N, PO4-P, TP, NO3-N, NO2-N and TN, respectively.

The samples for the 3DEEM and dissolved organic carbon (DOC) analyses were filtered through a 0.45 μm membrane filter (Millipore MCE 47 mm, USA). Dissolved organics were determined using a total organic carbon analyzer (TOC-V, Shimadzu, Japan). The fluorescence spectra were obtained by a fluorescence spectrometer (F-7000, Hitachi Corp., Japan), at excitation wave lengths of 220–450 nm and emission wave lengths of 230–650 nm. The widths of slit for excitation and emission were set to be 5 nm. The PMT voltage was 700 volt and the scan speed was 9,600 nm per second.

Statistical analysis of water quality data

Primary statistical analyses including average and correlation, as well as advanced analysis via variation (ANOVA), cluster analysis, and principal component analysis (PCA) were conducted using the tool boxes in the SPSS 13.0 software in the current study. ANOVA was used to identify the dependence of target variables on their grouping references, such as the sampling time or sites. Cluster analysis could classify the data group by calculating the distances between the groups. PCA is usually used to break complex data into several components, usually two to five, to simplify the interpretation of the dataset.

EEM interpretation by PARAFAC model

The 3DEEMs of pure water were acquired for each survey and used as blanks to be subtracted from the observed spectra of all investigated samples to remove the Raman scattering effects in the EEMs. The first- and second-order Rayleigh-Tyndall scattering effects were removed by simply cutting off specific regions in the spectra (Stedmon & Bro 2008). The fluorescence intensity was standardized using the peak height of the Raman scattering in the EEMs of pure water.

The PARAFAC analysis of the acquired EEMs for the Lake Taihu samples was conducted using the DOMFluor toolbox in a Matlab® environment contributed by previous researchers (Stedmon & Bro 2008). PARAFAC simplifies EEM data into three linear terms and residuals using the following equation.

\[ x_{ijk} = \sum a_{ij}b_{i}\cdot c_{kf} + e_{ijk} \]  

where \( x_{ijk} \) is the fluorescence intensity for the \( i \)th sample at emission wavelength \( j \) and excitation wavelength \( k \); \( a_{ij} \) is the intensity proportional to the concentration of the \( j \)th analyte; \( b_{i} \) and \( c_{k} \) are linearly related to the emission and excitation spectra at wavelengths \( j \) and \( k \) respectively; and \( e_{ijk} \) is the residual.

RESULTS

Water quality data

The water quality observations from all investigated samples are summarized in Table 1. Some WQI concentrations of the water samples fall within wide ranges of variation and are often higher than the grade III levels set by the China National Surface Water Quality Standard. This result indicates that the water body has not always been suitable as a drinking water source. The ANOVA significance of WQI concentrations by sampling site implies that the WQIs, except temperature and DOC, strongly depend on the locations (\( p < 0.005 \)). The ANOVA significance of the WQI concentrations by sampling month implies that TN and PO4-P concentrations in the water samples were independent of the sampling time, whereas other WQIs showed significant seasonal dependence (\( p < 0.05 \)).
Cluster analysis of water quality data

The hierarchical cluster analysis of the seven WQIs (DO, COD, DOC, TN, NH$_3$-N, TP, and PO$_4$-P) by sampling site demonstrates the spatial differences between the locations (Figure 2). Sites 3, 6 and 7, which are the current or past drinking water sources and represent good water quality in the sampling sites, showed nearly similar results and grouped as one cluster. Sites 4 and 5 were both strongly affected by municipal population and industries and thus classified into one cluster. The water quality in Site 1 was nearly the worst among all observations, and the cluster of Sites 1, 4 and 5 represents worse water quality than that of Sites 2, 3, 6 and 7, which were grouped in another cluster.

PCA analysis of water quality data

Bivariant correlation results show strong correlations between TN, NH$_3$-N, TP and PO$_4$-P as well as between COD, NH$_3$-N, TP and PO$_4$-P ($p < 0.01$), suggesting that the local organic and nutrient pollutants have similar sources. Given the strong bivariant correlation, PCA was conducted to estimate the water contamination. Table 2 shows the PCA component coefficients of the DO, COD, DOC, TN, NH$_3$-N, TP and PO$_4$-P WQIs. Three components were extracted from the dataset and designated as PCA-1, PCA-2, and PCA-3. According to the WQI coefficients for each component, PCA-1 represents pollution from both organics and nutrients, PCA-2 indicates strong pollution by organics but lower N content, and PCA-3 implies that the pollutants are nutrient-rich but contain fewer organics.

PCA-1 and PCA-2 were used to interpret the data. Figure 3 shows the PCA scores for every observation depending on the sampling site and season, respectively. Figure 3(a) shows that PCA-1 in Sites 1, 4 and 5 (solid points) has positive and higher scores than that in Sites 3, 6 and 7 (hollow points), indicating the relatively more serious pollution in Sites 1, 4 and 5. For PCA-2, Site 1 has...
Table 2 | PCA coefficients for the water quality indices (WQIs) of all samples

<table>
<thead>
<tr>
<th>PCA Componenta</th>
<th>Water quality indicesb</th>
<th>Extraction ratec</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DO</td>
<td>CODcr</td>
</tr>
<tr>
<td>PCA-1</td>
<td>-0.471</td>
<td>0.623</td>
</tr>
<tr>
<td>PCA-2</td>
<td>0.359</td>
<td>0.580</td>
</tr>
<tr>
<td>PCA-3</td>
<td>0.769</td>
<td>-0.040</td>
</tr>
</tbody>
</table>

aComponents extracted by principal component analysis (PCA), labeled as PCA-1, PCA-2 and PCA-3.  
bWater quality indices to calculate the component scores, which is the sum of individual production of each water quality indices in a row and its coefficient. DO: dissolved oxygen, CODcr: chemical oxygen demand, DOC: dissolved organic carbon, TN: total nitrogen, NH3-N: ammonia nitrogen, TP: total phosphorous, PO4-P: ortho phosphorous.  
cThe percentage of the dataset that could be explained by the additionally extracted components.

Figure 3 | PCA component scores by (a) sampling site, and (b) season.

A higher score than Sites 4 and 5, and Site 6 obtains a higher score compared with Sites 3 and 7, indicating organic matter contributed more to the pollution in Sites 1 and 6. According to Figure 3(b), although the difference was not significant as in sampling sites, wet season points obtained from May to October had generally lower scores than dry season points (November to April), indicating that surface water quality during the wet season was comparatively better than that in the dry season. This is expected because of higher precipitations in the wet season, thus higher flow rates of rivers result in greater flushing of the surface and lake water.

PARAFAC analysis for EEMs

A total of 83 acquired EEMs were processed using PARAFAC analysis (Stedmon & Bro 2008). Nine EEMs were considered as outliers and eliminated from the analysis according to the standard procedure. This may be caused by the complex sources and strong intensity of organic pollution in Site 1. Finally, the remaining 74 EEMs were decomposed into four major components (Figure 4). This result does not imply that only four components existed in the DOM; rather, the four-component model showed the best potential to explain the investigated dataset using the PARAFAC method.

Identification of CDOM fluorophores

Descriptions of the four extracted PARAFAC components (Table 3) are similar to those of CDOM previously reported in other water bodies. According to previous studies with identified CDOM sources, the four components in the current study can be amino acids or tyrosine, humic
substances associated with microbial metabolism, terrestrial humic-like organics, and protein-like materials. The difference between the peak wavelengths in the current study and references may be explained by two reasons. First, samples from very contaminated water (Site 1) and very clean water (Site 3) were included in the dataset, thereby introducing additional difficulties during iteration and model fitting. Second, the spectra may have been subjected to interference from dissolved organic N (Vasel & Praet 2002), especially in the contaminated water samples.

The EEM spectra of component C1 can be characterized by the peaks at the 280/230 nm excitation with a 320 nm emission wavelength, which have been attributed to an autochthonous tryptophan-like fluorescence. This component is similar to previously reported PARAFAC autochthonous tryptophan-like components, as well as amino acids that are free or bound in proteins, which are often found dominant in the fluorophores of surface water samples (Coble 1996; Murphy et al. 2008; Kowalczuk et al. 2009; Yao et al. 2011; Zhang et al. 2011).

Component C2 EEM could be related to the peaks at 295/235 nm excitation by 405 nm emission wavelength, which was similar to the terrestrial humic-like fluorescence peak A and the marine humic-like fluorescence peak M.

**Table 3** | Spectral characterization of the four components with previous identified sources

<table>
<thead>
<tr>
<th>Component</th>
<th>EX&lt;sub&gt;max&lt;/sub&gt;</th>
<th>EM&lt;sub&gt;max&lt;/sub&gt;</th>
<th>Reference</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>280 (230)</td>
<td>320</td>
<td>270/332 (Kowalczuk et al. 2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>275/305 (Coble 1996)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>270/305 (Zhang et al. 2011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>275/306 (Yao et al. 2011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>280/328 (Murphy et al. 2008)</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>295 (235)</td>
<td>405</td>
<td>250/400 (Kowalczuk et al. 2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>290/410 (Zhang et al. 2011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>300/404 (Yao et al. 2011)</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>240</td>
<td>455</td>
<td>240/456 (Holbrook et al. 2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>240/436 (Stedmon et al. 2003)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250/458 (Kowalczuk et al. 2005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250/450 (Hong et al. 2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>240/470 (Zhang et al. 2011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250/461 (Yao et al. 2011)</td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>245</td>
<td>340</td>
<td>230/340 (Holbrook et al. 2006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>230/344 (Yao et al. 2011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>240/338 (Murphy et al. 2008)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>230/342 (Wang et al. 2009)</td>
<td></td>
</tr>
</tbody>
</table>

*Component: the components obtained from PARAFAC analysis, C1–C4 represent four different fluorophores (Figure 4).

*EX<sub>max</sub>: the excitation wavelength in which the component shows a peak, the digits in brackets represents the excitation wavelength for the second and smaller peak.

*EM<sub>max</sub>: the emission wavelength in which the component shows a peak.
This component was assumed to be fulvic and humic substances strongly related to microbial reprocessing, such as algae generation and decay (Kowalczuk et al. 2009; Yao et al. 2011; Zhang et al. 2011).

Component C3 exhibited peaks at the 240 nm excitation with a 405 nm emission wavelength, similar to the traditional terrestrial humic-like peaks A and C (Coble 1996), and confirmed as a mixture of terrestrial-derived humic-like substances (Stedmon et al. 2003; Hong et al. 2005; Kowalczuk et al. 2005; Holbrook et al. 2006; Yao et al. 2011; Zhang et al. 2011).

For component C4, a peak at the 245 nm excitation with a 340 nm emission wavelength was observed, similar to the protein-like fluorescent compound confirmed by the autochthonous tryptophan-like fluorescence peak T (Coble 1996), which is usually strongly correlated with microbial metabolism in a water environment (Holbrook et al. 2006; Murphy et al. 2008; Wang et al. 2009; Yao et al. 2011).

Variations of CDOM fluorophores

Figures 5 and 6 show the fluorescence intensity loadings (\(a_{ij}\) in Equation (1)) for the obtained CDOM components (C1–C4) averaged by site and by sampling time, respectively. In general, the seasonal and spatial variations of the CDOM fluorophores of the investigated samples are similar to those of the common WQIs.

As shown in Figure 5, the lowest fluorescence intensity loading values for the CDOM components were obtained in Site 3, followed by Sites 6 and 7, indicating that fewer organic compounds existed in those areas. Sites 1, 2, 4 and 5 contained a higher amount of organic pollutants as indicated by their higher component loadings, which are coincident with the cluster analysis results shown in Figure 2. Therefore, the CDOM components showed a spatial variation trend similar to those of common WQIs.

In Figure 6, the CDOM components exhibited similar seasonal variations, with heavy pollution during the dry seasons and reduced pollution during the wet season (May to October). A similar trend was also observed in the common WQIs (e.g. CODcr and NH3-N), which can be attributed to higher precipitation and flushing during the wet season.

**DISCUSSION**

**Water quality evaluation**

Figure 7 shows a summary of the water quality grades according to the Chinese National Water Standards. Grade I indicates the highest water quality, grade V represents the lowest, and grade III is the minimal requirement for drinking water sources. According to the data, TN and TP can be classified as primary contaminants in Lake Taihu because of their high frequencies of occurrence at levels scoring above grade V. The water was also seriously polluted by CODcr and NH3-N, which also frequently occurred at levels above grade V.
Moreover, Table 1 shows that the TN concentrations for all investigated samples were higher than 1.0 mg N/L, which is the required standard for drinking water sources, implying a general unsuitability of the lake as a drinking water source. This result demonstrates a clear need to understand the origins and fate of dissolved N for the restoration of surface water quality in Lake Taihu. Second, the P discharge to the lake should be strictly controlled to alleviate the trends of eutrophication. Public data shows that ammonium concentrations in the river water have continuously decreased from 2.5 to 1.5 mg N/L from 2004 to 2010. However, the concentrations in the lake water were almost constant at approximately 0.5 mg N/L (Ministry of Environmental Protection 2011), indicating that although the contamination in the river water has gradually been controlled by strict management of the sewage system and industrial discharges, the lake water exhibited a very slow response to the measures adopted because of the organic and nutrient deposit in the lake sediments (Yu et al. 2007).

**Spatial variations of the water quality**

The spatial variations of the water quality in the lake area were characterized using the combined results of ANOVA (Table 1), cluster analysis (Figure 2) and PCA from the different sampling sites (Figure 3(a)). In general, the riverine samples contained more contaminants than the lake water ones, and upstream samples were cleaner than downstream samples. This finding is similar to those of other reports on Lake Taihu (Yao et al. 2011), Lake Tianmuhu (Zhang et al. 2011) and Lake Hongfeng (Wang et al. 2009) in South China, where riverine inputs contributed to the degradation of lake water quality.

The water quality of the samples from Site 3 was the highest among all investigated cases in the current study, not only in WQIs such as CODcr, TN and TP, but also in the CDOM fluorophores (Figure 5). Site 3 is located at the estuary of River Tiaoxi into Lake Taihu, which has reversing water flows (Yao et al. 2011). During the wet season, the water level in the lake is higher and the runoff flows from the lake to the river, and finally to the East China Sea. However, in the winter and spring dry seasons, the water flow direction at the estuary reverses. Given the better water quality of the lake water during the wet season and the clean water of the river during the dry season because of the geography of the land along the river (mountains and forest), water in Site 3 has been traditionally used as a drinking water source.

The downstream water quality (Sites 1 and 2) was the worst of all the investigated samples, as shown by the common WQIs (Figure 3) and the CDOM components (Figure 5). On the northern part of the lake, the river water is isolated by a number of sluice gates that prevent the influent loadings from the urbanized north shore. The Canal (Site 1) is utilized as a ditch to transport the urban municipal wastes and industrial discharges from the north lake shore to the sea. Thus, Site 1 samples typically contain more contaminants originating from diverse sources, including domestic wastewater treatment effluent, industrial wastewater discharges, agricultural runoffs, and others.

**Seasonal variations of the water quality**

The seasonal variation of the water quality in Lake Taihu area is clearly demonstrated by the results of ANOVA (Table 1), PCA by seasons (Figure 3(b)) and CDOM fluorophores variation (Figure 6). Based on the ANOVA results, the organic content, including DO, CODcr, and DOC, varies by season, whereas nutrients such as TN, NH3-N, and PO4-P show smaller seasonal changes. This difference may be explained by the nutrient deposit in the lake and river sediments (Yu et al. 2007). On the other hand, more precipitation during the wet season results in higher flow rates and pollutant dilution in the water system (Zhang...
However, in the shallow Lake Taihu with hypereutrophic conditions, the flush water likely resulted in the resuspension of the sediments containing decayed algae and microbes, which in turn resulted in the release of nutrients back to the water. The balance of the dilution, deposition and release of contaminants in Lake Taihu complicates the determination of the nutrient behavior in the lake and necessitates a more comprehensive and detailed investigation.

As components of DOM, the CDOM fluorophores should have similar responses to seasonal variations with common WQIs. The C1, C3 and C4 components had lower concentrations during the wet season and higher concentrations in the dry season (Figure 6), which is similar to the results of other studies in Lake Tianmuhu in South Eastern China (Zhang et al. 2011), with the exception of C2, which recorded higher loadings only in January and February in 2009. The reason for the abnormal seasonal variation of C2 is still unknown and requires further investigation. C2 was identified in previous studies as a partially marine humic-like source (Yao et al. 2011), which is assumed to be the result of a microbial DOM processing via internal cycling rather than via an external input (Yamashita et al. 2008).

**CDOM fluorophore properties**

The 3DEEM spectra results (Table 2) used as reference in previous studies worldwide are consistent with the component analyses in previous studies in China, such as those on Lake Taihu (Zhang et al. 2009; Yao et al. 2011), Lake Tianmuhu (Zhang et al. 2011), coastal rivers in the same Jiangsu province (Gao et al. 2010), and on Lake Hongfeng (Wang et al. 2009) and the coast of the Pearl River in nearby provinces (Callahan et al. 2004; Hong et al. 2005; Zhao et al. 2009). Free amino acids or bonded protein-like sources (C1 and C4 in the current study) were found dominant in the CDOM fluorophores of most of the previous studies on inland water in China. These substances represent protein-like components of biological processes assisted by microbial enzymes in the lake. This early observation is supported by results from the present study, with the fluorescence intensity average for all samples recorded as 1,367, 452, 508 and 1,300 RU (standardized using the peak height of Raman scattering of pure water as 10,000 RU) for components C1, C2, C3 and C4, respectively. The fluorescence intensities of C1 and C4 were 73.5% of the total amount, indicating that CDOM in the Lake Taihu area is dominated by free amino acids or bonded protein-like components related to local biological activities, such as the generation and decay of plankton and algae (Zhang et al. 2009).

Humic substances (C2 and C3) were ubiquitous in the water environment and might have diverse sources such as planktonic production, microbial degradation and external upstream input. Previous studies suggested that component C2 in Lake Taihu represents the humic-like fluorophores produced by non-fluorescent matter from phytoplankton (Yamashita et al. 2008; Zhang et al. 2009), whereas C3 is a typical terrestrial humic substance (Yao et al. 2011).

A positive relationship between the fluorescence intensities of the CDOM fluorophores was found (Figure 8). This relationship was observed between C1 to C4 and C3 for all investigated samples ($R^2 = 0.77$ and 0.47, respectively), supporting the hypothesis that they may have similar sources and exhibit similar environmental behavior. On the other hand, the correlation between C1 and C2 is generally negative, implying that the source and behavior of component C2 are quite different from those of the other components. As previously described, component C2 may be strongly associated with the death and decay of algae and plankton via internal substance cycling, whereas other components may have been affected by both internal microbial activity and external inputs.

**Correlations between WQIs**

Table 4 shows the bivariant correlation results used to analyze the pollutant sources. According to the data, strong correlations existed between C1, C3 and C4 ($p < 0.01$), indicating that a number of common factors controlled their concentrations. However, no correlations between the CDOM component loadings and organic concentrations such as COD and DOC were found, which can be explained by the large portion of non-fluorescent components in the organic matter.
By considering the CDOM sources of microbial activities in the lake, a positive correlation (Table 4) between the primary components (C1, C3 and C4) and the concentrations of nutrients such as TN, ammonia and TP was found, implying that the CDOM concentrations are associated with environmental N and P, which are essential for biological activities. Figure 9 shows the relationship between the spatially averaged fluorescence intensities of the sum of the primary components (C1 + C3 + C4) and the NH3-N and TP concentrations.

![Figure 8](image1.png)  
**Figure 8** | Relationship between the fluorescence intensities of CDOM components.

![Figure 9](image2.png)  
**Figure 9** | Relationship between the fluorescence intensities of (a) ammonia, and (b) total phosphorous.

### Table 4 | Bivariant correlation coefficients for the WQIs and CDOM component loadings (C1–C4)

<table>
<thead>
<tr>
<th></th>
<th>DO</th>
<th>CODc</th>
<th>DOC</th>
<th>TN</th>
<th>NH₃-N</th>
<th>NO₂-N</th>
<th>NO₃-N</th>
<th>TP</th>
<th>PO₄-P</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>-0.12</td>
<td>0.20</td>
<td>-0.01</td>
<td>0.29&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.37&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.16</td>
<td>-0.02</td>
<td>0.34&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>0.26&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.02</td>
<td>0.16</td>
<td>0.07</td>
<td>0.09</td>
<td>-0.01</td>
<td>-0.15</td>
<td>-0.09</td>
<td>0.12</td>
<td>-0.27&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>0.08</td>
<td>0.09</td>
<td>0.00</td>
<td>0.31&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.33&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.02</td>
<td>-0.01</td>
<td>0.20</td>
<td>0.04</td>
<td>0.69&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>-0.23</td>
<td>0.16</td>
<td>-0.03</td>
<td>0.22</td>
<td>0.29&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.10</td>
<td>-0.01</td>
<td>0.29&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.06</td>
<td>0.89&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.24&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.45&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Note: *p < 0.05; **p < 0.01.
For ammonia, a strong linearity was observed during the spring, summer and autumn seasons ($R^2 = 0.90$, Figure 9(a)), with a large diversity during the winter season. One explanation for this difference may be the reduced amount of precipitation during winter as well as possible external sources from discharged wastewater. Although the points are scattered and show only a weak linearity, a positive correlation between the average fluorescence intensity and TP was found ($R^2 = 0.44$, Figure 9(b)), indicating possible similar sources from biological processes in the lake. All these results demonstrate that the combined WQI analysis and 3DEEM spectrum interpretation can contribute to the understanding of the environmental behavior of organics and nutrients in inland surface water systems.

**CONCLUSIONS**

Sophisticated statistical analysis tools, including PARAFAC modeling of 3DEEM, were used to interpret the dataset of a 1 year seasonal survey and obtain a better understanding of organics and nutrients in the Lake Taihu area. The following conclusions are drawn:

1. The combination of ANOVA, cluster analysis and PCA of WQIs confirmed the spatial distribution and seasonal variation of pollutants in the surface water around Lake Taihu. Among all pollutants, N and P behaviors require further investigation to control the pollution and restore the local surface water environment.

2. Four CDOM fluorophores were extracted from acquired 3DEEMs via PARAFAC analysis, including two autochthonous biodegradation products (C1 and C4) and two humic-like substances (C2 and C3).

3. The CDOM fluorophores showed similar spatial and seasonal variations with common WQIs, with the exception of the seasonal response of C2. The abnormal behavior of C2 in winter requires further investigation.

4. Although the fluorescence intensities of the CDOM fluorophores were not correlated with the concentrations of the total organic matter, positive correlations between these fluorescence intensities and the concentrations of nutrients such as NH$_3$-N and TP in the surveys were found.

**ACKNOWLEDGEMENTS**

This study was supported by a project in Tsinghua-Toshiba Cooperation Frame and partially funded by the Major Projects on Control and Rectification of Water Body Pollution (2009ZX07313-003). The authors would like to thank the editor Dr Susan B. Watson and the three anonymous reviewers for their valuable time and full support in improving the quality of this paper.

**REFERENCES**


Kowalczyk, P., Durako, M. J., Young, H., Kahn, A. E., Cooper, W. J. & Gonsior, M. 2009 Characterization of dissolved organic matter fluorescence in the South Atlantic Bight with


First received 2 June 2011; accepted in revised form 12 January 2012