Available online at www.sciencedirect.com
ScienceDirect
Estuarine, Coastal and Shelf Science 72 (2007) 63–80
ESTUARINE COASTAL AND SHELF SCIENCE
www.elsevier.com/locate/ecss

Application of Bayesian structural equation modeling for examining phytoplankton dynamics in the Neuse River Estuary (North Carolina, USA)

G.B. Arhonditsis [a,*], H.W. Paerl [b], L.M. Valdes-Weaver [b], C.A. Stow [c], L.J. Steinberg [d], K.H. Reckhow [e]

[a] *Department of Physical and Environmental Sciences, University of Toronto, 1265 Military Trail, Toronto, ON, Canada M1C 1A4*
[b] *Institute of Marine Sciences, University of North Carolina at Chapel Hill, Morehead City, NC 28557, USA*
[c] *NOAA Great Lakes Environmental Research Laboratory, Ann Arbor, MI 48105, USA*
[d] *Department of Civil and Environmental Engineering, Southern Methodist University, Dallas, TX 75275-0339, USA*
[e] *Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708, USA*

Received 12 May 2006; accepted 29 September 2006
Available online 20 November 2006

Abstract

We introduce a Bayesian structural equation modeling framework to explore the spatiotemporal phytoplankton community patterns in the Neuse River Estuary (study period 1995–2001). The initial hypothesized model considered the influence of the physical environment (flow, salinity, and light availability), nitrogen (dissolved oxidized inorganic nitrogen, and total dissolved inorganic nitrogen), and temperature on total phytoplankton biomass and phytoplankton community structure. Generally, the model gave plausible results and enabled the identification of the longitudinal role of the abiotic factors on the observed phytoplankton dynamics. River flow fluctuations and the resulting salinity and light availability changes (physical environment) dominate the up-estuary processes and loosen the coupling between nitrogen and phytoplankton. Further insights into the phytoplankton community response were provided by the positive path coefficients between the physical environment and diatoms, chlorophytes, and cryptophytes in the down-estuary sections. The latter finding supports an earlier hypothesis that these three groups dominate the phytoplankton community during high freshwater conditions as a result of their faster nutrient uptake and growth rates and their tolerance on low salinity conditions. The relationship between dissolved inorganic nitrogen concentrations and phytoplankton community becomes more apparent as we move to the down-estuary sections. A categorization of the phytoplankton community into cyanobacteria, dinoflagellates and an assemblage that consists of diatoms, chlorophytes, and cryptophytes provided the best results in the upper and middle segments of the estuary. Finally, the optimal down-estuary grouping aggregates diatoms and chlorophytes, lumps together dinoflagellates with cryptophytes, while cyanobacteria are treated separately. These structural shifts in the temporal phytoplankton community patterns probably result from combined bottom-up and top-down control effects.

Keywords: phytoplankton dynamics; structural equation modeling; Neuse River Estuary; aggregate variability; compositional variability; eutrophication

1. Introduction

The explicit recognition of the dual—compositional and aggregate—nature of community variability is highlighted as an essential piece of knowledge when exploring ecological patterns (Micheli et al., 1999). For example, the study of both total phytoplankton biomass and community composition can be very useful to ecosystem management and restoration, can assist environmental monitoring and also provide early ''warning signs'' of eutrophic trends in aquatic environments (Cottingham and Carpenter, 1998; Arhonditsis et al., 2003; Paerl et al., 2003a; Arhonditsis and Brett, 2005). In practice,

* Corresponding author.
E-mail address: georgea@utsc.utoronto.ca (G.B. Arhonditsis).

0272-7714/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ecss.2006.09.022

the exploration of compositional and aggregate variability and the identification of the underlying ecological mechanisms can be problematic and raise several methodological issues. At higher aggregation levels, plankton communities exhibit satisfactory predictability, and are proposed by theoretical ecologists as a ''paradigm'' for elucidating spatiotemporal patterns of complex natural systems (McCauley and Murdoch, 1987; Scheffer et al., 2003). The aggregate plankton properties (e.g., total biomass, productivity), however, are also characterized by lower sensitivity to perturbations and are often considered unreliable indicators of changes in ecosystem functioning (Schindler, 1990; Frost et al., 1995). Unlike the aggregate variates, populations are more sensitive to external perturbations (e.g., nutrient enrichment, episodic meteorological events), but they also exhibit high natural variability across time and space that impedes the delineation of real effects and background noise (Cottingham and Carpenter, 1998).

The selection of the optimal resolution level is often encountered in phytoplankton ecology and several studies have examined the efficiency of population (e.g., species), community (e.g., genera, taxa), and aggregate (e.g., biomass, primary productivity, nutrient cycling) variables to illuminate different aspects of ecosystem dynamics (McCormick and Cairns, 1994; Cottingham and Carpenter, 1998). The crux of the ''optimal grouping'' problem is to obtain stable baseline variability by encompassing information from multiple populations, while maintaining the individual population sensitivity to external perturbations. In this regard, Reynolds et al. (2002) underscored the importance of clustering species on the basis of their general properties, i.e., morphological, physiological, and ecological characteristics. The ''functional grouping'' is an advancement that accounts for different patterns of adaptive specialism, and recognizes that the concerted effects of external (e.g., climatic conditions, trophic interactions) and internal (e.g., interspecific competition) factors obfuscate the expected signals of the phytoplankton community (Sommer, 1995; Huisman and Weissing, 2001; Scheffer et al., 2003).

By acknowledging that the study of natural systems at different resolution levels entails trade-offs in sensitivity and predictability, we implicitly recognize the need for flexible modeling approaches that can effectively depict the relationships between ecological habitats and both aggregate and compositional variability (Reynolds et al., 2002). There is a need for a methodology that has the ability to synthesize information from a variety of conditions, effectively depict the strength of cause–effect relationships and filter out the noise in ecological time-series. Our study introduces a Bayesian structural equation modeling methodology that has the flexibility to (a) translate fairly complicated ecological phenomena and express them as functions of several conceptual environmental factors; (b) link the conceptual factors of interest with observed variables by explicitly acknowledging that none of those perfectly reflects the underlying property; (c) test both direct and indirect paths of this ecological structure and identify the importance of their role (McCune and Grace, 2002); and (d) sequentially update the model, integrate information over time and space, and further consolidate community pattern delineation by treating aggregate and compositional variability as ''two complementary aspects of the same phenomenon'' (e.g., phytoplankton dynamics).

Our modeling framework is used to examine the role of the physical environment (flow, salinity, and light availability) versus the role of nitrogen (dissolved oxidized inorganic nitrogen, and total dissolved inorganic nitrogen) on the spatiotemporal phytoplankton biomass and community composition patterns in the Neuse River Estuary. Specifically, we focus on the question: what is the relative importance of these factors on phytoplankton growth and compositional alterations in this freshwater-marine continuum? Our work is founded on diagnostic photopigment data, extensively evaluated as indicators of coastal water quality and trophic state in previous studies (Paerl et al., 2003a, 2006). By using a comprehensive data set that includes quantitative information from the five major phytoplankton taxa (dinoflagellates, diatoms, cyanobacteria, cryptophytes, and chlorophytes) of the estuary, we aim to distinguish between opportunistic behaviors and regularities in phytoplankton dynamics and test the extent of which these patterns are expressed in our structural equation modeling framework.

2. Materials and methods

2.1. Study area

The Neuse River drains a 16,008 km^2 watershed and discharges into the Neuse Estuary (35°00′N; 76°45′W) and Pamlico Sound (Fig. 1). The major land uses in the basin are agriculture (35%) and forestry (34%) while the upper portion of the basin includes much of North Carolina's Research Triangle (defined by the cities of Raleigh, Durham, and Chapel Hill), an area that has experienced economic prosperity and rapid population growth during the last three decades (Stow and Borsuk, 2003a). Rapid development is also occurring in lower portions of the basin with an increasing coastal population and an expanding commercial hog-farming industry. The Neuse River Estuary (NRE) is an intermittently mixed and shallow (<4 m) system, where salinity varies with precipitation and wind conditions, river discharges, and saltwater influx from Pamlico Sound (Luettich et al., 2000; Borsuk et al., 2001). The system has a long history of algal blooms, bottom water hypoxia, and fishkills (Stow and Borsuk, 2003b; Paerl et al., 2004). Excessive chlorophyll *a* (Chl *a*) levels are generally attributed to high point and nonpoint source inputs of nitrogen, though historic and other lines of evidence suggest that phosphorus has also contributed to excessive algal production (Qian et al., 2000; Buzzelli et al., 2002; Paerl et al., 2004). These eutrophication problems led the Neuse River to be characterized as one of the 20 most threatened rivers in the United States in 1997 (American Rivers Foundation, 1997). The Neuse has also been listed as an impaired water body on the Federal 303(d) list because, in certain segments, more than 10% of water quality samples analyzed for Chl *a* exceeded the 40 $\mu g\, L^{-1}$ criterion. Water-quality conditions in the Neuse River show significant interannual variability and

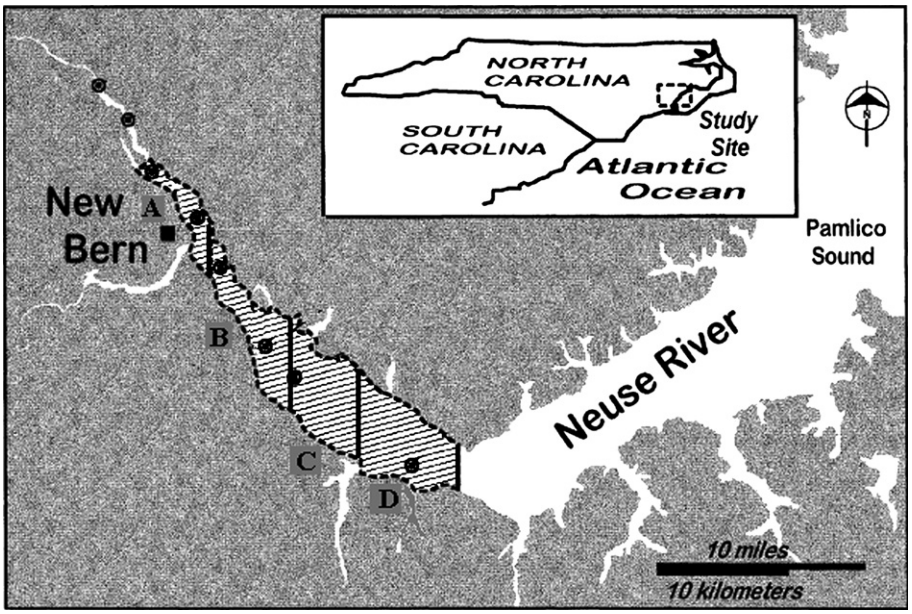

Fig. 1. The Neuse River Estuary; the vertical lines separate the four segments used for this study.

phytoplankton growth is regulated by a complex interplay between physical, chemical and biological factors (Rudek et al., 1991; Mallin and Paerl, 1994a; Borsuk et al., 2004), and is further influenced by a recent rise in the frequency of hurricanes and tropical storms (Paerl et al., 2003b; Paerl et al., 2006).

2.2. Data description

Daily mean flow rates were based on the United States Geological Survey streamflow gauging station at Fort Barnwell (Stow and Borsuk, 2003a), while the remaining data were provided from the UNC-CH Institute of Marine Sciences Neuse River Bloom Project, the Neuse River Estuary Modeling and Monitoring Project, ModMon, and the Atlantic Coast Environmental Indicators Consortium Project, ACE-INC (study period 1995–2001). Detailed information regarding the collection and analytical protocols and methods used in these programs can be found elsewhere (Pinckney et al., 1999; Luettich et al., 2000; Paerl et al., 2004). To assess spatial and temporal patterns of phytoplankton community structure, we also used diagnostic phytoplankton photopigment data (Paerl et al., 2003a). Phytoplankton photopigments (chlorophylls and carotenoids) representative of specific phytoplankton taxonomic groups were separated and quantified using high-performance liquid chromatography (HPLC) coupled to an inline photodiode array spectrophotometer (PDAS) (Tester et al., 1995; Jeffrey et al., 1997). Photopigment concentrations were subsequently analyzed using ChemTax, a matrix factorization program that determines the absolute and relative contribution of specific phytoplankton taxonomic groups to the total chlorophyll *a* pool based on an initial pigment ratio matrix specific to the study area (Mackey et al., 1997). Based on Mackey et al.'s (1996) suggestion to separate complex datasets into more homogeneous subsets and reduce the variation of pigment ratios due to large changes in phytoplankton species-composition, we defined homogenous data groupings (i.e., groups by season and salinity regime) of the HPLC-derived photopigment concentrations prior to running ChemTax (Pinckney et al., 1998). The initial pigment ratio matrix used in this study was obtained from a matrix characteristic of coastal phytoplankton groups and consisted of nine photopigments (alloxanthin, antheraxanthin, chlorophyll *b*, total chlorophyll *a* (chlorophyll *a* + chlorophyllide *a*), fucoxanthin, lutein, peridinin, violaxanthin, and zeaxanthin). These photopigments are indicative of five major algal taxonomic groups, chlorophytes, cryptophytes, cyanobacteria, diatoms, and dinoflagellates, the dominant phytoplankton classes found in the NRE. We also used a modification of the Pinckney et al. (1998) spatial segmentation by dividing the study area into four sections A, B, C, and D (Fig. 1), i.e., this study's first and third segments were grouped with the second and the fourth spatial compartments, respectively. For each segment, we calculated volume-weighted averages for all the environmental variables of the model based on the corresponding water volumes (m^3) for two depth intervals, i.e., surface to 2 m and 2 m to bottom.

2.3. Bayesian structural equation modeling

Structural equation modeling (SEM) is a multivariate statistical method that allows for evaluating a network of relationships between observed and latent variables. In this statistical technique, pre-conceptualizations that reflect research questions or existing understanding of system structure form the initial framework for model development. In contrast with multivariate regression, SEM can evaluate indirect effects between two explanatory variables, i.e., effects mediated by other intermediary variables (Bollen, 1989). Additionally,

SEM can explicitly incorporate uncertainty due to observation error or lack of validity of the observed variables. The latter aspect refers to the essential feature of representing variables of conceptual interest that are not directly measurable with multiple indicator (observed) variables. In this study, we also adopted a Bayesian approach to SEM that has several advantages over the classical methods (e.g., maximum likelihood, generalized and weighted least squares). For example, Bayesian SEM has the ability to incorporate prior knowledge about the parameters and more effectively treat unidentified models (Scheines et al., 1999; Congdon, 2003). In addition, the assumptions used to determine the latent variable metrics can be treated stochastically and can provide additional insight into the ecological structures (Arhonditsis et al., 2006). The modeling process does not rely on asymptotic theory, a feature that is particularly important when the sample size is small and the classical estimation methods are not robust (Congdon, 2003). Markov Chain Monte Carlo (MCMC) samples are taken from the posterior distribution, and as a result the procedure works for all sample sizes and various sources of non-normality (e.g., avoidance of multimodality problems; see Scheines et al., 1999). Finally, the Bayesian nature of our framework provides more realistic estimates of the existing knowledge/predictive uncertainty by taking into account both the uncertainty about the parameters and the uncertainty that remains when the parameters are known (posterior predictive distribution).

Our starting point is a ''conceptual/mental model'' that considers the effects of three latent variables, i.e., physical environment, nitrogen, and temperature on phytoplankton dynamics (as a general/abstract idea; Fig. 2). Each of these conceptual factors (latent variables) can be linked with observed variables (''what can be measured in the real world''), while it is explicitly acknowledged that none of these variables perfectly represents the underlying property (measurement errors; δ and ε in Fig. 2). Specifically, we hypothesized that the latent variable physical environment along with the three indicator variables attenuation coefficient (m^{-1}), salinity (‰), and daily flow rates ($m^3\,s^{-1}$) comprised the measurement model for the physical environment. The premise for the selection of the first two indicators was based on the findings of the principal factor analysis presented by Pinckney et al. (1997; see their Table 2), where salinity and attenuation coefficient had the highest positive and negative loadings on the first principal factor (31% of the observed variability), respectively. The inclusion of the flow rates builds upon the results of a recent study that highlights the regulatory role of flow on the spatiotemporal NRE phytoplankton dynamics (Borsuk et al., 2004). We also used the latent variable nitrogen and two surrogate variables: the first variable was total dissolved inorganic nitrogen (DIN) concentrations, while the second one included only the oxidized forms of inorganic nitrogen (nitrate + nitrite; NO_x). [It should be noted that the use of NH_4 and NO_x as surrogate variables for the latent variable nitrogen resulted in a model that was not supported by the data.] Temperature was considered as directly observable with no measurement error and the coefficient that relates temperature (expressed as water temperature deviance from 20 °C) to phytoplankton was not considered spatially constant. As a result, the respective path values are confounded with the effects of other drivers not explicitly accounted for by the model, and mainly aim to detect shifts in the seasonal phytoplankton patterns along the estuary.

The structural equation model that aimed to provide a quantitative description of total phytoplankton dynamics (aggregate phytoplankton SEM) had one endogenous latent variable phytoplankton combined with two indicator variables, i.e., chlorophyll *a* and primary productivity (Fig. 2). The basis for the selection of the two phytoplankton indicators was provided by the observed spatiotemporal similarity of the chlorophyll *a* and primary productivity NRE patterns (Pinckney et al., 1997). Note that this multivariate method accounts for two sources of error, i.e., measurement and structural error (its variance is denoted as ψ in Fig. 2). The latter error source reflects the latent variable model efficiency, i.e., ''how well can the physical environment, nitrogen, and temperature describe phytoplankton''. [The matrix presentation of the aggregate phytoplankton structural equation model is provided in Appendix A.]

Concurrently, we formulated and tested two alternative models to examine the relative importance of the various abiotic factors on the phytoplankton community composition (compositional phytoplankton SEMs). The development of these models was based on SEM's ability to use multiple observed variables for the representation of latent variables of conceptual interest (Bollen, 1989). Using this property, we constructed measurement models that allow several phytoplankton groups (e.g., species, genera or taxa) of variant degree of observed correlation to be amalgamated into single entities, i.e., latent variables that characterize phytoplankton functional groups. The first model considers the functional group A (PFG A) that comprises diatoms, cryptophytes and chlorophytes, and retains dinoflagellates and cyanobacteria as two further, distinct groups. The second model aggregates diatoms and chlorophytes (PFG B), lumps together dinoflagellates with cryptophytes (PFG C), while cyanobacteria are treated separately (Fig. 3). We tested the compatibility of the two different pre-conceptualizations with the observed ecological patterns, and selected the model with the higher performance in each spatial section to resolve the optimal aggregation level of the phytoplankton community structure.

Square root and natural logarithm transformations were used to achieve linearity among the observed indicators of the exogenous and the endogenous latent variables. Based on calculated residence times (Christian et al., 1991) and exploratory data analyses, mean daily flow rates were calculated for the two-day, one-week, two-week, and 25-day period preceding the sampling dates in the four spatial sections, respectively. Aside from the flow rate values, we used contemporaneous measurements from individual samplings for all the environmental variables, i.e., no time-averaging or lagged relationships were considered. The nature of the data used for model development (i.e., temporal resolution/aggregation and absence of lagged relationships) largely determined the

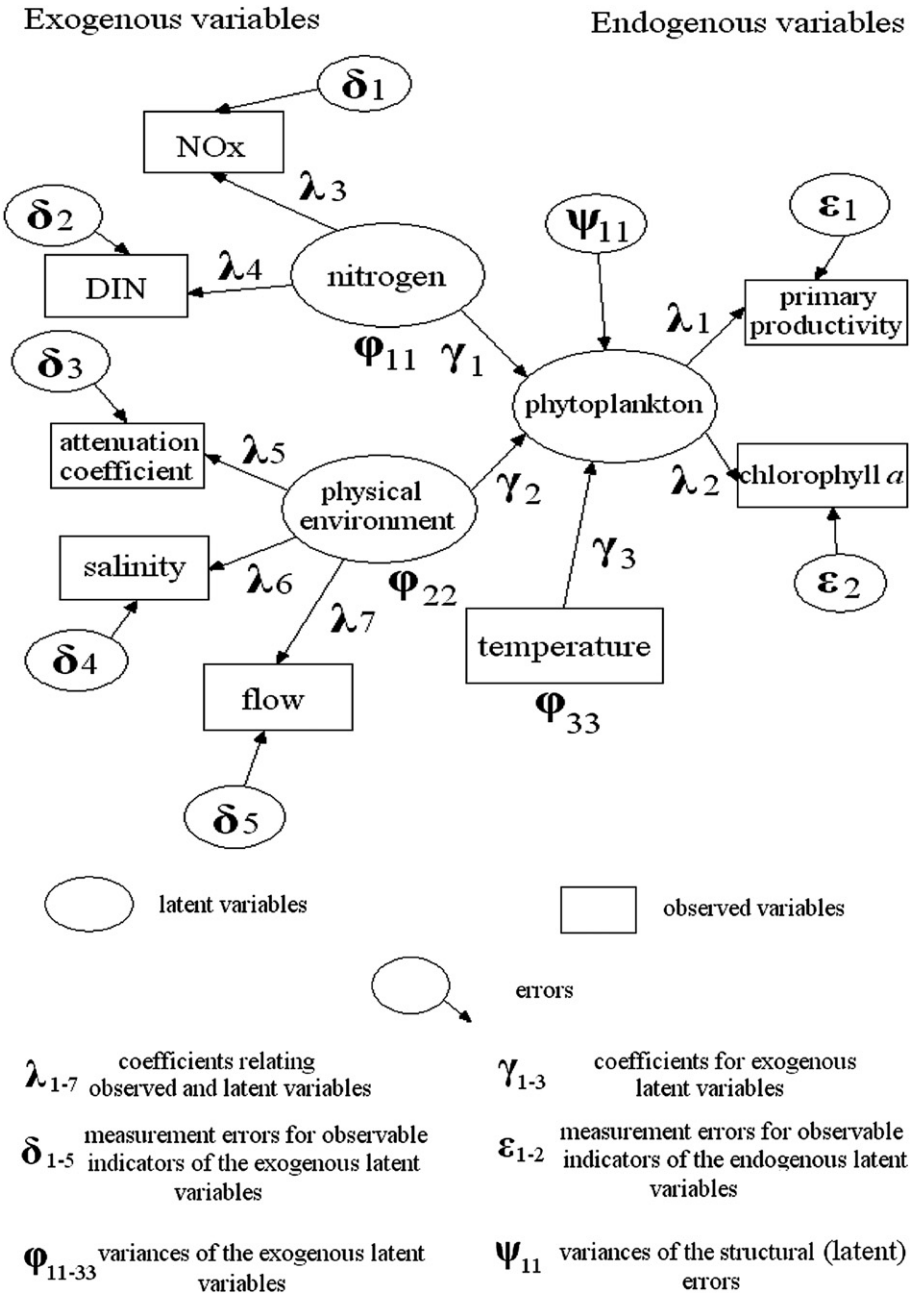

Fig. 2. The structural equation model used for predicting the Neuse River Estuary phytoplankton dynamics. The use of a rectangular box for the temperature implies that the variable was considered as directly observable with no measurement error ($\lambda_8 = 1.0$, and $\delta_6 = 0$). The metrics of the latent variables were set by fixing $\lambda_2 = \lambda_4 = \lambda_5 = 1.0$.

signs of the various paths of our structural equation models (e.g., negative dissolved inorganic nitrogen–phytoplankton relationships). In addition, even though the relative magnitudes of the nitrogen–phytoplankton path coefficients were used to draw inferences about the aggregate and compositional phytoplankton variability, our interpretation should not be considered as an evidence of nitrogen limitation in the Neuse River Estuary. Ambient nutrient concentrations partly reflect the residual nutrients from phytoplankton activity, but there are several other potentially influential factors that are not accounted for in our approach (e.g., phosphate, intracellular storage, zooplankton grazing). Thus, a combination of modeling and experimental approaches (e.g., bioassays) is certainly more appropriate to make a more definite statement regarding the status of nutrient limitation in the Neuse River Estuary (Granéli et al., 1990; Piehler et al., 2004). In a strict causal sense, the inclusion of external nutrient loading (instead of ambient nutrient concentrations) would have had a more unequivocal interpretation. However, given the spatially explicit character of our model, the consideration of this causal link also entails substantial increase of uncertainty and is beyond the scope of the present paper. Finally, the study by Borsuk et al.

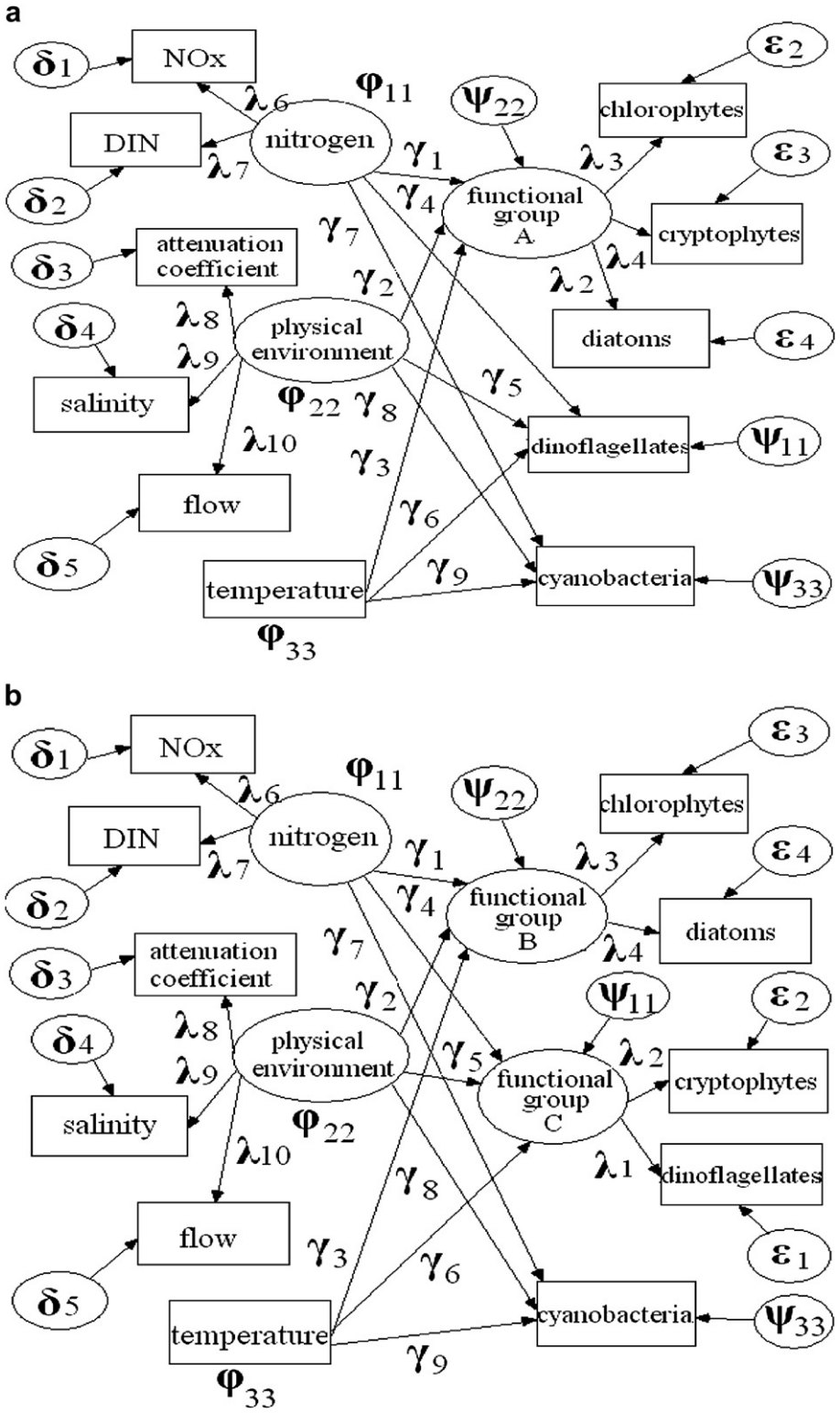

Fig. 3. The two alternative conceptualizations of the Neuse River Estuary phytoplankton community dynamics. (a) The use of a rectangular box for the temperature, dinoflagellates and cyanobacteria implies that the variables were considered as directly observable with no measurement error ($\lambda_1 = \lambda_5 = \lambda_{11} = 1.0$, and $\varepsilon_1 = \varepsilon_5 = \delta_6 = 0$). The metrics of the latent variables were set by fixing $\lambda_2 = \lambda_7 = \lambda_8 = 1.0$. (b) The use of a rectangular box for the temperature, and cyanobacteria implies that the variables were considered as directly observable with no measurement error ($\lambda_5 = \lambda_{11} = 1.0$, and $\varepsilon_5 = \delta_6 = 0$). The metrics of the latent variables were set by fixing $\lambda_1 = \lambda_3 = \lambda_7 = \lambda_8 = 1.0$.

(2004) found a non-monotonic relationship between the flow rates and DIN, chlorophyll *a*, and thus adopted a piecewise linear relationship (after the data-transformation) with a spatially varying breakpoint. Even though our approach assumes monotonicity, we did not find any evident trend in our predictive error distributions.

Assessment of the goodness-of-fit between the model outputs and the observed data was based on the posterior predictive *p*-value, i.e., the Bayesian counterpart of the classical *p*-value. In brief, the *p*-value is defined as the probability that the replicated data (i.e., the posterior predictive distribution) could be more extreme than the observed data. The null hypothesis H_0 is rejected if the tail-area probability is close to 0.0 or 1.0, whilst the model can be regarded as plausible if the *p*-value is near to 0.5. The discrepancy variable chosen for carrying out the posterior predictive model checks was the x^2 test (see also Gelman et al., 1996 for a detailed description of the posterior predictive *p*-value). Intersectional comparisons of the aggregate and compositional model performance ("Which spatial segment better supports our prior models?") along with the comparison of the two phytoplankton community composition pre-conceptualizations ("Which of the two phytoplankton groupings is better supported in each segment?") were based on the use of the Bayes factor (Kass and Raftery, 1995). When comparing two alternative models, the Bayes factor is the posterior odds of one model over the other (assuming the prior probability on either model is 0.5). If SEM_1 and SEM_2 denote the two alternative models and D corresponds to the observed data, the Bayes factor is

$$B_{12} = \frac{\mathrm{pr}(D|\mathrm{SEM}_1)}{\mathrm{pr}(D|\mathrm{SEM}_2)} \quad (1)$$

For model comparison purposes, the model likelihood ($\mathrm{pr}(D|\mathrm{SEM}_k)$; $k = 1,2$) is obtained by integrating over the parameter space:

$$\mathrm{pr}(D|\mathrm{SEM}_k) = \int \mathrm{pr}(D|\mathrm{SEM}_k, \theta_k)\pi(\theta_k|\mathrm{SEM}_k)\mathrm{d}\theta_k \quad (2)$$

where θ_k is the parameter vector under model SEM_k and $\pi(\theta_k|\mathrm{SEM}_k)$ is the prior density of θ_k. Using the MCMC method, we can estimate $\mathrm{pr}(D|\mathrm{SEM}_k)$ from posterior samples of θ_k. Letting $\theta_k^{(i)}$ be samples from the posterior density $\mathrm{pr}(\theta_k|\mathrm{SEM}_k)$, the estimated $\mathrm{pr}(D|\mathrm{SEM}_k)$ is:

$$\overline{\mathrm{pr}(D|\mathrm{SEM}_k)} = \left\{ \frac{1}{m} \sum_{i=1}^{m} \mathrm{pr}\left(D \middle| \mathrm{SEM}_k, \theta_k^{(i)}\right)^{-1} \right\}^{-1} \quad (3)$$

the harmonic mean of the likelihood values (Kass and Raftery, 1995).

3. Results

The aggregate phytoplankton SEM provided satisfactory results (see the posterior predictive *p*-values reported in Fig. 4). The root mean square error (RMSE) values for the modeled Chl *a* and primary productivity in the four segments varied between 3–8.4 $\mu g\,L^{-1}$ and 4–27 $mg\,C\,m^{-3}\,h^{-1}$. The RMSE values for the two inorganic nitrogen forms varied from 10 to 32 $\mu g\,L^{-1}$, while the respective salinity, attenuation coefficient, and flow ranges were 1.41–2.69‰, 0.37–0.48 m^{-1} and 38–61 $m^3\,s^{-1}$, respectively. Based on the Bayes factor values, we also infer that the model was better supported by the data in section D, while section B seems to be the least favorable for the underlying conceptualization of the total phytoplankton dynamics (Table 1). The direct path from nitrogen to phytoplankton is very weak in the upper part of the estuary (<-0.01) and becomes stronger as we move to the down-estuary, i.e., -0.14, -0.31, and -0.29 at the segments B, C and D, respectively. The physical environment and most importantly the flow rates play a significant role on phytoplankton dynamics in the up-estuary segments A and B, where the respective total standardized effects were -0.40 ($=-0.44 \times 0.91$) and -0.72 ($=-0.77 \times 0.93$). The path from the physical environment to phytoplankton was weaker in the down-estuary segments, while the posterior median effects were switched from negative (-0.13 in segment C) to positive (0.12 in segment D) between the two spatial sections. Interestingly, salinity has a slightly higher total effect than the other two indicator variables (attenuation coefficient and flow) of the physical environment measurement model. The temperature effects on phytoplankton become weaker as we move to the down-estuary sections (i.e., from 0.42 to 0.08). Finally, the comparison between the observed and predicted chlorophyll *a* values is presented in Fig. 5, where it can be seen that our structural modeling approach describes sufficiently the observed phytoplankton patterns and more than 98% of the data were included within the 95% credible intervals.

Generally, the SEM that categorizes the phytoplankton community into cyanobacteria, dinoflagellates and a diatom–chlorophyte–cryptophyte assemblage (PFG A) also provided satisfactory results (see the posterior predictive *p*-values reported in Fig. 6). The RMSE for the five group-specific Chl *a* values were lower in the upper part ($<1\ \mu g\,L^{-1}$) and the down-estuary section D (0.7–1.9 $\mu g\,L^{-1}$). The highest RMSE values were found for cryptophytes and cyanobacteria in section B ($>2\ \mu g\,L^{-1}$) and for cryptophytes (2.1 $\mu g\,L^{-1}$), diatoms (2.2 $\mu g\,L^{-1}$), dinoflagellates (3.6 $\mu g\,L^{-1}$), and cyanobacteria (2.7 $\mu g\,L^{-1}$) in section C of the estuary. [One point worth mentioning is that our reported measures of model performance are inflated by several observed phytoplankton (total and group-specific) peaks in the estuary, but we chose to keep the original information unaltered and thus no outlier exclusion was implemented.] The RMSE values for the two inorganic nitrogen forms varied from 4 to 45 $\mu g\,L^{-1}$, while the respective salinity, attenuation coefficient and flow ranges were 1.35–2.59‰, 0.38–0.44 m^{-1} and 44–65 $m^3\,s^{-1}$. The posterior median paths from nitrogen to the three phytoplankton functional groups of the model (dinoflagellates, cyanobacteria, and functional group A) were weak in the up-estuary section (A), where the strongest path was found between nitrogen and dinoflagellates (-0.18). A relatively strong coupling exists between dinoflagellates (-0.33) and cyanobacteria (-0.31) with nitrogen in the second NRE segment. In addition, the functional group A (from -0.25 to -0.34) and the

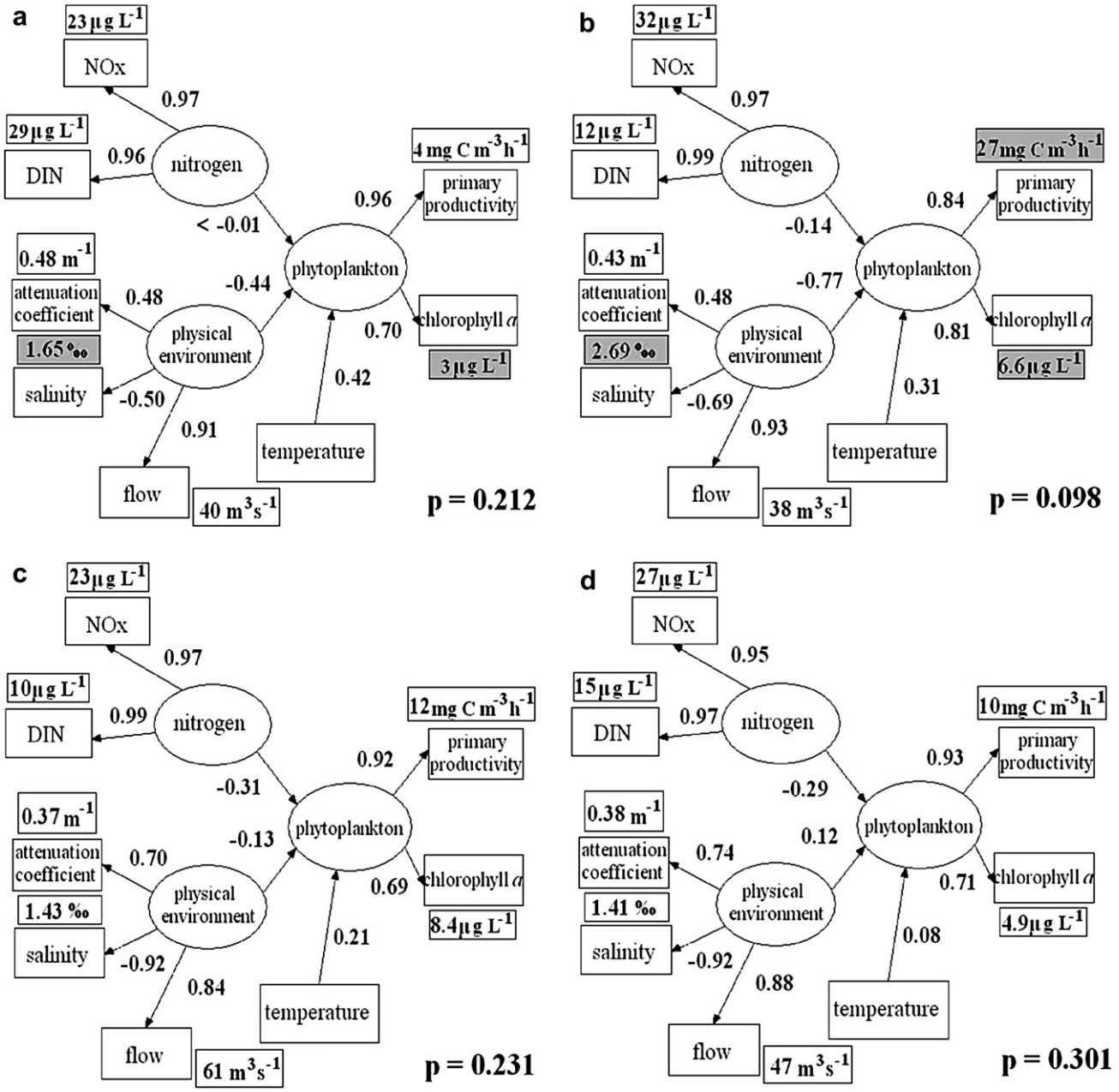

Fig. 4. Aggregate phytoplankton SEM for the four spatial segments of the Neuse River Estuary (see Fig. 1). The numbers correspond to the posterior predictive *p*-values, the posterior medians of the standardized path coefficients and the root mean square error (numbers in rectangles) between the observed values and the medians of the predictive posterior distributions. Shaded rectangular boxes correspond to RMSE values higher than 50% of the observed averages. The standardized coefficients correspond to the shift in standard deviation units of the dependent variable that is induced by shifts of one standard deviation units in the explanatory variables, and thus provide a means to assess the relative importance of the various model paths.

dinoflagellates (from −0.23 to −0.32) have the strongest association with nitrogen in the mid- and down-estuary section. The physical environment is the major regulatory factor of the phytoplankton community dynamics at the upper and middle NRE segments (A and B), where the absolute values of the respective posterior median paths were consistently higher than 0.40. Moreover, the path between the physical environment and the functional group A was switched from negative (−0.54) to positive (0.18) between the second and the third spatial section. This positive physical environment–functional group A relationship is more evident (0.34) in the down-estuary segment (D), and also characterizes the dynamics of dinoflagellates (0.14). As we move closer to the river (segments A and B), flow predominates over the other two indicator variables (attenuation coefficient and salinity) of the physical environment measurement model, and its total standardized effects on phytoplankton vary from −0.32 to −0.49. As was also indicated by the aggregated phytoplankton SEM, salinity has a slightly higher effect on the three PFGs (PFG A, dinoflagellates, cyanobacteria) in the down-estuary NRE segments (C and D). A relatively strong positive cyanobacteria-temperature path exists along the estuary (>0.25),

Table 1
Intersectional comparisons of the aggregate and compositional model performance using the Bayes factor. The likelihood of the models contained in the first column/row forms the numerator/denominator of the Bayes factor

	Section A	Section B	Section C	Section D
Aggregate phytoplankton model				
Section A	1	1.221	0.934	0.788
Section B		1	0.809	0.753
Section C			1	0.852
Section D				1
Compositional phytoplankton model				
Section A	1	1.542	1.656	1.412
Section B		1	1.071	0.911
Section C			1	0.848
Section D				1

while a weak up-estuary positive temperature relationship with the functional group A is switched to a stronger negative one in the down-estuary segments (C and D). The intersectional Bayes factor comparisons indicate that the underlying conceptualization of phytoplankton community dynamics was better supported by the observed patterns in the up-estuary section (A), and becomes weaker as we move to the lower parts of the estuary (Table 1). In fact, the comparison with the alternative phytoplankton community classification (i.e., cyanobacteria, functional groups B and C) indicates that the latter has better foundation in the down-estuary segment D (Table 2). However, it must be noted that none of the models was rejected on the basis of the posterior assessment (p-values), while the Bayes factor did not provide strong evidence in support of one of the two alternative conceptualizations of the NRE phytoplankton community (Kass and Raftery, 1995, p. 777). Finally, the comparison between the observed concentrations and the medians of the predictive chlorophyte, cryptophyte, cyanobacteria, diatom, and dinoflagellate distribution in the four spatial segments of the estuary illustrates the satisfactory description of the spatiotemporal phytoplankton community patterns from our structural equation model (Fig. 7).

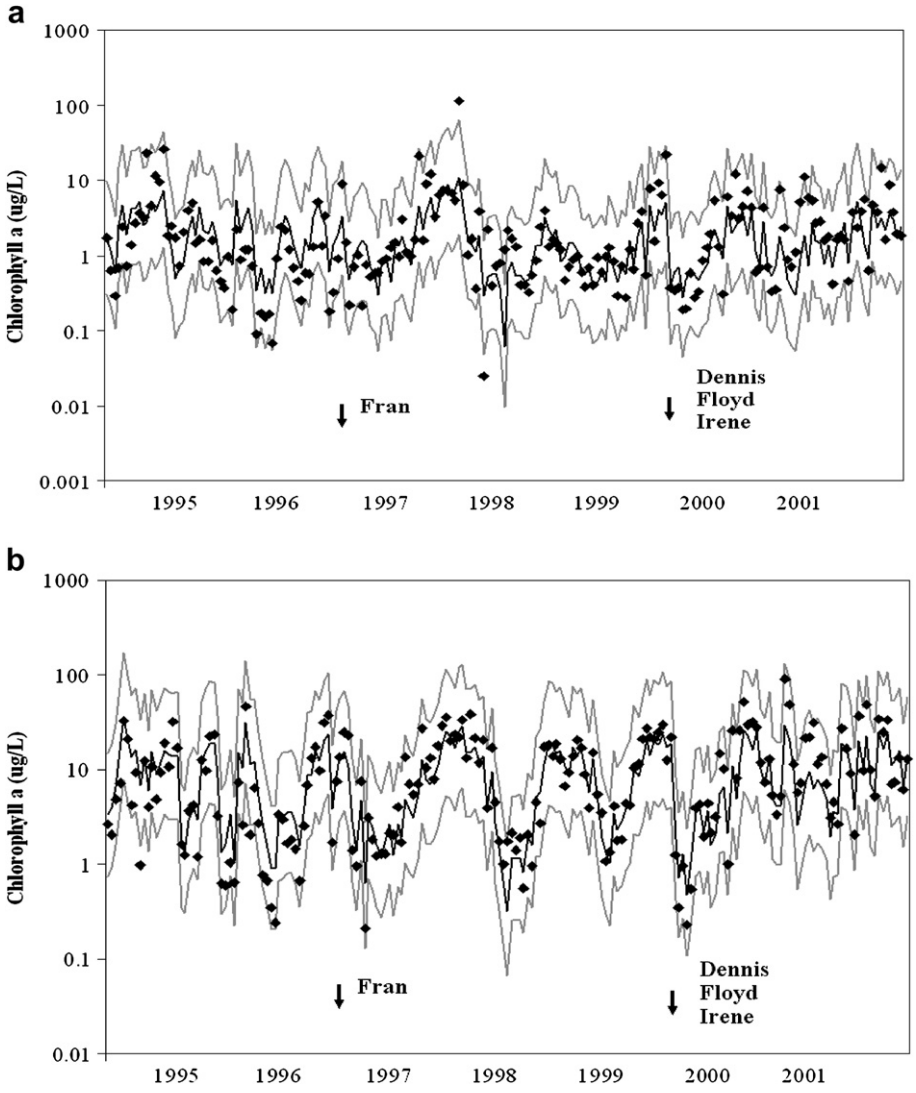

Fig. 5. Comparison of the observed (volume weighted) and mean predicted (along with 95% credible intervals) chlorophyll values in the four segments of the Neuse River Estuary. [Note the timing of the occurrence of four major hydrological events (hurricanes Fran, Dennis, Floyd, and Irene) during the study period.]

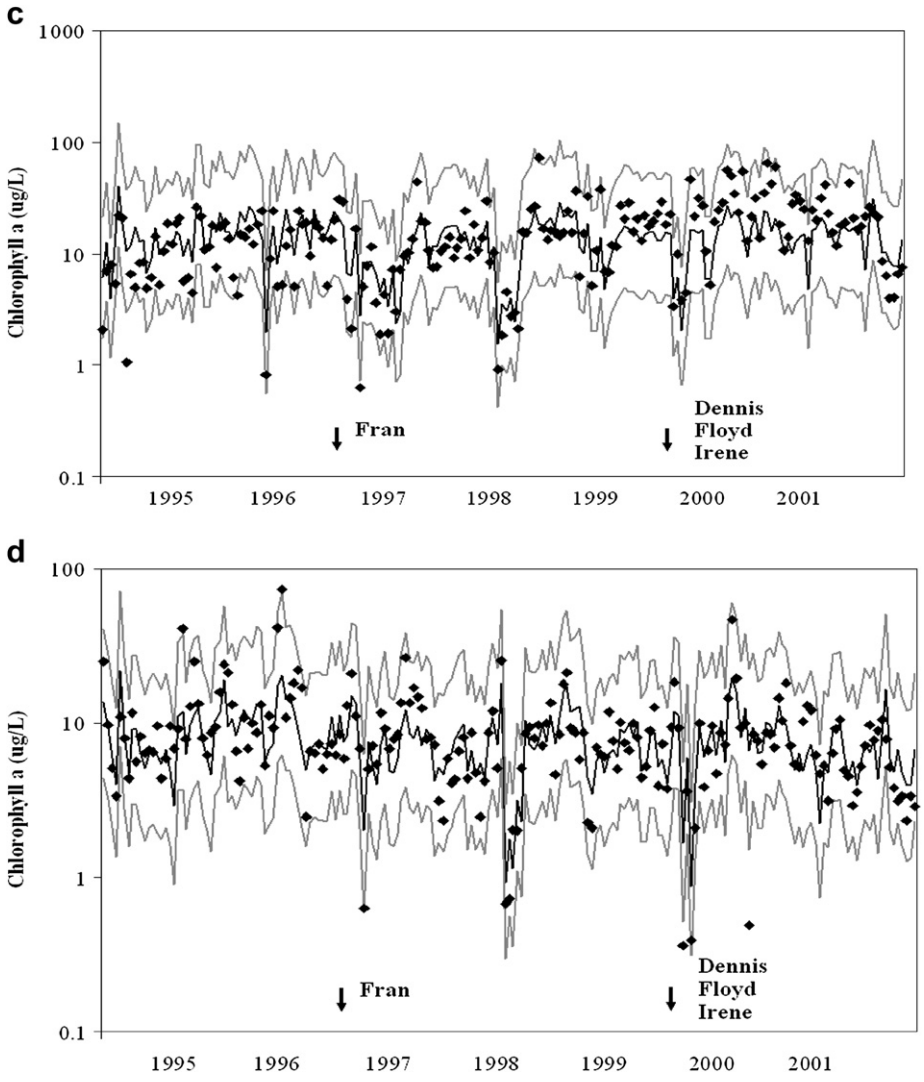

Fig. 5 (*continued*).

4. Discussion

Elucidating the—often contrasting—phytoplankton patterns and understanding the underlying cause–effect relationships require rigorous quantitative tools with the ability to analyze multiple (direct and indirect) causal pathways and consider both compositional and aggregate variability (Cottingham and Carpenter, 1998; Micheli et al., 1999). Our modeling framework can be particularly useful for this purpose. We used a Bayesian structural equation modeling approach that intended to identify recurrent patterns in phytoplankton dynamics, while the optimal phytoplankton grouping was selected on the basis of model performance and was amenable to the complex interplay among several abiotic (latent and observed) variables. Our approach has a ''conditional'' character; the interpretation of the system dynamics and the selection of the optimal phytoplankton grouping are based on the specific assumptions (scale, data aggregation, subset of abiotic causal variables). The robustness of our results is subject to further confirmation by both model updating as new data become available and by considering additional causal factors (e.g., phosphorus, herbivory). The Neuse River Estuary provided an ideal location for testing model efficiency because the phytoplankton spatiotemporal patterns are driven by several physical, chemical, and biological factors and are frequently masked by anthropogenic and climatic perturbations.

Model results underscore the influence of the physical environment on phytoplankton dynamics at the upper reaches of the estuary (section A), while the high standardized loading values of flow on this latent variable give insight into the nature of the underlying mechanism. Flow determines the water residence time, which in turn regulates the phytoplankton growth-minus-physical advection loss balance and controls the accumulation of phytoplankton biomass (Gallegos et al., 1992; Twomey et al., 2002; Borsuk et al., 2004). The other two environmental variables (salinity and light attenuation) associated with river discharge fluctuations seem to have a secondary role on phytoplankton dynamics. The dominance of the physical environment on the up-estuary processes also loosens the coupling between inorganic nitrogen and phytoplankton. This finding was evident in the results of both phytoplankton aggregate and compositional SEMs with the only exception

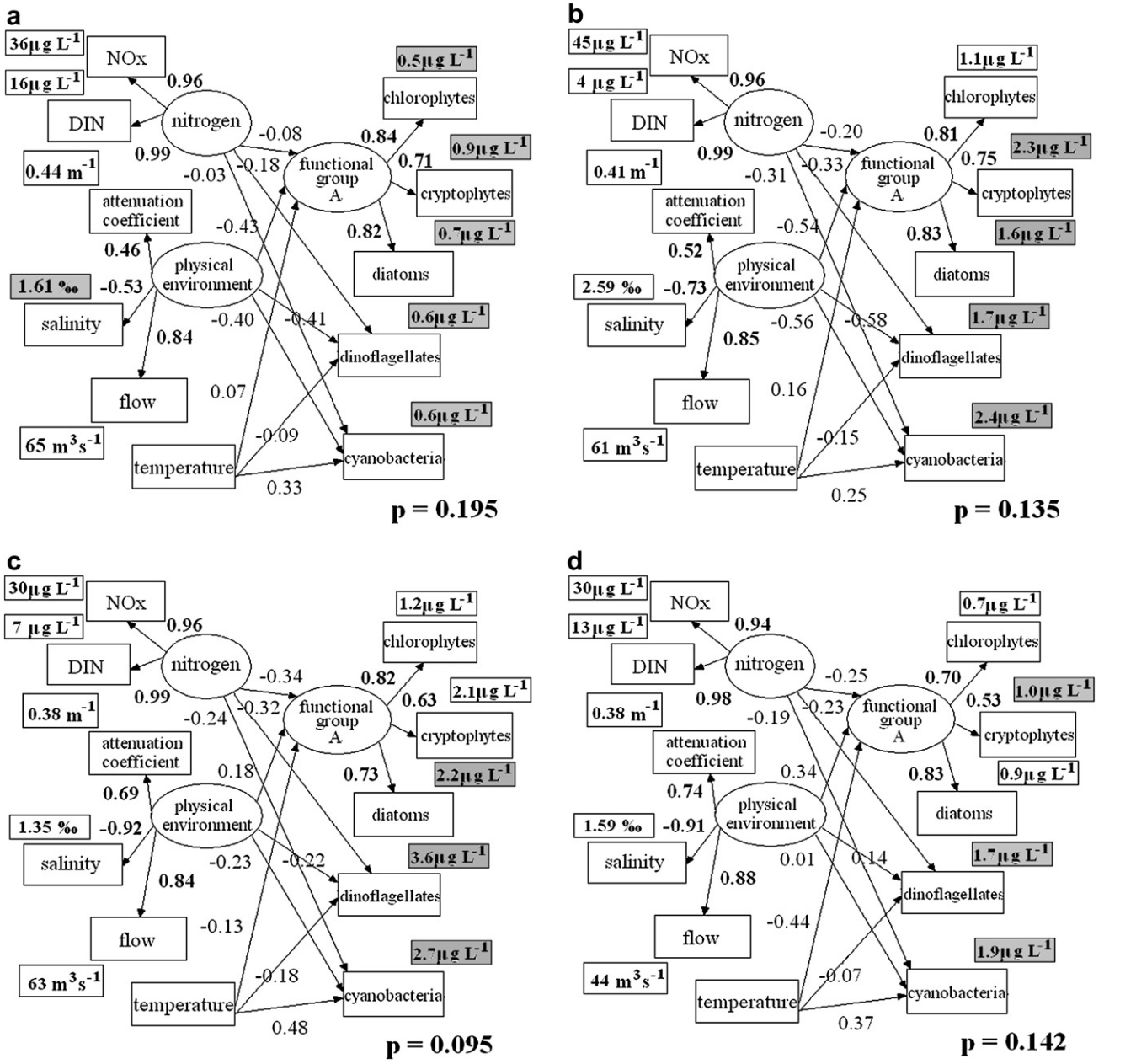

Fig. 6. Compositional phytoplankton SEM for the four spatial segments of the Neuse River Estuary (see Fig. 1). The numbers have the same interpretation as in Fig. 4. For the sake of consistency, the same compositional SEM (PFG A, dinoflagellates, and cyanobacteria) results are presented in the four segments of the estuary, although this categorization was not the most favorably supported by the data in the lower section (D).

being the relationship between nitrogen and dinoflagellates. Interestingly, the dinoflagellates are also the less abundant group in the upper estuary segments, since most of their morphological and physiological features do not permit them to compete successfully with the other major phytoplankton groups. For example, the dinoflagellates have slower growth rates, lower nutrient affinities, while their sensitivity to hydrologic forcing results in significantly reduced concentrations during elevated river flows (Pollingher, 1988; Paerl et al., 2003a, see also discussion following). The competitive handicap can partly be counterbalanced by their tolerance of cold water conditions and their motility; the latter enables vertical migration and screening of the water column for nutrients and optimal light conditions (Pollingher, 1988; Pinckney et al., 1998; Paerl et al., 2006). Thus, the relatively distinct negative nitrogen–dinoflagellate path reflects the almost

Table 2
Comparisons of the two alternative conceptualizations of the Neuse River Estuary phytoplankton community composition using the Bayes factor. The numerator and denominator of the Bayes factor correspond to the likelihood of the models A and B of Fig. 3, respectively

	Section A	Section B	Section C	Section D
Section A	1.227			
Section B		1.035		
Section C			1.056	
Section D				0.927

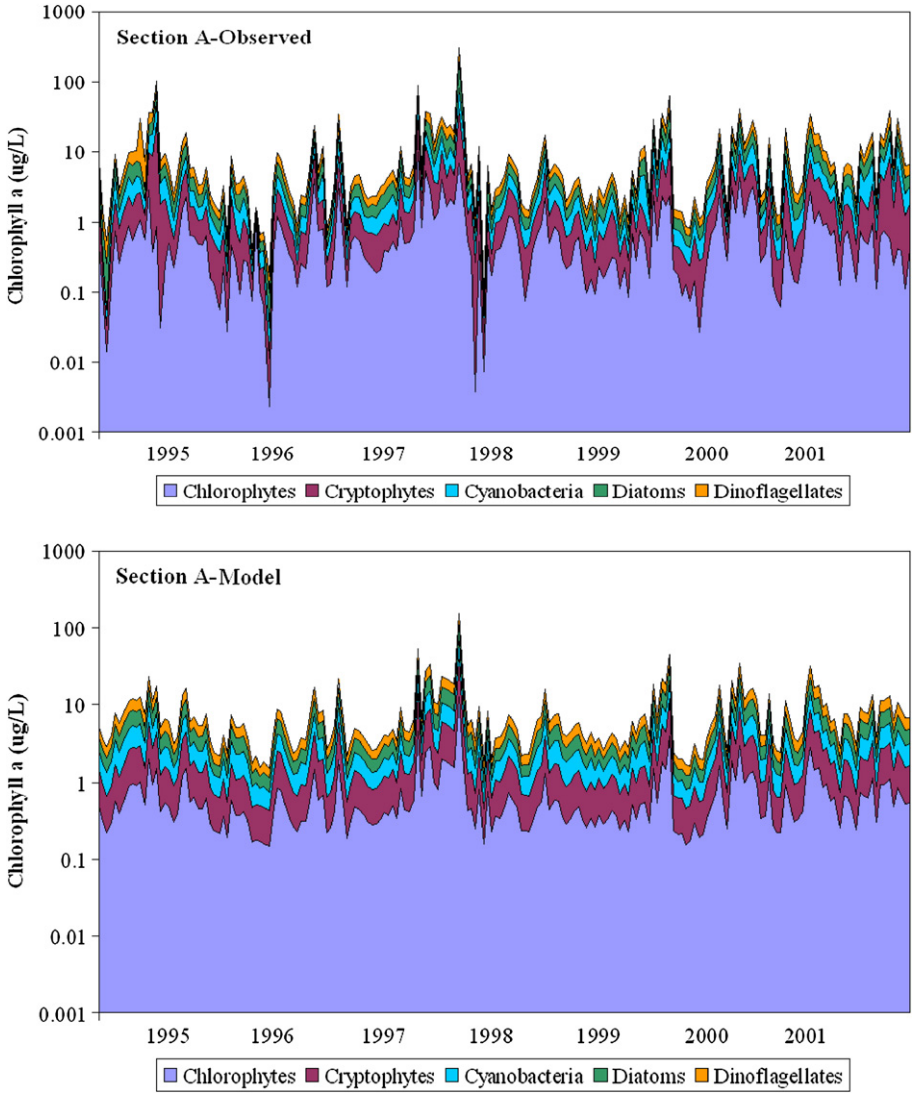

Fig. 7. Observed (volume weighted) and predicted phytoplankton community composition in the four segments of the Neuse River Estuary. Model predictions are based on the medians of the predictive chlorophyte, cryptophyte, cyanobacteria, diatom, and dinoflagellate distributions.

consistent occurrence of the dinoflagellate annual maxima during the fall months, which coincide with periods of longer residence time and the annual upstream DIN minima ($\approx 500\ \mu g\ L^{-1}$). The high standardized loading values of the chlorophytes, cryptophytes and diatoms on the latent functional group A indicate that these taxa are characterized by fairly uniform patterns; they can be treated as one entity without missing much of the information underlying their individual behavior. The latter assertion is further supported by the higher performance of the respective model in the upper part of the estuary. The relatively strong positive path between cyanobacteria and temperature reflects the dominance of this particular group during the summer months.

The physical environment and most importantly the flow is still the major driving force of phytoplankton dynamics, as we move to the middle portion of the estuary (section B). Even though this section is wider than the upper part, the flushing rates are still fast enough to induce significant advective losses. Nonetheless, both the aggregate and compositional models show that the association between nitrogen and phytoplankton dynamics becomes more evident in this section. Interestingly, the path between nitrogen and the functional group A was still weaker than those with dinoflagellates, cyanobacteria and invites further investigation. As is also indicated by the respective temperature paths, both dinoflagellates and cyanobacteria exhibit fairly regular patterns in the Neuse River Estuary, i.e., the dinoflagellate annual maxima are usually observed during the fall months and cyanobacteria dominate the phytoplankton community during the summer period when the dissolved inorganic nitrogen concentrations are relatively lower (Pinckney et al., 1998). The fast-growing functional group A is usually more abundant during the late spring–mid summer months and also shows more opportunistic behavior that allows dominating after episodic hydrologic perturbations characterized by high flow conditions, e.g., hurricanes, tropical storms (Paerl et al., 2006). Thus, the PFG A assemblage has less straightforward relationship with the contemporaneous dissolved inorganic nitrogen concentrations and the weak nitrogen paths

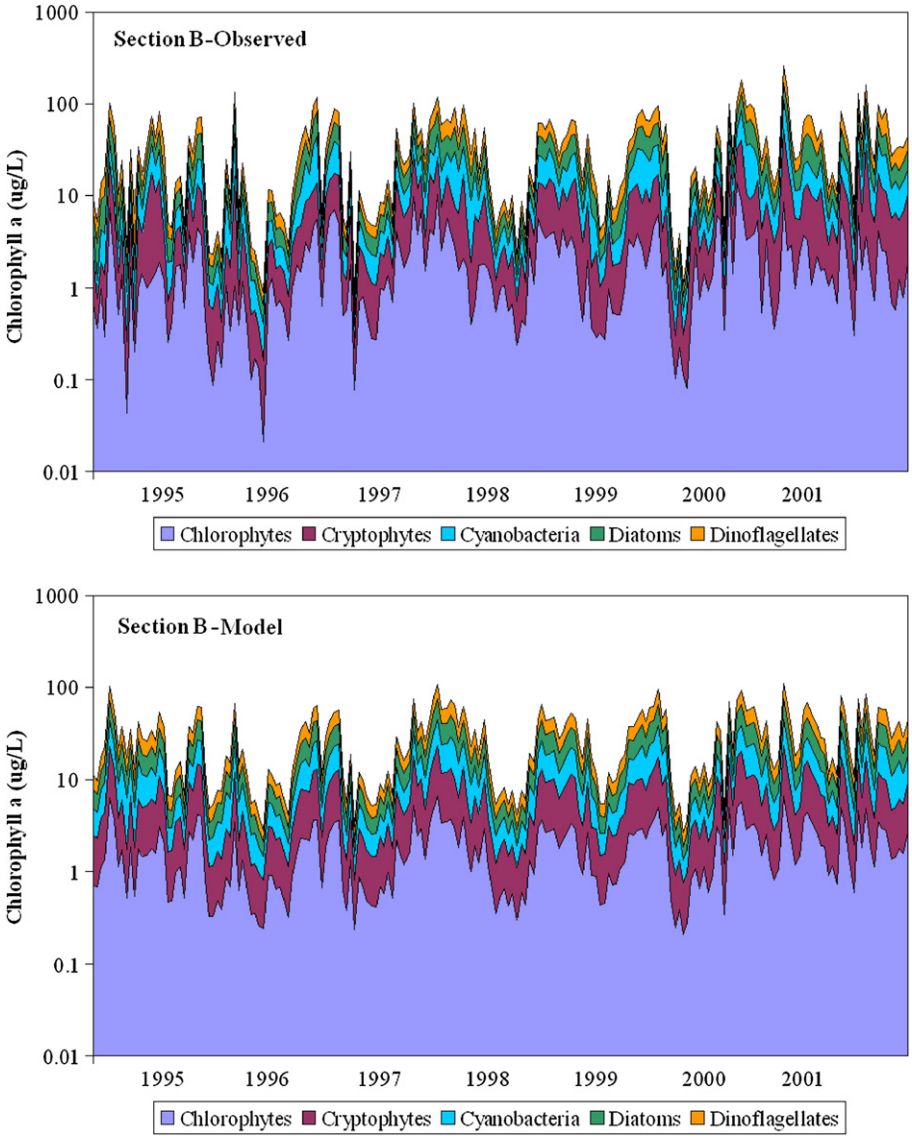

Fig. 7 (*continued*).

suggest that the periods when this functional group is more responsive are also associated with sufficient and relatively constant upstream DIN levels.

The nitrogen–functional group A path is stronger in the mid-lower segment of the estuary (section C), and the negative sign probably reflects their ability to respond to episodic hydrologic events, increase their abundance and consequently decrease the contemporaneous nitrogen levels; especially as we move downstream where nitrogen is in short supply (Pinckney et al., 1998; Paerl et al., 2003a). The physical environment–PFG A path also becomes positive, while the increased standardized loading of salinity on the respective latent variable highlights their ability to demonstrate optimal growth rates under reduced salinity conditions (Pinckney et al., 1999). The switch of the signs of the physical environment–PFG A path between upper and lower segments along with the increased mid-estuary chlorophyll levels (see Fig. 5) are probably related to the conceptualization proposed to explain the spatiotemporal disparity between elevated growth rates and phytoplankton productivity or biomass increase in the Neuse River Estuary (Pinckney et al., 1997). Specifically, sufficient nutrient levels and absence of density-dependent limitations (e.g., self-shading effects) stimulate algal growth in the upper reaches of the estuary, which, however, is usually masked by the increased flow rates. These elevated growth rates are maintained during horizontal transport, and under the most favorable physical conditions (higher residence time, low turbidity) in the mid-lower reaches lead to biomass accumulation and bloom manifestation. On the other hand, the paths from the physical environment to total phytoplankton, cyanobacteria, and dinoflagellates are still negative which provides evidence of a differential response between the overall phytoplankton assemblage and various groups, possibly because this conceptual pattern is more pronounced on the fast-growing taxa (chlorophytes, cryptophytes, and diatoms).

Consistent with the previous conceptualization, both the aggregate and compositional models highlight the positive

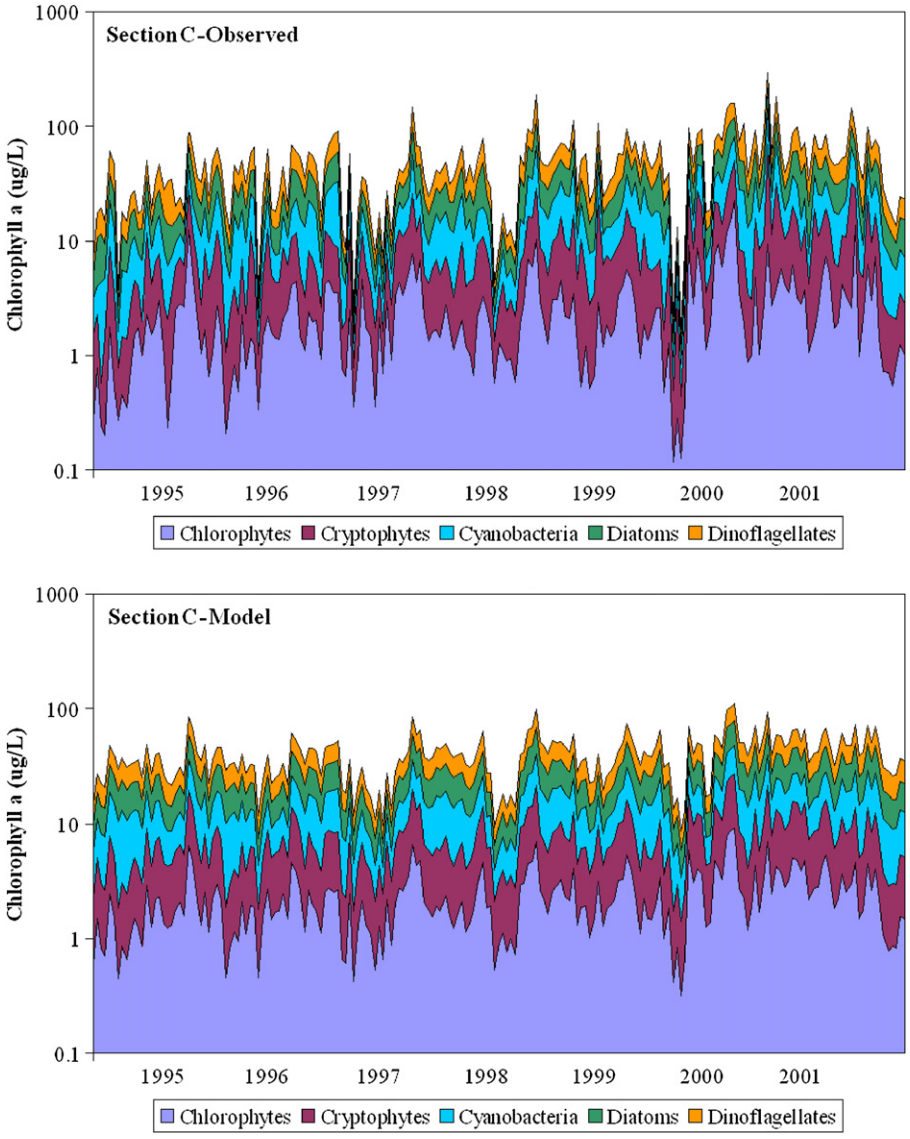

Fig. 7 (*continued*).

relationship between physical environment and phytoplankton community dynamics in the lower NRE section (section D), while the functional group A is the most responsive (see the corresponding standardized path values). Another interesting finding from the down-estuary phytoplankton community SEM was that the categorization of the phytoplankton community into cyanobacteria, dinoflagellates, and an assemblage that consists of diatoms, chlorophytes, and cryptophytes is no longer the most favorably supported by the data. The model that aggregates diatoms and chlorophytes (PFG B), lumps together dinoflagellates with cryptophytes (PFG C), and separately treats the cyanobacteria provides slightly better results (Fig. 3b). In fact, the relatively lower standardized loading of cryptophytes on the latent PFG A was also indicative of a more loose association between the cryptophyte dynamics and the diatoms/chlorophytes in the lower NRE section (Fig. 6d). We hypothesize that these structural shifts on the phytoplankton community temporal patterns may reflect the combined effects of both bottom-up and top-down control. The former effects probably result from the prevailing downstream conditions (warmer water temperatures, longer residence times, and lower DIN concentrations) that favor cyanobacteria dominance, and restrict the occurrence of diatom/chlorophyte blooms to shorter time windows over the annual cycle, i.e., mainly during the winter/early spring period (see also the respective strongly negative temperature paths in Fig. 6d). On the other hand, the signature of the zooplankton grazing is likely to become more apparent under the longer residence times in the lower NRE. For example, an earlier study by Mallin and Paerl (1994b) reported that both diatom and dinoflagellate abundance were correlated with the observed grazing rates in the down-estuary area, while it is now well established in the literature that cryptophytes and diatoms usually have high nutritional value for zooplankton (Brett et al., 2000).

In conclusion, we used a Bayesian structural equation modeling framework to explore the Neuse River Estuary spatio-temporal phytoplankton aggregate and compositional

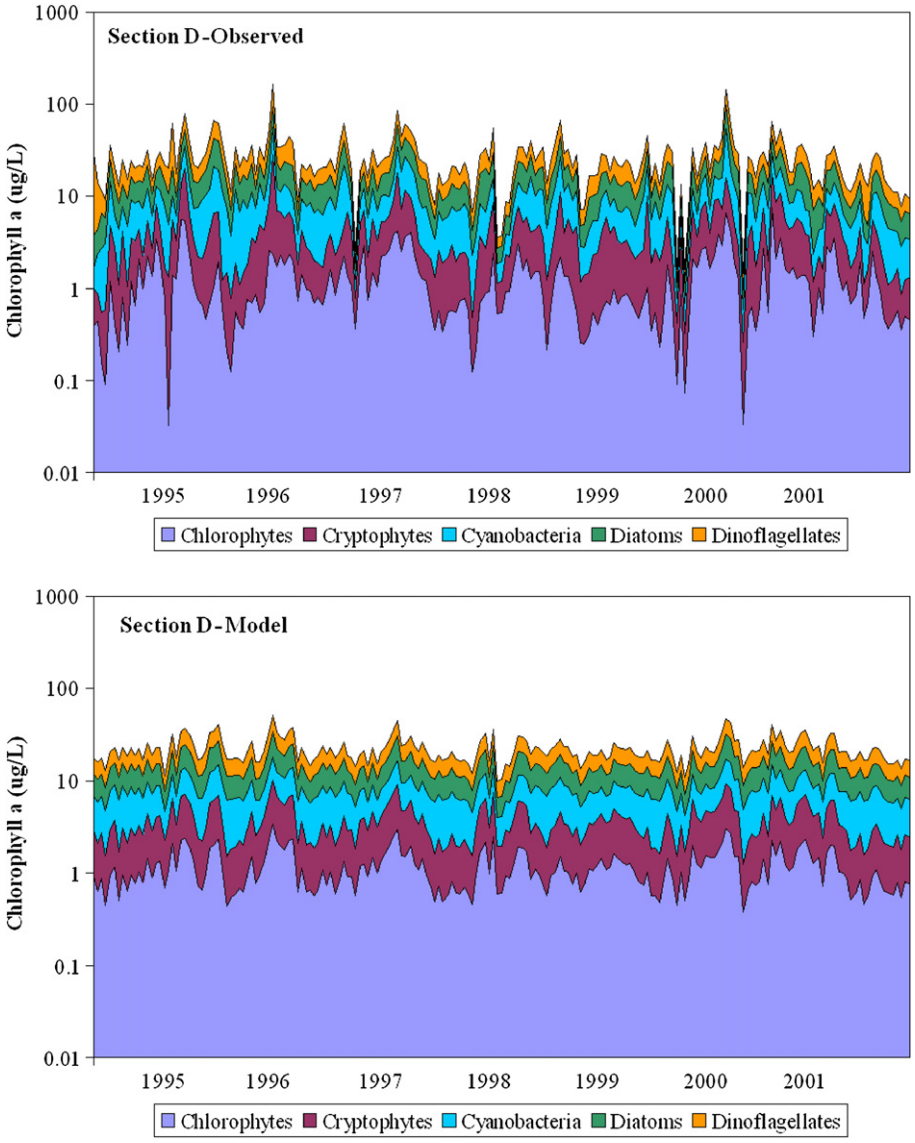

Fig. 7 (*continued*).

patterns. Our model highlighted the role of the physical environment, primarily by the river flow fluctuations and secondarily by the resulting salinity and light availability changes, which dominate the up-estuary processes and loosens the coupling between nitrogen and phytoplankton. The relationship between nitrogen and the phytoplankton community becomes more apparent as we move to the down-estuary sections. Further insights into the phytoplankton community response were provided by the positive paths between the physical environment and diatoms, chlorophytes, cryptophytes in the down-estuary sections. This finding supports a previous hypothesis by Paerl et al. (2003a) that these groups dominate the phytoplankton community during high freshwater conditions as a result of their faster nutrient uptake and growth rates and their tolerance on low salinity conditions. In addition, our modeling study provided evidence that the initiation of the opportunistic phytoplankton taxa blooms takes place in the upper parts of the estuary and due to the horizontal transport and the more favorable physical conditions is fully manifested (biomass accumulation) in the central segments. Finally, an appealing future application of the present methodological framework will be the consideration of smaller functional units (e.g., genera, species) and the evaluation of its ability to separate noise, identify the important cause–effect relationships, and define the optimal aggregation level in community variability studies.

Acknowledgments

This study was funded by grants from the Water Environment Research Foundation through the University of North Carolina Water Resources Research Institute, the U.S. EPA STAR EaGLe Program (R82867701), the Connaught Committee (University of Toronto, Matching Grants 2006–2007), the NC Department of Environment and Natural Resources, Neuse River Estuary Modeling and Monitoring Program, ModMon.

Appendix A

Using the classical SEM notation, we present an illustrative example of the matrices' forms and the specific assumptions made for the aggregate phytoplankton structural equation model. The extraction of the two compositional SEMs can be similarly obtained. The exogenous latent variable measurement model consists of four matrices; i.e., X is a $q \times 1$ vector of observable indicators of the independent latent variables ξ; Λ_X is a $q \times n$ matrix of coefficients relating X to ξ; ξ is a $n \times 1$ vector of independent (exogenous) latent variables; and δ is a $q \times 1$ vector of measurement errors for X. In the present model, we included three ($n = 3$) exogenous latent variables ξ which were described from six ($q = 6$) indicator variables; i.e., NO_x and DIN were used for the latent variable "Nitrogen"; salinity, attenuation coefficient, and flow for the latent variable "Physical Environment"; and the temperature for the respective latent variable. Thus,

$$X = \begin{bmatrix} X_1 = NO_x \\ X_2 = DIN \\ X_3 = \text{Attenuation coefficient} \\ X_4 = \text{Salinity} \\ X_5 = \text{Flow} \\ X_6 = \text{Temperature} \end{bmatrix}, \Lambda_X = \begin{bmatrix} \lambda_3 & 0 & 0 \\ \lambda_4 & 0 & 0 \\ 0 & \lambda_5 & 0 \\ 0 & \lambda_6 & 0 \\ 0 & \lambda_7 & 0 \\ 0 & 0 & \lambda_8 \end{bmatrix},$$

$$\xi = \begin{bmatrix} \xi_1 = \text{Nitrogen} \\ \xi_2 = \text{Physical environment} \\ \xi_3 = \text{Temperature} \end{bmatrix}, \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{bmatrix} \quad \text{(A1)}$$

The endogenous latent variable measurement model also consists of four matrices; i.e., Y is a $p \times 1$ vector of observable indicators of the dependent latent variables η; Λ_Y is a $p \times m$ matrix of coefficients relating Y to η; η is a $m \times 1$ vector of dependent (endogenous) latent variables; ε is a $p \times 1$ vector of measurement errors for Y. Here, two indicator variables ($p = 2$) were used for the representation of one ($m = 1$) endogenous latent variable; i.e., primary productivity and chlorophyll a were used as indicators for the latent variable "Phytoplankton". Thus, the exogenous latent variable measurement model can be described from the four matrices:

$$Y = \begin{bmatrix} Y_1 = \text{Primary productivity} \\ Y_2 = \text{Chlorophyll } a \end{bmatrix}, \Lambda_Y = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix},$$

$$\eta = [\eta_1 = \text{Phytoplankton}], \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad \text{(A2)}$$

The additional two matrices of the structural equation for the latent variable model are

$$\Gamma = [\gamma_1 \quad \gamma_2 \quad \gamma_3], \zeta = [\zeta_1] \quad \text{(A3)}$$

where, Γ is the matrix of the coefficients that relate latent exogenous to endogenous variables; ζ is the vector of latent (structural) errors. [Note that the inclusion of only one endogenous variable in the model implies that the matrix of the coefficients that relate latent endogenous variables—denoted as B in Eq. (A6)—is a zero matrix.] As it can be inferred from the path diagram (i.e., absence of double-headed arrows between the error terms in Fig. 2), the associated covariance matrices of the model; i.e., $\text{Cov}(\xi) = \Phi(n \times n)$: covariances between the independent variables ξ; $\text{Cov}(\varepsilon) = \Theta_\varepsilon(p \times p)$: covariances between the measurement errors in Y; $\text{Cov}(\delta) = \Theta_\delta(q \times q)$: covariances between the measurement errors in X; $\text{Cov}(\zeta) = \Psi(m \times m)$: covariances between the structural errors ζ, have the off-diagonal elements equal to zero:

$$\Theta_\varepsilon = \begin{bmatrix} \text{var}(\varepsilon_1) & \\ 0 & \text{var}(\varepsilon_2) \end{bmatrix},$$

$$\Theta_\delta = \begin{bmatrix} \text{var}(\delta_1) & & & & & \\ 0 & \text{var}(\delta_2) & & & & \\ 0 & 0 & \text{var}(\delta_3) & & & \\ 0 & 0 & 0 & \text{var}(\delta_4) & & \\ 0 & 0 & 0 & 0 & \text{var}(\delta_5) & \\ 0 & 0 & 0 & 0 & 0 & \text{var}(\delta_6) \end{bmatrix},$$

$$\Psi = [\psi_{11}], \Phi = \begin{bmatrix} \phi_{11} & & \\ 0 & \phi_{22} & \\ 0 & 0 & \phi_{33} \end{bmatrix} \quad \text{(A4)}$$

The metric of the latent variables was set by fixing one loading in each column of Λ_X and Λ_Y to 1.0. In this particular case, we assumed that $\lambda_2 = \lambda_4 = \lambda_5 = \lambda_8 = 1.0$. Moreover, implicit in the assumption that the latent variable temperature coincides with the respective observed variable temperature is: $\delta_6 = 0$.

The hierarchical Bayesian configuration of the phytoplankton SEM can be specified as follows:

$$Y_{1i} = \lambda_1 \eta_{1i} + \varepsilon_1, \; Y_{2i} = \lambda_2 \eta_{1i} + \varepsilon_2$$

$$\varepsilon \sim N(0, \Theta_\varepsilon)$$

$$X_{1i} = \lambda_3 \xi_{1i} + \delta_1, \; X_{2i} = \lambda_4 \xi_{1i} + \delta_2$$

$$X_{3i} = \lambda_5 \xi_{2i} + \delta_3, \; X_{4i} = \lambda_6 \xi_{2i} + \delta_4, \; X_{5i} = \lambda_7 \xi_{2i} + \delta_5$$

$$X_{6i} = \lambda_8 \xi_{3i} + \delta_6$$

$$\delta \sim N(0, \Theta_\delta), \; \xi \sim N(0, \Phi)$$

$$\eta_{1i} = \gamma_1 \xi_{1i} + \gamma_2 \xi_{2i} + \gamma_3 \xi_{3i} + \zeta_1$$

$$\zeta_1 \sim N(0, \Psi) \quad \text{(A5)}$$

Let $w_i = \{y_i, x_i, i = 1,\ldots,\nu\}$ be the joint vector of the observed variables (expressed as deviations from the respective means) for an arbitrary observation i. According to the model (A5), each observation i comes from a multivariate normal distribution $f(\mu(\theta)_i, \Sigma(\theta))$ where $\mu(\theta)_i$ is the conditional mean

(expected) vector, $\Sigma(\theta)$ is the conditional covariance matrix, given by (Bollen, 1989):

$$\Sigma(\theta) = \begin{bmatrix} \Lambda_Y (I-B)^{-1}(\Gamma\Phi\Gamma' + \Psi)\left[(I-B)^{-1}\right]' \Lambda_Y' + \Theta_\varepsilon & \Lambda_Y (I-B)^{-1}\Gamma\Phi\Lambda_X' \\ \Lambda_X \Phi\Gamma' \left[(I-B)^{-1}\right]' \Lambda_Y' & \Lambda_X \Phi \Lambda_X' + \Theta_\delta \end{bmatrix} \quad \text{(A6)}$$

and θ is the vector of the unknown model parameters. The likelihood of $w = (w_1,\ldots,w_\nu)$ is

$$p(w|\theta) = \prod_{i=1}^{n} (2\pi)^{-(p+q)/2} |\Sigma(\theta)|^{-1/2} \times \exp\left[-\frac{1}{2}\left[w_i - \mu(\theta)_i\right]' \Sigma(\theta)^{-1}\left[w_i - \mu(\theta)_i\right]\right] \quad \text{(A7)}$$

where $q = 6$ and $p = 2$ are the number of exogenous and endogenous observed variables, respectively. In the context of the Bayesian statistical inference, the focus is on the posterior density of θ given the observed data w, which is defined as:

$$p(\theta|w) = \frac{p(w|\theta)p(\theta)}{\int p(w|\theta)p(\theta)\mathrm{d}\theta} \propto p(w|\theta)p(\theta) \quad \text{(A8)}$$

where $p(\theta)$ is the prior density of θ which is required to be specified for each of the unknown model parameters. Aside from the cases where no measurement error was assumed between the latent and indicator variables (i.e., temperature), we used independent non-informative conjugate gamma priors (0.1, 0.1) for the elements of the matrices Θ_δ^{-1}, Θ_ε^{-1}, Φ^{-1} and Ψ^{-1} (Spiegelhalter et al., 1996; p. 39). Effectively "flat" normal prior distributions with means equal to 0 and precisions (1/variance) equal to 0.0001 were used for the structural parameters (i.e., the parameters that relate latent variables) and the factor loadings ($\lambda_2 = \lambda_3 = \lambda_6 = \lambda_8$ were kept fixed and equal to 1). [Note that a methodology to test the sensitivity of the model results to these assumptions was presented in Arhonditsis et al. (2006)]. MCMC simulation was used as the computation tool implemented in the WinBUGS software (Spiegelhalter et al., 2003). We used three chain runs of 30,000 iterations and samples were taken every 50th iteration to avoid serial correlation. Convergence was assessed using the modified Gelman–Rubin convergence statistic. Generally, the sequences converged rapidly ($\approx$2000 iterations), while the summary statistics reported in this study were based on the last 7500 draws. The accuracy of the posterior estimates was inspected by assuring that the Monte Carlo error (an estimate of the difference between the mean of the sampled values and the true posterior mean) for all the parameters was less than 5% of the sample standard deviation. Finally, all the material (WinBUGS codes and data) pertinent to this study is available upon request from the first author.

References

American Rivers Foundation, 1997. Report on the 20 Most Threatened American River. American Rivers Foundation, New York, NY.

Arhonditsis, G., Karydis, M., Tsirtsis, G., 2003. Analysis of phytoplankton community structure using similarity indices: a new methodology for discriminating among eutrophication levels in coastal marine ecosystems. Environmental Management 31, 619–632.

Arhonditsis, G.B., Brett, M.T., 2005. Eutrophication model for Lake Washington (USA): part II – model calibration and system dynamics analysis. Ecological Modelling 187, 179–200.

Arhonditsis, G.B., Stow, C.A., Steinberg, L.J., Kenney, M.A., Lathrop, R.C., McBride, S.J., Reckhow, K.H., 2006. Exploring ecological patterns with structural equation modeling and Bayesian analysis. Ecological Modelling 192, 385–409.

Bollen, K.A., 1989. Structural Equations with Latent Variables. Wiley and Sons, New York.

Borsuk, M.E., Stow, C.A., Luettich, R.A., Paerl, H.W., Pinckney, J.L., 2001. Modelling oxygen dynamics in an intermittently stratified estuary: estimation of process rates using field data. Estuarine, Coastal and Shelf Science 52, 33–49.

Borsuk, M.E., Stow, C.A., Reckhow, K.H., 2004. Confounding effect of flow on estuarine response to nitrogen loading. Journal of Environmental Engineering, ASCE 130, 605–614.

Brett, M.T., Müller-Navarra, D.C., Park, S.K., 2000. Empirical analysis of mineral P limitation's impact on algal food quality for freshwater zooplankton. Limnology and Oceanography 47, 1564–1575.

Buzzelli, C.P., Powers, S.P., Luettich, R.A., McNinch, J.E., Peterson, C.H., Pinckney, J.L., Paerl, H.W., 2002. Estimating the spatial extent of bottom water hypoxia and benthic fishery habitat degradation in the Neuse River Estuary, NC. Marine Ecology Progress Series 230, 103–112.

Christian, R.R., Boyer, J.N., Stanley, D.W., 1991. Multiyear distribution patterns of nutrients within the Neuse River estuary, North-Carolina. Marine Ecology Progress Series 71, 259–274.

Congdon, P., 2003. Applied Bayesian Modelling. Wiley Series in Probability and Statistics, West Sussex, England.

Cottingham, K.L., Carpenter, S.R., 1998. Population, community, and ecosystem variates as ecological indicators: phytoplankton responses to whole-lake enrichment. Ecological Applications 8, 508–530.

Frost, T.M., Carpenter, S.R., Ives, A.R., Kratz, T.K., 1995. Species compensation and complementarity in ecosystem function. In: Jones, C., Lawton, J. (Eds.), Linking Species and Ecosystems. Chapman and Hall, New York, USA, pp. 224–239.

Gallegos, C.L., Jordan, T.E., Correl, D.L., 1992. Event-scale response of phytoplankton to watershed inputs in a subestuary: timing, magnitude, and location of blooms. Limnology and Oceanography 37, 813–828.

Gelman, A., Meng, X.L., Stern, H., 1996. Posterior predictive assessment of model fitness via realized discrepancies. Statistica Sinica 6, 733–807.

Granéli, E., Wallström, K., Larsson, U., Granéli, W., Elmgren, R., 1990. Nutrient limitation of primary production in the Baltic Sea area. Ambio 19, 142–151.

Huisman, J., Weissing, F.J., 2001. Fundamental unpredictability in multispecies competition. The American Naturalist 157, 488–494.

Jeffrey, S.W., Mantoura, R.F.C., Wright, S.W., 1997. Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods. UNESCO.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of the American Statistical Association 90, 773–795.

Luettich, R.A., McNinch, J.E., Paerl, H., Peterson, C.H., Wells, J.T., Alperin, M., Martens, C.S., Pinckney, J.L., 2000. Neuse River Estuary Modeling and Monitoring Project Stage 1: Hydrography and Circulation, Water Column Nutrients and Productivity, Sedimentary Processes and Benthic–Pelagic Coupling, and Benthic Ecology. Report No. 325B. University of North Carolina, Water Resources Research Institute, Raleigh, NC.

Mackey, M.D., Mackey, D.J., Higgins, H.W., Wright, S.W., 1996. CHEMTAX – a program for estimating class abundances from chemical markers: application to HPLC measurements of phytoplankton. Marine Ecology Progress Series 144, 265–283.

Mackey, M.D., Higgins, H.W., Mackey, D.J., Wright, S.W., 1997. CHEMTAX User's Manual: A Program for Estimating Class Abundances from Chemical Markers – Application to HPLC Measurements of Phytoplankton Pigments. Marine Lab Report No. 229. CSIRO.

Mallin, M.A., Paerl, H.W., 1994a. Commentary on primary productivity and nutrient limitation in the Neuse River Estuary, North Carolina, USA. Marine Ecology Progress Series 111, 311–312.

Mallin, M.A., Paerl, H.W., 1994b. Planktonic trophic transfer in an estuary-seasonal, diel, and community structure effects. Ecology 75, 2168–2184.

McCauley, E., Murdoch, W.W., 1987. Cyclic and stable populations: plankton as a paradigm. The American Naturalist 129, 97–121.

McCormick, P.V., Cairns Jr., J., 1994. Algae as indicators of environmental change. Journal of Applied Phycology 6, 509–526.

McCune, B., Grace, J.B., 2002. Analysis of Ecological Communities. MjM Software Design, Oregon, USA.

Micheli, F., Cottingham, K.L., Bascompte, J., Bjørnstad, O.N., Eckert, G.L., Fischer, J.M., Keitt, T.H., Kendall, B.E., Klug, J.L., Rusak, J.A., 1999. The dual nature of community variability. Oikos 85, 161–169.

Paerl, H.W., Valdes, L.M., Pinckney, J.L., Piehler, M.F., Dyble, J., Moisander, P.H., 2003a. Phytoplankton photopigments as indicators of estuarine and coastal eutrophication. BioScience 53, 953–964.

Paerl, H.W., Dyble, J., Moisander, P.H., Noble, R.T., Piehler, M.F., Pinckney, J.L., Steppe, T.F., Twomey, L., Valdes, L.M., 2003b. Microbial indicators of aquatic ecosystem change: current applications to eutrophication studies. FEMS Microbiology Ecology 46, 233–246.

Paerl, H.W., Valdes, L.M., Joyner, A.R., Piehler, M.F., Lebo, M.E., 2004. Solving problems resulting from solutions: evolution of a dual nutrient management strategy for the eutrophying Neuse River Estuary, North Carolina. Environmental Science and Technology 38, 3068–3073.

Paerl, H.W., Valdes, L.M., Adolf, J., Peierls, B.L., Harding Jr., L.W., 2006. Anthropogenic and climatic influences on the eutrophication of large estuarine ecosystems. Limnology and Oceanography 51, 448–462.

Piehler, M.F., Twomey, L.J., Hall, N.S., Paerl, H.W., 2004. Impacts of inorganic nutrient enrichment on phytoplankton community structure and function in Pamlico Sound, NC, USA. Estuarine, Coastal and Shelf Science 61, 197–209.

Pinckney, J.L., Millie, D.F., Vinyard, B.T., Paerl, H.W., 1997. Environmental controls of phytoplankton bloom dynamics in the Neuse River Estuary, North Carolina, USA. Canadian Journal of Fisheries and Aquatic Sciences 54, 2491–2501.

Pinckney, J.L., Paerl, H.W., Harrington, M.B., Howe, K.E., 1998. Annual cycles of phytoplankton community structure and bloom dynamics in the Neuse River Estuary, North Carolina. Marine Biology 131, 371–381.

Pinckney, J.L., Paerl, H.W., Harrington, M.B., 1999. Responses of the phytoplankton community growth rate to nutrient pulses in variable estuarine environments. Journal of Phycology 35, 1455–1463.

Pollingher, U., 1988. Freshwater armored dinoflagellates: growth, reproduction strategies and population dynamics. In: Sandgren, C.D. (Ed.), Growth and Reproductive Strategies of Freshwater Phytoplankton. Cambridge University Press, pp. 134–174.

Qian, S.S., Borsuk, M.E., Stow, C.A., 2000. Seasonal and long-term nutrient trend decomposition along a spatial gradient in the Neuse River watershed. Environmental Science and Technology 34, 4474–4482.

Reynolds, C.S., Huszar, V., Kruk, C., Naselli-Flores, L., Melo, S., 2002. Towards a functional classification of the freshwater phytoplankton. Journal of Plankton Research 24, 417–428.

Rudek, J., Paerl, H.W., Mallin, M.A., Bates, P.W., 1991. Seasonal and hydrological control of phytoplankton nutrient limitation in the lower Neuse River Estuary, North Carolina. Marine Ecology Progress Series 75, 133–142.

Scheffer, M., Rinaldi, S., Huisman, J., Weissing, F.J., 2003. Why plankton communities have no equilibrium: solutions to the paradox. Hydrobiologia 491, 9–18.

Scheines, R., Hoijtink, H., Boomsma, A., 1999. Bayesian estimation and testing of structural equation models. Psychometrika 64, 37–52.

Schindler, D.W., 1990. Experimental perturbations of whole lakes as tests of hypotheses concerning ecosystem structure and function. Oikos 57, 25–41.

Sommer, U., 1995. An experimental test of the intermediate disturbance hypothesis using cultures of marine phytoplankton. Limnology and Oceanography 40, 1271–1277.

Spiegelhalter, D., Thomas, A., Best, N., Gilks, W., 1996. Bayesian version using Gibbs sampling manual, version ii. Available at <http://www.mrc-bsu.cam.ac.uk/bugs>.

Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2003. WinBUGS User Manual, version 1.4. Available at <http://www.mrc-bsu.cam.ac.uk/bugs>.

Stow, C.A., Borsuk, M.E., 2003a. Assessing TMDL effectiveness using flow-adjusted concentrations: a case study of the Neuse River, North Carolina. Environmental Science and Technology 37, 2043–2050.

Stow, C.A., Borsuk, M.E., 2003b. Enhancing causal assessment of estuarine fishkills using graphical models. Ecosystems 6, 11–19.

Tester, P.A., Geesey, M.E., Guo, C.Z., Paerl, H.W., Millie, D.F., 1995. Evaluating phytoplankton dynamics in the Newport River Estuary (North-Carolina, USA) by HPLC-derived pigment profiles. Marine Ecology Progress Series 124, 237–245.

Twomey, L., John, J., Thompson, P., 2002. Seasonal succession of diatoms and other phytoplankton in a bar built estuary, Wilson Inlet, Western Australia. In: John, J. (Ed.), Proceedings of the 15th International Diatom Symposium, Perth, Australia. Koeltz Scientific Books, Königstein, pp. 395–420.