Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Journal of Sea Research 72 (2012) 69–76

Contents lists available at SciVerse ScienceDirect

Journal of Sea Research

journal homepage: www.elsevier.com/locate/seares

Environmental controlling mechanisms on bacterial abundance in the South China Sea inferred from generalized additive models (GAMs)

Bingzhang Chen [a,*], Hongbin Liu [b], Bangqin Huang [a]

[a] State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen, PR China
[b] Division of Life Science, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 13 December 2011
Received in revised form 3 May 2012
Accepted 18 May 2012
Available online 29 May 2012

Keywords:
Heterotrophic bacteria
Generalized additive model
South China Sea

ABSTRACT

We modeled the abundance distribution of heterotrophic bacteria collected from 4 cruises in the northern South China Sea using generalized additive models to infer the underlying mechanisms controlling bacterial abundance and to predict bacterial abundance using environmental parameters that can be easily obtained. We incorporated spatial coordinates, depth, month, chlorophyll (Chl) *a* concentration, temperature, salinity, nutricline and mixed layer depth in the model, which captures the main features of the observations and explains 88% of the total variation of bacterial abundance. The most important factor affecting bacterial abundance is chlorophyll, followed by salinity and nutricline depth, reflecting the importance of carbon and nutrient sources to bacteria. Bacterial abundance shows a unimodal relationship with temperature and decreases with depth. All the functions are nonlinear. After controlling environmental parameters, bacterial abundances are higher in fall and winter than in spring and summer and usually show an onshore–offshore decreasing gradient, which probably signify transportation pathways of terrestrial organic matter to the sea *via* atmospheric deposition. Comparisons of variograms between raw data and residuals of the model show that positive autocorrelation at small scales is induced by both environmental similarity and geographic proximity, while the negative autocorrelation at large scales is mostly contributed by environmental similarity in remote water masses.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The South China Sea (SCS) is the second largest marginal sea in the world and plays important roles in regulating regional climate and carbon budget owing to its vast area and volume (Kienast et al., 2001). The carbon flux in upper ocean of the SCS is being investigated by a National Basic Research Program of China named "Carbon Cycling in China Seas—Budget, Controls and Ocean Acidification (CHOICE-C)" and this paper reports the first series of results from studies on heterotrophic bacteria, which are important players (*e.g.*, decomposers of dissolved organic matter, the foundation of the microbial food web, and major contributors to respiration) in the carbon system (Azam et al., 1983; Del Giorgio and Williams, 2005). Here we loosely define heterotrophic bacteria as prokaryotic cells without autofluorescence that can be detected by flow cytometer and may include some Archaea (Li et al., 2004). The aim of this paper is toward understanding environmental controlling mechanisms on bacterial abundance.

The mechanisms controlling bacterial abundance are complex. Bacterial growth in the ocean can be limited by temperature, organic carbon, and inorganic nutrients (Pomeroy and Deibel, 1986; Rivkin and Anderson, 1997). Bacterial abundance is also controlled by bacterivorous grazers and bacteriophages (Azam et al., 1983; Proctor and Fuhrman, 1990). In the field, it is difficult to quantify the availability of the dissolved organic matter readily available to heterotrophic bacteria in oligotrophic environments because of the high uptake rate by microbes and low ambient level of nutrients that can be directly measured (Zubkov et al., 2008). Determining the response of vital rates of diverse bacteria (most of which cannot be cultured) to environmental factors also poses a challenge to biologists but is essential in mechanistic modeling. These problems make mechanistic modeling of bacterial abundance more difficult than on phytoplankton (Barton et al., 2010; Fasham et al., 1990; Follows et al., 2007). Conventional nitrogen–phytoplankton–zooplankton–detritus (NPZD) models do not contain a bacterial compartment (Fasham et al., 1990; Gan et al., 2010). While the compartment of detritus is assigned the role as 'decomposer', it is not allowed to grow actively as bacteria do.

Estimation of phytoplankton biomass in terms of chlorophyll concentration at large scales can be achieved by retrieving and analyzing satellite images of ocean color (McClain, 2009), which however can hardly work for heterotrophic bacteria. Heterotrophic bacteria show a strong absorption peak in the red band which overlaps with the absorption spectrum of water; while based on backscattering that can be linked with particle size *via* Mie theory, a serious problem is to separate heterotrophic bacteria from other similar-sized particles such as cyanobacteria and detritus especially in productive waters (Kostadinov et al., 2009).

* Corresponding author at: Rm. B624, Zeng Cheng Kui Building, Xiamen University, 182 Daxue Rd., Xiamen, Fujian 361005, PR China. Tel./fax: +86 592 2182978.
E-mail address: bzchen2011@xmu.edu.cn (B. Chen).

1385-1101/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.seares.2012.05.012

Author's personal copy

Modern statistical techniques provide an alternative approach to understanding environmental effects on bacterial dynamics and predicting bacterial parameters using easily-obtainable parameters. Although they do not provide explicit controlling mechanisms, they may give useful hints in revealing the underlying mechanisms and their mathematical predictions could be approximate to or even better than those from the mechanistic models.

Generalized additive models (GAMs) allow one response variable being fitted by several predictors in an additive manner (Hastie and Tibshirani, 1989; Wood, 2006). The partial term of one predictor is an unspecified function constructed from regression splines, which have the advantage of having minimal integrated square secondary derivatives (*i.e.*, being smooth) and do not need to assume *a priori* function. A penalty is added to the regression to control the degree of smoothness of the fitting curve. GAMs thus attempt to attain the conflicting goal of minimizing the square deviations and maximizing the smoothness of the fit. Further developments using thin plate regression splines and tensor product splines allow inclusion of several covariates in one function although, practically, computation cost limits the numbers of interacting covariates (Wood, 2006). The functional forms of partial effects of individual covariates generated by GAMs can help construct the mechanistic models. In marine science, GAMs have been used in modeling phytoplankton biomass (Irwin and Finkel, 2008; Llope et al., 2009) and fish catch (Lorance et al., 2010).

The main physical factors affecting the upper ocean of the SCS include the monsoon-affected spatial circulation patterns, changes in water column vertical structure due to solar radiation and wind stress, river plume and groundwater discharge from the north, tropical typhoons, mesoscale features such as eddies and internal waves (Hu et al., 2000; Wong et al., 2007). The majority of these factors have a strong seasonal pattern. Together with variations in geographical topography and bathymetry, these factors cause salient spatial patterns in chemical and biological environments (Gan et al., 2010). It is a usual practice to incorporate spatiotemporal components (*e.g.*, geographic coordinates and sampling time) in GAMs, which can capture the effects of some unidentified environmental factors (*e.g.*, the inventory of organic carbon readily available to bacteria) with significant spatiotemporal trends (Irwin and Finkel, 2008; Llope et al., 2009). Although biologists would prefer the percentage of the variance explained by purely spatiotemporal parameters to be negligible, spatiotemporal patterns may also be caused by biotic interactions, which, sometimes chaotic, have spatiotemporal trends but are not simply correlated with environmental parameters (Benincà et al., 2009; Borcard et al., 2011). For the sake of prediction, the incorporation of easily-obtainable spatiotemporal components is more preferable. In the following, we also partition the full model into purely spatiotemporal and environmental sub-models to assess the contribution of each factor in explaining the variations of bacterial abundance.

2. Methods

2.1. Sample collection and analysis

Two milliliter seawater samples were collected at 3 to 12 depths from 0 to 150 m using Niskin bottles attached to a CTD rosette system in four cruises (summer: July 18 to August 16, 2009, 79 stations; winter: January 6 to 30, 2010, 63 stations; fall: October 26 to November 24, 46 stations; spring: April 30 to May 24, 92 stations) in the northern South China Sea (north of 18° N; Fig. 1). The samples were fixed with seawater buffered paraformaldehyde (0.5% final concentration) and stored at −80 °C until analysis. Upon return to the lab, cell abundances of autotrophic picoplankton were enumerated using a Becton–Dickson FACSCalibur cytometer, with different populations distinguished based on side-scattering, orange and red fluorescence (Olson et al., 1993). Yellow-green fluorescent beads (1 μm, Polysciences) were added to each sample as an internal standard. The exact flow rate was calibrated by weighing a tube filled with distilled water before and after running for certain time intervals and the flow rate was estimated as the slope of a linear regression curve between elapsed time and weight differences (Li and Dickie, 2001). For counting of heterotrophic bacteria, SYBR Green I was added to the sample at the final concentration of 0.01% of the original stock (Molecular Probes Inc.) and the samples were stained in the dark at 37 °C for 1 h before being analyzed on the cytometer (Marie et al., 1997). The cells were distinguished based on side-scattering and green fluorescence. The abundance of heterotrophic bacteria was calculated as the difference between the abundance of total bacteria and that of autotrophic bacteria estimated without staining. As counting by the cytometer has been shown very accurate (coefficient of variation < 1%), we did not run replicates during analysis.

Temperature (*Temp*), salinity (*Sal*), and pressure were determined by a CTD system (Seabird 911). Mixed layer depth (*MLD*) was determined as the first depth where the temperature difference with that at 5 m exceeded 0.2 °C (Steinhoff et al., 2010). Nitrate and phosphate concentrations were measured following the methods of Parsons et al. (1984). The detection limits ranged from 0.1 to 0.3 μmol L^{-1} for nitrate and were 0.08 μmol L^{-1} for phosphate. The depth of nutricline (*Nutricline*) was defined as the shallowest depth where nitrate concentrations exceeded 0.5 μmol L^{-1}. Total Chl *a* concentrations (*Chl*) including monovinyl and divinyl Chl *a* were measured by high performance liquid chromatography according to the methods of Furuya et al. (1998).

2.2. GAM modeling

We used the function 'gam' in the R package 'mgcv' developed by Wood (2006) to model the functional response of abundances of heterotrophic bacteria (*HB*) on the geographic (longitude (*Lon*), latitude (*Lat*), and sampling depth (*Depth*)), temporal (*Month*), and environmental parameters (*Chl*, *MLD*, *Sal, Nutricline*). In general, *Chl* can be regarded as a proxy for autochthonous organic carbon source for bacteria although sometimes there might exist a time lag between labile dissolved organic carbon and *Chl* due to lysis of phytoplankton cells (*e.g.*, post-bloom condition). As we mainly focus on the spatial gradient instead of temporal variation in this study, this should not be a serious problem. *MLD* describes the vertical mixing regime of the water column. *Sal* can be considered as a proxy for freshwater input. Spatial coordinates (*Lon* and *Lat*) were included in a single function using thin plate regression splines. We further tested whether allowing other parameters such as *Month* and *Depth* interacting with *Lon* and *Lat* improves the model. As these parameters have different units with *Lon* and *Lat*, a tensor product spline was used for these 3-dimensional interactions (Wood, 2006).

As the distributions of *Chl* and *HB* are highly right-skewed, we $\log_e$-transformed them to log (*Chl* + 0.001) and log (*HB*), respectively, to satisfy a roughly normal distribution and to deal with the problem of zero value. The new variables were named "*logChl*" and "*logHB*". To control the degree of smoothing and minimize overfitting to the data, a penalty term was added in the regression and we set the gamma = 1.4 in gam, which forces each model effective degree of freedom to count as 1.4 degrees of freedom in the generalized cross-validation (GCV) score. Graph plotting and statistical analysis were conducted using R 2.13.1 (R Development Core Team 2011).

3. Results

3.1. Spatial and temporal patterns

Physical parameters differed remarkably among seasons. The ranges of sea surface temperatures were 25.3 °C to 30.5 °C (median = 29.6 °C) in summer, 16.2 °C to 25.9 °C (median = 22.7 °C) in winter, 20.4 °C to 28.0 °C (median = 25.3 °C) in fall, and 23.3 °C to 29.5 °C (median = 27.9 °C) in spring. Sea surface temperatures often showed onshore–offshore increasing gradients due to coastal upwelling (induced by the southwest monsoon and bottom topography) and the influence

Author's personal copy

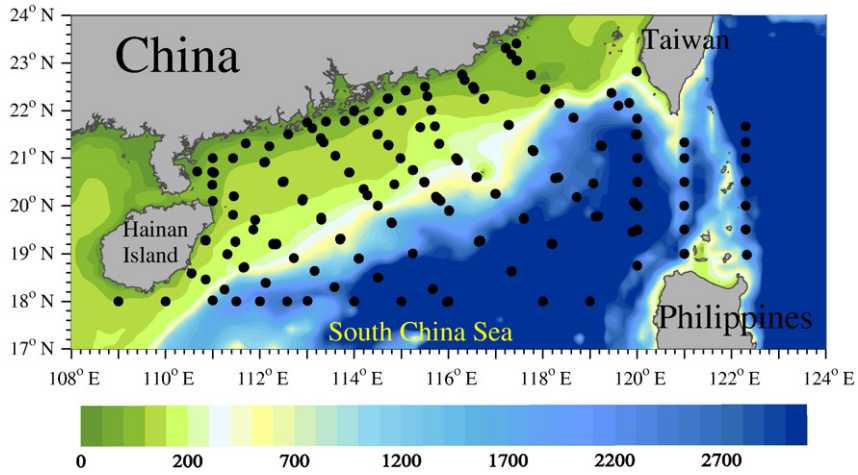

Fig. 1. A geographic map in the northern South China Sea showing coastlines and topography. The dots represent all sampling stations. The color key shows bottom depth (m). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

of China coast current (induced by the northeast monsoon). In summer, the lowest salinity values were found east of the Pearl River estuary, indicating the eastward direction of the Pearl River plume induced by the southwest monsoon. In other seasons, the surface salinity fields showed that the Pearl River plume went westward along the coast.

Chl *a* concentrations also differed among seasons. The ranges of surface Chl *a* concentrations were 0.03 to 3.71 μg L^{-1} (median = 0.12 μg L^{-1}) in summer, 0.05 to 3.79 μg L^{-1} (median = 0.58 μg L^{-1}) in winter, 0.16 to 3.00 μg L^{-1} (median = 0.57 μg L^{-1}) in fall, and 0.05 to 1.47 μg L^{-1} (median = 0.13 μg L^{-1}) in spring. Onshore–offshore decreasing trends were evident in all seasons. High concentrations were also found in areas affected by the Pearl River discharge.

The spatial distributions of bacterial abundances in surface waters (5 m) are shown in Fig. 2A, C, E, G. A common decreasing trend of bacterial abundances existed from onshore to offshore in all four cruises. Except in summer where the highest bacterial abundances were found east of Pearl River estuary, the highest bacterial abundances were usually found near the coast west of Pearl River estuary, consistent with the seasonal variations of directions of Pearl River plume. A crude comparison of ln bacterial abundances among seasons shows that bacterial abundances were significantly higher in winter and fall than in summer and spring (t-tests, $p<0.001$).

The vertical patterns of chlorophyll concentrations and bacterial abundances are shown in Fig. 3. In summer and spring, there were increasing trends of chlorophyll concentrations and bacterial abundances from surface to a depth of 50 m or 75 m (*i.e.*, depth of chlorophyll maximum); while this increasing trend was not evident in winter and fall due to more thorough mixing. Both chlorophyll concentrations and bacterial abundances decreased remarkably from the depth of chlorophyll maximum to deeper layers.

3.2. Model selection

Allowing interactions among *Lon*, *Lat*, and *Month* using a 3-dimensional tensor product spline significantly improves the model (compare model a and b in Table 1), while allowing interactions among *Lon*, *Lat*, and *Depth* just slightly improves the model. The use of a 4-dimensional tensor product with *Lon*, *Lat*, *Month*, and *Depth* not only substantially increases computation time, but does not improve the model. The results of full model a capture the main features of real observations with slight mismatches (Fig. 2).

A model consisting of pure environmental parameters (model e) reduces the R^2 to 0.76 and increases the GCV value compared with model a (Table 1). In this environmental model, *logChl* is the most important, followed by *Sal*, *Nutricline*, *MLD* and *Temp*. Used as a single predictor, *logChl* can explain 62% of the total variability of *logHB*. A model consisting of pure spatiotemporal parameters (model c) is better than the environmental model e, accounting for more than 80% of the variance observed. The interaction between *Month* and spatial coordinates is again highly significant (compare models c and d).

3.3. Partial effects of individual predictors

Fig. 4 shows the functional relationships between *logHB* and each individual predictor, leaving other parameters fixed. *logHB* increases with *Temp* in an approximately linear fashion from 15 °C to 23 °C and decreases with increasing temperature thereafter. *logHB* slowly increases with *logChl* at low chlorophyll levels (e^{-6}–e^{-3} = 0.002–0.05 mg m^{-3}), but increases roughly linearly with *logChl* at higher chlorophyll levels, which are more typical in the euphotic zone. *logHB* is usually invariable with *Depth* at 0–50 m and decreases with depth from 50 to 100 m. Below 100 m, *logHB* slightly increases with depth. *logHB* decreases with increasing *Sal* at relevant salinity ranges (>29). *logHB* also decreases slightly with increasing *MLD* when *MLD*<100 m and increases with *MLD* when *MLD*>100 m. *logHB* decreases with increasing *Nutricline* when *Nutricline* increases from 20 m to 50 m and then increases slightly with *Nutricline* when *Nutricline* becomes deeper.

As the interactions between spatial locations and seasons are highly significant, we plot in Fig. 5 the spatial distributions of *logHB* in each cruise after controlling other factors (*i.e.*, *logChl*, *Temp*, *Depth*, *MLD*, *Sal*, *Nutricline*). On average, *logHB* tends to be higher in January (winter) and November (fall) and lower in August (summer) and May (spring) (Fig. 5E). Spatially, decreasing trends from coastal to offshore regions (particularly in May and August) and mesoscale features are evident. Higher *logHB* northeast of the Hainan Island is also prevalent in the three seasons. Note that since *Sal* has been controlled, the effects of Pearl River plume are not shown on the figures.

3.4. Spatial autocorrelations

We show in Fig. 6 the effects of GAM modeling on spatial autocorrelations of *logHB*. For the un-modeled data, the semivariance increases sharply with distance until roughly 6° and then decreases with distance, showing a positive autocorrelation at short distances and a negative autocorrelation at large distances. The variograms for the residuals of the full model and the environmental model give a relatively flat variogram at small scales (<6°) and an increasing variogram at large scales (>6°), suggesting that mechanisms for the unexplained variance might work at large scales. The differences among the three variograms suggest

Author's personal copy

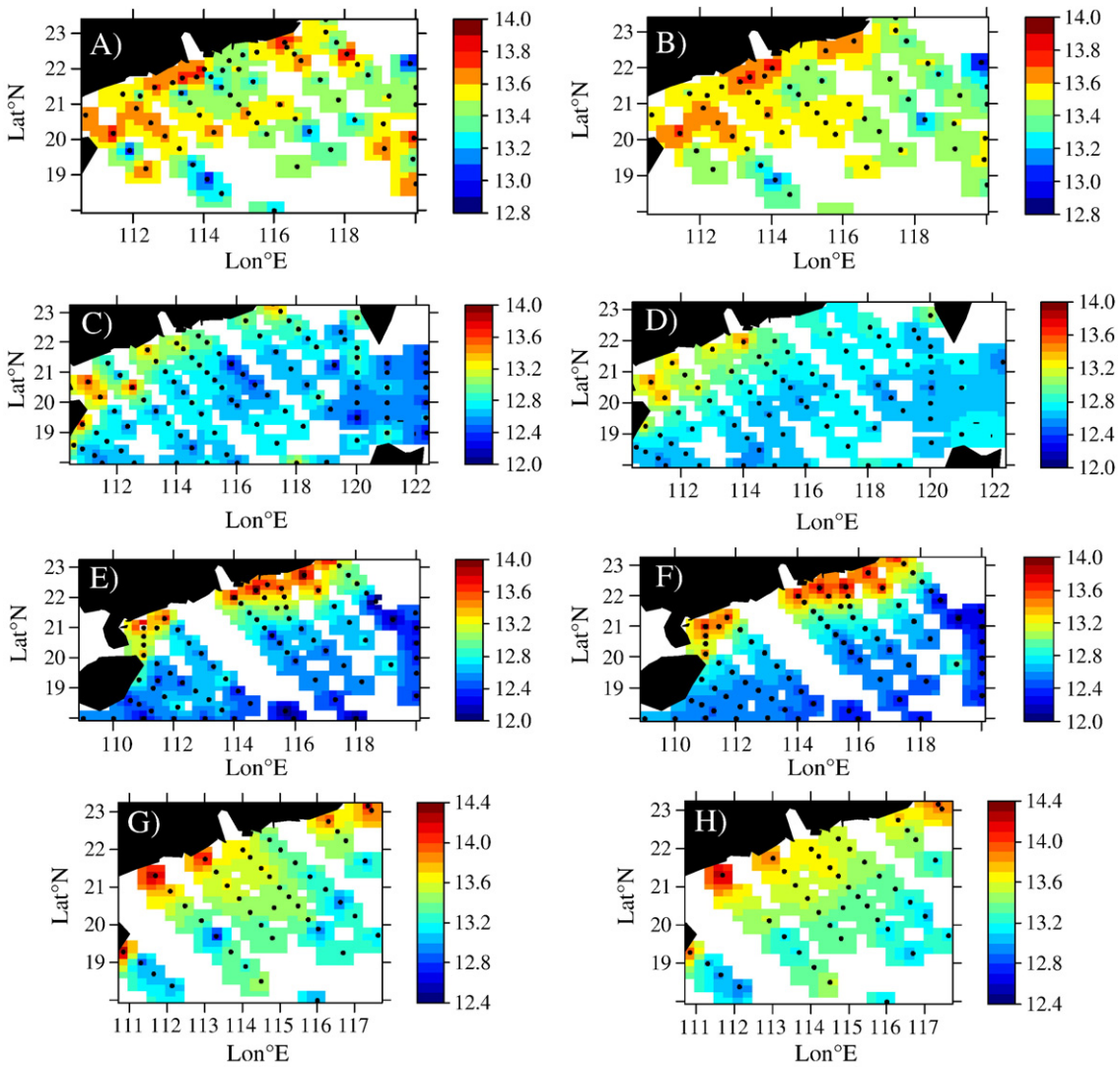

Fig. 2. Comparisons between real observations (A, C, E, G) and modeled data (B, D, F, H) of ln bacterial abundances (cells mL^{-1}) at 5 m. A, B: winter; C, D: spring; E, F: summer; G, H: fall. Black points show the sampling stations.

that the positive autocorrelation at small scales is induced by both environmental similarity and geographic proximity, while the negative autocorrelation at large scales is mostly contributed by environmental similarity in remote water masses.

4. Discussion

4.1. Bottom-up vs. top-down controls on bacterial abundance

Dynamic variations of bacterial abundance are governed by both growth and mortality terms. Bacterial growth rate can be affected by dissolved organic matter (Church et al., 2000; Pomeroy and Deibel, 1986), inorganic nutrients (Rivkin and Anderson, 1997), temperature (Lopez-Urrutia and Moran, 2007; Rivkin et al., 1996), and light (Church et al., 2004; Mary et al., 2008). Bacterial mortality induced by bacterivory and viral lysis (Azam et al., 1983; Proctor and Fuhrman, 1990; Vaque et al., 2008) can be as important as bacterial growth in regulating bacterial abundance. Thus, it is a little surprising for the relative success of our GAM model ($R^2 = 0.88$; Fig. 2) as it is mainly built on bottom-up factors. The same situation can be also applied to modeling chlorophyll (Irwin and Finkel, 2008). The high R^2 using bottom-up factors as predictors does not deny the importance of top-down factors (*i.e.*, bacterivory and viral lysis), which are also dependent on the same set of environmental factors. We briefly illustrate below how environmental dependency of rates can be translated into environmental dependency of biomass.

Both bacterial growth rate μ and mortality rate m can be expressed as functions of temperature (T), *Chl*, nutrients (N), as well as bacterial abundance (HB).

$$\mu = f(T, Chl, N, HB),$$

$$m = g(T, Chl, N, HB)$$

in which f and g are functions remaining to be determined.

At steady state when $\mu = m$, B can be rearranged as functions of T, *Chl*, and N. For example, if we assume

$$\mu = e^{\alpha_1 T}\frac{Chl}{Chl + A_1}\left(1 - \frac{HB}{C_1}\right),$$

$$m = e^{\alpha_2 T}\frac{Chl}{Chl + A_2}\frac{HB}{C_2},$$

in which α_1, α_2, A_1, A_2, C_1, C_2 are constants. When $\mu = m$,

$$HB = \frac{e^{\alpha_1 T}\frac{Chl}{Chl+A_1}}{e^{\alpha_1 T}\frac{Chl}{Chl+A_1}\frac{1}{C_1} + e^{\alpha_2 T}\frac{Chl}{Chl+A_2}\frac{1}{C_2}} = \frac{C_1}{1 + e^{(\alpha_2-\alpha_1)T}\frac{Chl+A_1}{Chl+A_2}\frac{C_1}{C_2}}, \quad (1)$$

Author's personal copy

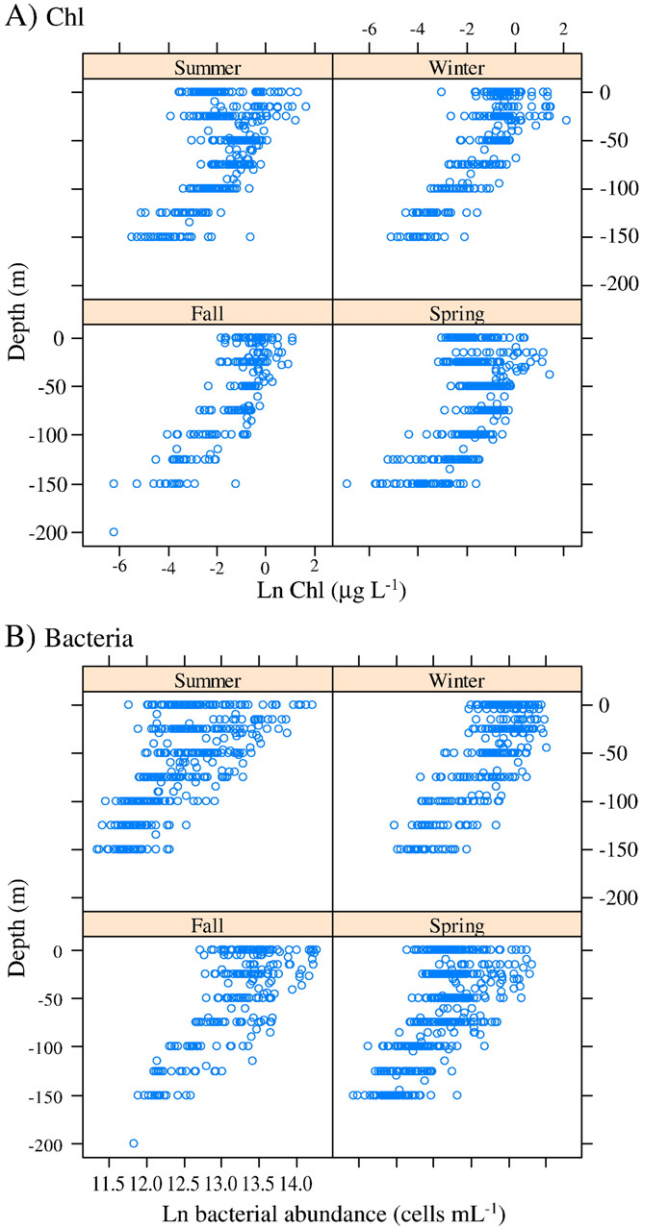

Fig. 3. Vertical distributions of (A) Chl *a* and (B) bacterial abundances in four cruises.

describing the dependency of bacterial abundance on temperature and chlorophyll concentration. Thus, the functional relationships between environmental parameters and bacterial abundance obtained from GAM modeling not only reflect environmental dependency of growth rate but also of mortality rate.

4.2. Functional relationships between environmental parameters and bacterial abundance

GAM allows the effects of individual environmental parameters to be examined and the underlying controlling mechanisms being inferred. The positive correlation between bacterial abundance and temperature at low temperatures has been shown in a bivariate plot with the data compiled globally by Li (1998) and also in local studies (Shiah and Ducklow, 1994). Controlling other parameters, the positive relationship between bacterial growth rate and temperature can be translated into that between bacterial abundance and temperature when the effect of temperature is less in mortality than in growth rates (*e.g.*, $\alpha_2 < \alpha_1$ in Eq. (1)). If the effect of T is greater in mortality than in growth terms, bacterial abundance will decrease with increasing T even if growth rate still increases with temperature. Following this logic, the unimodal relationship between temperature and bacterial abundance in Fig. 4 suggests that, at low temperatures, the effect of temperature is greater in growth than in mortality terms, while the converse is true at higher temperatures. Current literature reviews are not able to fully support or reject this hypothesis. The analysis by Lopez-Urrutia and Moran (2007) implies that the temperature coefficient of bacterial specific growth rate does not vary along the whole temperature gradient. It is worth mentioning that they assumed that bacterial specific growth rate is density independent, which remains to be confirmed. Another problem in calculating bacterial specific growth rate from bacterial production data in literature surveys is the choice of conversion factors that convert leucine/thymidine incorporation to carbon assimilation. While Lopez-Urrutia and Moran (2007) used a constant conversion factor, it is actually a variable value that may relate with bacterial growth efficiency (Alonso-Sáez et al., 2007). An early synthesis study on bacterivory rate (Vaqué et al., 1994) reported that the community bacterivory rate increases with temperature at low temperatures, but reaches a plateau at high temperatures, which, at face value, contradicts with our hypothesis. However, in order to fit into our model, their community bacterivory rate (bacteria $mL^{-1}\,h^{-1}$) needs to be converted to mortality rate ($mL^{-1}\,h^{-1}$) by dividing by bacterial abundance. The bacterial community mortality rate also needs to be explicitly expressed as functions of *T, Chl, N*, and *B*.

Similarly, the strong positive relationship between chlorophyll and bacterial abundance suggests that chlorophyll plays a more important role in affecting growth than mortality. A Michaelis-Menton function relating chlorophyll and bacterial growth rate like in Eq. (1) has been constructed by Lopez-Urrutia and Moran (2007). Again, for the purpose of inferring rate–biomass relationships, bacterial abundance needs to be incorporated into their model. Similar functions relating bacterial mortality rate with *Chl* (and bacterial abundance), to our knowledge, have not been shown in literature.

Table 1
Statistical summary of generalized additive models. R^2 is the adjusted proportion of total variability explained by the model. GCV: generalized cross validation score. *n*: the total number of samples. s: thin plate regression spline. te: tensor product spline. *b* is a mean constant.

Model	R^2	GCV	*n*
(a) $logHB = s(logChl) + s(Temp) + te(Lon, Lat, Month) + s(Depth) + s(Sal) + s(MLD) + s(Nutricline) + b$	0.88	0.042	1437
(b) $logHB = s(logChl) + s(Temp) + s(Lon, Lat) + s(Month) + s(Depth) + s(Sal) + s(MLD) + s(Nutricline) + b$	0.82	0.053	1437
(c) $logHB = te(Lon, Lat, Month) + s(Depth) + b$	0.81	0.064	1604
(d) $logHB = s(Lon, Lat) + s(Month) + s(Depth) + b$	0.68	0.079	1604
(e) $logHB = s(logChl) + s(Temp) + s(MLD) + s(Sal) + s(Nutricline) + b$	0.76	0.08	1437
(f) $logHB = s(logChl) + b$	0.62	0.14	1500
(g) $logHB = s(Sal) + b$	0.32	0.22	1604
(h) $logHB = s(Nutricline) + b$	0.32	0.24	1534
(i) $logHB = s(Temp) + b$	0.13	0.29	1604
(j) $logHB = s(MLD) + b$	0.13	0.30	1604

Author's personal copy

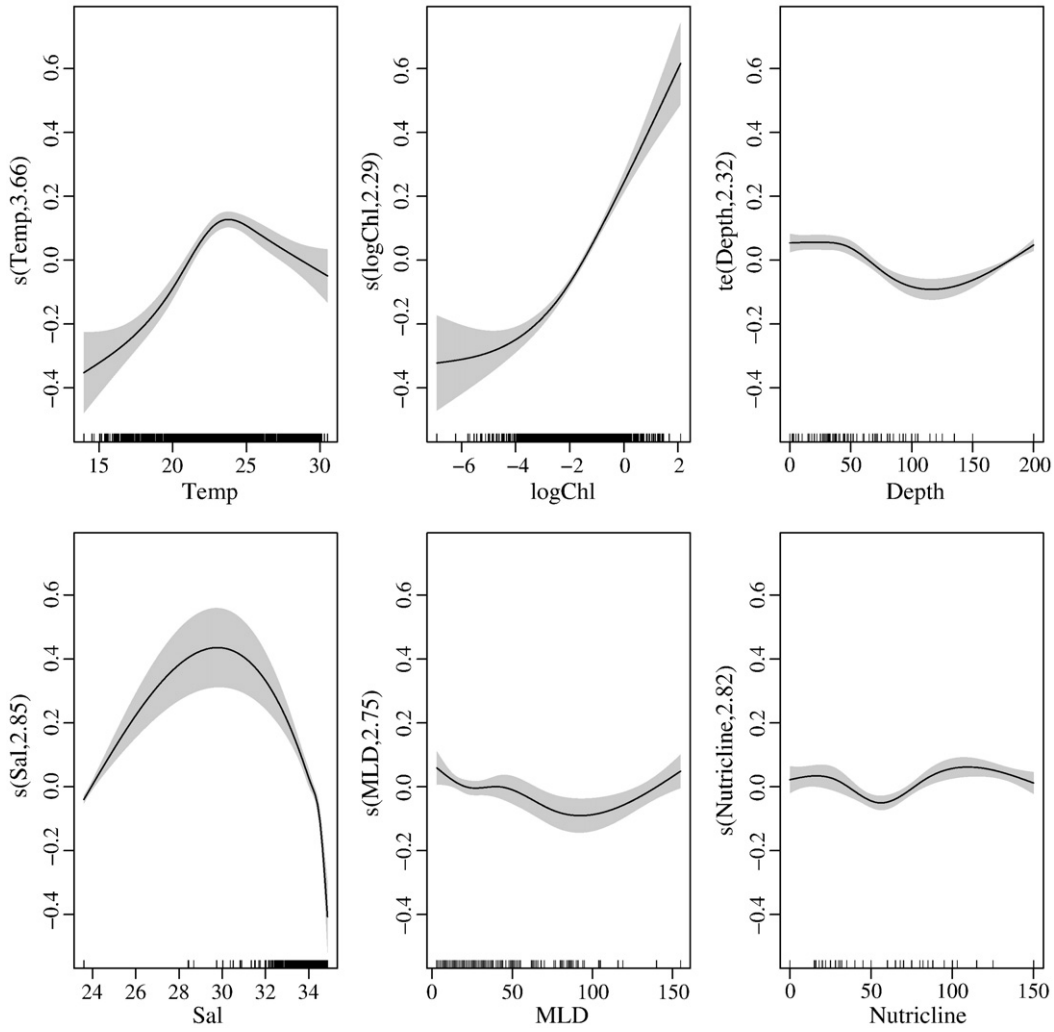

Fig. 4. Partial effects of temperature (*Temp*), log-transformed chlorophyll *a* (*logChl*), *Depth*, salinity (*Sal*), mixed layer depth (*MLD*), and nutricline depth (*Nutricline*). The solid line is the fitted line, while the shaded areas represent 95% confidential intervals. The numbers in the labels of y-axis denote the effective degrees of freedom.

Salinity is the second important parameter in the environmental model and can be considered as a proxy for freshwater discharge (including riverine discharge and precipitation) that delivers terrestrial organic matters and nutrients to the ocean. Organic matters within the Pearl River discharge can significantly stimulate bacterial biomass and respiration in coastal waters (Zhai et al., 2005).

The raw data show that, vertically, bacterial abundances tend to peak at the depth of chlorophyll maximum. After controlling *Chl* and other environmental factors, the decrease of bacterial abundance with increasing depth might result from the stimulatory effect of light on bacterial growth (Mary et al., 2008) or on growth rate and excretion of dissolved organic matter of phytoplankton. It is also possible that grazers shift to selectively graze on heterotrophic bacteria in deeper waters as autotrophic preys become less abundant. The slight decreasing trend of bacterial abundance with mixed layer depth might suggest a physical diluting effect, which is comparatively small. The decreasing trend of bacterial abundance with increasing nutricline depth when nutricline depth ranges from 20 m to 50 m may signify the effect of inorganic nutrient limitation. The increasing trend of bacterial abundance with increasing nutricline depth at stations with deeper nutricline depths is hard to interpret, but may relate with variations of chlorophyll-to-carbon ratios with nutrient availability. In nutrient limited waters, chlorophyll-to-carbon ratios tend to be high. A same amount of *Chl* is associated with higher concentrations of phytoplankton carbon, which may correspond to higher bacterial abundance, in nutrient limited waters than in nutrient replete waters.

The most intriguing part of the GAM model is to interpret the spatiotemporal patterns of residuals after controlling environmental parameters. The residuals were significantly greater in fall and winter than in summer and spring (Fig. 5E). Within each season, onshore–offshore decreasing gradients also existed, which leads us to conjecture that it is something of terrestrial origin that generated these patterns. In winter and fall, it is also possible that the prevailing northwest monsoon carried some terrestrial organic matters to the SCS, leading to greater abundances of bacteria. Therefore, we propose that one probable factor leading to the residual patterns might relate to atmospheric deposition, which is stronger in winter (Gao et al., 1997). The pronounced high anomalies northeast of the Hainan Island (Fig. 5) may relate with the local upwelling events (Jing et al., 2011), during which the organic matters from the sediment re-suspension might support plethora bacterial biomass. It will be useful in future studies to identify the 'missing' key variables that can account for these residual patterns.

4.3. Projection of bacterial abundance with future climate change

Our model might be useful for predicting future changes of bacterial abundance under the scenario of global warming. Based on the unimodal relationship between temperature and bacterial abundance (Fig. 4), increases in summer temperature would reduce bacterial abundances in surface waters (but may increase bacterial abundance

Author's personal copy

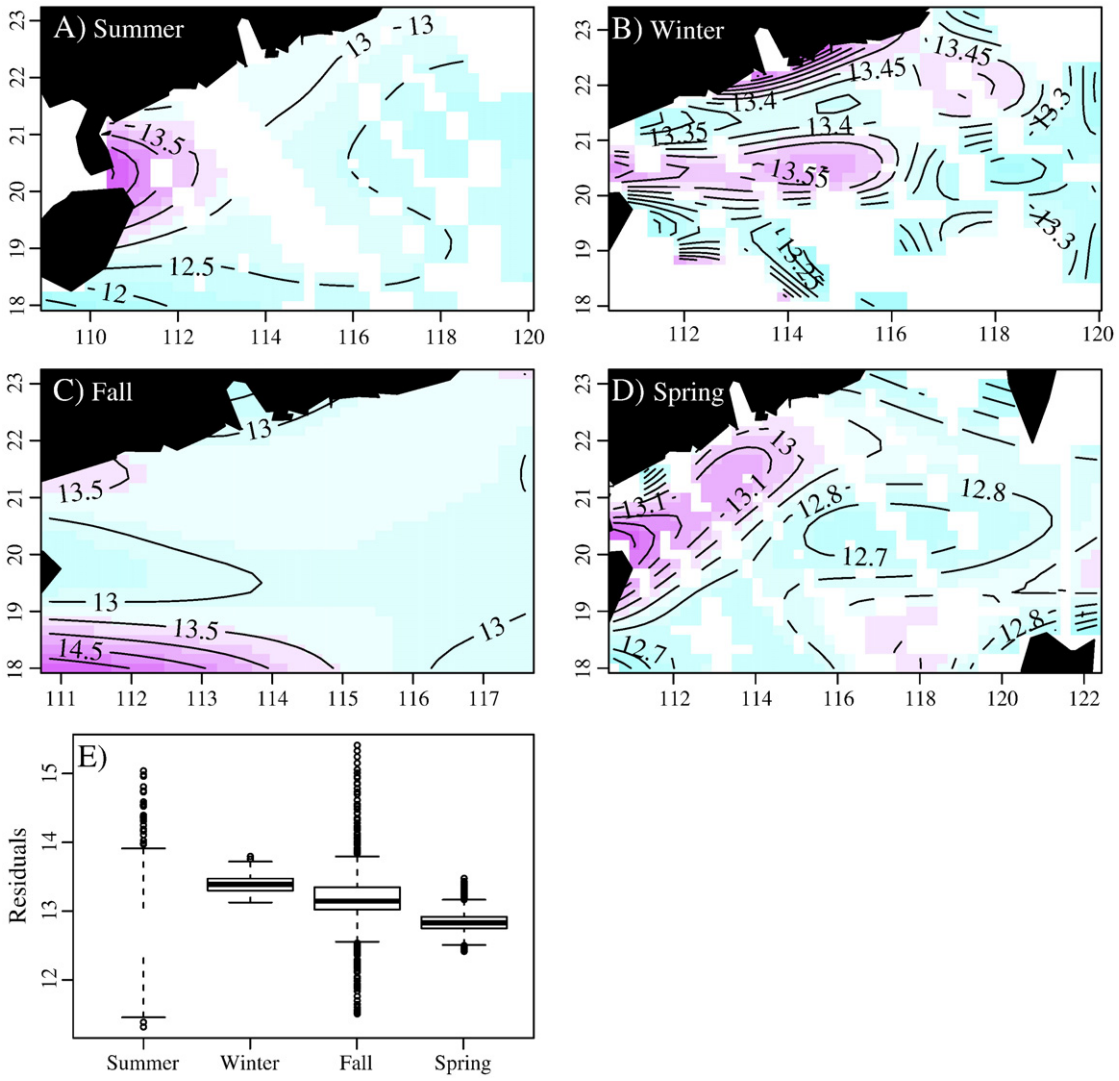

Fig. 5. (A–D) Interactions between spatial locations and season after controlling *Chl* (=0.3 mg m^{-3}), *Depth* (=5 m), *Temp* (=25 °C), *Sal* (=34 g kg^{-1}), *MLD* (=50 m), and *Nutricline* (=50 m). The number labels on contour lines denote ln abundance (cells mL^{-1}) of heterotrophic bacteria. (E) A "box-and-whiskers" plot of residuals in each season after controlling *Chl*, *Depth*, *Temp*, *Sal*, *MLD*, and *Nutricline*.

below the surface mixed layer), while increases of winter temperature would probably lead to abundance increases.

If we assume that warming promotes stratification and reduces *Chl* in the oligotrophic part of the SCS (Doney, 2006), bacterial abundances would be reduced accordingly, other things being equal. It is noteworthy that the first derivative of the fitting curve of *logHB* against *logChl* tends to be lower than unity (Fig. 4), which suggests that the decreasing rate of bacterial abundance will be lower than that of chlorophyll concentration and leads to the possibility that warming would likely lead to higher bacteria to phytoplankton biomass ratios. In addition, we need to assess how the riverine discharge and precipitation regimes will change in the future, as they are expected to affect bacterial biomass and the metabolic

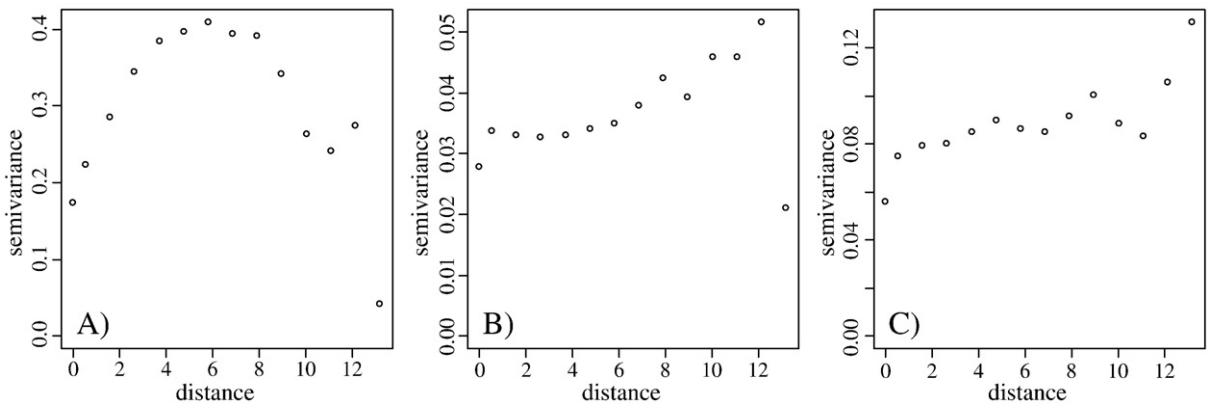

Fig. 6. Variograms of ln bacterial abundance of (A) raw data, (B) residuals of the full model, (C) residuals of the environmental model. The unit of distance is degree.

Author's personal copy

state of the northern SCS (Zhang et al., 2008). The river discharge might decrease in the future due to enhanced evaporation and human water demand. Or conversely, increasing melting of glaciers in the upstream might increase river discharge.

In summary, our exercise on modeling bacterial abundances in the SCS using GAM provides a relatively simple and straightforward approach for understanding environmental controlling mechanisms on bacterial abundance and for predicting future trends in bacterial abundance affected by global change. The partial effect of individual variable can be directly examined and can be related with its underlying controlling mechanism. It is desirable to incorporate variables (such as concentrations of labile or semilabile dissolved organic matters) that are directly related with bacterial growth. While this cannot be achieved right now, the use of surrogates (*e.g.*, chlorophyll and salinity) can also be used for prediction, but the interpretations are not so straightforward. After controlling known environmental parameters (*e.g.*, chlorophyll and temperature), the spatiotemporal patterns of the residuals provide hints on how some unmeasured factors affect bacterial distribution, which can stimulate new questions for future research. The model can be incorporated into more complex models dealing with carbon fluxes and ecosystem dynamics.

Acknowledgments

We sincerely thank the captain and crew of the research vessel *Dongfanghong2* and the chief scientists M. Dai, P. Cai, and W. Zhai for organizing the cruises. We also thank J. Hu, J. Zhu, and Z. Sun for providing the CTD data, L. Wang and C. Zhong for providing the chlorophyll data, and S. Song for assisting sampling on the sea. This study is supported by National Basic Research Program ("973" Program) of China through grant 2009CB421203 provided to B. H. and H. L. and Fundamental Research Funds for the Central Universities (2011121007 and 2012121058) of Xiamen University provided to B. C.. B.C. is also supported by State Key Laboratory of Tropical Oceanography (South China Sea Institute of Oceanology, Chinese Academy of Sciences; LTO1103). B. H. also acknowledges funding support from the National Science Foundation of China (40925018 and 41176112). H. L. is also supported by Hong Kong University Grant Council through the Area of Excellence program (AoE/P-04/04), Hong Kong Research Grant Council General Research Fund grants (661809, 661610 and 661911) and the TUYF Charitable Trust (TUYF10SC08).

References

Alonso-Sáez, L., Gasol, J.M., Arístegui, J., Vilas, J.C., Vaqué, D., Duarte, C.M., Agustí, S., 2007. Large-scale variability in surface bacterial carbon demand and growth efficiency in the subtropical northeast Atlantic Ocean. Limnology and Oceanography 52, 533–546.

Azam, F., Fenchel, T., Field, J.G., Gray, J.S., Meyer-Reil, L.A., Thingstad, F., 1983. The ecological roles of water-column microbes in the sea. Marine Ecology Progress Series 10, 257–263.

Barton, A.D., Dutkiewicz, S., Flierl, G., Bragg, J., Follows, M.J., 2010. Patterns of diversity in marine phytoplankton. Science 327, 1509–1511.

Benincà, E., Jöhnk, K.D., Heerkloss, R., Huisman, J., 2009. Coupled predator–prey oscillations in a chaotic food web. Ecological Letters 12, 1367–1378.

Borcard, D., Gillet, F., Legendre, P., 2011. Numerical Ecology with R. Springer, New York.

Church, M.J., Hutchins, D.A., Ducklow, H.W., 2000. Limitation of bacterial growth by dissolved organic matter and iron in the southern ocean. Applied and Environmental Microbiology 66, 455–466.

Church, M.J., Ducklow, H.W., Karl, D.K., 2004. Light dependence of [^{3}H]Leucine incorporation in the oligotrophic north Pacific Ocean. Applied and Environmental Microbiology 70, 4079–4087.

Del Giorgio, P.A., Williams, P.J.L., 2005. Respiration in aquatic ecosystems. Oxford University Press, New York.

Doney, S.C., 2006. Oceanography: plankton in a warmer world. Nature 444, 695–696.

Fasham, M.J.R., Ducklow, H., McKelvie, S.M., 1990. A nitrogen-based model of plankton dynamics in the oceanic mixed layer. Journal of Marine Research 48, 591–639.

Follows, M.J., Dutkiewicz, S., Grant, S., Chisholm, S.W., 2007. Emergent biogeography of microbial communities in a model ocean. Science 315, 1843–1846.

Furuya, K., Hayashi, M., Yabushita, Y., 1998. HPLC determination of phytoplankton pigments using N,N-dimethylfolmamide. Journal of Oceanography 54, 199–203.

Gan, J.P., Lu, Z.M., Dai, M.H., Cheung, A.Y.Y., Liu, H.B., Harrison, P.J., 2010. Biological response to intensified upwelling and to a river plume in the northeastern South China Sea: a modeling study. Journal of Geophysical Research 115, C09001, http://dx.doi.org/10.1029/2009JC005569.

Gao, Y., Arimoto, R., Duce, R.A., Zhang, X.Y., Zhang, G.Y., An, Z.S., Chen, L.Q., Zhous, M.Y., Gu, D.Y., 1997. Temporal and spatial distributions of dust and its deposition to the China Sea. Tellus B 49, 172–189.

Hastie, T., Tibshirani, R., 1989. Generalized additive models. Statistical Science 1, 297–318.

Hu, J., Kawamura, H., Hong, H., Qi, Y., 2000. A review on the currents in the South China Sea: seasonal circulation, South China Sea warm current and Kuroshio intrusion. Journal of Oceanography 56, 607–624.

Irwin, A.J., Finkel, Z.V., 2008. Mining a sea of data: deducing the environmental controls of ocean chlorophyll. PloS One 3, e3836 (3810.1371/journal.pone.0003836).

Jing, Z., Qi, Y., Du, Y., 2011. Upwelling in the continental shelf of northern South China Sea associated with 1997–1998 El Niño. Journal of Geophysical Research 116, C02033, http://dx.doi.org/10.1029/2010JC006598.

Kienast, M., Steinke, S., Stattegger, K., Calvert, S.E., 2001. Synchronous tropical South China Sea SST change and greenland warming during deglaciation. Science 291, 2132–2134.

Kostadinov, T.S., Siegel, D.A., Maritorena, S., 2009. Retrieval of the particle size distribution from satellite ocean color observations. Journal of Geophysical Research 114, C09015, http://dx.doi.org/10.1029/2009JC005303.

Li, W.K.W., 1998. Annual average abundance of heterotrophic bacteria and *Synechococcus* in surface ocean waters. Limnology and Oceanography 43, 1746–1753.

Li, W.K.W., Dickie, P.M., 2001. Monitoring phytoplankton, bacterioplankton, and virioplankton in a coastal inlet (Bedford Basin) by flow cytometry. Cytometry 44, 236–246.

Li, W.K.W., Head, E.J.H., Harrison, W.G., 2004. Macroecological limits of heterotrophic bacterial abundance in the ocean. Deep-Sea Research I 51, 1529–1540.

Llope, M., Chan, K.S., Ciannelli, L., Reid, P.C., Stige, L.C., Stenseth, N.C., 2009. Effects of environmental conditions on the seasonal distribution of phytoplankton biomass in the North Sea. Limnology and Oceanography 54, 512–524.

Lopez-Urrutia, A., Moran, X.A.G., 2007. Resource limitation of bacterial production distorts the temperature dependence of oceanic carbon cycling. Ecology 88, 817–822.

Lorance, P., Pawlowski, L., Trenkel, V.M., 2010. Standardizing blue ling landings per unit effort from industry haul-by-haul data using generalized additive models. ICES Journal of Marine Science 67, 1650–1658.

Marie, D., Partensky, F., Jacquet, S., Vaulot, D., 1997. Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. Applied and Environmental Microbiology 63, 186–193.

Mary, I., Tarran, G.A., Warwick, P.E., Terry, M.J., Scanlan, D.J., Burkill, P.H., Zubkov, M.V., 2008. Light enhanced amino acid uptake by dominant bacterioplankton groups in surface waters of the Atlantic Ocean. FEMS Microbiology Ecology 63, 36–45.

McClain, C.R., 2009. A decade of satellite ocean color observations. Annual Review of Marine Science 1, 19–42.

Olson, R.J., Zettler, E.R., DuRand, M.D., 1993. Phytoplankton analysis using flow cytometry. In: Kemp, P., Sherr, B.F., Sherr, E.B., Cole, J.J. (Eds.), Handbook of methods in aquatic microbial ecology. Lewis Publishers, New York, pp. 175–186.

Parsons, T.R., Maita, Y., Lalli, C.M., 1984. A manual of chemical and biological methods for seawater analysis. Pergamon, New York.

Pomeroy, L.R., Deibel, D., 1986. Temperature regulation of bacterial activity during the spring bloom in Newfoundland coastal waters. Science 233, 359–361.

Proctor, L.M., Fuhrman, J.A., 1990. Viral mortality of marine-bacteria and cyanobacteria. Nature 343, 60–62.

Rivkin, R.B., Anderson, M.R., 1997. Inorganic nutrient limitation of oceanic bacterioplankton. Limnology and Oceanography 42, 730–740.

Rivkin, R.B., Anderson, M.R., Lajzerowicz, C., 1996. Microbial processes in cold oceans. 1. Relationship between temperature and bacterial growth rate. Aquatic Microbial Ecology 10, 243–254.

Shiah, F.K., Ducklow, H.W., 1994. Temperature regulation of heterotrophic bacterioplankton abundance, production, and specific growth-rate in Chesapeake Bay. Limnology and Oceanography 39, 1243–1258.

Steinhoff, T., Friedrich, T., Hartman, S.E., Oschlies, A., Wallace, D.W.R., Kortzinger, A., 2010. Estimating mixed layer nitrate in the North Atlantic Ocean. Biogeosciences 7, 795–807.

Vaqué, D., Gasol, J.M., Marrasé, C., 1994. Grazing rates on bacteria: the significance of methodology and ecological factors. Marine Ecology Progress Series 109, 263–274.

Vaque, D., Guadayol, O., Peters, F., Felipe, J., Angel-Ripoll, L., Terrado, R., Lovejoy, C., Pedros-Alio, C., 2008. Seasonal changes in planktonic bacterivory rates under the ice-covered coastal Arctic Ocean. Limnology and Oceanography 53, 2427–2438.

Wong, G.T.F., Ku, T.L., Mulholland, M., Tseng, C.M., Wang, D.P., 2007. The SouthEast Asian time-series study (SEATS) and the biogeochemistry of the South China Sea — an overview. Deep-Sea Research II 54, 1434–1447.

Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman and Hall, Boca Raton.

Zhai, W., Dai, M., Cai, W.-J., Wang, Y., Wang, Z., 2005. High partial pressure of CO_2 and its maintaining mechanism in a subtropical estuary: the Pearl River estuary, China. Marine Chemistry 93, 21–32.

Zhang, S., Lu, X.X., Higgitt, D.L., Chen, C.-T.A., Han, J., Sun, H., 2008. Recent changes of water discharge and sediment load in the Zhujiang (Pearl River) Basin, China. Global and Planetary Change 60, 365–380.

Zubkov, M.V., Tarran, G.A., Mary, I., Fuchs, B.M., 2008. Differential microbial uptake of dissolved amino acids and amino sugars in surface waters of the Atlantic Ocean. Journal of Plankton Research 30, 211–220.