

Interaction among Non-toxic Phytoplankton, Toxic Phytoplankton and Zooplankton: Inferences from Field Observations

Shovonlal Roy · Sabyasachi Bhattacharya ·
Partha Das · Joydev Chattopadhyay

Abstract We explore the mutual dependencies and interactions among different groups of species of the plankton population, based on an analysis of the long-term field observations carried out by our group in the North–West coast of the Bay of Bengal. The plankton community is structured into three groups of species, namely, non-toxic phytoplankton (NTP), toxic phytoplankton (TPP) and zooplankton. To find the pair-wise dependencies among the three groups of plankton, Pearson and partial correlation coefficients are calculated. To explore the simultaneous interaction among all the three groups, a time series analysis is performed. Following an Expectation Maximization (E-M) algorithm, those data points which are missing due to irregularities in sampling are estimated, and with the completed data set a Vector Auto-Regressive (VAR) model is analyzed. The overall analysis demonstrates that toxin-producing phytoplankton play two distinct roles: the inhibition on consumption of toxic substances reduces the abundance of zooplankton, and the toxic materials released by TPP significantly compensate for the competitive disadvantages among phytoplankton species. Our study suggests that the presence of TPP might be a possible cause for the generation of a complex interaction among the large number of phytoplankton and zooplankton species that might be responsible for the prolonged coexistence of the plankton species in a fluctuating biomass.

Keywords Coexistence · Paradox of plankton · Toxin-producing plankton · Partial correlation · Time series · E-M algorithm · VAR model

1 Introduction

Exploration of the dynamics of phytoplankton and zooplankton is a central theme in marine ecology. However, because of the dazzling diversity of plankton species in marine ecosystems, plankton dynamics is hardly predictable. In most aquatic ecosystems, the prolonged coexistence of a large number of phytoplankton species on a seemingly limited variety of resources is paradoxical according to the well-known competitive exclusion principle [1]. Various competition models as well as competition experiments in the laboratory have established that the coexistence of more species than the limiting resources allow is possible only when some additional mechanisms are involved [1–12]. A brief overview of these mechanisms can be found in the reviews by [13] and [14]. In general, the mechanisms describing the coexistence of many phytoplankton species rely on a ‘non-equilibria’ hypothesis [13], which suggests that, due to several external and internal factors, the plankton species never settle to a steady state. On the other hand, analyzing a realistic mathematical model including predation and competition mechanisms in a plankton community, Gagnani et al. [15] have suggested that the selective predation by zooplankton on nutritionally-differentiable plankton species, and the food chain structures of aquatic ecosystems, play key roles in the coexistence of competing phytoplankton species. In particular, the nutrient bound of prey regulates the stable or oscillatory dynamics of predator-mediated competitive-coexistence [16]. However, among the phytoplankton community, some species have an ability to liberate “toxic” or “allelopathic agents” that affect the growth of other micro algae [17, 18]. Among marine algae, allelopathy was observed both *in vitro* and *in situ* (e.g., [19–24]). Studies indicate that the toxic substances have a great impact on phytoplankton–zooplankton interactions [25]. However, the chemical nature and rigorous role of allelopathic or toxic compounds has been poorly understood [26]. For the biochemical activity of the toxic substances released, the presence of toxic species is likely to influence competitive interactions and predator–prey relationships. Although a one-to-one interaction with a toxic species harms a non-toxic species, the role of toxic species in a combination of a great number of both species, in a well-mixed environment, is yet to explore extensively. The dynamics of the interacting species becomes even more unpredictable when a number of common grazers (zooplankton) are present.

In this article, based on an analysis of field data collected from the North–West coast of the Bay of Bengal, we investigate the interactive role of toxin-producing phytoplankton (TPP) on overall plankton populations. To explore the resultant effect of species level interaction among the phytoplankton and zooplankton species, we structure the plankton community into three functional groups: (1) non-toxic phytoplankton (NTP) – the species of phytoplankton that do not produce any toxic chemicals, (2) toxin-producing phytoplankton (TPP) – the phytoplankton species that have the ability to produce toxic or inhibitory substances, and (3) zooplankton – the common grazers of both NTP and TPP. On one hand, a non-toxic and a toxic species of phytoplankton (NTP, TPP) are competitors for the common resources, and on the other hand, zooplankton is the common grazer of both non-toxic and toxic phytoplankton. A simple competition model predicts that the competition for a single resource [27] would allow the persistence of either a toxic or a non-toxic species

of phytoplankton, and thus a strong negative correlation between the biomasses of toxic and non-toxic phytoplankton is expected. Moreover, simple predator-prey interaction [28] might show a periodic oscillation between a zooplankton and a phytoplankton species. Here, we investigate whether the resultant interaction among all the species of plankton, that come under the three groups described, deviates from the predictions of simple competition and predator-prey models. Using a set of our field-collected samples we explore by suitable statistical analysis the mutual dependencies and interactions among these three groups of plankton, and investigate the role of TPP on the overall interaction.

Ideally, the mutual dependency between two groups of plankton should be investigated when any third group is absent in the system. However, because the species-abundance data that has been used for our analysis is obtained from a field study and not from any laboratory experiment, it is impossible to physically isolate any of the pairs of plankton from the influences of the third one. To overcome this limitation, we have calculated both Pearson correlation coefficients and partial correlation coefficients. Statistically, Pearson correlation represents the mutual dependencies of the plankton groups in the live data, while partial correlation depicts the dependencies of any two groups when the effect of the third one is removed. We find that the nature of the partial correlations calculated for the above three pairs are similar to that of normal Pearson correlations. Now, for the analysis of the field-data under our consideration, an application of classical statistics such as partial and Pearson correlation might be controversial due to the autocorrelation structure. To overcome this situation we perform a time series analysis to explore the simultaneous interaction among all the three groups of plankton. Time series models are quite useful as potential prediction models. Using this property of the model, we find the correlation of species abundances of each group of plankton at any time point with the past abundances of the same group as well as that of the other groups. Unfortunately the data series under consideration are comparatively short. Moreover, due to limitations of the field work for various reasons (such as bad weather and limited resources), the abundance data obtained are not equally spaced in time, making it more difficult to use this data to approximate a continuous time series. All these limitations pose significant challenges for time series analysis of the data. Treating the unavailable samples as missing data, we have estimated the missing values by an imputation technique, namely, the Expectation-Maximization (E-M) algorithm. Incorporating the estimated missing values in the observed data set, we have fitted a "Vector-Autoregressive" (VAR) model. Because our objective is to study the simultaneous interactive effects of all the three groups of plankton species, we have used a Vector-Autoregressive model instead of a general "Autoregressive" model. The outcomes of the correlation analysis are compared with those of the time series analysis. The overall analysis establishes that the presence of TPP contributes to a significant effect on the overall plankton dynamics that might be favourable in preserving the coexistence of the plankton species in a marine ecosystem.

The organization of the paper is as follows. Section 2 gives a description of sample collection and the structure of the collected samples. Section 3 deals with the statistical models that have been used for analysis of the data. The detailed results of the analysis are given in Section 4. Finally in Section 5, implications of the overall results of the study is discussed.

2 Field Observation

2.1 Sampling

Monitoring and identification [29] of marine plankton population has been carried out by our group in the North–West coast of the Bay of Bengal since January, 1999 (for details of the study area and sampling procedure, see [30]). Frequency of the sampling was every two weeks except the months of September and October each year when, because of bad weather, the sampling program had to be suspended. (Local disturbances and financial constraints also limited our ability to collect data every fortnight.) Plankton samples were collected both from the surface and subsurface water (1–2 m depth) by a horizontal plankton tow with a 20 micron mesh net 0.3 m in diameter. Samples were collected from seven stations distributed at various distances from each other. The collected samples were preserved in 3% formaldehyde in seawater. Identification of plankton community was performed following the method of [29]. Counting of phytoplankton was made under the microscope using a Sedgewick–Rafter counting cell and expressed in number of samples per liter (no./l). The cell numbers were averaged over the number of replicates to get a single value of the species biomass at a given time point at the collection point.

2.2 Description of the Sample

A total of 115 species of phytoplankton have been identified in the water sample collected from the study area. Out of the 115 phytoplankton species, 65 species of diatom group, 19 of green algae (Chlorophyceae), nine of blue greens (*Cyanophyceae*) and 22 of Dinoflagellates have been identified. We have identified 13–15 species of phytoplankton [30] that have the ability to produce toxic or inhibitory compounds [31–34]. In general, the TPP group contains (1) planktonic or benthic microalgae that produce toxins (e.g., the motile stage of *Alexandrium*, the benthic *Gambierdiscus*), (2) other toxic dinoflagellates (e.g., *Pfiesteria*), (3) macroalgae that result in noxious smells (e.g., *Pilayella*), (4) a few species of Cyanobacteria or blue algae (e.g., *Microcystis*), (5) non-toxic microorganisms that result in hypoxic conditions (e.g., *Chaetoceros*, *Mesodinium*). For a detailed list of TPP species that have been observed by our group, see [30]. Sample data on plankton species for the period 2002–2004 have been considered for the present study.

3 Description of Statistical Techniques

3.1 Correlation and Partial Correlation Analysis

Let ρ_{NT} , ρ_{ZN} and ρ_{ZT} respectively be the sample Pearson correlation coefficients of the three pairs of plankton species, namely, TPP–NTP, NTP–zooplankton and TPP–zooplankton respectively. Similarly $\rho_{NT.Z}$, $\rho_{ZN.T}$ and $\rho_{ZT.N}$ are sample partial correlation coefficients of the same pairs when the effect of the third member is eliminated.

The partial correlations in terms of simple correlations are represented by $\rho_{NT,Z} = \frac{\rho_{NT} - \rho_{NZ}\rho_{TZ}}{\sqrt{1-\rho_{NZ}^2}\sqrt{1-\rho_{TZ}^2}}$, $\rho_{ZN,T} = \frac{\rho_{ZN} - \rho_{ZT}\rho_{NT}}{\sqrt{1-\rho_{ZT}^2}\sqrt{1-\rho_{NT}^2}}$ and $\rho_{ZT,N} = \frac{\rho_{ZT} - \rho_{ZN}\rho_{NT}}{\sqrt{1-\rho_{ZN}^2}\sqrt{1-\rho_{NT}^2}}$.

For each of the three pairs of plankton species, the tests of significance of the correlations are carried out through a t -test using the test statistic represented by $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$, where r is the sample correlation (Pearson/partial) of any two plankton species and n is the degrees of freedom. The test is rejected if $t > t_{\alpha,n-2}$, where α is the level of significance of the test.

3.2 Time Series Analysis

The data obtained from the field study are not equally spaced due to limitations of sampling in some time periods for each of the year. Usually the field experiments are conducted with an approximate interval of 15 days or a multiple of that interval. The incomplete data set showing the fluctuations of the abundance of different species is illustrated in Fig. 1. The gaps in the time axis show the positions of the missing of data. We estimate these missing data using an imputation method called Expectation-Maximization algorithm (E-M algorithm). To avoid numerical complications in the algorithm, we transform the original time scale to a scale of (0–1). However, this scaling does not affect the output of the estimation procedure. With the completed

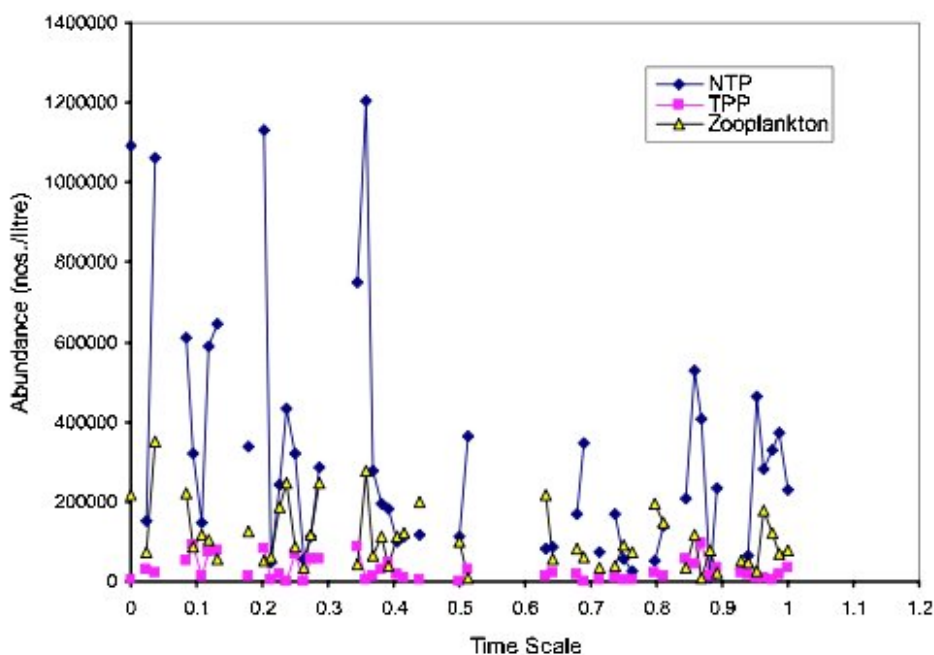


Fig. 1 Plot of NTP, TPP and zooplankton biomass against a normalized time-scale. In the actual time-scale, the time difference between two sampling points is measured in days. Samplings were conducted with an approximate interval of 15 days. The original time-scale is transformed to a scale of (0–1). The gaps in the time axis shows the breaks in data collection. The plot shows fluctuating abundances of all the three groups of plankton species

data set obtained on incorporation of the missing values in the observed data set, we analyze a Vector-Autoregressive model.

3.2.1 Estimation of Missing Values

To estimate the missing values we apply an Expectation Maximization (E-M) algorithm (for detail see, [35–37]), an iterative method that deals with the prediction of the missing values, and with the estimation of the model parameters, simultaneously.

Suppose the values of the variables are $Y_0, Y_1, Y_2, \dots, Y_{T-1}$ corresponding to times $\tau = \{0, 1, 2, \dots, T-1\}$. Define the time space $S = \{i : Y_i \text{ not missing}\}$. Let Y_S and $Y_{\bar{S}}$ (where, $\bar{S} = \tau \setminus S$) be the vector of observed values and missing values respectively, i.e., $Y_S = \{Y_i : i \in S\}$ and $Y_{\bar{S}} = \{Y_i : i \in \bar{S}\}$.

Keeping in mind that the data is fluctuating (see Fig. 1), and is neither completely cyclic nor of any other specific pattern, we consider the following non-stationary Auto-Regressive model (non-stationary $AR(p)$),

$$Y_t = \mu_t + \sum_{i=1}^p \beta_i Y_{t-i} + \epsilon_t, \quad (1)$$

where t indicates the transform time point on a scale of 0 to 1. $E[\epsilon_t] = 0$, $Var[\epsilon_t] = \sigma^2$ and μ_t is given by

$$\mu_t = a + bt + ct^2 + \sum_{i=1}^p \{\alpha_i \cos(2\pi(4i)t) + \gamma_i \sin(2\pi(4i)t)\}. \quad (2)$$

The parameter p is selected using an optimization technique following [38] with the help of ITSM software.

The E-M algorithm works iteratively by alternatively applying two steps : the E-Step (expectation) and M-Step (maximization). Formally, let $\hat{\theta}^{(i)}$, for $i = 0, 1, 2, \dots$, denote the successive parameter estimates; the E and M steps are defined as :

E-step: Compute the conditional expectation (with respect to the missing values $Y_{\bar{S}}$ of the logarithm of the complete *a posteriori* probability function, $\log p(Y_{\bar{S}}, \theta | Y_S)$, given the observed data Y_S and the current parameter estimate $\hat{\theta}^{(i)}$ (usually called the Q-function):

$$Q(\theta | \hat{\theta}^{(i)}) \equiv E[\log p(Y_{\bar{S}}, \theta | Y_S)] \quad (3)$$

M-step: Update the parameter estimate according to

$$\hat{\theta}^{(i+1)} = \arg \max_{\theta} Q(\theta | \hat{\theta}^{(i)}). \quad (4)$$

The process continues until some stopping criterion is met. The detailed estimation procedure adopted from [35] and [36] is described in the [Appendix](#).

3.2.2 VAR Model

Considering that the abundance of any particular group of plankton depends on the abundances of all the three groups of species present at previous time

points, we explore the simultaneous interaction of the three plankton groups by a Vector-Autoregressive model (VAR). Let us define $P_t = (NTP_t, ZOO_t, TPP_t) = (N_t, Z_t, T_t)$ to be a vector variable. The VAR model we consider with a lag p , is denoted by

$$P_t = \phi(0) + \phi(1) P_{t-1} + \dots + \phi(p) P_{t-p} + \epsilon_t, \quad (5)$$

where ϵ_t is a white noise component. Here, $\phi(0)$ is a column vector with dimension (3×1) and the $\phi(i)$ are coefficient matrices of dimension (3×3) .

Now, from (5), each element of the vector P_t can be written as

$$N_t = \phi_N(0) + \sum_{i=1}^p \phi_{NN}(i) N_{t-i} + \sum_{i=1}^p \phi_{NZ}(i) Z_{t-i} + \sum_{i=1}^p \phi_{NT}(i) T_{t-i} + \epsilon_t, \quad (6)$$

$$Z_t = \phi_Z(0) + \sum_{i=1}^p \phi_{ZZ}(i) Z_{t-i} + \sum_{i=1}^p \phi_{ZN}(i) N_{t-i} + \sum_{i=1}^p \phi_{ZT}(i) T_{t-i} + \epsilon_t, \quad (7)$$

$$T_t = \phi_T(0) + \sum_{i=1}^p \phi_{TT}(i) T_{t-i} + \sum_{i=1}^p \phi_{TZ}(i) Z_{t-i} + \sum_{i=1}^p \phi_{TN}(i) N_{t-i} + \epsilon_t, \quad (8)$$

where

$$\phi(0) = \begin{pmatrix} \phi_N(0) \\ \phi_Z(0) \\ \phi_T(0) \end{pmatrix} \text{ and } \phi(i) = \begin{pmatrix} \phi_{NN}(i) & \phi_{NZ}(i) & \phi_{NT}(i) \\ \phi_{ZN}(i) & \phi_{ZZ}(i) & \phi_{ZT}(i) \\ \phi_{TN}(i) & \phi_{TZ}(i) & \phi_{TT}(i) \end{pmatrix} \text{ and } i = 1, \dots, p.$$

The correlation between any two variables at any two time points (say, N_t and Z_t) is determined by the magnitude and asymptotic covariances of the estimated coefficients (i.e., $\phi_{NZ}(i)$ or $\phi_{ZN}(i)$) that are statistically significant at a desired level. The model estimation given above is based on the minimum AICC criterion.

4 Results and Discussion

The abundances of NTP, TPP and zooplankton show fluctuations over the observed time points (Fig. 1). The simple scatter plots show a trend of positive correlations for the pairs TPP-NTP and NTP-zooplankton; however, a negative correlation is observed for the pair TPP-zooplankton (Fig. 2). The linear regression lines on the scatter plots exhibit a negative slope for the pair zooplankton and TPP (see Fig. 2a), whereas, a positive slope is observed for the other two pairs, namely, (TPP, NTP) (see Fig. 2b) and (zooplankton, NTP) (see Fig. 2c). The regression lines fitted are respectively $Z_t = 126.012 - 0.6342T_t$, $Z_t = 77.680 + 0.0975N_t$ and $N_t = 216.967 + 3.4484T_t$. The sample normal and partial correlations for the pair (TPP, zooplankton) are -0.2270 and -0.3958 respectively, where the observed test statistic value for partial correlations is -2.9546 , which is significant at the 1% level. This result indicates that TPP biomass has a significant antagonistic correlation with zooplankton biomass. The correlation and partial correlations between NTP and zooplankton biomass is positive ($+0.3793$ and $+0.5008$ respectively), and significant

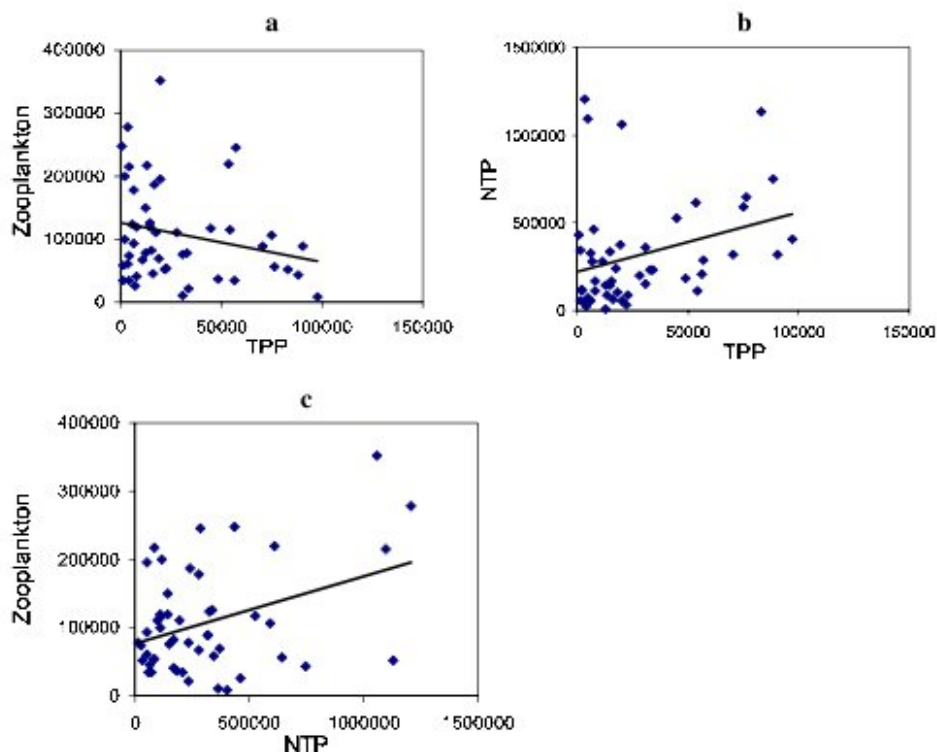


Fig. 2 Pair-wise dependencies among the species of NTP, TPP and zooplankton groups. Scatter plot between (a) TPP-zooplankton: the data is fit with a trend line of negative slope, $y = -0.6342x + 126.012$, ($R^2 = 0.312$) (b) TPP-NTP: the data is fit with a trend line of positive slope, $y = 3.4484x + 216.967$ ($R^2 = 0.4165$), and (c) NTP-zooplankton: the data is fit with a trend line of positive slope, $y = 0.0975x + 77.680$ ($R^2 = 0.437$)

at the 1% level (Table 1). This result, as one may expect, suggests that 1% the non-toxic species of phytoplankton as a whole is a favourable food for zooplankton. On the other hand, the biomass of NTP with TPP shows a significant positive Pearson and partial correlation (+0.3172, +0.4476 respectively, Table 1). This correlation structure indicates that the presence of TPP might be favourable for the existence of NTP. Our result here differs from a possible prediction of simple competition models.

Table 1 Pearson and partial correlations and test statistics

Variable	Zooplankton-NTP	Zooplankton-TPP	NTP-TPP
Correlation	0.3793	-0.2270	0.3172
Partial correlation	0.5008	-0.3958	0.4476
Test statistic	3.9667	-2.9546	3.4313
$p =$ values	0.00012	0.00243	0.00063
Comments	Significant at 1% level	Significant 1% level	Significant 1% level

Although, at the species level, two phytoplankton species exhibit resource competition, the biomasses of two types of phytoplankton when conceived as two groups of species show significant positive dependency. However, this correlation structure between NTP and TPP may also be due to the following reason. Because both NTP and TPP groups are sensitive in a similar manner to the environmental factors (e.g., temperature of water, inorganic materials), they may react to the environment similarly. Consequently the biomass at a given time point of both NTP and TPP might show a positive correlation. But the outcome of the autoregressive analysis described in the following paragraphs make this explanation improbable.

As mentioned earlier, to perform a time series analysis for exploring the simultaneous interactions (over the observed time points) among the species of non-toxic phytoplankton, toxic phytoplankton and zooplankton, the missing part of the data is estimated by an E-M algorithm (described in Section 3). The observed and predicted values for the three variables, zooplankton, NTP and TPP, are illustrated in Fig. 3a-c. In VAR analysis the estimated coefficient vectors and matrices for (6), (7) and (8) are represented in Table 2. Adopting the method due to [38] and using the ITSM software we obtain an optimum value of $p = 5$ for our data set. Most of the estimated coefficients $\phi_{NZ}(i)$ (or $\phi_{ZN}(i)$), $\phi_{ZT}(i)$ (or $\phi_{TZ}(i)$) and $\phi_{NT}(i)$ (or $\phi_{TN}(i)$) are significant at a 5% level of significance. The elements of the asymptotic variance-covariance matrix of the estimates have the same signs as the correlations, and the estimated asymptotic variances and covariances are also significant at a 5% level of significance. The statistically significant coefficients $\phi_{NZ}(i)$ are all positive, while the significant coefficients $\phi_{ZT}(i)$ (or $\phi_{TZ}(i)$) are all negative (see Table 2). It follows that on one hand the biomass of NTP in the past influences positively the present biomass of zooplankton. This positive correlation might be due to a simple predator-prey relationship. However, it is clear from the field data that the biomass of NTP and zooplankton do not show periodic cycles (Fig. 1) as predicted by a Lotka-Volterra interaction. One can also confirm this result from the shape of the limit cycles and by measuring the correlation between them.

On the other hand, the biomass of TPP in the past influences negatively the biomass of zooplankton in the present, suggesting a possible inhibitory effect of TPP on the growth and predatory activity of zooplankton. Some previous studies suggest that inhibitory substances released by TPP reduce the grazing pressure of zooplankton (e.g., [22, 39, 40]). Moreover, in their field studies, [41] have demonstrated that *micro* and *meso* zooplankton populations are reduced during the blooms of the chrysophyte *Aureococcus anophagefferens* on the southern Texas coast. The results of the time series analysis of our field samples resemble the previous results. Moreover, this analysis indicates a uniformity of the influence of toxic species identified in other studies and that of our study, on the predatory activity of zooplankton. We note that the inhibitory activity, which may be expected from a species-level interaction, holds even if we consider the species under groups in the case of multi-species interaction. Through analysis of a mathematical model, Gragnani et al. [15] have demonstrated that the difference in edibility between a filamentous cyanobacteria and green algae has widely different effects on phytoplankton-zooplankton community dynamics. Moreover, under certain conditions of nutrient condition and fish stock, the *Daphnia*-cyanobacteria-green algae systems may even settle to cyanobacteria, green algae or a mixed equilibrium, depending on the history [15]. We also find through time series analysis that the past values of NTP and TPP significantly affect the zooplankton

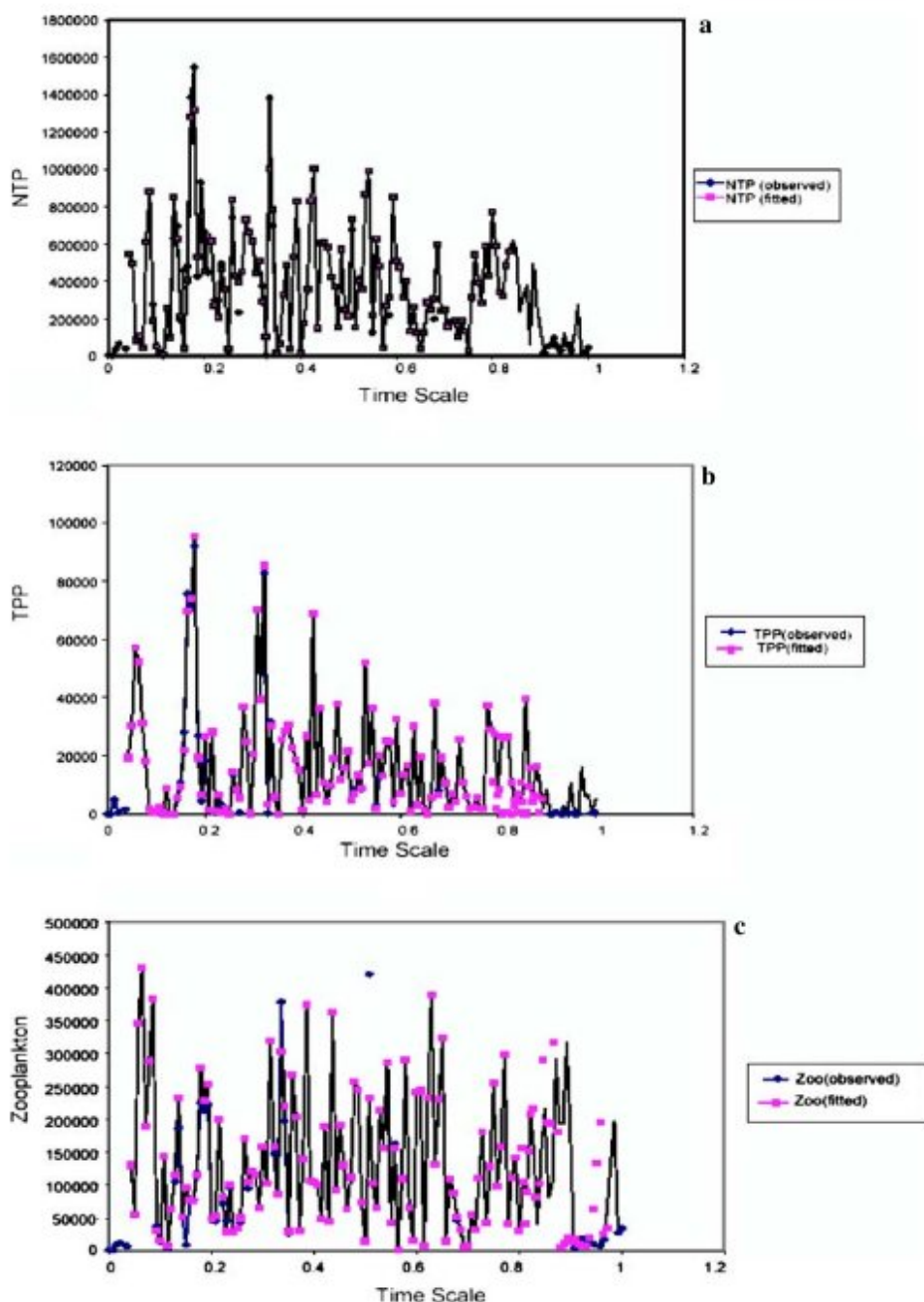


Fig. 3 Plot of observed and imputed values of (a) NTP, (b) TPP and (c) zooplankton

biomass, thereby affecting the overall community dynamics. Additionally, the analysis of a mathematical model of NTP-TPP-zooplankton suggests that the rate of toxin inhibition of TPP on zooplankton acts as a driving force for determining the

Table 2 Table showing the estimated coefficients for the VAR model

$\phi(0)$	0.600340E+06	0.633578E+06	0.301695E+06
$\phi(1)$	0.255	-0.109	0.293
	0.101	0.265	-0.117
	0.175	-0.199	0.290
$\phi(2)$	-0.221	0.234	0.172
	0.033	-0.065	-0.220
	0.087	-0.154	0.393
$\phi(3)$	0.299	-0.042	+0.022
	0.012	0.408	-0.227
	0.315	-0.293	-0.349
$\phi(4)$	-0.170	-0.009	0.147
	0.027	-0.031	-0.034
	0.064	-0.069	-0.216
$\phi(5)$	0.270	-0.138	0.128
	0.025	-0.286	-0.169
	0.016	-0.154	-0.047

stability and oscillatory dynamics of the overall plankton community; the details of this analysis has been reported in [42].

The coefficients $\phi_{NT(i)}$ (or $\phi_{TN(i)}$) are positive (Table 2), suggesting that the entire biomass of the TPP species in the past has a significant positive influence in determining the NTP biomass in the present. This result differs from what one may expect from a species-level interaction between a toxic and a non-toxic phytoplankton. Our results suggest that in a group of a number of non-toxic and toxic phytoplankton, toxicity indirectly helps the survival of the total biomass of non-toxic species. The mutual dependencies between NTP and TPP groups is thus different from that, for example, between blue-green algae and green algae (e.g., [8]). The mechanism for such dependency between NTP and TPP groups might be that the level of dominance between two non-toxic phytoplankton is reduced, when a third species that is toxic to both is present. This interaction can easily be demonstrated with the help of a simple three-species Lotka–Volterra competition model where two species are non-toxic and the third one is blindly toxic to both; the details of this model-analysis are reported in [43].

5 Concluding Remarks

The present article is an effort to explore the interactions among the overall plankton population, based on an analysis of the long-term field observations carried out in the North-West coast of the Bay of Bengal. Structuring the overall plankton community into nontoxic phytoplankton, toxic phytoplankton and zooplankton, we explore through statistical analysis the mutual dependencies and simultaneous interactions within the plankton populations. The pair-wise dependencies and simultaneous interactions among the three groups of species are explored with Pearson and partial correlation analysis and by a Vector-Auto Regressive model (VAR). The overall analysis suggests that, similar to a species-level interaction, the entire biomass of non-toxic phytoplankton is favourable for the growth of grazer zooplankton; however, because the toxic phytoplankton species liberate inhibitory compounds

that reduce the grazing pressure of zooplankton (see also [22, 39]), these species as a group are unfavourable for the growth of zooplankton. On the other hand, the interaction among the species of non-toxic and toxic species as two groups differs from a species-level competition. The analysis suggests that when considered as two groups in a combination of a large number of non-toxic and toxic species, instead of simple competition, an indirect mutualism exists between these two groups of phytoplankton. The effect of toxin-allelopathy is a potential candidate for such indirect mutualism, and thus the presence of TPP is favourable for the existence of NTP. Thus, the overall analysis suggests that toxin-producing phytoplankton play two distinct roles: the inhibition of consumption of toxic substances reduces the abundance of zooplankton, and the toxic materials released by TPP significantly compensate for the competitive disadvantages among phytoplankton species. The mutual dependencies at a group level suggests that, among the plankton community, the interactions present might be different from simple predator-prey relationships and competition. On one hand, the species of non-toxic phytoplankton that have a good dietary value enhances the growth of zooplankton; however, on the other hand, the species of toxic phytoplankton reduces the zooplankton growth (a situation similar to the interaction among cyanobacteria, green algae and *Daphnia* analyzed by [15]). Again, by reducing the competition among other phytoplankton, toxic species favours the survival of the weak competitors. In this way, a complex interaction as a combination of predator-prey relationship, competition-mutualism and toxin inhibition, is generated among the large number of phytoplankton and zooplankton species that might be a possible cause for the prolonged coexistence of plankton species in a fluctuating biomass. Although our field studies were concentrated in marine ecosystems, the mode of interaction among non-toxic phytoplankton, toxic phytoplankton and zooplankton is very likely to be similar for freshwater ecosystems also. However, further investigations will be necessary to extrapolate these results in the natural waters.

Acknowledgements The authors are grateful to Professor D. Sengupta, Stat-Math Division of the Indian Statistical Institute, and to Dr. Sugata Sen Roy of the University of Calcutta for useful discussions and comments. We are thankful to two learned referees for their valuable comments and suggestions, which improved the content of the paper. The research was supported by a project fund of the Indian Statistical Institute.

Appendix

Theory and Estimation Procedure

We have adopted a suitable estimation procedure for an E-M algorithm from [35] and [36]. We denote

$$X_{T \times 15} = \begin{pmatrix} 1 & \frac{0}{T} & \left(\frac{0}{T}\right)^2 & \cos\left(2\pi \cdot 4 \cdot \frac{0}{T}\right) & \dots & \cos\left(2\pi \cdot 24 \cdot \frac{0}{T}\right) & \dots & \sin\left(2\pi \cdot 24 \cdot \frac{0}{T}\right) \\ 1 & \frac{1}{T} & \left(\frac{1}{T}\right)^2 & \cos\left(2\pi \cdot 4 \cdot \frac{1}{T}\right) & \dots & \cos\left(2\pi \cdot 24 \cdot \frac{1}{T}\right) & \dots & \sin\left(2\pi \cdot 24 \cdot \frac{1}{T}\right) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \frac{T-1}{T} & \left(\frac{T-1}{T}\right)^2 & \cos\left(2\pi \cdot 4 \cdot \frac{T-1}{T}\right) & \dots & \cos\left(2\pi \cdot 4 \cdot \frac{T-1}{T}\right) & \dots & \sin\left(2\pi \cdot 24 \cdot \frac{T-1}{T}\right) \end{pmatrix}. \quad (9)$$

Let x'_i denote the i^{th} row of the matrix X . Then we partition X as follows:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \quad (10)$$

such that X_1 is a (6×15) and X_2 is a $(\overline{T-6} \times 15)$ matrix. Denote

$$\eta = (a \ b \ c \ \alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4 \ \alpha_5 \ \alpha_6 \ \gamma_1 \ \gamma_2 \ \gamma_3 \ \gamma_4 \ \gamma_5 \ \gamma_6)', \quad (11)$$

and

$$\beta = (\beta_1 \ \beta_2 \ \beta_3 \ \beta_4 \ \beta_5 \ \beta_6)'. \quad (12)$$

Following the least-squares estimation method, we minimize

$$\sum_{i=p}^{T-1} \left(Y_i - x'_i \eta - \sum_{i=1}^p \beta_i Y_{i-i} \right)^2. \quad (13)$$

The normal equations with respect to η are

$$\sum_{i=p}^{T-1} \left(Y_i - x'_i \eta - \sum_{i=1}^p \beta_i Y_{i-i} \right) x_i = 0. \quad (14)$$

The normal equations with respect to β_k are

$$\sum_{i=p}^{T-1} \left(Y_i - x'_i \eta - \sum_{i=1}^p \beta_i Y_{i-i} \right) Y_{i-k} = 0 \quad \forall k = 1(1)p. \quad (15)$$

Combining the set of equations in (27) we can write

$$\sum_{i=p}^{T-1} \left(Y_i - x'_i \eta - \sum_{i=1}^p \beta_i Y_{i-i} \right) v_{i-p} = 0 \quad (16)$$

where

$$v = (Y_{i-1} \ Y_{i-2} \ Y_{i-3} \ \dots \ Y_{i-p})'. \quad (17)$$

We define v_j as follows:

$$v_j = (Y_{p+j-1} \ Y_{p+j-2} \ \dots \ Y_j)'. \quad (18)$$

Now,

$$Y_i - x'_i \eta - \sum_{i=1}^p \beta_i Y_{i-i} = \sum_{i=0}^{T-1} a_{il} Y_i = \sum_{i \in S^c} a_{il} Y_i + \sum_{i \in S} a_{il} Y_i = b'_i Y_S - c'_i Y_S \quad (19)$$

where

$$b'_i = (a_{il}, i \in S^c), c'_i = (a_{il}, i \in S), Y_S = (Y_i, i \in S^c), Y_S = (Y_i, i \in S). \quad (20)$$

Under the normality assumption of the error terms (ϵ_i), we get the likelihood contribution of the t^{th} data as follows,

$$L_t(\eta, \beta_t) \propto \exp \left\{ - \left(b'_t Y_{\bar{S}} - x'_t \eta - c'_t Y_S \right)^2 \right\} \quad (21)$$

$$\implies \log L_t(\eta, \beta_t) \propto \left\{ - \left(b'_t Y_{\bar{S}} - x'_t \eta - c'_t Y_S \right)^2 \right\} \quad (22)$$

So the total log-likelihood of the data is

$$\log L(\eta, \beta_i : i = 1(1)p) = - \sum_{t=0}^{T-1} \left(b'_t Y_{\bar{S}} - x'_t - c'_t Y_S \right)^2 \quad (23)$$

i.e.,

$$- \log L(\eta, \beta_i : i = 1(1)p) = \sum_{t=0}^{T-1} \left(b'_t Y_{\bar{S}} - x'_t - c'_t Y_S \right)^2 \quad (24)$$

where the b_t 's are functions of the β_t 's, thus this log-likelihood is a function of η and β_t . We maximize likelihood L , (i.e., maximize log-likelihood $\log L$). Clearly, $\log L$ is maximized when $\sum_{t=0}^{T-1} \left(b'_t Y_{\bar{S}} - x'_t - c'_t Y_S \right)^2$ is minimized. So the least square estimates are also maximum likelihood estimates of the parameters. Now we have to solve the following minimization problem,

$$\min_{\eta, \beta_t} \sum_{t=0}^{T-1} \left(b'_t Y_{\bar{S}} - x'_t - c'_t Y_S \right)^2. \quad (25)$$

In the first step of E-M method, we estimate $Y_{\bar{S}}$ by minimizing $E \left[\sum_{t=p}^{T-1} \left(T_t - x'_t \eta - \sum_{i=1}^p \beta_i Y_{t-i} \right)^2 \mid Y_S \right]$. The distribution of $Y_{\bar{S}} \mid Y_S$ is described in the following lemma.

Lemma *The random variable $Y_{\bar{S}} \mid Y_S$ is normally distributed.*

Proof The expression $\sum_{t=0}^{T-1} \left(b'_t Y_{\bar{S}} - x'_t \eta - c'_t Y_S \right)^2$ can be written as follows,

$$\| B Y_{\bar{S}} - C Y_S - X \eta \|^2, \quad (26)$$

where,

$$B = \begin{pmatrix} b'_{p-1} \\ b'_{p-2} \\ \dots \\ b'_{T-1} \end{pmatrix}, \quad C = \begin{pmatrix} c'_p \\ c'_{p-1} \\ \dots \\ c'_{T-1} \end{pmatrix}. \quad (27)$$

Let

$$C Y_S + X \eta = d = d_B + d_B^\perp$$

where

$$d_B = B \left(B' B \right)^{-1} B' d, \quad d_B^\perp = d - d_B,$$

i.e., d_B is the projection of d on $C(B)$ (the column space of B), which implies that $B Y_{\bar{S}} - d_B \in C(B)$ and is orthogonal to d_B^\perp . So,

$$\begin{aligned} \| B Y_{\bar{S}} - C Y_S - X \eta \|^2 &= \| B Y_{\bar{S}} - d_B - d_B^\perp \|^2 \\ &= \| B Y_{\bar{S}} - d_B \|^2 + \| d_B^\perp \|^2 \\ &\cong \| B Y_{\bar{S}} - d_B \|^2 \\ &= \| B Y_{\bar{S}} - (B' B)^{-1} B' d \|^2 \\ &= \left(Y_{\bar{S}} - (B' B)^{-1} B' d \right)' (B' B) \left(Y_{\bar{S}} - (B' B)^{-1} B' d \right). \end{aligned} \quad (28)$$

The above expression is in the exponent of the probability density function of $Y_{\bar{S}} | Y_S$, thus it follows that

$$Y_{\bar{S}} | Y_S \sim N \left((B' B)^{-1} B' d, \hat{\sigma}^2 (B' B)^{-1} \right). \quad (29)$$

Next, we estimate $Y_{\bar{S}}$ by $(B' B)^{-1} B'$, where B and d are based on the current estimates of b_t, c_t and η . □

The normal equations (32) and (36) can be rewritten as follows,

$$X_2' (U \beta + X_2 \eta - Y_2) = 0, \quad (30)$$

$$V (U \beta + X_2 \eta - Y_2) = 0, \quad (31)$$

where X_2 has been defined in (20),

$$Y_2 = (Y_p \ Y_{p-1} \ \dots \ Y_{T-1})', \quad (32)$$

$$U = (u_1 \ u_2 \ \dots \ u_p), \quad u_i = (Y_{p-i} \ Y_{p-i+1} \ \dots \ Y_{T-i+1})', \quad i = 1(1)p, \quad (33)$$

$$V = (v_0 \ v_1 \ \dots \ v_j) \quad v_j = (Y_{p+j-1} \ Y_{p+j-2} \ \dots \ Y_j)', \quad j = 0(1)\overline{T-p-1}. \quad (34)$$

Combining (30) and (14) we get,

$$\begin{pmatrix} X_2' \\ V \end{pmatrix} (Y_2 - \sum \zeta) = 0, \quad (35)$$

where

$$\sum = (X_2 \ U), \quad \zeta = \begin{pmatrix} \eta \\ \beta \end{pmatrix}. \quad (36)$$

By definitions of U and V ,

$$U = V' \Rightarrow \sum' (Y_2 - \sum \zeta) = 0 \Rightarrow \sum' Y_2 = \sum' \sum \zeta \Rightarrow \hat{\zeta} = \left(\sum' \sum \right)^{-1} \sum' Y_2. \quad (37)$$

Note that estimation of ζ is an iterative procedure, because at each step we require to have estimates of \sum and Y_2 which are, in turn, the functions of the previous set of estimates of ζ .

References

- Hutchinson, G.E.: The paradox of the plankton. *Am. Nat.* **95**, 137–145 (1961)
- Chesson, P.: Multispecies competition in variable environments. *Theor. Popul. Biol.* **45**, 227–276 (1994)
- Chesson, P.L.: Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000)
- Huisman, J., Weissing, F.J.: Biodiversity of plankton by species oscillation and chaos. *Nature* **402**, 407–410 (1999)
- Richerson, P.J., Armstrong, R., Goldman, C.R.: Contemporaneous disequilibrium: a new hypothesis to explain the paradox of plankton. *Proc. Natl. Acad. Sci. U. S. A.* **67**, 1710–1714 (1970)
- Rothhaupt, K.O.: Mechanistic resource competition theory applied to laboratory experiments with zooplankton. *Nature* **333**, 660–662 (1988)
- Rothhaupt, K.O.: Laboratory experiments with a mixotrophic and phototrophic competitors. *Ecology* **77**, 716–724 (1996)
- Scheffer, M., Rinaldi, S., Gragnani, L., Mur, R., van Nes, E.H.: On the dominance of filamentous cyanobacteria in shallow, turbid lakes. *Ecology* **78**, 272–282 (1997)
- Sommer, U.: Comparison between steady state and non-steady state competition: experiments with natural phytoplankton. *Limnol. Oceanogr.* **30**, 335–346 (1985)
- Sommer, U.: Nitrate- and silicate-competition among Antarctic phytoplankton. *Mar. Biol.* **91**, 345–351 (1986)
- Tilman, D.: Resource competition between planktonic algae: an experimental and theoretical approach. *Ecology* **58**, 338–348 (1977)
- Tilman, D.: Test of resource competition theory using four species of Lake Michigan algae. *Ecology* **62**, 802–815 (1981)
- Scheffer, M., Rinaldi, S., Huisman, J., Weissing, F.J.: Why phytoplankton communities have no equilibrium: solutions to the paradox. *Hydrobiologia* **491**, 9–18 (2003)
- Roy, S., Chattopadhyay, J.: Towards a resolution of ‘the paradox of the plankton’: a brief overview of the proposed mechanism. *Ecol. Complex.* **4**, 26–33 (2007)
- Gragnani, A., Scheffer, M., Rinaldi, S.: Top-down control of cyanobacteria: a theoretical analysis. *Am. Nat.* **153**(1), 59–72 (1999)
- Roy, S., Alam, S., Chattopadhyay, J.: Role of nutrient bound of prey on the dynamics of predator-mediated competitive-coexistence. *BioSystems* **82**, 143–153 (2005)
- Arzul, G., Seguel, M., Guzman, L., Denn, E.E.-L.: Comparison of allelopathic properties in three toxic alexandrium species. *J. Exp. Mar. Biol. Ecol.* **232**(C11), 285–295 (1999)
- Hallam, T., Clark, C., Jordan, G.: Effects of toxicants on populations: a qualitative approach. II. First order kinetics. *J. Theor. Biol.* **18**, 25–37 (1983)
- Chan, A., Andersen, R., Blanc, M.L., Harrison, P.: Algal planting as a tool for investigating allelopathy among marine microalgae. *Mar. Biol.* **84**, 287–291 (1980)
- Fistarol, G., Legrand, C., Graneli, E.: Allelopathic effect of prymnesium parvum on a natural plankton community. *Mar. Ecol.-Prog. Ser.* **255**, 115–125 (2003)
- Fistarol, G., Legrand, C., Selander, E., Hummert, C., Stolte, W., Graneli, E.: Allelopathy in alexandrium spp.: effect on a natural plankton community and on algal monocultures. *Aquat. Microb. Ecol.* **35**, 45–56 (2004)
- Nielsen, T.G., Kiørboe, T., Bjørnsen, P.K.: Effects of a *Chrysochromulina polylepis* sub surface bloom on the plankton community. *Mar. Ecol. Prog. Ser.* **62**, 21–35 (1990)
- Schmidt, L., Hansen, P.: Allelopathy in the prymnesiophyte *Chrysochromulina polylepis*: effect of cell concentration, growth phase and pH. *Mar. Ecol., Prog. Ser.* **216**, 67–81 (2001)
- Tillmann, U., John, U.: Toxic effects of alexandrium spp. on heterotrophic dinoflagellates: an allelochemical defence mechanism independent of psp-toxin content. *Mar. Ecol. Prog. Ser.* **230**, 47–58 (2002)

25. Kozłowski-Suzuki, B., et. al.: Feeding, reproduction and toxin accumulation by the copepods *Acartia biflosa* and *Eurytenora affinis* in the presence of the toxic cyanobacterium *Nodularia spumigena*. *Mar. Ecol. Prog. Ser.* **249**, 237–249 (2003)
26. Sole, J., García-Ladona, E., Ruardij, P., Estrada, M.: Modelling allelopathy among marine algae. *Ecol. Model.* **183**, 373–384 (2005)
27. Hardin, G.: The competitive exclusion principle. *Science* **131**, 1292–1298 (1960)
28. Odum, E.P.: *Fundamentals of Ecology*. Saunders, Philadelphia, PA (1971)
29. Tomas, C.R.: *Identifying Marine Diatoms and Dinoflagellates*, p. 858. Academic Press, San Diego, CA (1997)
30. Chattopadhyay J., Sarkar R.R., Mandal S.: Toxin producing plankton may act as a biological control for planktonic blooms-field study and mathematical modeling. *J. Theor. Biol.* **215**, 333–344 (2002a)
31. Chattopadhyay, J., Sarkar, R.R., Elabdlou, A.: A delay differential equation model on harmful algal blooms in the presence of toxic substances. *IMA J. Math. Appl. Med. Biol.* **19**, 137–161 (2002b)
32. Hallegraeff, G.M.: A review of harmful algae blooms and the apparent global increase. *Phycologia* **32**, 79–99 (1993)
33. Sarkar, R.R., Chattopadhyay, J.: The role of environmental stochasticity in a toxic phytoplankton–non-toxic phytoplankton–zooplankton system. *Environmetrics* **14**, 775–792 (2003)
34. Steidinger, K.A., et. al.: *Pfiesteria piscicida*, a new toxic dinoflagellate genus and species of the order Dinamoebales. *J. Phycol.* **32**, 157–164 (1996)
35. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc., Ser. B* **39**(1), 1–38 (1977)
36. Hartley, H.: Maximum likelihood estimation from incomplete data. *Biometrics* **14**, 174–194 (1958)
37. McLachlan, G., Krishnan, T.: *The EM algorithm and extensions*. In: *Wiley Series in Probability and Statistics*. Wiley, New York (1997)
38. Brockwell, P.J., Davis, R.A.: *Time Series: Theory and Methods*, 2nd edn. Springer, New York (1991)
39. Ives, J.D.: Possible mechanism underlying copepod grazing responses to levels of toxicity in red tide dinoflagellates. *J. Exp. Mar. Biol. Ecol.* **112**, 131–145 (1961)
40. Kirk, K., Gilbert, J.: Variation in herbivore response to chemical defence: zooplankton foraging on toxic cyanobacteria. *Ecology* **73**, 2208 (1992)
41. Buskley, E.J., Stockwell, A.J.: Effect of persistent “brown tide” on zooplankton population in the Laguna Madre of Southern Texas. In: Smayda, T.J., Shimuzu, Y. (Eds.), *Toxic Phytoplankton bloom in the Sea*. Elsevier, Amsterdam, Netherlands, pp. 659–666 (1997)
42. Roy, S., Alam, S., Chattopadhyay, J.: Competitive effects of toxin-producing phytoplankton on the overall plankton populations in the Bay of Bengal. *Bull. Math. Biol.* **68**(8), 2303–2320 (2006)
43. Roy, S., Chattopadhyay, J.: Toxin-allelopathy among phytoplankton species prevents competitive exclusion. *J. Biol. Syst.* **15**(1), 1–21 (2007)