ORIGINAL ARTICLE

Identification of temporal and spatial variations of water quality in Sanya Bay, China by three-way principal component analysis

Jun-De Dong · Yan-Ying Zhang · Si Zhang · You-Shao Wang · Zhi-Hao Yang · Mei-Lin Wu

Abstract A large data set obtained by a 1-year monthly determination of water quality from Sanya Bay, South China Sea, was treated by three-way principal component analysis aimed at exploring the spatial and temporal patterns of water quality in Sanya Bay. Tucker3 model of optimum complexity (2, 2, 1) explaining 33.18% of the data variance, allowed interpretation of the data information in three modes. The model explained spatial and temporal variation trends in terms of water quality variables during the study period. Water quality in sampling station (S2) Sanya River was mainly influenced by Sanya River, and water quality in other stations (S1, S3–S10) were mainly influenced by the waters in South China Sea. The results delineated the mouth of Sanya River as critical from pollution point of view. The dry season from October to the next April and rainy season from May to September have different influences on water quality in Sanya Bay. The information extracted by the three-way models would be very useful to regional agencies in developing a strategy to carry out scientific plans for resource use based on marine system functions.

J.-D. Dong (✉) · Y.-Y. Zhang · S. Zhang · Z.-H. Yang
LMB, South China Sea Institute of Oceanology,
Chinese Academy of Sciences, 164 West Road Xingang,
510301 Guangzhou, Guangdong, China
e-mail: dongjunde@vip.163.com

Y.-S. Wang · M.-L. Wu
LED, South China Sea Institute of Oceanology,
Chinese Academy of Sciences, 164 West Road Xingang,
510301 Guangzhou, Guangdong, China

J.-D. Dong · Y.-Y. Zhang · S. Zhang · Y.-S. Wang · Z.-H. Yang
Tropical Marine Biological Research Station in Hainan,
Chinese Academy of Sciences, 572000 Sanya, China

Introduction

With the sudden increase of population and rapid economic development in littoral area, coastal water has received large amounts of pollution from a variety of sources such as recreation, fish culture, toilet flushing, and the assimilation and transport of pollution effluents (Bowen and Depledge 2006; Kuppusamy and Giridhar 2006). The coastal water faces many ecological problems, such as eutrophication and environmental pollution (Huang et al. 2003). It is therefore essential and urgent to prevent and control marine water pollution, and regularly implement monitoring programs which help to understand the spatial and temporal variations in coastal water quality. Coastal water quality is largely determined by a number of factors, such as climatic conditions, interaction between land and ocean and anthropogenic activities. Water quality monitoring programs have generated huge databases describing spatial and temporal variations of water quality. The large and complicated data sets consisting of water quality parameters are often difficult to analyze for meaningful interpretation and require data reduction methods to simplify the data structure, in order to extract useful and interpretable information which could explain the spatial and temporal variation patterns of water quality.

Multivariate statistical methods provide powerful tools for extracting useful and interpretable information from large environmental multivariate data arrays. Multivariate methods, such as principal component analysis, have been widely employed in identifying temporal and spatial variation and sources of pollution in coastal water (Yeung 1999;

Yung et al. 2001; Simeonov et al. 2003; Singh et al. 2004; Kuppusamy and Giridhar 2006; Wang et al. 2006; Chau and Muttil 2007; Wu and Wang 2007; Zhou et al. 2007a, b; Wu et al. 2009a). In fact, the water quality monitoring data sets, exhibiting multi-dimensional structure (space, time, variables), require multi-way analysis methods, e.g., three-way principal component analysis to explore and extract the hidden data structure and their relationships. Three-way principal component analysis allows a much easier interpretation of the information contained in the data set, since it directly takes into account its three-way structure. It is impossible for the classical principal component analysis to do it. Three-way principal component analysis has successfully been employed to interpret multi-array data sets in different area, such as food chemistry (Morais et al. 2001; Pravdova et al. 2001, 2002) and environmental studies (Grotti et al. 1999; Leardi et al. 2000; Singh et al. 2006; Giussani et al. 2008; Pardo et al. 2008).

In this paper, three-way principal component analysis has been applied to identify anthropogenic effects and natural character on water quality, to get the environmental information contained in a wide data set of the waters in Sanya Bay. These results could be helpful for regional environmental protection agencies to assess the marine water quality and enhance their pollution control actions in Sanya Bay.

Materials and methods

Study area

Sanya Bay is in the southern part (from $109°20'$ to $109°30'$E, $18°11'$ to $18°18'$N) of Hainan Island, with a water area of 120 km^2 and an average depth of 16 m. It is a typical tropical bay in China. Dongmao Island, Ximao Island and Luhuitou possess mostly coastal coral reefs. The Sanya River, is in the eastern part of the bay (length 31.3 km, drainage area 337 km^2 and annual flow of 2.11×10^9 m^3) (Huang et al. 2003). The wet and warm southwest monsoon prevails in the wet season from April to September, which brings humid air from low latitudes, resulting in gentle monsoonal rainfall in spring and heavy rainfall in summer. In contrast, a dry and cold northeast monsoon predominates in dry season from October to the following March.

In order to evaluate the anthropogenic and nature effects in this bay, there are 11 monitoring stations in Sanya Bay (Fig. 1).

Sampling and analytical method

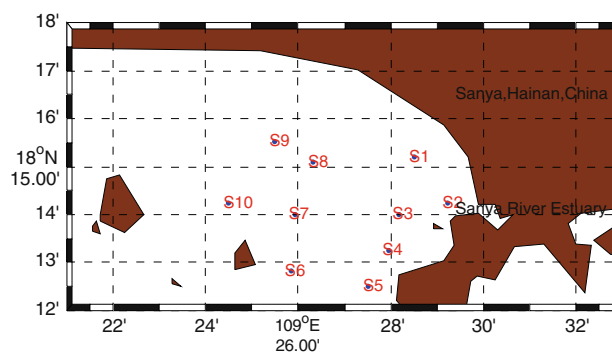Water samples were taken at the surface layer of all stations at month intervals in 2003. A Quanta® Water Quality

Fig. 1 Monitoring stations in Sanya Bay

Monitoring System (Hydrolab Corporation, USA) was employed to collect the data for temperature ($T°C$), pH, salinity (S PSU), specific conductivity (SPC) in the surface layer. Seawater samples for analysis of nutrients were taken using 5-L GO FLO bottles at surface layer, according to the methods and sampling tools of "The specialties for oceanography survey" (GB12763-91, China). Water samples from the surface layer were analyzed for nitrate concentration (NO_3-N/µmol L^{-1}), nitrite concentration (NO_2-N/µmol L^{-1}) and silicate concentration (SiO_3-Si/µmol L^{-1}) with a SKALAR auto-analyzer (Skalar Analytical B.V. SanPlus, Holland). Ammonium concentration (NH_4-N/µmol L^{-1}) was analyzed with methods of oxidized by hypobromite. Phosphorus concentration (PO_4-P/µmol L^{-1}) was analyzed with methods of oxidized by molybdophosphoric blue. Dissolved oxygen concentration (DO/mg L^{-1}) was determined with the method of Winkler titration.

Three-way modeling

Classical PCA could be applied to the data. Results could be difficult to interpret because the information of the three modes can be mixed. As an example, a score plot obtained by a PCA performed on the unfolded matrix 120 (10 stations × 12 months) × 10 (parameters) concerning the data set is reported in Fig. 2. It is evident that the interpretation of this plot is very difficult due to the high overlapping of samples. Information about sampling stations and sampling time is indeed mixed in the score plot.

A three-way modeling was preferred because it directly takes into account the three-way structure of the data, allowing an easy interpretation of the results. The final result is given by three sets of loadings together with a core array describing the relationship among them.

Such a model can be formulated as follows:

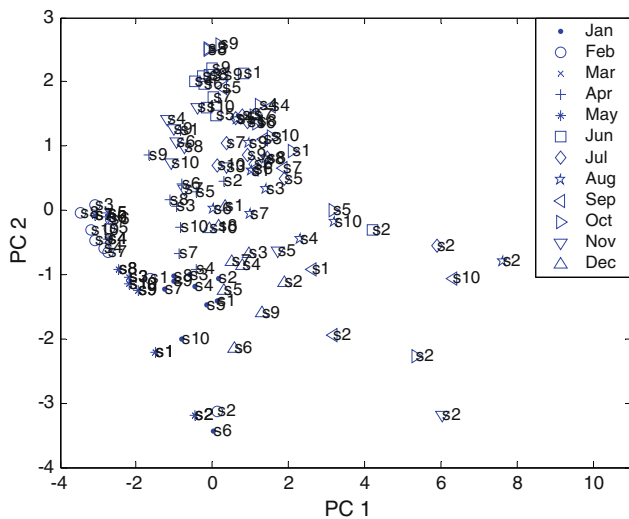$$x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} g_{pqr} + e_{ijk}$$

Fig. 2 Principal component analysis for the unfolded matrix 120 (10 stations × 12 months) × 10 (parameters)

where a_{ip}, b_{jq} and c_{kr} denote elements of the component matrices A, B and C of orders $I \times P$, $J \times Q$ and $K \times R$, respectively. Each of these matrices can be interpreted as a loading matrix in the classical principal component analysis. g_{pqr} denotes the elements (p, q, r) of the $P \times Q \times R$ core array G, and e_{ijk} denotes the error term for element x_{ijk} and is an element of the $I \times J \times K$ array E. Where I is the number of sampling sites (objects), J is the number of variables and K is the number of sampling times (conditions).

The collected data is arranged in three-dimensional arrays with the dimension of 10 (stations) × 10 (parameters) × 12 (months) for the matrix SS.

Before three-way principal component analysis, the normal distribution of each variable should be tested due to the evident outliers around pollution sources.

Shapiro–Wilk test was applied to check the distribution pattern of the variables. Variables temperature, DO, pH and SiO_3-Si concentrations had normal distributions ($P < 0.05$). Other variables showed a skewed distribution, and therefore a logarithmic transformation was applied on them.

In three-way modeling methods, scaling and centering of the data are often crucial. Several pretreatments can be applied owing to the possibility of scaling and/or centering along or across the different modes. A j-scaling (Leardi et al. 2000) was performed in order to remove the differences among the variables arising from their different ranges and magnitudes, allowing all of them to have the same possibilities to contribute to the model. With this approach, differences between objects (sampling station) and conditions (sampling time) have been retained. In order to perform it, the three-way array X was matricized to a two-way matrix with 10 (stations) × 12 (months) rows and 10 (parameters) columns; autoscaling was then carried out.

Results

Environment factors

The water temperature ranged between 23.17 and 30.51°C during the study period. The mean water temperature was 26.31°C. The salinity varied narrowly from 32.59 to 34.79 PSU. The concentration of SiO_3-Si fluctuated widely from 1.71 to 17.47 μmol L^{-1}. The pH varied from 8.07 to 8.41.

The temporal and horizontal distributions of NO_3-N and NH_4-N concentrations in the surface water are shown in Fig. 3. The horizontal distribution of NO_3-N surface concentration revealed that the NO_3-N concentration decreased from inner bay to outer bay in January, April, August and October. The surface concentration of NO_3-N in Sanya River estuary is higher than that in the bay. The spatial distribution of NH_4-N concentration is similar to that of the NO_3-N concentration.

Linear correlation coefficient between variables shows in Fig. 4. As expected, dissolved oxygen is negatively correlated with temperature because the solubility of oxygen in water decreases with increasing temperature. There was significant negative correlation between salinity and temperature. The salinity was significantly correlated to DO, SPC, SiO_3-Si, NO_2-N and NO_3-N. The concentration of nutrients decreased greatly with higher salinity.

Three-way principal component analysis

Generally, the optimal complexity of the Tucker3 model is the one that requires the smallest number of factors, but still describes relatively high fraction of data variance. All possible models, with different number of factors in each mode ($I = 1, 2,…, 12$; $J = 1, 2,…, 13$; $K = 1, 2,…, 4$), have been evaluated. In order to visualize how many percent of explained variance is gained by adding factors, the explained variance was plotted versus increasing value of the product ($I \times J \times K$) (Fig. 5). A compromise between models describing a high percentage of variation and models with fewer components was thus sought.

The optimal components of the Tucker3 model was considered to be two components in mode A (sampling site), two component in mode B (the variable) and one component in mode C (the sampling time), shortly indicated as [2 2 1], and explaining 33.18% of the total variance of the data. Such a low variance with environmental data is not unusual due to very high noise related to the great variability of weather conditions (Leardi et al. 2000).

In the plot of the stations (Fig. 6a), all the stations show negative values in the first component (A1), and are spread along the second component (A2). A2 is related to nutrients with high values. If also the map of the bay is taken into
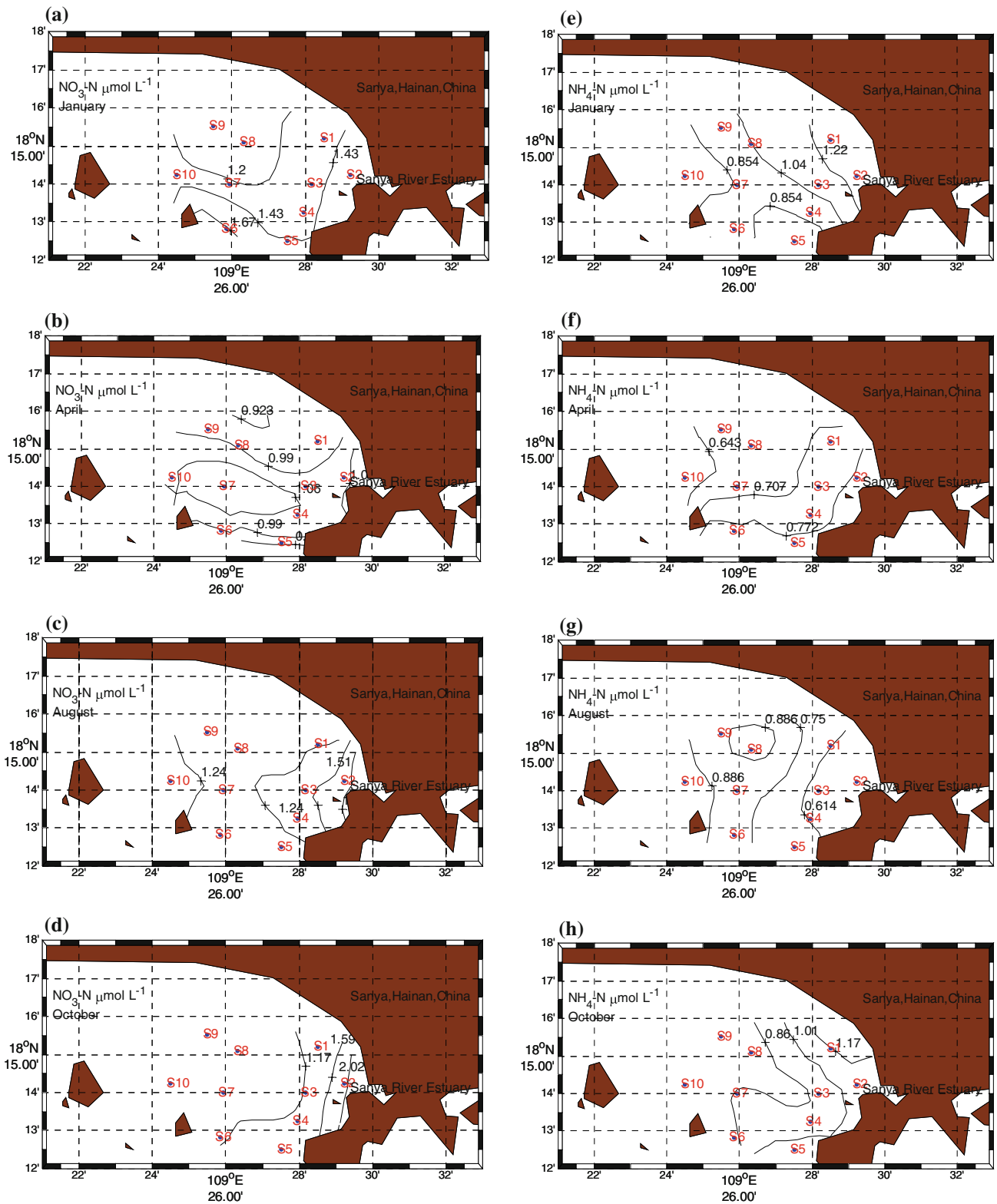
Fig. 3 a–h Horizontal distribution for NO₃-N, NH₄-N surface concentrations in Sanya Bay. **a** Nitrate surface distribution in January. **b** Nitrate surface distribution in April. **c** Nitrate surface distribution in August. **d** Nitrate surface distribution in November. **e** Ammonia surface distribution in January. **f** Ammonia surface distribution in April. **g** Ammonia surface distribution in August. **h** Ammonia surface distribution in November
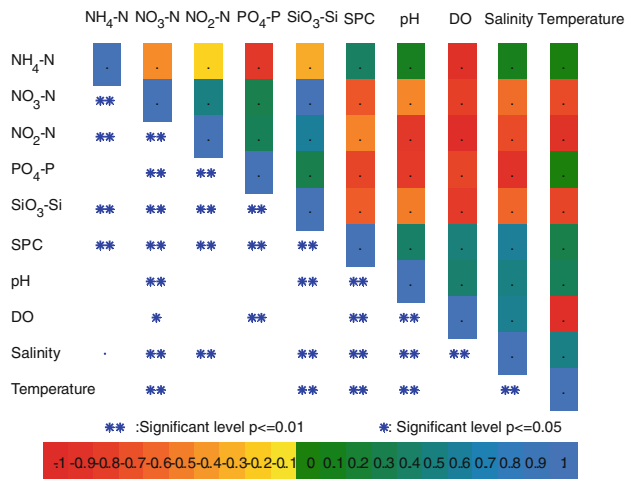
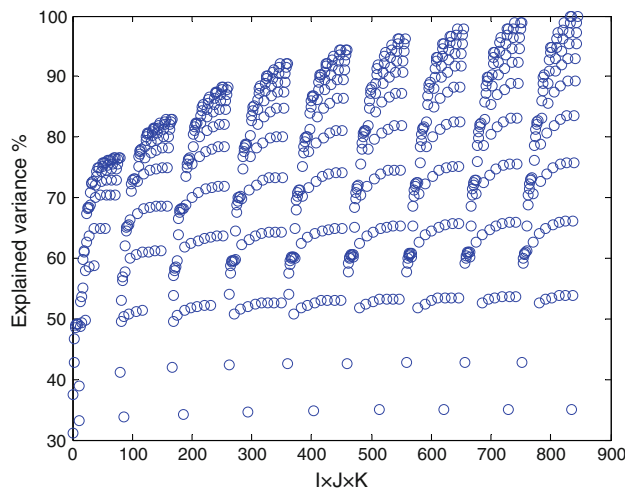**Fig. 4** Linear correlation coefficients of 10 parameters

Fig. 5 Scatter plot of modeled sum of squares (%) as function of the product of the number of components in different modes for the considered Tucker models

account, one can see that the distribution of the stations along A2 has a very strong correspondence with the geographical location, with the direction low values–high values (low nutrients–high nutrients) roughly corresponding to the direction inside–outside.

In more detail, the S2 in the Sanya River estuary area is away from other stations (S3–S10) where concentrations of nutrients are low. This direction also corresponds to a decrease in salinity.

The loading plot of variables in the first two components (B1 and B2) is shown in Fig. 6b. The loading of salinity in modes B1 and B2 is negative. The loading of temperature is opposite to the salinity in mode B1. The temperature and salinity may be important indicators of climate and marine
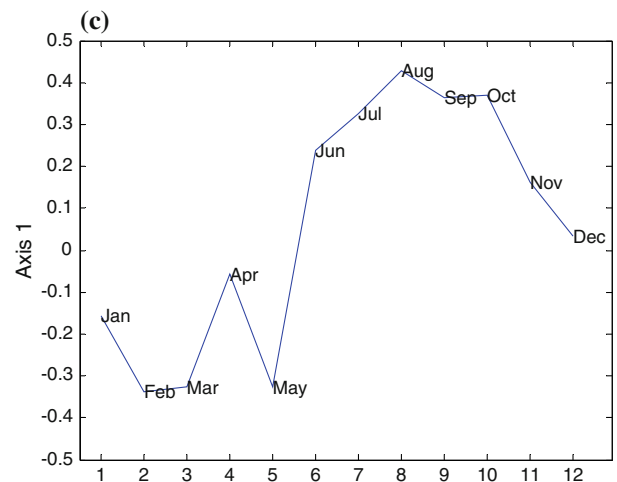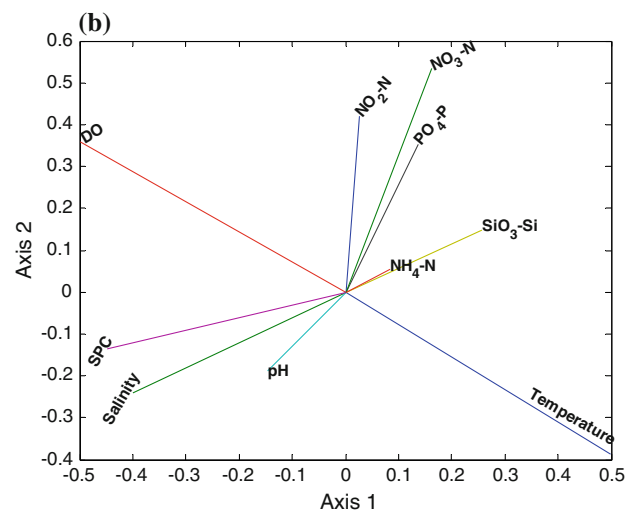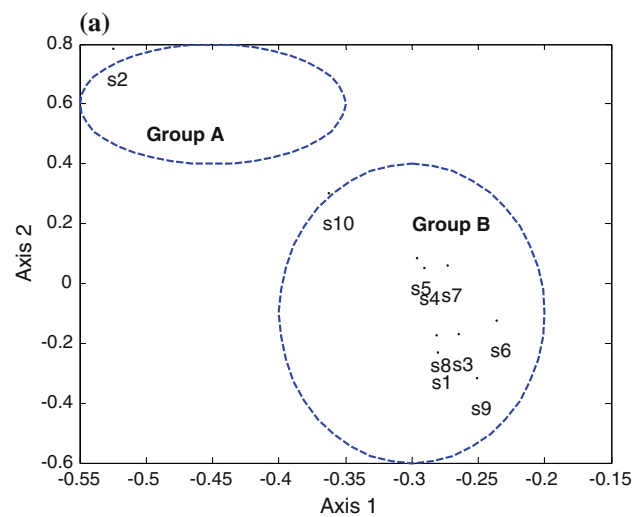

Fig. 6 Plots of the first mode (sampling stations), the second mode (variables) and the third mode (sampling times). **a** Plot of the first mode (sampling stations). **b** Plot of the second mode (variables). **c** Plot of the third mode (sampling times)

character. The loadings of the nutrients in the first two components (B1 and B2) are positive. High negative loadings on pH and DO in mode B1 can be explained with high levels of dissolved organic matter which consumes large amounts of oxygen, which undergoes anaerobic fermentation processes leading to formation of ammonia and organic acids. Hydrolysis of these acidic materials causes a decrease of pH values (Vega et al. 1998; Singh et al. 2004). Therefore, this component represents variables indicating anthropogenic pollution. On the other hand, pollution indicator variables (PO_4-P, NH_4-N, NO_3-N and DO) exhibit positive loadings in mode B2, but temperature has negative loading. Therefore, this component represents variables indicating nutrient character.

The temporal information is described in terms of loadings of each sampling month on the single component of the sampling time mode, shown in Fig. 6c.

All the months are spread along the first component (C1). Loadings of the months (June–October) are more than 0.2. In fact, these months have heavy rainfall. A wet and warm southwest monsoon prevails in the rainy season from April to September, which brings humid air from low latitudes, resulting in gentle monsoonal rainfall in spring and heavy rainfall in summer. In contrast, a dry and cold northeast monsoon predominates in dry season from October to the following March. In this region, the Southeast Asian monsoons, northeasterly from October to the next April and southwesterly from May to September have important effects on biogeochemical cycles in South China Sea waters (Chen et al. 2006).

Discussion

Spatial pattern

In spatial pattern, the result from the loadings of the station in the first two axis demonstrated that the two different regions of stations were well distinguished (Fig. 6a). The concentration of nutrients had an important contribution to the loading of S2 in mode A2. S2 is in the mouth of Sanya River which is only river entry into the bay. The average concentrations of nutrients show similar spatial variations, decreasing from the eastern to the western part of the bay (Fig. 7).

The domestic and industrial wastewater discharge enter into the Sanya River, and then into the bay. S2 was mainly influenced by the domestic wastewater and runoff of freshwater of Sanya River. Thus, the municipal waste water and runoff of Sanya River play an important role in determining the local water quality. Both PO_4-P and SiO_3-Si concentrations decrease from east to west probably as a result of the effects of land sources and the Sanya River

(Huang et al. 2003). Nutrients' concentrations increased shoreward and clearly demonstrated the impact from the terrestrial input and the Sanya River (Zhou et al. 2009). These results showed that human activities have an important effect on water quality in the bay.

In this study, salinity was important marine factor, which had important effect on the water quality characters. Salinity has strong positive and negative loadings in modes A1 and A2, respectively. The annual salinity in Group A was significantly different from that in Group B by one-way analysis of variance ($P < 0.05$). Fresh water from the Sanya River diminishes the surface salinity in Group A (S2). This result showed that the discharge of Sanya River has an important effect on Group A (S2). Correlation coefficients analysis showed that nutrients were negatively correlated to salinity. They indicate the importance of mixing between polluted freshwaters and costal saline waters. Nutrients were introduced in the bay by river and sewage discharges (Paranhos et al. 1998). Negative correlation between nutrients and salinity demonstrate that land sources are the main reason for high levels of nutrients (Liu et al. 2005).

On the other hand, Sanya Bay is a permanently open bay. The waters from South China Sea play an important role on renewing waters of the bay. The annual salinity distribution show that it increases from the eastern to the western part of the bay (Fig. 7). The salinity varies from 33.0 to 34.5 PSU in South China Sea (Han 1998). The salinity in the bay waters lies in this range. The bay waters renewal promoted by the waters in South China Sea is an important mechanism in diluting the concentration of nutrients (pollution). The concentration of both PO_4-P and SiO_3-Si decrease from the mouth of Sanya River to offshore waters, probably as a result of the effects of discharge from the Sanya River, and diluting concentration of nutrients (pollution) from the waters in South China Sea.

NO_3-N is the main form of dissolved inorganic nitrogen, which contributes over 50% of dissolved inorganic nitrogen; the second one is the ammonia; NO_2-N is often less than 20% of dissolved inorganic nitrogen throughout the whole year. Composition of dissolved inorganic nitrogen in waters is similar to that in South China Sea. NO_3-N is the main form of dissolved inorganic nitrogen (75%), the second one is NH_4-N (20%), and the third one is NO_2-N (5%) (Han 1998). It is not in agreement with that in many coastal waters, such as Daya Bay. NH_4-N is the main form of dissolved inorganic nitrogen in winter. The second one is NO_3-N, and the third one is NO_2-N (Wu and Wang 2007). NH_4-N (c. 49%) and NO_3-N (c. 43%) were the dominant total dissolved inorganic nitrogen (TIN) forms, accounting for about 90% of the TIN during 1999–2002; NO_2-N was only about 8% (Wang et al. 2006).Thus, the bay water is also mainly influenced by the waters in South China Sea.
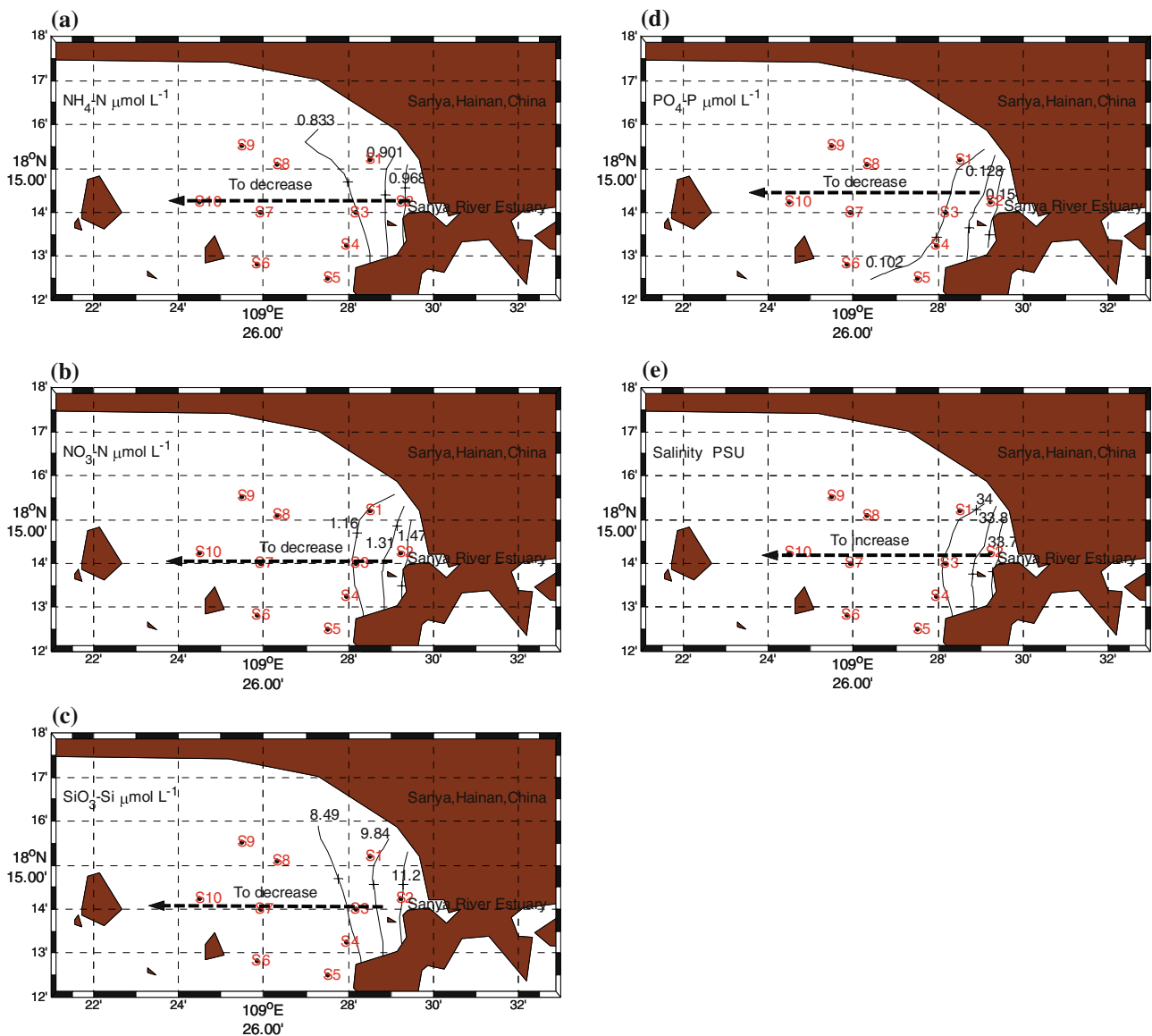
Fig. 7 The annual concentrations of nutrients and salinity for surface concentrations in Sanya Bay. **a** Spatial distribution of the annual concentration of NH₄-N. **b** Spatial distribution of the annual concentration of NO₃-N. **c** Spatial distribution of the annual concentration of SiO₃-Si. **d** Spatial distribution of the annual concentration of PO₄-P. **e** Spatial distribution of the annual concentration of salinity

Temporal pattern

In temporal pattern, seasonal character of the water quality was investigated through loadings of the monthly samples in mode C1 (Fig. 6c). The first group includes the samples in dry season (January, February, March, April, November and December). The second group includes the samples in rainy season from May to October. From the loading plot of variables in Fig. 6b, the temperature and salinity may be important indicators of climate and marine character. The average temperature of sea surface was higher in rainy season (29.15°C) than that in dry season (24.76°C), with significant difference by one-way analysis of variance ($P < 0.05$). The inverse relationship between temperature

and DO is a natural process because warm water easily becomes saturated with oxygen, and thus, can hold less DO.

The average salinity of sea surface was 33.89 and 34.16 PSU in rainy season and dry season, respectively. The salinity between two groups was significantly different ($P < 0.05$). In rainy season, rain is an important factor, which diminishes the surface salinity. In addition, fresh water from the Sanya River Estuary input diminishes the surface salinity. The lowest salinity appeared in autumn and the highest appeared in winter (Zhou et al. 2009).

The concentration of SiO₃-Si is 9.31 and 7.32 μmol L⁻¹ in rainy and dry season, respectively. The concentration of SiO₃-Si between two groups was significantly different

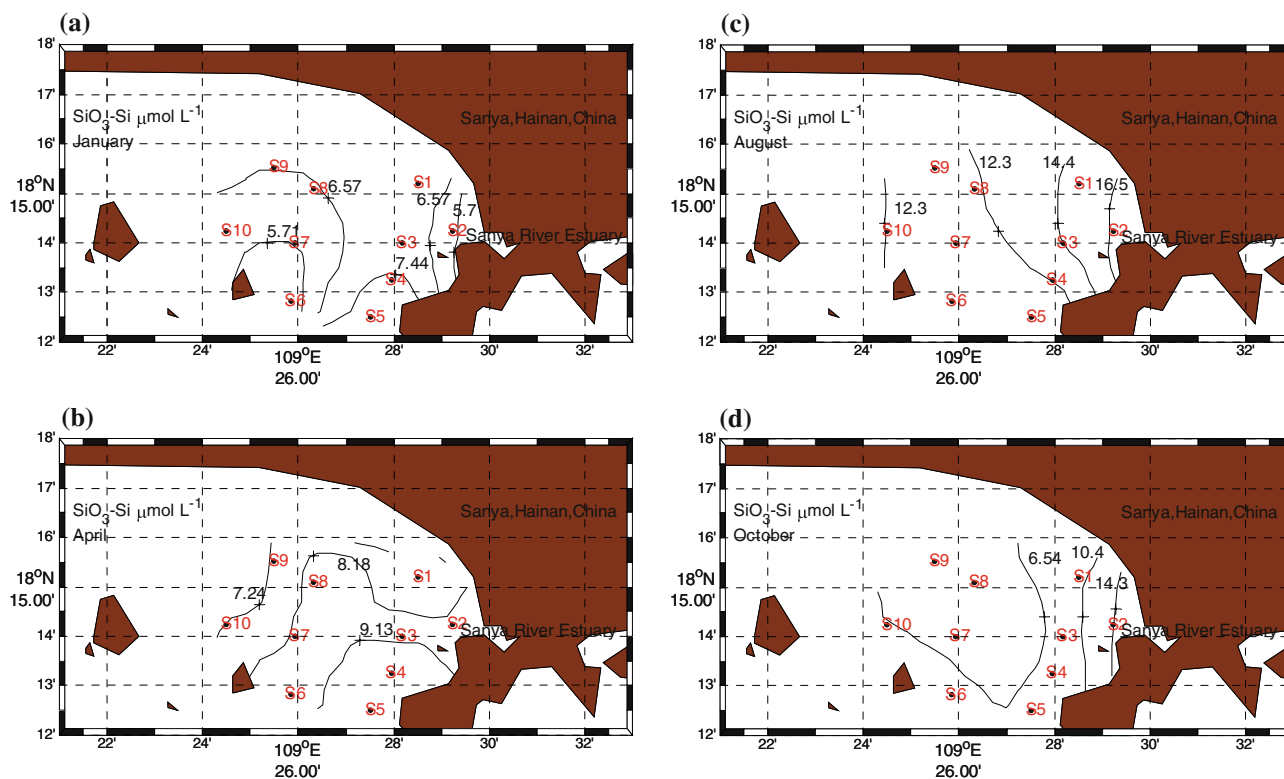**Fig. 8** Spatial distribution for SiO₃-Si surface concentrations in January, April, August and October. **a** Spatial distribution of the concentration of SiO₃-Si in January. **b** Spatial distribution of the concentration of SiO₃-Si in April. **c** Spatial distribution of the concentration of SiO₃-Si in August. **d** Spatial distribution of the concentration of SiO₃-Si in October

($P < 0.05$). This result is similar to that in Daya Bay. Silicate is from land-based resources, especially during rainfall (Zhu 1999). It exhibits the seasonal changes, and with the concentration in rainy season higher than in dry season (Huang et al. 2003), which is related to the organic matter decomposition with the temperature increase (Wu and Wang 2007). The concentration of SiO₃-Si decreased from the eastern to the western part of the bay in August and October (Fig. 8). A large amount of rainfall occurs usually during the southwest monsoon period (Jeong et al. 2008). There is plenty of rainfall in Sanya Bay during the southwest monsoon period from May to October, and there is less during the northeast monsoon period from November to the next April. Higher concentration of SiO₃-Si is introduced in the bay by rivers and sewage discharges during the southwest monsoon period. The low concentration of SiO₃-Si decreased from the southern to the northern part of the bay in January and April (Fig. 8). The waters were mixed well in winter owing to the Coriolis effect (force) and the northeastern monsoon winds (Zhou et al. 2009). The different monsoons may result in the different spatial distribution of SiO₃-Si concentration.

Precipitation is an important climate indicator of season change characters. In order to verify the season change, the average precipitation data in Haikou (110°35′E, 20°03′N)

and Dong fang (108°62′E, 19°10′N), Hainan, from 1971 to 2000 (China Meteorological Administration 1971–2000) are used to support and clarify the season change characters (Fig. 9). The monthly precipitation during rainy season is more than 100 nm, except May in Dongfang, and less than 100 nm during dry season. Rainfall and land-based water input may cause the difference of the water quality between dry season and rainy season. The season character is in agreement with that in a subtropical bay, Daya Bay (Wu and Wang 2007; Wu et al. 2009b).

Conclusion

Results of classical principal component analysis could be difficult to interpret because the information of the three modes can be mixed. However, three-way principal component analysis was preferred because it directly takes into account the three-way structure of the data, allowing an easy interpretation of the results.

The presented results show that three-way principal component analysis is useful for the study of these types of data. Three-way principal component analysis offers detailed information about the data set and allows visualization of the data structure to obtain an easy
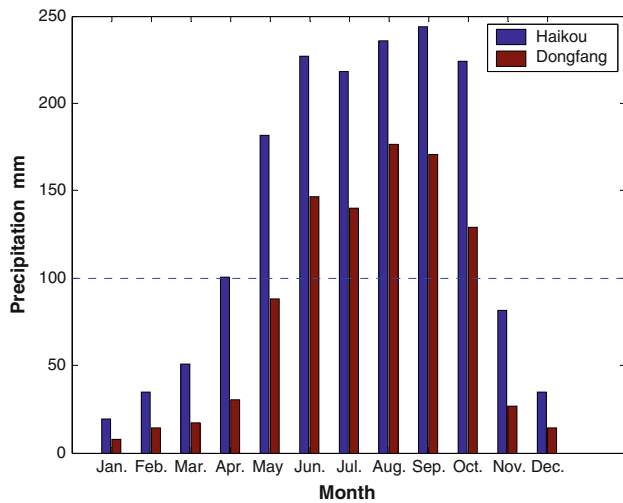
Fig. 9 Monthly precipitation in Haikou and Dongfang

interpretation of spatial and temporal phenomena taking place in the region. As a result, it was clear that the main difference among the stations was related to human activities and marine characteristics. The dry season from October to the next April and rainy season from May to September have been distinguished. The Southeast Asian monsoons, northeasterly from October to the next April and southwesterly from May to September have important effects on biogeochemical cycles in South China Sea waters. All of this information would be very useful to regional agencies in developing a strategy to carry out scientific plans for resource use based on marine system functions.

References

Bowen RE, Depledge MH (2006) Rapid assessment of marine pollution (RAMP). Mar Pollut Bull 53:631–639

Chau KW, Muttil N (2007) Data mining and multivariate statistical analysis for ecological system in coastal waters. J Hydroinformatics 9:305–317

Chen CC, Shiah FK, Chung SW, Liu KK (2006) Winter phytoplankton blooms in the shallow mixed layer of the South China Sea enhanced by upwelling. J Mar Syst 59:97–110

Giussani B, Monticelli D, Gambillara R, Pozzi A, Dossi C (2008) Three-way principal component analysis of chemical data from Lake Como watershed. Microchem J 88:160–166

Grotti M, Rivaro P, Leardi R, Magi E (1999) Three-way principal component analysis as a powerful tool to process marine environmental data sets. Annal Chim 89:591–600

Han WY (1998) Marine chemistry in South China Sea. Science Publishing House, Beijing, China, pp. 199–211 (in Chinese)

Huang LM, Tan YH, Song XY, Huang XP, Wang HK, Zhang S, Dong JD, Chen RY (2003) The status of the ecological environment and a proposed protection strategy in Sanya Bay, Hainan Island, China. Mar Pollut Bull 47:180–186

Jeong KS, Kim DK, Pattnaik A, Bhatta K, Bhandari B, Joo GJ (2008) Patterning limnological characteristics of the Chilika lagoon (India) using a self-organizing map. Limnology 9:231–242

Kuppusamy MR, Giridhar VV (2006) Factor analysis of water quality characteristics including trace metal speciation in the coastal environmental system of Chennai Ennore. Environ Int 32:174–179

Leardi R, Armanino C, Lanteri S, Alberotanza L (2000) Three-mode principal component analysis of monitoring data from Venice lagoon. J Chemom 14:187–195

Liu SM, Zhang J, Chen HT, Zhang GS (2005) Factors influencing nutrient dynamics in the eutrophic Jiaozhou Bay, North China. Prog Oceanogr 66:66–85

Morais H, Ramos C, Forgacs E, Cserhati T, Oliviera J, Illes T (2001) Three-dimensional principal component analysis employed for the study of the beta-glucosidase production of *Lentinus edodes* strains. Chemom Intell Lab Syst 57:57–64

Paranhos R, Pereira AP, Mayr LM (1998) Diel variability of water quality in a tropical polluted bay. Environ Monit Assess 50:131–141

Pardo R, Vega M, Deban L, Cazurro C, Carretero C (2008) Modelling of chemical fractionation patterns of metals in soils by two-way and three-way principal component analysis. Anal Chim Acta 606:26–36

Pravdova V, Walczak B, Massart DL, Robberecht H, Van Cauwenbergh R, Hendrix P, Deelstra H (2001) Three-way principal component analysis for the visualization of trace elemental patterns in vegetables after different cooking procedures. J Food Comp Anal 14:207–225

Pravdova V, Boucon C, de Jong S, Walczak B, Massart DL (2002) Three-way principal component analysis applied to food analysis: an example. Anal Chim Acta 462:133–148

Simeonov V, Stratis JA, Samara C, Zachariadis G, Voutsa D, Anthemidis A, Sofoniou M, Kouimtzis T (2003) Assessment of the surface water quality in Northern Greece. Water Res 37:4119–4124

Singh KP, Malik A, Mohan D, Sinha S (2004) Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. Water Res 38:3980–3992

Singh KP, Malik A, Singh VK, Basant N, Sinha S (2006) Multi-way modeling of hydro-chemical data of an alluvial river system—a case study. Anal Chim Acta 571:248–259

Vega M, Pardo R, Barrado E, Deban L (1998) Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Res 32:3581–3592

Wang YS, Lou ZP, Sun CC, Wu ML, Han SH (2006) Multivariate statistical analysis of water quality and phytoplankton characteristics in Daya Bay, China, from 1999 to 2002. Oceanologia 48:193–211

Wu ML, Wang YS (2007) Using chemometrics to evaluate anthropogenic effects in Daya Bay, China. Estuar Coast Shelf Sci 72:732–742

Wu ML, Wang YS, Sun CC, Wang H, Dong JD (2009a) Using chemometrics to identify water quality in Daya Bay, China. Oceanologia 52:217–232

Wu ML, Wang YS, Sun CC, Wang H, Dong JD, Han SH (2009b) Identification of anthropogenic effects and seasonality on water quality in Daya Bay, South China Sea. J Environ Manag 90:3082–3090

Yeung IMH (1999) Multivariate analysis of the Hong Kong Victoria Harbour water quality data. Environ Monit Assess 59:331–342

Yung YK, Wong CK, Yau K, Qian PY (2001) Long-term changes in water quality and phytoplankton characteristics in Port Shelter, Hong Kong, from 1988–1998. Mar Pollut Bull 42:981–992

Zhou F, Guo HC, Liu Y, Jiang YM (2007a) Chemometrics data analysis of marine water quality and source identification in Southern Hong Kong. Mar Pollut Bull 54:745–756

Zhou F, Huang GH, Guo HC, Zhang W, Hao ZJ (2007b) Spatio-temporal patterns and source apportionment of coastal water pollution in eastern Hong Kong. Water Res 41:3429–3439

Zhou WH, Li T, Cai CH, Huang LM, Wang HK, Xu JR, Dong JD, Zhang S (2009) Spatial and temporal dynamics of phytoplankton and bacterioplankton biomass in Sanya Bay, northern South China Sea. J Environ Sci 21:595–603

Zhu ZH (1999) Phosphorates and silicate. In: Pan JP (ed) Researches on ecosystem of Daya Bay. China Meteorological Press, Beijing, pp 25–29