Diatoms from the Pearl River estuary, China and their suitability as water salinity indicators for coastal environments

Yongqiang Zong a,*, Andrew C. Kemp b, Fengling Yu c, Jeremy M. Lloyd c, Guangqing Huang d, Wyss W.-S. Yim a

a Department of Earth Sciences, The University of Hong Kong, Hong Kong, SAR China
b Department of Earth and Environmental Science, University of Pennsylvania, Philadelphia, PA, USA
c Department of Geography, University of Durham, Durham, UK
d Guangzhou Institute of Geography, Guangzhou, China

A R T I C L E  I N F O
Article history:
Received 7 May 2009
Received in revised form 10 February 2010
Accepted 15 February 2010

Keywords:
Diatoms
Salinity
Transfer function
Pearl River estuary
Freshwater discharge

A B S T R A C T
We collected 77 modern diatom samples from sites across the Pearl River estuary, China to analyze the relationship between diatom assemblages and environmental parameters including water salinity, water depth and sediment particle size. Results showed that marine diatoms were dominant in the high salinity environment around Hong Kong and the outer part of the estuary. Brackish water diatoms were found in high abundance in the central part of the estuary. Both marine and brackish water diatoms were predominantly planktonic taxa. Freshwater diatoms dominated in low salinity environments, with planktonic taxa in the deep tidal channel and benthic species in the shallow deltaic distributaries. Statistical tests indicated that the modern diatom distribution is strongly correlated with salinity but is also influenced by several other environmental variables including sand content and water depth. Transfer functions relating diatom assemblages and water salinity had high $r^2$ (0.94–0.98 for WA-PLS, 0.95 for MA and 0.98 for MAT) and relatively low RMSEP (2.66–1.63% for WA-PLS, 2.35% for MA and 2.70% for MAT). Due to the geographical distribution of samples, some spatial autocorrelation is likely present in the dataset. When this effect is considered, $r^2$ decreases to 0.90 and RMSEP increases to 5.41%, although the diatom–salinity relationship remains appropriate for reconstructing paleosalinity. Based on this estimate, a diatom-based salinity transfer function with high accuracy and precision is developed and successfully applied to a sediment core for quantitative reconstruction of the Holocene paleosalinity in the Pearl River estuary.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Variability in freshwater river discharge is an important factor that causes changes in the sedimentary regime, salinity and bio-productivity within estuarine environments (Chen et al., 2007; Zhang et al., 2008; Zong et al., 2009a). Since freshwater discharge is closely related to precipitation patterns within a given river catchment, changing freshwater flux can be used to infer changes in climate conditions. In the case of Asian Monsoon, long-term changes in summer precipitation drive fluctuations in the freshwater flux entering an estuary (Clift and Plumb, 2008). During stronger summer monsoons a greater amount of freshwater is discharged into an estuary, and vice versa (Zong et al., 2006). Holocene changes in summer monsoon strength have also resulted in modified terrestrial sediment supply, and directly impacted the rate of deltaic progradation (Zong et al., 2009a). Between 6800 and 2000 cal. years BP, for example, the rate of deltaic shoreline progradation decreased gradually as a result of a weakening summer monsoon (Zong et al., 2009a). Such reconstructions of temporal changes in paleo-monsoon strength are important for improving the predictive power of global and regional climate models. Given the recent problems with winter freshwater supply in the Pearl River region, accurate knowledge of variability in freshwater discharge in the recent past is critical for assisting local communities to plan better for their future freshwater consumption. Such knowledge will also be useful to societies living on the coasts of East Asia and beyond, because water supply is becoming an important issue for human society in the 21st century (Zhang et al., 2008).

Most recent studies regarding Holocene Asian summer monsoon strength are based on loess sequences (e.g. An et al., 2004), marine deposits for Sea Surface Temperature (SST; e.g. Wang et al., 2005), lake sequences for local productivity and vegetation history (e.g. Mingram et al., 2004; Shen et al., 2006), and stalagmites (e.g. Wang et al., 2008). Few attempts have been made to reconstruct the history of monsoonal driven freshwater flux. Zong et al. (2006) first explored the possibility of using diatoms as an indicator of freshwater flux in a monsoon region, and examined the qualitative relationship between diatom assemblages and water salinity. Zong et al. (2009a) further successfully applied the diatom technique to reconstruct the sedimentary history of the Pearl River delta.

* Corresponding author. Tel.: +852 22194815; fax: +852 25176912.
E-mail address: yqzong@hkuc.hku.hk (Y. Zong).

0377-8398/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.marmicro.2010.02.004
Fig. 1. Maps show the Pearl River catchment (a), the Pearl River estuary and sampling locations (b) as well as winter and summer water salinity across the estuary (c and d).
Elsewhere (in the UK, USA, Japan and Southeast Asia), diatoms have been widely used to reconstruct sea-level history (e.g. Long, 1992; Shennan et al., 1993, 1994; Zong and Tooley, 1996; Zong, 1997b, 1998), ice sheet dynamics (e.g. Stabell, 1980; Long et al., 1999, 2008), earthquake induced land movements (Shennan et al., 1996, 1999; Long and Shennan, 1998; Sawai, 2001, 2009) and coastal evolution (Juggins, 1992; Zong, 1992, 1997a,b, 1998; Zong and Tooley, 1999). The success of these studies was based upon understanding the modern relationship between diatoms and coastal environments (e.g. Hendey, 1964; Sullivan, 1975, 1978; Denys, 1991/2; Vos and de Wolf, 1993; Nelson and Kashima, 1993; Hemphill-Haley, 1995; Zong, 1997a; Zong and Horton, 1998; Zong and Kamaludin, 2004; Sawai, 2001). Furthermore, reconstruction of coastal environmental change using fossil diatoms has become more precise and accurate when statistical methods such as transfer functions are used to establish the diatom-environment relationship quantitatively (e.g. Zong and Horton, 1999; Zong et al., 2003; Hamilton and Shennan, 2005; Horton et al., 2007).

In order to improve the reconstruction of monsoon-induced freshwater discharge using diatoms as explored by Zong et al. (2006), it is important to test the relationship between diatoms and a number of environmental factors including water salinity, water depth and the nature of substrate, as these factors influence population change and spatial distribution patterns of diatoms. For instance, planktonic diatoms distribute according to water salinity from tidal influenced river channels through estuarine environments to marine waters (Juggins, 1992; Zong et al., 2006), whilst benthic diatoms are sensitive to the combined effect of water salinity, substrate sediment type and ground altitude, the latter in the form of tidal inundation frequency (e.g. Nelson and Kashima, 1993; Hemphill-Haley, 1995; Zong and Horton, 1998; Zong and Kamaludin, 2004). The relationship between diatoms and water depth has not been examined, but preliminary results from Zong et al. (2006) suggested that the percentage of benthic diatoms increased toward the distributaries where water is shallower. Therefore, prior to reconstructing water salinity and monsoonal freshwater discharge in the Pearl River using diatom-based salinity transfer functions it is necessary to examine modern diatom distribution patterns and the relative influence of other environmental factors. In this paper we document the distribution patterns of modern diatoms across the Pearl River estuary and analyze the statistical relationship between diatom assemblages and a number of other appropriate environmental variables. We develop and test the applicability of the first diatom-based salinity transfer functions to aid reconstruction of freshwater flux history from estuarine sediments of the Pearl River estuary and beyond.

2. Study area

The Pearl River drains a latitude-orientated catchment between 26°N and 22°N (Fig. 1a), which is a transitional region between tropical and temperate zones under the influence of humid summer monsoon from the south and dry winter monsoon from the north. Three main rivers (the East River, the North River and the West River) drain into the drowned coastal basin (Zong et al., 2009b) and create two deltaic complexes which are separated by the estuary of about 1740 km² in area (Fig. 1). The estuary is protected from storm waves by a cluster of offshore rocky islands. Wave energy within the estuary is low, except during the passage of a typhoon. The main tidal channel runs directly south between Hong Kong and Macau into the South China Sea. A secondary tidal channel runs southeast, through Hong Kong, and turns south into the sea.

At present the annual average precipitation in the catchment area is between 1600 and 2000 mm/yr, of which more than 80% falls during spring and summer, indicative of a warm humid summer and a dry cool winter. Such seasonal differences in precipitation result in freshwater discharges as low as 2000 m³/s during a dry winter and up to 46,300 m³/s in a 100-year flood event (Huang et al., 2004). On average, the Pearl River discharges 302,000 × 10⁶ m³ of water and 83.4 × 10⁶ tons of suspended sediment each year (Zong et al., 2009a). However, recent construction of reservoirs in the catchment area has reduced the sediment load to only 54.0 × 10⁶ tons a year since the mid-1990s (Zhang et al., 2008). Due to the easterly offshore currents, freshwater runoff tends to concentrate on the western side of the estuary and flows southwestwards when it reaches offshore Macau. The range of spring tides within the estuary varies from 0.86 m at the mouth to 1.57 m at the head of the estuary (Huang et al., 2004). Despite the small tidal range, the average volume of flood tides is as high as 73,500 m³/s, which is nearly 13 times the average freshwater discharge.

3. Methods

3.1. Modern sediment samples and environmental variables

Surface sediment samples were collected from 77 locations across the estuary (Fig. 1b) for the analysis of diatom assemblages and particle size distribution through two winters and one summer between 2004 and 2006. At each sampling location, a grab sampler was used to capture the upper 10 cm of sediment. Water depth at each sampling site was obtained in relation to the national datum (YSD; Yellow Sea Datum) using a measuring rope and onboard echo-sounding readings where water depth was greater than 5 m, taking into account the timing of tides. Corrections to water depth measurements were made according to tidal data from the nearest tide gauges (there are eight tide gauges within the estuary). As the tidal range across the estuary is relatively small, the accuracy of water depth measurements was not compromised. Water salinity was recorded using an electrical salinity probe. The water salinity measurements of surface water from both winter and summer field work were compared with the winter and summer values produced for the estuarine and deltaic parts in 1980–1983 by the Guangzhou Institute of Geography which conducted repeated measurements of surface water salinity in both winter and summer seasons, and for Hong Kong waters in 2004–2006 by the Environmental Protection Department of the Government of Hong Kong who made repeated surveys at over 40 locations four times a year. Based on these data, we developed the salinity contours for winter and summer seasons (Fig. 1c, d) using ArcGIS and estimated the average annual water salinity for each sampling site, which is used for the statistical analysis. Details of the environmental variables are presented in Appendix A.

3.2. Diatom data

Samples for diatom analysis were mixed to average the seasonal variability (e.g. Horton et al., 2007; Zong, 1997a,b; Zong and Horton, 1999; Zong et al., 2003) and prepared following standard procedures. Diatom valves and frustules were identified under a microscope with 1000× magnification to the species level according to key references (e.g. van der Werff and Huls, 1958–1966). The taxonomic and ecological information of key taxa are listed in Appendix B. Diatom counts were expressed as percentage of total diatom valves (Appendix A). The diatom assemblages at each site are likely to be a mixture of locally

---

Table 1: Summary of the environmental variables recorded for each sampling sites.

<table>
<thead>
<tr>
<th>Group</th>
<th>Water salinity (%e)</th>
<th>Water depth (m)</th>
<th>Sand content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (5 samples)</td>
<td>3.7±2.0</td>
<td>2.2±2.6</td>
<td>340±6.8</td>
</tr>
<tr>
<td>B (9 samples)</td>
<td>3.7±3.7</td>
<td>2.6±0.7</td>
<td>32.6±21.5</td>
</tr>
<tr>
<td>C (10 samples)</td>
<td>8.0±4.8</td>
<td>3.6±1.2</td>
<td>23.5±11.1</td>
</tr>
<tr>
<td>D (13 samples)</td>
<td>17.3±4.1</td>
<td>6.4±3.1</td>
<td>20.3±15.6</td>
</tr>
<tr>
<td>E (7 samples)</td>
<td>21.7±3.1</td>
<td>10.3±6.1</td>
<td>20.3±10.6</td>
</tr>
<tr>
<td>F (15 samples)</td>
<td>25.6±4.3</td>
<td>10.1±4.2</td>
<td>16.4±2.9</td>
</tr>
<tr>
<td>G (7 samples)</td>
<td>30.6±0.9</td>
<td>14.2±11.3</td>
<td>21.3±14.6</td>
</tr>
<tr>
<td>H (7 samples)</td>
<td>31.8±1.1</td>
<td>13.7±4.0</td>
<td>15.0±6.5</td>
</tr>
<tr>
<td>I (4 samples)</td>
<td>33.4±0.2</td>
<td>27.0±3.2</td>
<td>13.7±7.2</td>
</tr>
</tbody>
</table>
Fig. 2. The diatom diagram (a) shows the frequencies of 17 most abundant diatom taxa. Graph (b) shows results of the Detrended Correspondence Analysis (DCA) on the modern diatom data.
produced taxa and transported taxa (e.g. Horton et al., 2007; Zong, 1997a,b; Zong and Horton, 1998). Such a mixture of allochthonous and autochthonous diatom valves would also occur in sediments that were deposited in the past (Zong, 1998; Sawai, 2001). Thus, we did not attempt to separate the allochthonous component from the diatom assemblages.

Core V37 was collected at the mouth of the Pearl River (Fig. 1b) using a 10-cm-diameter push corer onboard of a floating platform, the seabed altitude was 1.5 m below mean sea level. Core samples were prepared for diatom analysis using the same methods as modern sediment samples. AMS radiocarbon dates were obtained from Beta Analytic, Florida, using benthic foraminifera assemblages from four different depths (Zong et al., 2006).

To display the spatial distribution of diatoms across the estuary, we used the kriging function in ArcGIS to perform spatial interpolation based on the diatom data. The outputs show relative abundance of key diatom taxa. To detect, describe and classify the distribution of diatoms across the estuary in comparison to environmental variables, we used unconstrained cluster analysis based on unweighted Euclidean distance (CONISS) within the TILIA package (Grimm, 1993) and Detrended Correspondence Analysis (DCA) (Hill and Gauch, 1980). These two methods are complementary (Birks, 1992) and help group similar samples together and set apart dissimilar samples. The cluster analysis helped identify groups of samples having similar diatom assemblages, whilst DCA provides further information about the pattern of variation within and between groups (ter Braak and Smilauer, 1997–2003).

To test the relative strength of relationships between diatoms and environmental variables (salinity, water depth and sand content), we applied canonical correspondence analysis (CCA) with Monte Carlo tests of 199 random permutations (ter Braak and Smilauer, 1997–2003). We also applied Detrended Canonical Correspondence Analysis (DCCA) to determine if the diatom data set was unimodal. To develop a diatom-based water salinity transfer function, we used unimodal-based regressions (Birks, 1995; Gasse et al., 1997; Juggins, 1992) known as Weighted Averaging with Partial Least Squares (WA-PLS) (ter Braak and Juggins, 1993) using the program C2 (Juggins, 2004). This analysis used only species with an abundance of over 2%. The diatom data set was not transformed or standardized (Horton et al., 2007).

Our samples were purposely collected at sites close to one another in order to capture salinity variations across a known environmental gradient from distributaries through the estuary and into the shallow marine environment. Such non-random sampling regimes are appropriate given the purpose of this study, but are likely to include the influence of spatial autocorrelation. This may produce unrealistically optimistic estimates of transfer function performance (Telford and Birks, 2005). We applied the graphical method developed by Telford and Birks (2009) to investigate the influence of spatial autocorrelation on our dataset. This procedure compares cross-validated transfer function performance when a subset of sites is deleted randomly (the most common method of evaluation) against performance when samples are deleted geographically or based upon environmental similarity. We investigated the result of removing samples at 5 km intervals from 0–50 km (the maximum distance between any two samples in the study area was approximately 110 km).

4. Results

4.1. Environmental variables

As shown in Fig. 1c, d, water salinity is higher on the eastern side of the estuary. At the estuary mouth higher salinities were recorded in the waters east of Hong Kong, and decreased westward. This spatial pattern was particularly clear during summer as a result of the easterly offshore current. It was less clear in winter because the northeasterly winds push the offshore current further south. At the head of the estuary, salinity was generally lower, and close to zero within the deltaic distributaries. On annual average, water salinity within distributaries and around the head of the estuary (sample groups A, B and C) was generally low but varied between seasons (Fig. 3a). In the middle part and the mouth area of the estuary (sample groups D, E and F), water salinity varied between c. 15 and 30‰ depending on locality, and fluctuated between seasons. In the shallow marine environment of Hong Kong (sample groups G, H and I), water salinity changed little between sites or between seasons.

Waters within the deltaic distributaries and around the head of the estuary were shallow (mostly less than 10 m; Fig. 3b). Water depths became deeper seawards across the estuary but varied greatly depending on locality, i.e. subtidal flats or tidal channels, particularly within sample group G (Table 1). Water depths were deepest at sample locations east of Hong Kong (Fig. 3b). Sand content was generally higher in samples from the deltaic distributaries (groups A and B) and lower in locations of shallow marine environment (groups H and I), however, it varied greatly within each group of samples (Appendix A).
4.2. Spatial distribution of modern diatoms

From the surface sediment samples (Fig. 1), we found 73 diatom taxa in total, with 57 species being abundant over 2% and with 17 taxa exceeding 10% in abundance. The frequencies of these 17 diatom species are presented in Fig. 2, and the 77 sediment samples were divided into 9 groups based on the cluster analysis (CONISS). Among these dominant taxa, freshwater benthic diatoms (Cymbella affinis, Fallacia subhamulata, Gomphonema parvulum and Syndra ulna) dominated group B of sediment samples (Fig. 2a) which were from the shallow sections of the deltaic distributaries (Figs. 1b, 4b). Freshwater planktonic diatoms (Cyclotella meneghiniana, Cyl. radiosa and Aulacoseira granulata), however, featured strongly in groups A and C (Fig. 2a) or the inner part of the estuary and the main tidal channel that reaches Guangzhou (Fig. 4b). It is noted that A. granulata also appeared in high numbers in the central part of the inner estuary (group E). This species seems able to follow the outflow of freshwater further into the brackish water area. Thus, the spatial distribution patterns of the freshwater diatoms suggest that they are strongly associated with water salinity. However, to some extent, water depth is also a determining factor in the distributional patterns of benthic and planktonic taxa. This is illustrated by the average water depth for each zone (Table 1, Fig. 3b).

The most abundant brackish water diatoms were Coscinodiscus blandus, Coscinodiscus divisus, Cyclotella striata and Cyl. stylum. These taxa are planktonic and were concentrated in groups D, E, F and G (Fig. 2a), in the central and outer parts of the estuary (Fig. 4c, d). The dominant marine species were also planktonic, with Chaetoceros radians, Paralia sulcata, Skeletonema costatum and Thalassiosira nitzschioides concentrating in groups F, G, H and I (Fig. 2a). In particular, C. radians and S. costatum appeared most abundant in the area of southern Hong Kong (Fig. 4e). T. nitzschioides was most abundant in the four sediment samples southeast of Hong Kong, and its abundance decreased towards the inner estuary (Fig. 4f).

CCA results showed clear divisions of modern diatom assemblages. In the DCA bi-plot (Fig. 2b), sediment samples from groups A, B and C with dominant freshwater diatoms are plotted on the left hand side, with the two brackish water groups (D and E) plotted in the central part, followed by the two brackish-marine groups (F and G) on the right hand side, and finally the two marine groups (H and I) on the far right of the diagram. Overall, the spatial distributions of the diatom samples showed a strong correlation with water salinity from the freshwater to marine parts of the study area, as shown by axis 1 which was 3.60 in standard deviation unit length and represented 29.7% of total variance in the diatom data. Axis 2 represented 11.2% of total variance in the diatom data. The contribution from the inter-correlation between environmental variables was also low. However, about two-thirds of the variance (57.6%) was unexplained and it is unclear what other factors are involved. It is noted that there is a certain degree of variation along the DCA axis 2 among the sediment samples which contain dominantly freshwater diatoms (Fig. 2b). Thus part of the unexplained variance in the diatom data may involve nutrient supply from agricultural and industrial influence. Nevertheless, the unexplained percentage is considerably lower than reported from other published biological data (Gasse et al., 1997; Jones and Juggins, 1995; Zong and Horton, 1999).

6. The diatom-based salinity transfer function

In order to determine whether or not a unimodal-based technique such as WA-PLS should be used, Detrended Canonical Correspondence Analysis (DCCA) was performed. The DCCA result indicated that axis 1 is highly correlated with salinity (weighted correlation, r = 0.97) and represents 85.7% of total variance in the species-environment relationship. The length of gradient was long (3.14 standard deviations). Based on such a strong relationship between diatoms and salinity recorded in the Pearl River estuary, a transfer function was developed. Results from WA-PLS with cross validation indicated that prediction \( r^2 \) was high for the five components produced (between 0.94 and 0.98, Table 3). Such high correlation between the recorded diatom assemblages and the observed salinity without outliers (Fig. 6a) suggested a strong predictive power for the transfer function developed. The prediction \( r^2 \) reported here is similar to, or a little higher than, many comparable studies, e.g. diatom-based transfer functions for estuarine salinity (e.g. Juggins, 1992), nutrient level (e.g. Anderson and Vos, 1992), lake water salinity (e.g. Fritz et al., 1991), acidity (e.g. Battarbee et al., 2005) and sea level (e.g. Ng and Sin, 2003; Zong and Horton, 1999; Zong et al., 2003; Horton et al., 2007).

The Root Mean Squared Error of Prediction (RMSEP) varied between 2.66 and 1.63% according to the WA-PLS model (Table 3), suggesting that the transfer function was capable of producing predictions for paleosalinity reconstruction with reasonable precisions (Fig. 6b). The predicted error range was comparable with many other diatom-based transfer functions (e.g. Battarbee et al., 2005, 2008; Horton et al., 2007; Zong and Horton, 1999; Zong et al., 2003) in terms of the error range proportional to the range of the environmental gradient studied. Estuaries are dynamic environments where fluvial and marine waters are mixing and exchanging. As such, diatoms are transported between environments by tidal currents and floodwater and introduce an allochthonous component to recorded assemblages (Figs. 3, 4). In this context average error of c. 2% was considered as a very good result.

These results must be considered in light of the potential influence of spatial autocorrelation, which may produce optimistic estimates of the
criteria used for assessing transfer function performance. We used the graphical method proposed by Telford and Birks (2009) to investigate the influence of spatial autocorrelation on our dataset (Fig. 7). The first test shows the effect on $r^2$ of deleting samples at random during cross validation (the most common approach) against removing samples geographically or on the basis of environmental similarity. Divergence between $r^2$ values estimated using these approaches is indicative of spatial autocorrelation. For the dataset we present in this study, random deletion causes a clear decrease in $r^2$ when more than 75% of samples were removed (Fig. 7a). Exclusion of samples on the basis of geographic location or environmental similarity shows divergence in $r^2$ values from the random curve (Fig. 7a) after 5 km, indicating that some degree of spatial autocorrelation is present in the dataset. At distances greater than 20 km there is an additional divergence between $r^2$ values estimated by excluding samples on the basis of environmental similarity rather than distance (Fig. 7a), with the former showing a poorer performance.

If spatial autocorrelation is recognized in a dataset, Telford and Birks (2009) proposed using a modified $h$-block cross-validation scheme to provide more robust estimates of transfer function performance. In this approach samples within $h$-km are excluded where $h$ is estimated from the range of a fitted variogram model. We visually estimated a value of 25 km from the variogram for our dataset as the fitted range (1857 km) was not ecologically plausible and left no samples to utilize in cross validation (Fig. 7b). This distance also corresponds approximately to the radius of the diatom groups identified previously (Fig. 1). Using $h$-block

---

**Table 2**

Summary of CCA results from modern diatom assemblages.

<table>
<thead>
<tr>
<th>Axis</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total inertia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvalues</td>
<td>0.522</td>
<td>0.046</td>
<td>0.017</td>
<td>0.300</td>
<td>1.807</td>
</tr>
<tr>
<td>Species-environment correlations</td>
<td>0.974</td>
<td>0.570</td>
<td>0.496</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Cumulative percentage variance of species data</td>
<td>28.9</td>
<td>31.4</td>
<td>32.4</td>
<td>49.0</td>
<td></td>
</tr>
<tr>
<td>Cumulative percentage variance of species-environment relationship</td>
<td>89.2</td>
<td>97</td>
<td>100.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Sum of all unconstrained eigenvalues</td>
<td>1.807</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum of all canonical eigenvalues</td>
<td>0.586</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Table 3**

Correlation and prediction errors based on jack-knifed cross validation of transfer functions and $h$-block cross validation.

<table>
<thead>
<tr>
<th>Method</th>
<th>Prediction ($r^2$ jack-knifed)</th>
<th>RMSEP (‰)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WA-PLS (component 1)</td>
<td>0.940</td>
<td>2.66</td>
</tr>
<tr>
<td>(component 2)</td>
<td>0.967</td>
<td>1.96</td>
</tr>
<tr>
<td>(component 3)</td>
<td>0.973</td>
<td>1.80</td>
</tr>
<tr>
<td>(component 4)</td>
<td>0.976</td>
<td>1.68</td>
</tr>
<tr>
<td>(component 5)</td>
<td>0.977</td>
<td>1.63</td>
</tr>
<tr>
<td>MAT</td>
<td>0.976</td>
<td>2.70</td>
</tr>
<tr>
<td>MA</td>
<td>0.953</td>
<td>2.35</td>
</tr>
<tr>
<td>$h$-block (25 km)</td>
<td>0.897</td>
<td>5.41</td>
</tr>
</tbody>
</table>

---

**Fig. 5.** Results of the Canonical Correspondence Analysis are shown as bi-plots of species against environmental variables (a) and samples against environmental variables (b).

**Fig. 6.** Results of the diatom-based salinity transfer function (component 5) are shown as: (a) the diatom-predicted salinity against the observed salinity and (b) the residuals of the regression using weighted averaging partial least squares (WA-PLS).
cross validation of a MAT transfer function, we estimated $r^2$ to be 0.90 and RMSEP to be 5.41‰. This test demonstrates that spatial autocorrelation within the dataset does cause an overestimate of $r^2$ (compared to 0.98) and particularly RMSEP (compared to 2.7‰), although this trend may have been exaggerated through use of MAT rather than WA-PLS during h-block cross validation (Telford and Birks, 2009). Additional cross-validation results using values from 0–50 km for $h$ are presented in Fig. 7c. Despite apparent spatial autocorrelation, this dataset remains a powerful tool for reconstructing former salinities as the $r^2$ remains high and RMSEP is still appropriate over an environmental gradient of 33.5‰ (standard deviation 10.9‰). The application of this dataset to reconstructing paleosalinity in the Pearl River estuary using a transfer function likely remains accurate, although an additional error term should be included in salinity reconstructions to account for the influence of spatial autocorrelation.

7. Application of the transfer function to core V37

To demonstrate the usefulness of the transfer function, core V37 was selected for paleosalinity reconstruction. The core was recovered from the eastern side of the mouth of the Pearl River estuary (Fig. 1b), in a prodelta environment where continuous sediment accretion over the Holocene was recorded (Zong et al., 2009a). Holocene sediments overlay a thin layer of gravel and silt at 10.1 m and consisted of soft, dark greenish-grey silt and clay, with occasional shell fragments. A chronology was established with radiocarbon dates on calcareous foraminifera from four depths (Zong et al., 2006). For this study, diatom samples were taken from the core at 0.20 m intervals, except for the second and sixth meters which were removed for engineering tests.

Diatom techniques have been used for qualitative reconstruction of paleo-environmental history and coastal changes in the Pearl River delta area (e.g. Zong et al., 2006, 2009a,b). These studies relied on the classification of diatom species according to their salinity preferences (e.g. Denys, 1991;2; Vos and de Wolf, 1993) and interpretation of changing ratios between marine, brackish water and freshwater diatoms. From the base of the core to 7.0 m (Fig. 8), the percentages of marine diatoms reduced and is reflected particularly in the frequencies of P. sulcata, a marine planktonic species (Zong, 1997a), whilst brackish water diatoms (including Cos. divisus and Cyc. striata) increased, indicating a change from marine to brackish water conditions. This change (from 8550 to 7620 cal. years BP) coincided with a period of rapid sea-level rise (Zong, 2004; Bird et al., 2007) and the Pearl River mouth retreating landwards (Zong et al., 2009a). If the amount of freshwater discharge in this period remained constant, an increase in water salinity would be expected. Thus the decrease in marine diatoms for this period can be interpreted as a result of an increase in freshwater discharge. From 7.0 m to 2.0 m depth, all major taxa fluctuate in their abundance, with total marine diatoms increasing gradually (Fig. 8). This is a period of relatively stable sea level and deltaic shoreline progradation (Zong et al., 2009a) making the core site closer to the river mouth. The gradual increase in marine diatoms therefore suggests a decrease in freshwater discharge from 7620 to 3490 cal. years BP. The top meter of the core recorded an increase in marine diatoms, T. nitzschoides (a marine planktonic species) in particular, and a decline of Cyc. striata and Cyc. stylorum, indicating an increase in salinity. This final period saw rapid advances of the deltaic shoreline due to human activity (Zong et al., 2009a), resulting in the core site becoming much closer to the river mouth. Thus the increase in marine diatoms can be attributed to a further decline in freshwater discharge.

We applied the WA-PLS transfer function to diatom assemblages preserved in core V37 to estimate paleosalinity (Fig. 8). Water salinity estimates decreased from 29.0‰ at the base of the core to 20.0‰ at

---

**Fig. 7.** Results from graphical assessment of spatial autocorrelation following the method of Telford and Birks (2009). (a) Effect on $r^2$ of removing samples during cross validation randomly (open circles), geographically (closed circles) and on the basis of environmental similarity (open diamonds). The distances used for exclusion are shown. (b) Variogram used to visually estimate a value of 25 km (grey dashed line) for h-block cross validation. (c) Estimates of MAT transfer function performance using h-block cross validation for values of h between 0 and 50 km. RMSEP is shown by filled circles (left axis) and $r^2$ is represented by open circles on the right axis. The 0 km value is equivalent to standard random deletion of samples.
Fig. 8. Application of the WA-PLS transfer function to a Holocene sedimentary sequence taken from the mouth area of the Pearl River estuary (Fig. 1). There are 7 diatom taxa which are over 10% in abundance. The summary column shows the total marine, brackish water and freshwater diatoms. The reconstruction is reliable according to ‘goodness-of-fit’ (Birks, 1998), i.e. percentages of total diatoms from core sediment that are poorly represented (<5% occurrences) in the modern data set. Reconstructed paleosalinity is presented with sample-specific error terms from random cross validation (open boxes) and a modified estimate of uncertainty based upon analysis of spatial autocorrelation (solid lines).
7.0 m depth. This change is equivalent to the salinity change from the present mouth area to the inner part of the estuary. Such a salinity decrease would have required an amount of freshwater discharge much greater than the present, suggesting strong summer monsoon precipitation. From 7.0 m to 2.0 m in the core, transfer function estimates of paleosalinity show a gradual increase in salinity reflecting a decrease in freshwater discharge associated with a weakened summer monsoon (Wang et al., 2005; Zong et al., 2006) and deltaic shoreline progradation (Zong et al., 2009a). The sharp increase in salinity from 22.5% at 0.8 m to 28.5% at the top of the core may have recorded the recent decline in freshwater discharge due to human activity in the catchment area (Zhang et al., 2007).

Here, both the qualitative reconstruction based on the percentages of major diatom taxa and the quantitative reconstruction using the transfer function show a similar history of salinity change in the Pearl River estuary. However the former can only suggest changes in freshwater discharge in relative terms, yet the latter provides precise estimates of salinity in absolute terms. The transfer function method also assesses whether or not the modern diatom data is suitable for reconstruction of the fossil (core) diatom data according to ‘goodness-of-fit’ (Birks, 1998). Results confirm that the reconstruction is reliable because the percentages of total diatoms from core sediment that are poorly represented (<5% occurrences) in the modern data set are mostly below 10% (Fig. 8). Furthermore, the statistical procedures mentioned in this paper can help to produce estimates of potential reconstruction errors, one from the transfer function and the other from the spatial autocorrelation. The trends in paleosalinity estimates are not compromised by inclusion of an expanded error term associated with spatial autocorrelation (Fig. 8). Thus the diatom-based transfer function established here can potentially help precise and accurate reconstruction of paleosalinity of estuarine environments.

8. Conclusion

We have examined the modern diatom assemblages from 77 sites across the Pearl River estuary. Among these sediment samples, we have identified 73 diatom taxa, with 17 species of over 10% abundance. The majority of marine and brackish water diatoms are planktonic taxa. Marine diatoms appear mostly in the Hong Kong waters and the outer part of the estuary, whilst brackish water diatoms concentrate in the central part of the estuary. Both planktonic and benthic freshwater diatoms appear prominently in samples collected from the low salinity environments. Freshwater planktonic diatoms are recorded from the deep channel, whilst freshwater benthic diatoms are found from two deltaic distributaries. Statistical analyses confirm the strong correlation between diatom assemblages and water salinity. Further analysis indicates a degree of spatial autocorrelation due to the close sampling of sediment within the estuary, and such autocorrelation has shown influence in estimates of the strength of the relationship between diatoms and salinity. Taking into account of the effects of spatial autocorrelation, however, the relationship is still strong enough for accurate reconstruction of paleosalinity. Accordingly, a transfer function is developed and successfully applied to a sediment core.

Acknowledgements

This research is supported by a research grant from the National Science Foundation of China (no. 40771218) to Huang and Zong, two research grants from the Research Grants Council of the Hong Kong SAR, China (no. HKU7058/06P and HKU7052/08P) to Yim and a NERC/EPSRC 05-08 (UK) PhD studentship from the Dorothy Hodgkin Postgraduate Award to Yu. The authors thank the director of the Environmental Protection Department, Hong Kong SAR for the collection of surface sediment samples and water salinity in the Hong Kong area.

Appendix A and B. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.marmicro.2010.02.004.

References


