Delineation of the role of nutrient dynamics and hydrologic forcing on phytoplankton patterns along a freshwater–marine continuum

George B. Arhonditsis[a,*], *Craig A. Stow*[b], *Hans W. Paerl*[c], *Lexia M. Valdes-Weaver*[c], *Laura J. Steinberg*[d], *Kenneth H. Reckhow*[e]

[a] *Department of Physical & Environmental Sciences, University of Toronto, Toronto, Ontario, Canada M1C 1A4*
[b] *NOAA Great Lakes Environmental Research Laboratory, Ann Arbor, MI 48105, USA*
[c] *Institute of Marine Sciences, University of North Carolina at Chapel Hill, Morehead City, NC 28557, 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*

ARTICLE INFO

ABSTRACT

We examined the spatiotemporal phytoplankton community patterns and identified the nature of the underlying causal mechanisms in a freshwater–saltwater continuum, the Neuse River Estuary (North Carolina, USA). We used a Bayesian structural equation modeling (SEM) approach that considers the regulatory role of the physical environment (flow, salinity, and light availability), nitrogen (dissolved oxidized inorganic nitrogen and total dissolved inorganic nitrogen), phosphorus, and temperature on total phytoplankton biomass and phytoplankton community composition. Hydrologic forcing (mainly the river flow fluctuations) dominates the up-estuary processes and loosens the coupling between nutrients and phytoplankton. The switch from an upstream negative to a downstream positive phytoplankton–physical environment relationship suggests that the elevated advective transport from the upper reaches of the estuary leads to a phytoplankton biomass accumulation in the mid- and down-estuary segments. The positive influence of the physical environment on the phytoplankton community response was more evident on diatom, chlorophyte and cryptophyte dynamics, which also highlights the opportunistic behavior of these taxa (faster nutrient uptake and growth rates, tolerance on low salinity conditions) that allows them to dominate the phytoplankton community during high freshwater conditions. Model results highlight the stronger association between phosphorus and total phytoplankton dynamics at the upstream freshwater locations; both nitrogen and phosphorus played a significant role in the middle section of the estuary, while the nitrogen–phytoplankton relationship was stronger in the downstream meso-polyhaline zone. Finally, our analysis provided evidence of a protracted favorable environment (e.g., longer residence times, low DIN concentrations and relaxation of the phosphorus limitation) for cyanobacteria dominance as we move to the down-estuary area, resulting in structural shifts on the phytoplankton community temporal patterns.

* *Corresponding author.* Tel.: +1 416 208 4858; fax: +1 416 287 7279.
E-mail address: georgea@utsc.utoronto.ca (G.B. Arhonditsis).

1. Introduction

In his 2001 review paper, J.E. Cloern highlighted as one of the major advancements of our current conceptual model of coastal eutrophication the explicit recognition that significant variability exists in terms of the inter- and intrasystem responses to nutrient enrichment. Coastal research has provided overwhelming evidence that the operational characteristics of individual ecosystems – the so-called "filter" in Cloern's (2001) conceptual diagrams – determine their carrying capacity and largely regulate their sensitivity to the anthropogenically induced nutrient stressors. For example, physiographical properties (bathymetry, basin geography), circulation patterns (tidal and wind mixing, river flows, density stratification), physico-chemical attributes of the water column (light availability, salinity) and "top–down" processes (herbivory, benthic–pelagic coupling) modulate ecosystem dynamics and, thus, can explain the large differences in their responses to quantitatively similar nutrient loadings (Nixon, 1995; Cloern, 2001; Paerl et al., 2006). A second key advancement in eutrophication research is the realization that nutrient enrichment should not be analyzed as an isolated causative factor, because the manifestation of its signals can be amplified/dampened by complicated interactions with several other anthropogenic or natural perturbations (Boesch et al., 2001; Paerl et al., 2006). Consequently, understanding the individual and synergistic effects of the various components of the ecosystem filtering function and their interactions with the external multiple stressors is one of the major future challenges for the development of effective coastal management schemes (Cloern, 2001; Paerl et al., 2006).

Hydrological forcing, resulting from both intra- and inter-annual climatic variability, strongly drives estuarine and coastal ecosystem dynamics (Goldenberg et al., 2001; Cloern, 2001; Paerl et al., 2006). Intense, short-term episodic events (hurricanes, tropical storms) and seasonal hydrologic variability trigger biogeochemical and trophic responses and change ecosystem functional properties. The phytoplankton community is particularly responsive to these climatic perturbations (Paerl et al., 2006). For example, river flow fluctuations alter several physical (water residence times and stratification patterns) and chemical (salinity variations, nutrient loads and concentrations) waterbody attributes, which, in turn, can directly affect phytoplankton growth rate and spatiotemporal distribution (Harding, 1994; Paerl et al., 2003; Borsuk et al., 2004). The hydrological forcing is also likely to induce structural shifts in phytoplankton community, since phytoplankton taxonomic groups are differentially affected by the resulting physico-chemical changes (Paerl et al., 2006). It is expected that reduced freshwater discharges associated with drought conditions (long water residence time and reduced nutrient concentrations) favor slower growing taxa, such as dinoflagellates and cyanobacteria. Other phytoplankton groups with the ability to exhibit optimal growth rates under reduced salinity conditions, competitively utilize nutrients and grow rapidly, will be favored during periods of increased river flow rates (reduced water residence times and salinity and increased nutrient concentrations) associated with wet years and episodic precipitation events (Paerl et al., 2003).

Thus, considering the large amounts of nutrient and energy flows mediated through the primary producer level, phytoplankton can provide sensitive "warning signs" of the climate system effects on coastal ecosystem integrity (Arhonditsis et al., 2003, 2004; Paerl et al., 2006).

Resolution of spatiotemporal nutrient limitation patterns and the relative macronutrient importance on algal growth along a freshwater–marine continuum is another important factor for our basic understanding of the coastal/estuarine ecosystem structure and functioning (Downing, 1997; Paerl et al., 2004). More importantly, the elucidation of the resource limitation variability can directly aid eutrophication control and assist water quality management (Malone et al., 1996). Using micro- and mesoscale bioassays, numerous studies have concluded that nitrogen and phosphorus limitation are closely controlled by several factors (hydrology, geography and physiography) and can show wide seasonal and spatial variations (Paerl et al., 2006). Generally, phosphorus and nitrogen seem to control freshwater and saltwater primary production, respectively, but these effects vary seasonally and can further be masked by light availability (Rudek et al., 1991; Fisher et al., 1999; Elmgren and Larsson, 2001). Given the intricate nature of resource limitation in estuarine environments, sound scientific foundation is urgently needed to guide the costly implementation of nutrient reduction strategies and optimize their effectiveness (Fisher et al., 1999; Paerl et al., 2004).

We introduce a Bayesian structural equation modeling approach to evaluate the relative importance of the physical environment (flow, salinity, temperature, and light availability) versus nutrient concentration (inorganic nitrogen and phosphate) effects on the spatiotemporal phytoplankton biomass and community composition patterns in the Neuse River Estuary. Specifically, our modeling framework delineates the hydrological forcing vis-à-vis nutrient role on phytoplankton growth and compositional alterations along the estuary. By resolving the particularly problematic relationship between factors that often have confounding/conflicting influence on the waterbody properties (Borsuk et al., 2004), our intention is to provide insight into the complex interactions that drive phytoplankton community dynamics in estuarine ecosystems and influence their responses to nutrient input reductions.

2. Materials and methods

2.1. Study area

The Neuse River drains a 16,008 km^2 watershed, which includes the highly urbanized North Carolina's Research Triangle (defined by the cities of Raleigh, Durham, and Chapel Hill) upstream and the intensive row crop agriculture, silviculture and animal farming industry in the lower portions of the basin (Stow and Borsuk, 2003). The Neuse River discharges into the Neuse Estuary (35°00′N; 76°45′W) and Pamlico Sound (Fig. 1). Water-quality conditions in the Neuse River show significant interannual variability and phytoplankton growth is regulated by a complex interplay between physical, chemical and biological factors (Rudek et al., 1991; Pinckney et al., 1998;
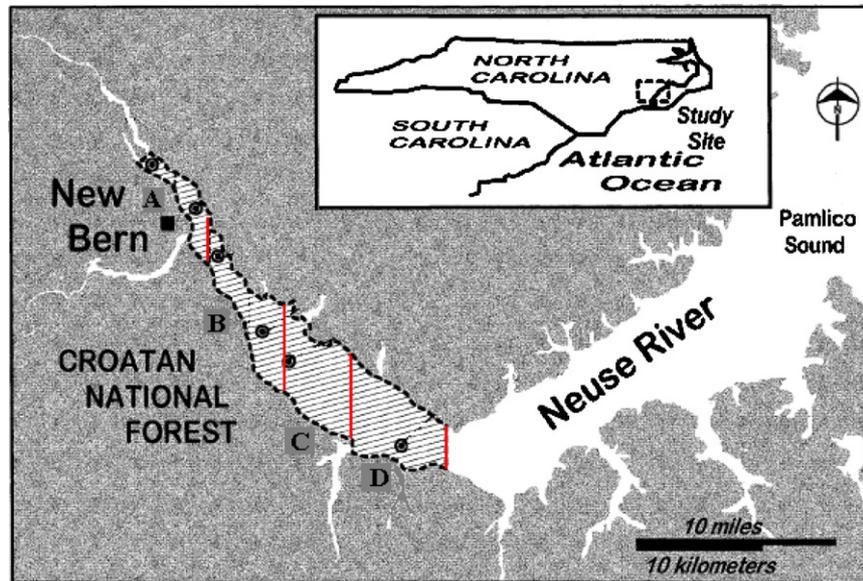
Fig. 1 – The Neuse River Estuary; the vertical lines separate the four segments used for this study.

Borsuk et al., 2004; Paerl et al., 2006), and is further influenced by a recent rise in the hurricane and tropical storm frequency (Paerl et al., 2003, 2006). The Neuse Estuary is an intermittently mixed and shallow (<4 m) system, where salinity varies with precipitation, river discharges, and saltwater inflows from Pamlico Sound (Luettich et al., 2000; Borsuk et al., 2001). The shallow average depth combined with the protection from tides offered by the Outer Banks results in a wind-controlled vertical mixing (McNinch and Luettich, 2000). The estuary has been experiencing characteristic symptoms of nutrient over-load including excessive algal blooms, low levels of dissolved oxygen, and large fish kills (Paerl et al., 2004). The eutroph-ication problems led the Neuse River to be characterized as one of the 20 most threatened rivers in the United States in 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 chloro-phyll a exceeded the state of North Carolina 40 μg/L criterion. These problems are generally attributed to high nitrogen load-ing, though historic and other lines of evidence suggests that phosphorus has also contributed to excessive algal pro-duction (Qian et al., 2000; Buzzelli et al., 2002; Paerl et al., 2004).

2.2.   Data description

Daily mean flow rates were based on the United States Geological Survey streamflow gauging station at Fort Barn-well (Stow and Borsuk, 2003). Nutrient concentrations, salinity vertical attenuation coefficient values and photopig-ments representative of the five dominant algal taxonomic groups (chlorophytes, cryptophytes, cyanobacteria, diatoms and dinoflagellates) in the NRE, were provided from the UNC-CH Institute of Marine Sciences Neuse River Bloom Project, the Neuse River Estuary Modeling and Monitoring Project, MonMod, and the Atlantic Coast Environmental Indi-cators 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 (Luettich et al., 2000; Paerl et al., 2003, 2004). 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 seg-ments were grouped with the second and the fourth spatial compartment, respectively. For each segment, we calculated volume-weighted averages for all the environmental vari-ables of the model based on the corresponding water volumes (m^3) for two depth intervals, i.e., surface to 2 and 2 m to bottom.

2.3.   Bayesian structural modeling

We selected a structural equation modeling approach for three basic reasons: (i) allows the evaluation of a network of relation-ships between observed and latent variables, (ii) enables to explicitly test indirect effects between two explanatory vari-ables, i.e., effects mediated by other intermediary variables, and (iii) explicitly incorporates uncertainty due to observation error or lack of validity of the observed variables, i.e., vari-ables of conceptual interest that are not directly measurable can be represented by multiple indicator (observed) variables (Bollen, 1989; Pugesek et al., 2003). We also adopted a Bayesian approach to SEM that has several advantages over the clas-sical methods (e.g., maximum likelihood, generalized and weighted least squares). A Bayesian SEM can incorporate prior knowledge about the parameters and more effectively treat unidentified models (Scheines et al., 1999; Congdon, 2003). In addition, the modeling process does not rely on asymptotic theory, a feature that is particularly important when the sam-ple 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 conse-quently the procedure works for all sample sizes and various sources of non-normality. The assumptions used to deter-
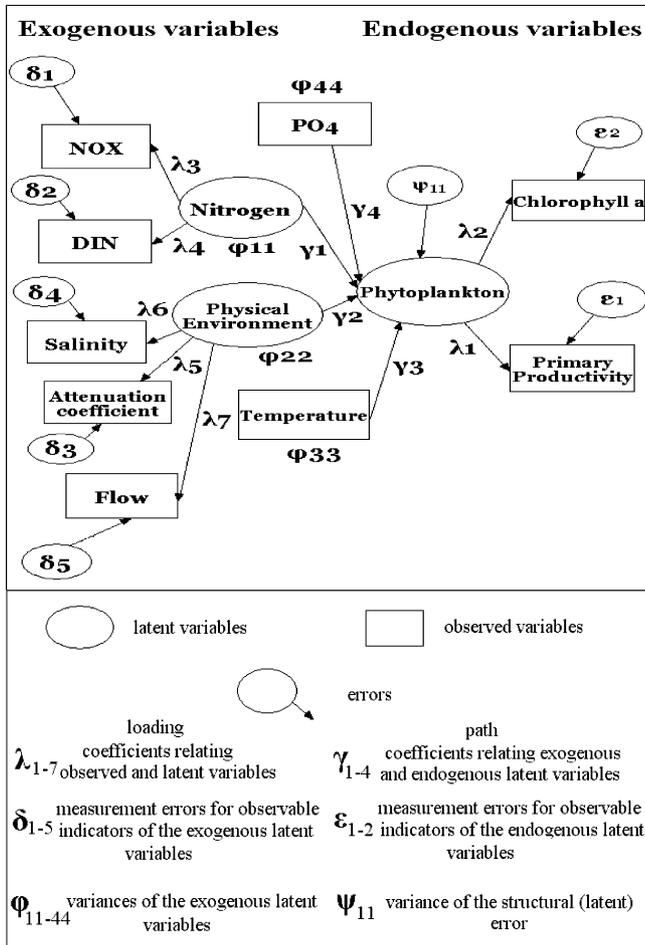
Fig. 2 – The structural equation model used for predicting the Neuse River Estuary phytoplankton dynamics. The use of rectangular boxes for temperature and phosphate implies that the variable was considered as directly observable with no measurement error ($\lambda_8 = \lambda_9 = 1.0$ and $\delta_6 = \delta_7 = 0$). The metrics of the latent variables were set by fixing $\lambda_2 = \lambda_4 = \lambda_5 = 1.0$. Variables that act only as predictors for other variables are referred to as exogenous variables, while those that are dependent are referred to as endogenous. Arrows correspond to the coefficients relating observed and latent variables (loadings) or exogenous and endogenous variables (paths).

phytoplankton dynamics (as a general/abstract idea). Each of these conceptual factors (latent variables) can be linked with observed variables (rectangular boxes in Fig. 2, i.e., "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 included only the oxidized forms of inorganic nitrogen (nitrate + nitrite; NO_x) (Arhonditsis et al., 2007). The phosphorus role was solely described by phosphate (the latent phosphorus coincided with the observed variable phosphate; PO_4). It should be noted that 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. In contrast with Borsuk et al. (2004) study, 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.

To provide a quantitative description of total phytoplankton dynamics (aggregated SEM), we formulated a model using one endogenous latent variable (phytoplankton) combined with two indicator variables, i.e., chlorophyll *a* and primary productivity (Fig. 2). 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, phosphorus, and temperature describe phytoplankton?" Alternatively, we tested four different models to examine the relative importance of the various abiotic factors on the phytoplankton community composition (compositional SEMs). The first model considers the functional group A (PFG A) comprised of diatoms, cryptophytes and chlorophytes, and also retains dinoflagellates and cyanobacteria as two further, distinct groups. The second model aggregates cryptophytes and chlorophytes (PFG B), lumps dinoflagellates with diatoms (PFG C), while cyanobacteria are treated separately. The third model combines chlorophytes with diatoms (PFG D), dinoflagellates with cryptophytes (PFG E) and

mine the latent variable metrics can be treated stochastically and provide additional insight into the ecological structures (Congdon, 2003). Detailed presentation of Bayesian structural equation modeling along with illustrative applications for exploring ecological patterns (e.g., epilimnetic phytoplankton dynamics in different trophic status lakes) can be found in a recent study by Arhonditsis et al. (2006).

The initial framework for model development was based on pre-conceptualizations that reflect our research questions or existing knowledge for the system structure. Our starting point is a "conceptual/mental model" that considers the effects of four latent variables (oval shapes in Fig. 2), i.e., physical environment, nitrogen, phosphorus and temperature, on

Fig. 3 – (a–d) The four alternative conceptualizations of the Neuse River Estuary phytoplankton community dynamics.

again cyanobacteria are considered a third distinct functional group. Finally, the fourth model considers two functional groups, i.e., the functional group A (PFG A) and the functional group F (PFG F) that aggregates dinoflagellates with cyanobacteria (Fig. 3). Thus, based on existing literature information or observed correlations from the study system, we constructed measurement models that allow several phytoplankton groups to form single entities, i.e., latent variables that characterize phytoplankton functional groups. We tested the compatibility of the 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.

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 (Gelman et al., 1996). The compatibility of the alterna-tive phytoplankton community pre-conceptualizations with the observed ecological patterns ("Which of the four phyto-plankton groupings is better supported in each segment?"), and the model performance with or without phosphorus (N + P versus N models) was evaluated by the use of the Bayes factor (Kass and Raftery, 1995). Linearity among the observed indi-cators of the exogenous and the endogenous latent variables was obtained by square root and natural logarithm transfor-mations. Based on calculated residence times (Christian et al., 1991) and exploratory data analyses, mean daily flow rates were calculated for the 2-day, 1-week, 2-week, and 25-day period preceding the sampling dates in the four spatial sec-tions, 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. [See also the Appendix A for further methodological details pertinent to Bayesian structural equation modeling.]

Table 1 – Average chlorophyll *a* and phytoplankton taxa values in the four segments of the study				
	Section A	Section B	Section C	Section D
Chlorophyll *a*	3.29	11.74	16.71	9.32
Chlorophytes	0.80 (35%)	2.14 (23%)	2.73 (19%)	1.31 (16%)
Cryptophytes	1.03 (32%)	3.47 (32%)	3.69 (23%)	1.66 (18%)
Cyanobacteria	0.35 (12%)	2.52 (18%)	4.10 (26%)	2.74 (32%)
Diatoms	0.86 (18%)	2.52 (20%)	3.49 (21%)	1.95 (20%)
Dinoflagellates	0.25 (3%)	1.39 (7%)	2.36 (11%)	1.76 (14%)

Numbers in parenthesis indicate the average percentage contribution of the five phytoplankton taxa.

3. Results

Average chlorophyll *a* levels were lower in the up-estuary segment ($\approx 3.29\,\mu g/L$), where chlorophytes and cryptophytes were the predominant groups (>30%) followed by diatoms (18%) and cyanobacteria (12%) (Table 1). The mid-estuary segments B and C exhibited higher average chlorophyll *a* concentrations (11.74 and $16.71\,\mu g/L$), while diatoms, chlorophytes and cryptophytes consistently comprised 65–70% of the total phytoplankton community biomass. Both cyanobacteria and dinoflagellate concentrations and average percentage contribution show increasing trends moving downstream, while the former group dominates the segment D with an average concentration of $2.74\,\mu g/L$ ($\approx 32\%$). Upstream DIN concentrations show a fairly steady annual cycle with an approximate median value of $600\,\mu g/L$ (Fig. 4). The DIN annual median levels were considerably lower in the other three segments (≈ 379, 104 and $46\,\mu g/L$), where the annual minima usually

occur during the summer months (July–August). In particular, down-estuary DIN concentrations were almost consistently lower than $50\,\mu g/L$ from April to October. In contrast, phosphate concentrations decrease from the up-estuary ($42\,\mu g/L$) to the down-estuary section ($18\,\mu g/L$) was less pronounced, and the annual patterns were characterized by distinct late summer–early fall annual maxima (Fig. 5). The DIN:DIP molar ratios were higher in the upstream area (annual median values in sections A and B >24) and dropped below the Redfield ratio (16) after the middle part of the estuary (annual median values C and D <10). In addition, the DIN:DIP patterns were fairly similar along the estuary with wide seasonal variation and annual minimum values usually observed by the end of summer (Fig. 6).

The Bayes factor values for both the total phytoplankton and phytoplankton community composition models indicate that phosphorus inclusion results in a higher model performance (Table 2). We note, however, that the reported Bayes factor values are lying in the range (1–3) where the support of the alternative hypothesis (N + P models) is "not worth more than a bare mention" (Kass and Raftery, 1995, page 777). The root mean square error (RMSE) values for the modeled chlorophyll *a* and primary productivity in the four segments ranged from 2.5–$7.6\,\mu g/L$ to 4–$23\,mg\,C\,m^{-3}\,h^{-1}$ (Fig. 7). In addition, the RMSE values for the two inorganic nitrogen forms in the total phytoplankton biomass structural equation model varied from 11 to $30\,\mu g/L$. The salinity, attenuation coefficient and flow RMSE value ranges were 1.32–2.65‰, 0.38–$0.47\,m^{-1}$ and 38–$57\,m^3/s$, respectively. The direct standardized path from nitrogen to phytoplankton is weak in the upper part of the estuary (-0.02) and becomes stronger moving to the down-estuary, i.e., -0.08, -0.25 and -0.26 in the segments B, C and
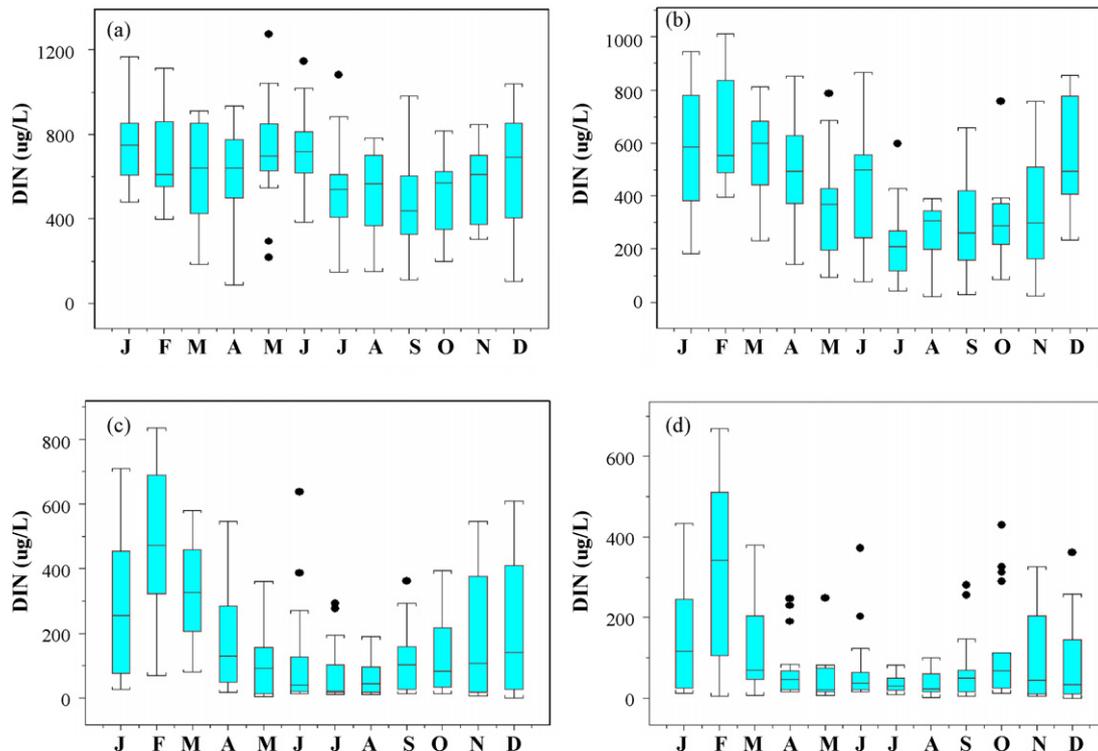
Fig. 4 – (a–d) Seasonal variability of the dissolved inorganic nitrogen concentrations ($\mu g/L$) in the four segments of the Neuse River Estuary.
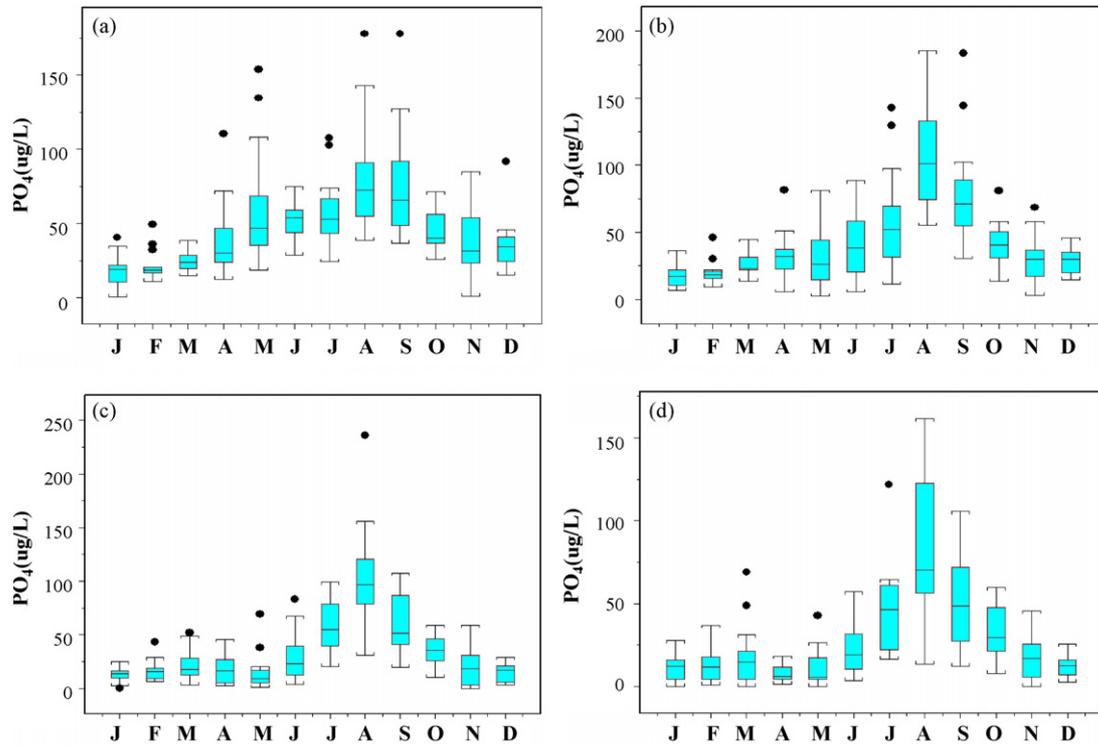
Fig. 5 – (a–d) Seasonal variability of the phosphate concentrations (μg/L) in the four segments of the Neuse River Estuary.
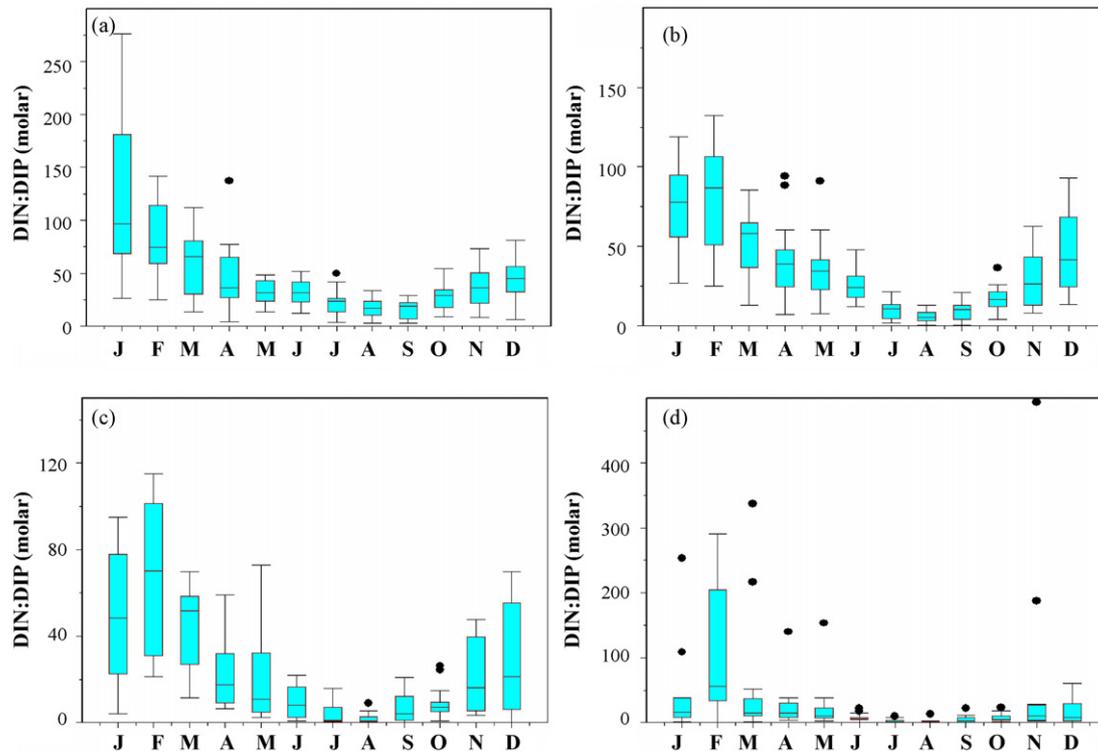
Fig. 6 – (a–d) Seasonal variability of the DIN/DIP (molar) ratio in the four segments of the Neuse River Estuary. The respective median annual values were 31, 24, 10, and 7.

(a)

27 µg/l — NOX — 0.98 — Nitrogen
28 µg/l — DIN — 0.97
PO4 — −0.17
−0.02
Nitrogen → Phytoplankton
2.5 µg/l — Chlorophyll a
0.73
1.65 ‰ — Salinity — −0.50 — Physical Environment — −0.43
0.47
0.91
0.96 — Primary Productivity — 4 mg C m⁻³h⁻¹
0.49
Attenuation coefficient — 0.47 m⁻¹
Temperature
Flow — 38 m³/sec
p=0.235

(b)

30 µg/l — NOX — 0.97 — Nitrogen
15 µg/l — DIN — 0.99
PO4 — −0.28
−0.08
Nitrogen → Phytoplankton
5.8 µg/l — Chlorophyll a
0.86
2.65 ‰ — Salinity — −0.71 — Physical Environment — −0.69
0.49
0.93
0.83 — Primary Productivity — 23 mg C m⁻³h⁻¹
0.45
Attenuation coefficient — 0.43 m⁻¹
Temperature
Flow — 38 m³/sec
p=0.112

(c)

25 µg/l — NOX — 0.97 — Nitrogen
11 µg/l — DIN — 0.99
PO4 — −0.16
−0.25
Nitrogen → Phytoplankton
7.6 µg/l — Chlorophyll a
0.71
1.32 ‰ — Salinity — −0.92 — Physical Environment — −0.13
0.69
0.84
0.91 — Primary Productivity — 14 mg C m⁻³h⁻¹
0.30
Attenuation coefficient — 0.38 m⁻¹
Temperature
Flow — 57 m³/sec
p=0.258

(d)

27 µg/l — NOX — 0.95 — Nitrogen
13 µg/l — DIN — 0.97
PO4 — −0.09
−0.26
Nitrogen → Phytoplankton
5.1 µg/l — Chlorophyll a
0.73
1.40 ‰ — Salinity — −0.92 — Physical Environment — 0.11
0.74
0.87
0.92 — Primary Productivity — 11 mg C m⁻³h⁻¹
0.13
Attenuation coefficient — 0.38 m⁻¹
Temperature
Flow — 48 m³/sec
p=0.317

Fig. 7 – (a–d) Aggregate phytoplankton SEM for the four spatial segments of the Neuse River Estuary. The numbers correspond to 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. 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.

D, respectively. In contrast, the direct path from phosphorus to phytoplankton was stronger in the upper and middle estuary segments (−0.17, −0.28 and −0.16 in the segments A–C), and weaker (−0.09) in the down-estuary section D. Temperature effects on phytoplankton become weaker as we move to the down-estuary sections (i.e., from 0.49 to 0.13). The physical environment plays a significant role on phytoplankton dynamics in the up-estuary segments A and B. In particular, the flow rates have the stronger total standardized effects −0.39 (= −0.43 × 0.91) and −0.64 (= −0.69 × 0.93) among the three-indicator variables used for the physical environment measurement model. The path from the physical environment to phytoplankton was weaker in the segment C (−0.13), and the posterior median effects were switched from negative to positive (0.11) in the spatial section D. Salinity has a higher (absolute) loading on the respective latent variable (physical environment) in the middle and down-estuary segments C

and D. Finally, the comparison between the observed and predicted chlorophyll a/primary productivity values is presented in Fig. 8, where it can be seen that our structural modeling approach describes sufficiently the observed phytoplankton patterns and more than 97% of the data were included within the 95% credible intervals (see also the posterior predictive p-values reported in Fig. 7).

Generally, the structural equation model that considers a classification of the phytoplankton community into the PFG A assemblage (diatoms, chlorophytes and cryptophytes), cyanobacteria, and dinoflagellates provided satisfactory results (see the posterior predictive p-values in Fig. 9). The RMSE for the five group-specific chlorophyll a values were lower in the upper part (<1 µg/L) of the estuary. The highest RMSE values were found for cryptophytes and cyanobacteria in the section B (2.2 µg/L), for dinoflagellates (3.2 µg/L), cyanobacteria (2.5 µg/L) in the section C of
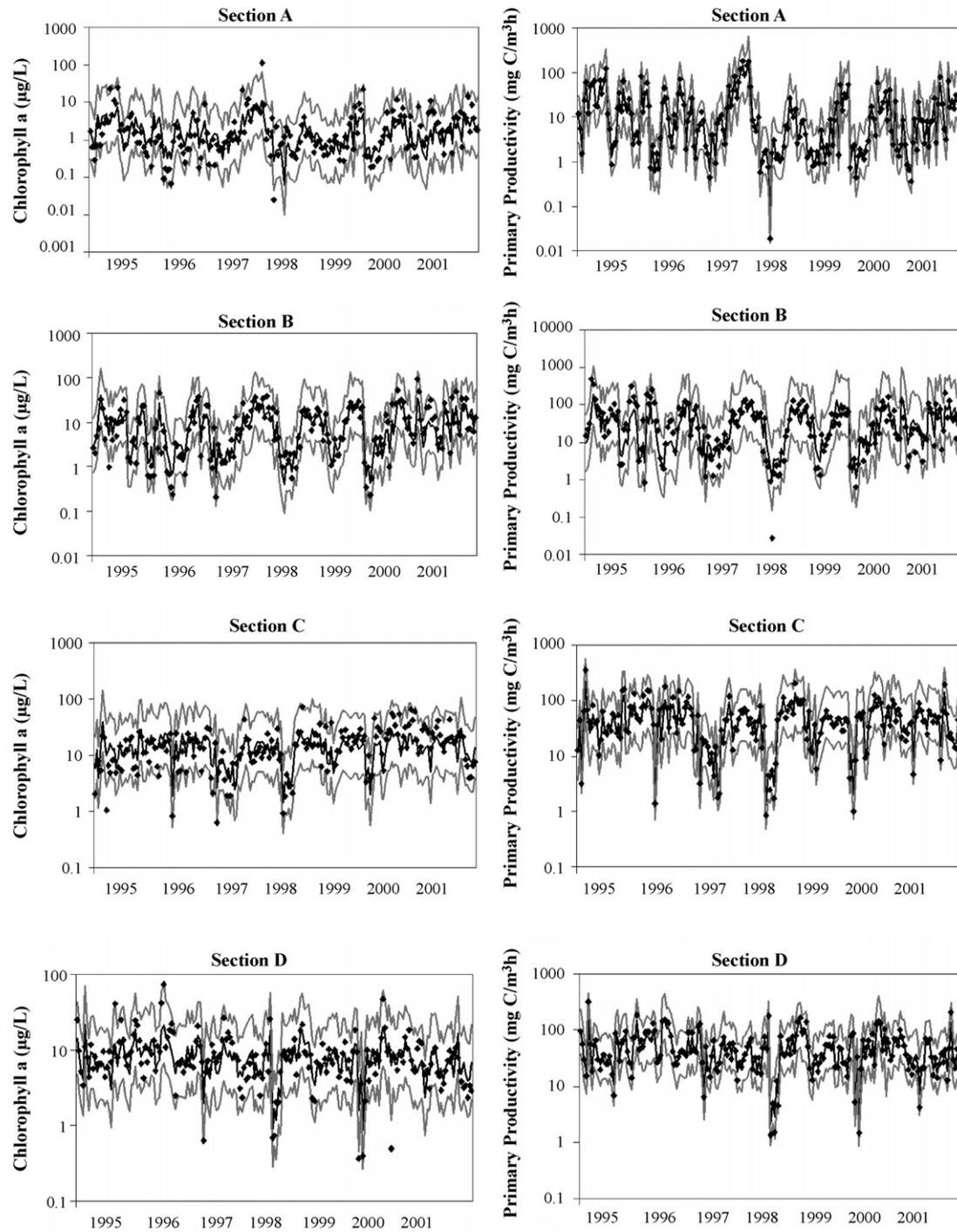
Fig. 8 – Comparison of the observed (volume weighted) and mean predicted (along with 95% credible intervals) chlorophyll and primary productivity values in the four segments of the Neuse River Estuary.

the estuary, and for dinoflagellates (1.8 µg/L), cyanobacteria (1.9 µg/L) in the down-estuary section D. [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 were varying from 5 to 38 µg/L,

while the respective salinity, attenuation coefficient and flow ranges were 1.42–2.52‰, 0.39–0.43 m^{-1} and 43–63 m^3/s. Among the three phytoplankton functional groups of the model (dinoflagellates, cyanobacteria and PFG A), dinoflagellates had the relatively stronger posterior median path with nitrogen (−0.17) in the up-estuary section (A), whereas the respective cyanobacteria and PFG A–nitrogen paths were fairly weak (< −0.09). The coupling between dinoflagellates
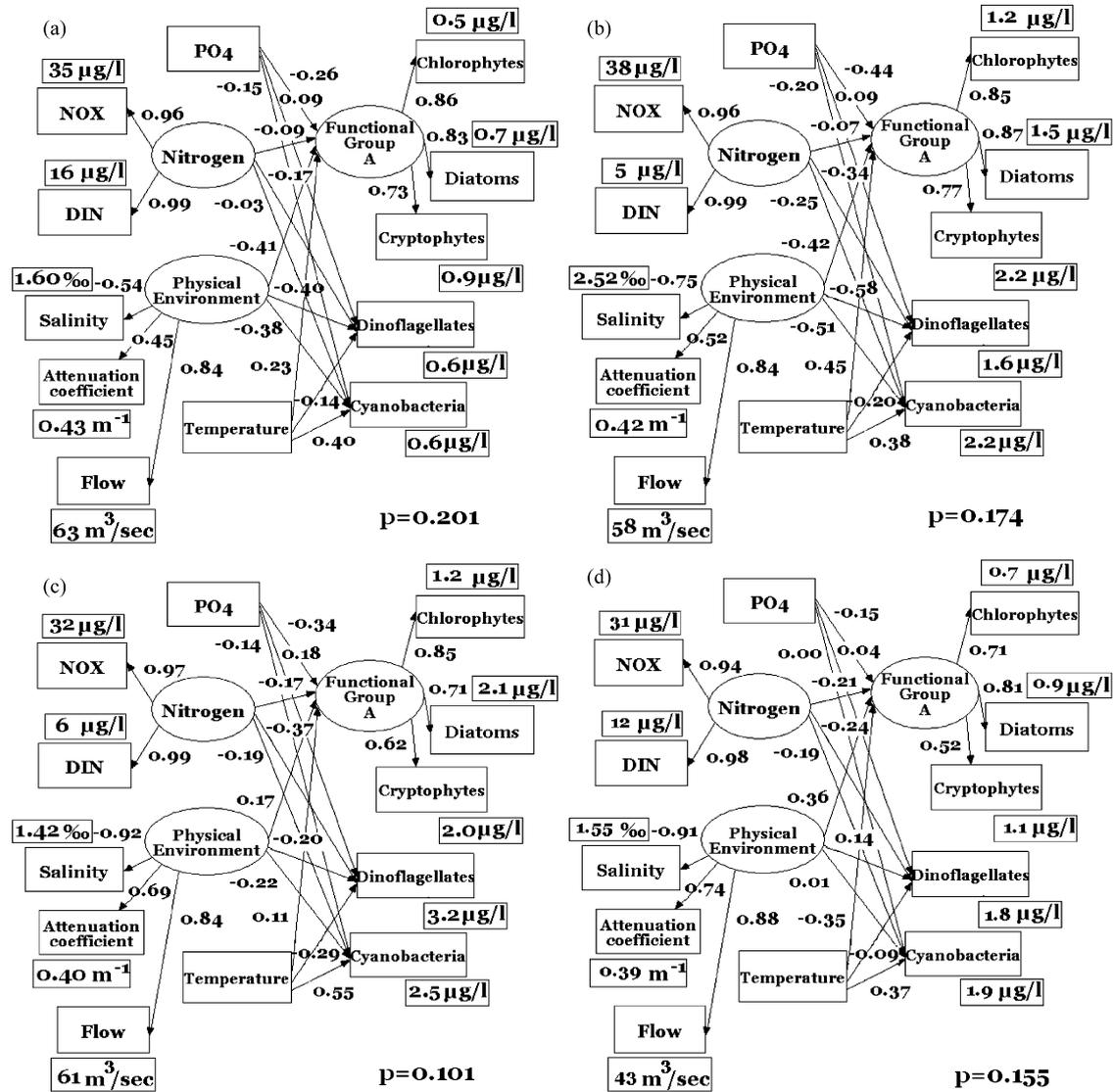
Fig. 9 – (a–d) Compositional phytoplankton SEM for the four spatial segments of the Neuse River Estuary. The numbers have the same interpretation as in Fig. 7. For the sake of consistency, the same compositional SEM (PFG A, dinoflagellates, and cyanobacteria) results are presented in the four segments of the estuary, although this categorization was not the most favorably supported by the data in the lower section (D).

(−0.34) and cyanobacteria (−0.25) with nitrogen was strong in the second NRE segment. In addition, the dinoflagellates had the strongest association with nitrogen in the mid- and down-estuary sections C and D (−0.37 and −0.24), where the respective PFG A and cyanobacteria path values were relatively similar and constant ($\approx$ −0.20). The posterior median standardized phosphate–PFG A paths were strongly negative in the upper and middle sections of the estuary (< −0.26), and relatively weaker in the downstream section D (−0.15). Likewise, the up- and mid-estuary cyanobacteria–phosphate paths were negative (<−0.14), whereas their coupling with phosphorus is minimized as we move downstream. Interestingly, the dinoflagellates seem to have a positive – but relatively weak – relationship with phosphate along the estuary (0.04–0.18). The physical environment is the major regulatory factor of the phytoplankton community dynamics in the upper and

middle NRE segments (A and B), where the absolute values of the respective posterior median paths were usually higher than 0.40. The path between the physical environment and the functional group A was switched from negative (−0.42) to positive (0.17) between the second and the third spatial section, while the positive relationship was further manifested (0.36) in the down-estuary segment (D). Similarly, the dinoflagellate dynamics are also characterized by a positive – but weaker – relationship with the physical environment (0.14) in the fourth spatial section. Flow rates are the predominant indicator variable of the physical environment measurement model with total standardized effects on phytoplankton that varied from −0.32 to −0.49. On the other hand, the (absolute) salinity loading values are slightly higher (>0.91) in the down-estuary NRE segments C and D. The functional group A has a relatively strong positive relationship with temperature in the up- and
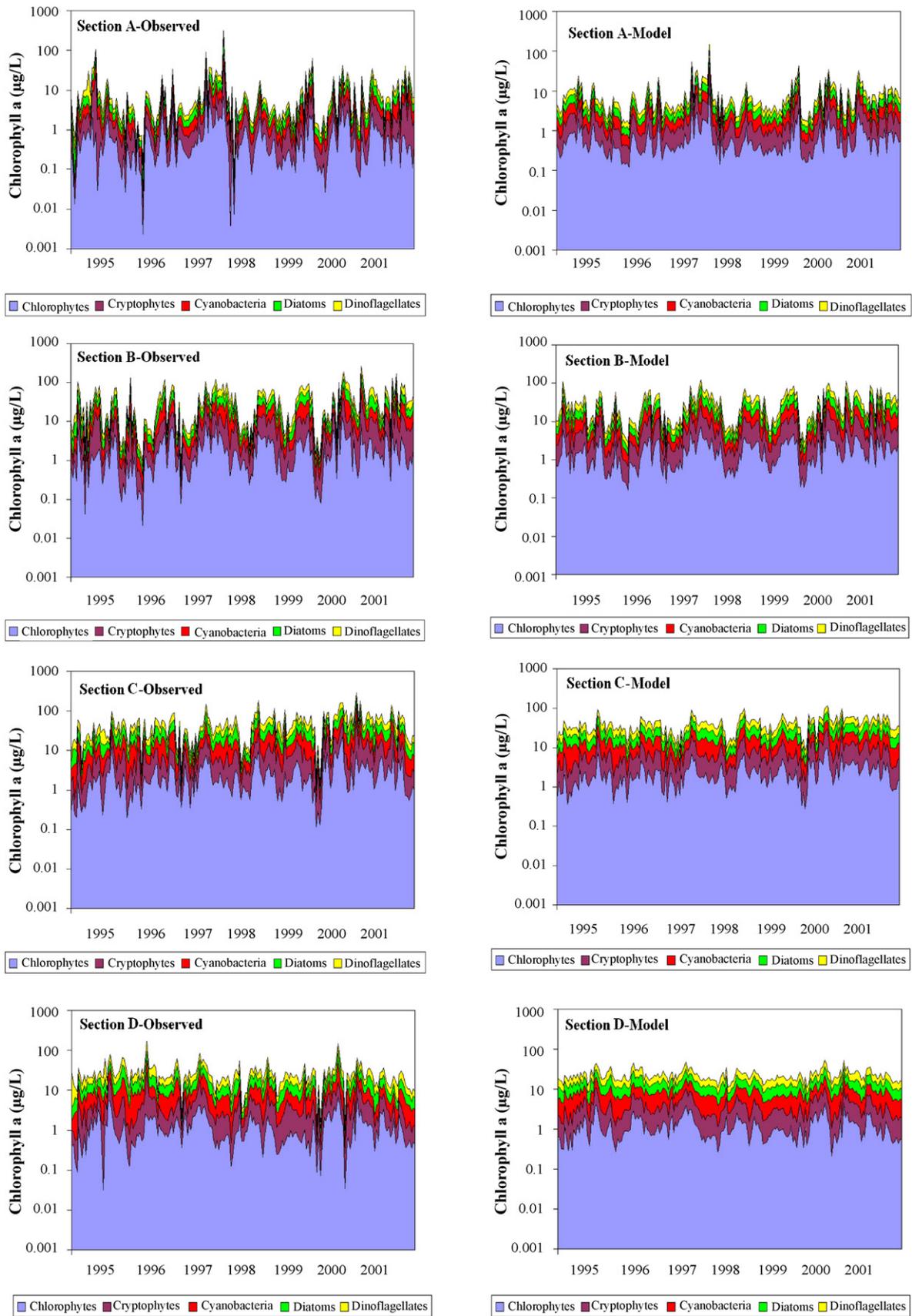
Fig. 10 – Observed (volume weighted) and predicted phytoplankton community composition in the four segments of the Neuse River Estuary. Model predictions are based on the medians of the predictive chlorophyte, cryptophyte, cyanobacteria, diatom, and dinoflagellate distributions.

	Section A	Section B	Section C	Section D
Phytoplankton model				
Section A	1.281			
Section B		1.192		
Section C			1.015	
Section D				1.018
Phytoplankton community model				
Section A	1.075			
Section B		1.085		
Section C			1.038	
Section D				1.301

The likelihood of the (N + P)/(N) model forms the numerator/denominator of the Bayes factor (see also Arhonditsis et al., 2007).

Table 3 – Bayes factor comparisons of the two phytoplankton community conceptualizations with the highest performance

	Section A	Section B	Section C	Section D
Section A	1.255			
Section B		1.082		
Section C			1.013	
Section D				0.945

The numerator of the Bayes factor is the likelihood of the model A, while the denominators correspond to the models C, D, C and C for the Neuse River Estuary sections A, B, C and D, respectively.

mid-estuary segments A and B, which progressively weakens and switches to a negative relationship (−0.35) as we move to the down-estuary area. The cyanobacteria–temperature path is consistently strong and positive along the estuary (>0.37). The comparisons between the alternative phytoplankton community classifications indicated that the first model (Fig. 3a) that considers dinoflagellates and cyanobacteria as two distinct functional groups and lumps together diatoms, cryptophytes and chlorophytes, outperforms in the sections A–C (Table 3). In contrast, the model that combines chlorophytes with diatoms, dinoflagellates with cryptophytes and treats cyanobacteria separately has a better foundation in the down-estuary segment D. However, it should be noted again that none of the Bayes factor values provided strong evidence in support of one of the alternative conceptualizations of the NRE phytoplankton community. 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. 10).

4.   Discussion

Elucidation of the fundamental causal connections and mechanistic understanding are often proposed as the "ulti-mate objective" for the development of effective coastal eutrophication management schemes. It is claimed that the consideration of the most important causal relationships is essential for a reliable model foundation that addresses the problem of the unwieldy ecological complexity and has the potential to provide quantitative tools with general applicability (Shipley, 2000). However, our current experience from the coastal ecosystem functioning does not justify the latter argument and highlights the necessity to adopt site-specific approaches in eutrophication research (Cloern, 2001). Furthermore, earlier work by Peters (1991) questions the actual meaning of the concept "cause" and the feasibility of identifying causal relationships in ecological systems, where everything is connected to everything else. In this study, although we do not agree with Peters' point of view that most of the "thorny" issues in ecology will disappear if cause, mechanism and explanation are ignored, we do recognize that even in our relatively simple conceptual model several meaningful paths are missing (e.g., connections between physical environment and nutrients or water temperature). In addition, the magnitude or the sign of the included ecological paths might differ if we consider lagged instead of contemporaneous measurements or a different temporal resolution. Our intention is neither to illustrate the enormous interconnectivity in nature nor to identify the "optimal model". In this study, we simply formulated a model based on our best knowledge of the system, which through a reasonable representation of the observed variability and conditional on our assumptions (scale, data aggregation) aims to illuminate the relative impact of the considered causal factors on the Neuse River Estuary spatiotemporal phytoplankton dynamics.

The physical environment is the predominant factor that regulates phytoplankton dynamics at the upper reaches of the estuary (sections A and B), where the respective posterior median paths for both the aggregate and compositional models were consistently strong (< −0.40). Given the high-standardized flow rate loading values on the latent physical environment, we also infer that the phytoplankton dynamics in the relatively narrow upper NRE area are primarily driven by the control of flow on the phytoplankton growth-minus-physical advection loss balance (Borsuk et al., 2004). As a result, the up-estuary total phytoplankton maxima usually occur during periods of low flow conditions, i.e., late summer–early fall (section A) and late spring–mid summer (section B), which is also reflected by the positive temperature path values (>0.45). In addition, the chlorophyll *a* concentrations are notably higher as we move downstream and further insight into the mechanisms that drive this gradient was provided by Pinckney et al. (1997). Specifically, this study proposed a conceptual pattern that explains the spatiotemporal disparity between elevated growth rates and phytoplankton productivity or biomass increase in the Neuse River Estuary. During periods of low phytoplankton biomass, favorable density-independent factors (meteorological conditions, nutrient and salinity levels) and the absence of density-dependent limitations (e.g., self-shading effects) stimulate algal growth in the upper reaches of the estuary. Phytoplankton blooms, however, are usually manifested in the central parts of the estuary, where the combination of elevated growth

rates and favorable physical conditions (low turbidity, slower current velocities and longer residence time) lead to biomass accumulation (Pinckney et al., 1997).

The fast-growing PFG A (diatoms, chlorophytes and cryptophytes) is the most abundant group in the upper portions of the estuary. [Note that the high-standardized loadings of the three phytoplankton taxa on the respective latent variable justify their treatment as one entity without missing much of the information underlying their individual patterns.] Local PFG A maxima are usually observed during the late spring–mid summer months (see also the respective temperature paths), but this group's opportunistic behavior also allows dominating the system after episodic hydrologic perturbations and showing irregular peaks throughout the annual cycle (Paerl et al., 2003). In addition, although the amplitude of the up-stream DIN:DIP ratio seasonal variations is quite wide; its values almost consistently lie above the Redfield ratio (16) and, thus, support stoichiometric predictions of phosphorus limitation (Redfield, 1958). The close coupling between phosphorus availability and PFG A dynamics is also indicated by the strongly negative posterior median path values (−0.26 and −0.44). It should be noted that the negative signs probably stem from the adopted temporal resolution of the study (data from individual samplings with no time averaging) and also reflect their ability to promptly respond to externally induced nutrient pulses (affinity/velocity strategists), increase their abundance and, consequently, decrease the contemporaneous nutrient levels; especially those in shortest supply (Happey-Wood, 1988; Klaveness, 1988; Hecky and Kilham, 1988). In contrast, the weak nitrogen paths suggest that the periods when the PFG A assemblage is more responsive (winter, spring and fall) are also associated with sufficient and relatively constant upstream DIN levels.

Being weak phosphorus competitors and having slower growth rates, cyanobacteria can only gain competitive advantage in warmer water temperatures and low nitrogen concentrations (Paerl, 1988; Andersson et al., 1994; Piehler et al., 2002). Thus, they usually dominate the upstream phytoplankton community during the mid-late summer period (see the strongly positive temperature paths), when: (i) the water column conditions are more stable, (ii) the DIN levels are lower, especially in section B where the respective nitrogen path is relatively strong (−0.25), and (iii) the DIN:DIP ratios reach their annual minima and, thus, their phosphorus competitive handicap is relaxed. Finally, the dinoflagellates also have slower growth rates, lower nitrogen and phosphorus affinities and high sensitivity to hydrological forcing (Pollingher, 1988; Paerl et al., 2003, 2006). The main dinoflagellate competitive advantages are their tolerance on cold water conditions and their flagellar mobility that allows screening the water column for nutrients and optimal light conditions (Pollingher, 1988). Thus, the dinoflagellates exhibit a fairly regular upstream annual cycle with fall maxima, associated with longer residence times, the minimum – but still sufficiently high – upstream DIN concentrations and relatively high DIP concentrations. This conceptual pattern probably explains the distinct negative temperature – (−0.14 and −0.20) and nitrogen – dinoflagellate paths (−0.17 and −0.34) and the weakly positive relationship with phosphorus (≈0.09).

The relationship between physical environment and total phytoplankton is weaker and also changes sign in the transitional area between the upstream freshwater environment and the downstream meso-polyhaline zone, i.e., the respective path switches from weakly negative (segment C) to positive (segments D). By inspecting the paths of the phytoplankton composition model, we infer that the three functional groups of the highest performing model (PFG A, dinoflagellates and cyanobacteria) show a differential response to the signals of the physical forcing. The positive physical environment–PFG A paths (particularly in section D) along with the higher standardized salinity loadings on the respective latent variable underscore this group's ability to exhibit optimal growth rates and dominate the phytoplankton community during high freshwater discharge/reduced salinity conditions (Paerl et al., 2003). Further insight into the PFG A dynamics can be obtained by the temperature posterior median paths, which became weakly positive (0.11) and negative (−0.35) in the spatial segments C and D, respectively. These path values probably reflect a progressive spatiotemporal shift of the PFG A patterns from the upstream summer maxima to a more uniform seasonal cycle in the mid-estuary, and more pronounced responses during the colder period of the year as we move to the downstream area. Aside from the physical environment effects that are no longer restrictive for winter bloom development, these changes should also be driven by the concurrent nutrient dynamics. Both DIN concentrations and DIN:DIP seasonal variations indicate transition towards a more prolonged status of nitrogen limitation in the lower estuarine locations. As a result, the duration of the suite of conditions (e.g., longer residence times, warm water temperatures, low DIN concentrations and relaxation of the phosphorus limitation) that can cause cyanobacteria dominance is longer, whereas the rest of the phytoplankton taxa experience a favorable environment for their competitive potential for a shorter period of the year (late fall until mid–spring). For the latter period, stoichiometry usually predicts a phosphorus and/or transition towards nitrogen limitation status, which probably explains the relatively distinct nitrogen and phosphorus signatures on the PFG A dynamics. Thus, this functional group's patterns are likely to be controlled by the synergistic nitrogen and phosphorus effects in the freshwater–saltwater transition zone, i.e., likelihood of N and P co-limitation instead of a single nutrient limitation. Interestingly, the nutrient–PFG A interactions are also depicted in the total phytoplankton model (see the respective nitrogen and phosphorus path values in Fig. 9c), since the three taxa that comprise this functional group account for a significant portion of the total phytoplankton biomass.

Similarly, the physical environment/temperature–dinoflagellate paths indicate a temporal shift towards the colder months of the year in the middle-lower estuary area. As it was also suggested from past NRE studies, this change is mainly due to the hospitable environment (sufficient nutrient levels, longer residence times and usually oligo- to mesohaline conditions) that this estuarine location provides for the fairly regular occurrence of late winter/early spring dinoflagellate blooms (e.g., *Heterocapsa triquetra* blooms, see Mallin, 1994; Paerl et al., 2006). Moreover, based on the existing NRE literature, we can also hypothesize that the signature of the

top–down control is likely to become more apparent under the longer residence times of the lower estuarine area and play a role on the phytoplankton community pattern formation. For example, Mallin and Paerl (1994) provided evidence that the abundance of both centric diatoms and dinoflagellates was positively correlated with the grazing rates, while the principal species (*Chroomonas amphioxiae*, *Chroomonas minuta* and *Cryptomonas testaceae*) of the local cryptophyte community are usually considered as high food quality for zooplankton (Brett et al., 2000). Generally, a complex interplay between physical, chemical and biotic factors probably induces the structural shifts on the phytoplankton community temporal patterns and, thus, can explain the change of the optimal phytoplankton community categorization in the spatial segment D; the model that aggregates diatoms with chlorophytes (PFG D), lumps dinoflagellates with cryptophytes (PFG E), and separately treats cyanobacteria provides slightly better results (Fig. 3c and Table 3).

One point for careful consideration is that the identification of the nutrient limitation status can be a formidable task and that inherent weaknesses characterize both the nutrient addition bioassays and the stoichiometric predictions based on the ambient DIN:DIP ratio (Hecky and Kilham, 1988; Carpenter, 1996). In this study, we interpreted our model results within the context of knowledge of the system; our inference was based on the relative magnitudes/signs of the structural paths vis-à-vis the ambient nutrient concentrations and ratios. However, there are several other mechanisms that were not considered in our analysis, and their explicit recognition is likely to put our results into perspective. More specifically, phytoplankton have the ability to physiologically adapt (e.g., luxury uptake and intracellular storage, alkaline phosphatase production) and alleviate nutrient limitation, while increasing evidence from theoretical and experimental studies indicates that the canonical Redfield ratio of 16 is likely to reflect an overall community average rather than a "universal optimum" (Sanudo-Wilhelmy et al., 2004; Piehler et al., 2004; Klausmeier et al., 2004). In addition, our modeling study did not consider the dissolved organic nitrogen (e.g., urea, nucleic and amino acids) role, which can become important in the middle and lower estuary area and supply a large proportion of the phytoplankton requirements (Twomey et al., 2005). Finally, several NRE studies have discussed the close coupling between phytoplankton community and lateral/sediment nutrient fluxes, emphasizing the importance of the bacterial-mediated nutrient recycling on the local phytoplankton patterns (Rudek et al., 1991; Twomey et al., 2005; Paerl et al., 2006). It is worth mentioning that, although implicitly, the fine temporal resolution of our models is likely to have captured the rapid phytoplankton-microbial consortia dynamics and, perhaps, the higher primary productivity standardized loadings on the latent phytoplankton of the sections C and D can be interpreted as an indication of a phytoplankton community intermittently dependent on nutrient recycling (i.e., carbon fixation with little net gain in biomass).

Undoubtedly, the derivation of phytoplankton paradigms entails compromises between simplistic approaches, where phytoplankton dynamics are perceived as an operation mainly driven by the availability of few natural resources, and complex approaches, where factors, such as adaptive strategies, niche behavior and colonization abilities come into play (Smayda, 2002; Cloern and Dufford, 2005). We believe that the former strategy can be a better starting point; as this study showed, the reduction of the hyperdimensional space of factors (main and interactive effects) that underlie phytoplankton dynamics to a low dimensional problem of several important causal connections does facilitate meaningful insights.

Acknowledgments

Appendix A

Using the classical SEM notation, we present an illustrative example of the matrices' forms and the specific assumptions made for the first phytoplankton community composition structural equation model (model A, Fig. 3). The development of the other three SEMs along with one used for exploring the total phytoplankton patterns is similar. 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 four ($n = 4$) exogenous latent variables ξ which were described from seven ($q = 7$) 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"; phosphate for the latent variable "Phosphorus" and temperature for the respective latent variable. Thus,

$$X = \begin{bmatrix} X_1 = NO_x \\ X_2 = DIN \\ X_3 = \text{Attenuation coefficient} \\ X_4 = \text{Salinity} \\ X_5 = \text{Flow} \\ X_6 = \text{Temperature} \\ X_7 = PO_4 \end{bmatrix},$$

$$\Lambda_X = \begin{bmatrix} \lambda_6 & 0 & 0 & 0 \\ \lambda_7 & 0 & 0 & 0 \\ 0 & \lambda_8 & 0 & 0 \\ 0 & \lambda_9 & 0 & 0 \\ 0 & \lambda_{10} & 0 & 0 \\ 0 & 0 & \lambda_{11} & 0 \\ 0 & 0 & 0 & \lambda_{12} \end{bmatrix},$$

$$\xi = \begin{bmatrix} \xi_1 = \text{Nitrogen} \\ \xi_2 = \text{Physical environment} \\ \xi_3 = \text{Temperature} \\ \xi_4 = \text{Phosphorus} \end{bmatrix},$$

$$\delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \end{bmatrix} \tag{A1}$$

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; for the model A, five indicator variables ($p = 5$) were used for the representation of three ($m = 3$) endogenous latent variable, i.e., chlorophytes, cryptophytes, diatoms were used as indicators for the latent variable "Functional Group A", and dinoflagellates, cyanobacteria for the respective latent variables. Thus, the exogenous latent variable measurement model can be described from the four matrices:

$$Y = \begin{bmatrix} Y_1 = \text{Dinoflagellates} \\ Y_2 = \text{Diatoms} \\ Y_3 = \text{Chlorophytes} \\ Y_4 = \text{Cryptophytes} \\ Y_5 = \text{Cyanobacteria} \end{bmatrix}, \quad \Lambda_Y = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & \lambda_3 & 0 \\ 0 & \lambda_4 & 0 \\ 0 & 0 & \lambda_5 \end{bmatrix},$$

$$\eta = \begin{bmatrix} \eta_1 = \text{Dinoflagellates} \\ \eta_2 = \text{Functional Group A} \\ \eta_3 = \text{Cyanobacteria} \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix} \tag{A2}$$

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

$$\Gamma = \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ \gamma_5 & \gamma_6 & \gamma_7 & \gamma_8 \\ \gamma_9 & \gamma_{10} & \gamma_{11} & \gamma_{12} \end{bmatrix}, \quad \zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} \tag{A3}$$

where Γ is the matrix of coefficients for the latent exogenous variables; ζ is the vector of latent (structural) errors. Note that the absence of direct cause–effect relationships between the three phytoplankton functional groups (PFG A, dinoflagellates and cyanobacteria) implies that the matrix of the coefficients that relate latent endogenous variables (B) is a zero matrix. As it can be inferred from the path diagram (Fig. 3), the associated covariance matrices of the model, $\text{Cov}(\xi) = \Phi(n \times n)$: 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 ele-

ments equal to zero:

$$\Theta_\varepsilon = \begin{bmatrix} \text{var}(\varepsilon_1) \\ 0 & \text{var}(\varepsilon_2) \\ 0 & 0 & \text{var}(\varepsilon_3) \\ 0 & 0 & 0 & \text{var}(\varepsilon_4) \\ 0 & 0 & 0 & 0 & \text{var}(\varepsilon_5) \end{bmatrix},$$

$$\Theta_\delta = \begin{bmatrix} \text{var}(\delta_1) \\ 0 & \text{var}(\delta_2) \\ 0 & 0 & \text{var}(\delta_3) \\ 0 & 0 & 0 & \text{var}(\delta_4) \\ 0 & 0 & 0 & 0 & \text{var}(\delta_5) \\ 0 & 0 & 0 & 0 & 0 & \text{var}(\delta_6) \\ 0 & 0 & 0 & 0 & 0 & 0 & \text{var}(\delta_7) \end{bmatrix},$$

$$\Psi = \begin{bmatrix} \psi_{11} \\ 0 & \psi_{22} \\ 0 & 0 & \psi_{33} \end{bmatrix},$$

$$\Phi = \begin{bmatrix} \phi_{11} \\ 0 & \phi_{22} \\ 0 & 0 & \phi_{33} \\ 0 & 0 & 0 & \phi_{44} \end{bmatrix} \tag{A4}$$

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 $\lambda_1 = \lambda_2 = \lambda_5 = \lambda_7 = \lambda_8 = \lambda_{11} = \lambda_{12} = 1.0$. Moreover, implicit in the assumption that the latent variables temperature, phosphorus, dinoflagellates, and cyanobacteria coincide with the respective observed variables, is: $\delta_6 = \delta_7 = \varepsilon_1 = \varepsilon_5 = 0$.

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

$$\begin{aligned} & Y_{1i} = \lambda_1 \eta_{1i} + \varepsilon_1, \\ & Y_{2i} = \lambda_2 \eta_{2i} + \varepsilon_2, \quad Y_{3i} = \lambda_3 \eta_{2i} + \varepsilon_3, \quad Y_{4i} = \lambda_4 \eta_{2i} + \varepsilon_4, \\ & Y_{5i} = \lambda_5 \eta_{3i} + \varepsilon_5 \\ & \varepsilon. \sim N(0, \Theta_\varepsilon) \\ & X_{1i} = \lambda_6 \xi_{1i} + \delta_1, \quad X_{2i} = \lambda_7 \xi_{1i} + \delta_2 \\ & X_{3i} = \lambda_8 \xi_{2i} + \delta_3, \quad X_{4i} = \lambda_9 \xi_{2i} + \delta_4, \quad X_{5i} = \lambda_{10} \xi_{2i} + \delta_5 \\ & X_{6i} = \lambda_{11} \xi_{3i} + \delta_6 \\ & X_{7i} = \lambda_{12} \xi_{4i} + \delta7 \\ & \delta. \sim N(0, \Theta_\delta), \quad \xi. \sim N(0, \Phi) \\ & \eta_{1i} = \gamma_1 \xi_{1i} + \gamma_2 \xi_{2i} + \gamma_3 \xi_{3i} + \gamma_4 \xi_{3i} + \zeta_1 \\ & \eta_{2i} = \gamma_5 \xi_{1i} + \gamma_6 \xi_{2i} + \gamma_7 \xi_{3i} + \gamma_8 \xi_{3i} + \zeta_2 \\ & \eta_{3i} = \gamma_9 \xi_{1i} + \gamma_{10} \xi_{2i} + \gamma_{11} \xi_{3i} + \gamma_{12} \xi_{3i} + \zeta_3 \\ & \zeta. \sim N(0, \Psi) \end{aligned} \tag{A5}$$

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):

$$\Sigma(\theta) = \begin{bmatrix} \Lambda_y(I - B)^{-1}(\Gamma\Phi\Gamma' + \Psi)[(I - B)^{-1}]'\Lambda_y' + \Theta_\varepsilon & \Lambda_y(I - B)^{-1}\Gamma\Phi\Lambda_x' \\ \Lambda_x\Phi\Gamma'[(I - B)^{-1}]'\Lambda_y' & \Lambda_x\Phi\Lambda_x' + \Theta_\delta \end{bmatrix} \tag{A6}$$

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

$$p(w|\theta) = \prod_{i=1}^{n}(2\pi)^{-(p+q)/2}|\Sigma(\theta)|^{-1/2}$$

$$\times \exp\left[-\frac{1}{2}[w_i - \mu(\theta)_i]'\Sigma(\theta)^{-1}[w_i - \mu(\theta)_i]\right] \qquad (A7)$$

where $q = 7$ and $p = 3$ are the number of exogenous and endogenous manifest (observed) variables. 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:

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

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, we used independent non-informative conjugate gamma priors (0.01, 0.01) for the elements of the matrices Θ_δ^{-1}, Θ_ε^{-1}, Φ^{-1} and Ψ^{-1}. Effectively "flat" normal prior distributions with means equal to 0 and precisions equal to 0.0001 were used for the structural parameters and the factor loadings. 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, and convergence was assessed using the modified Gelman–Rubin convergence statistic (Brooks and Gelman, 1998). 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; see Spiegelhalter et al., 2003) for all the parameters was less than 5% of the sample standard deviation. [All the material (e.g., model codes, data) pertinent to this study is available upon request from the first author.]

references

Andersson, A., Haecky, P., Hagström, Å., 1994. Effect of temperature and light on the growth of micro-, nano-, and pico-plankton: impact on algal succession. Mar. Biol. 120, 511–520.

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

Arhonditsis, G.B., Winder, M., Brett, M.T., Schindler, D.E., 2004. Patterns and mechanisms of phytoplankton variability in Lake Washington (USA). Water Res. 38, 4013–4027.

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. Ecol. Model 192, 385–409.

Arhonditsis, G.B., Paerl, H.W., Valdes, L.M., Stow, C.A., Steinberg, L.J., Reckhow, K.H., 2007. Application of Bayesian Structural Equation Modelling for examining the Neuse River Estuary (NC, USA) phytoplankton dynamics. Estuar. Coast. Shelf S 73, 63–80.

Boesch, D., Burreson, E., Dennison, W., Houde, E., Kemp, M., Kennedy, V., Newell, R., Paynter, K., Orth, R., Ulanowicz, R., 2001. Factors in the decline of coastal ecosystems. Science 293, 1589–1590.

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. Estuar. Coast. Shelf S 52, 33–49.

Borsuk, M.E., Stow, C.A., Reckhow, K.H., 2004. Confounding effect of flow on estuarine response to nitrogen loading. J. Environ. Eng. 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. Limnol. Oceanogr. 47, 1564–1575.

Brooks, S.P., Gelman, A., 1998. Alternative methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. 7, 434–455.

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. Mar. Ecol. Prog. Ser. 230, 103–112.

Carpenter, S.R., 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. Ecology 77, 677–680.

Christian, R.R., Boyer, J.N., Stanley, D.W., 1991. Multiyear distribution patterns of nutrients within the Neuse River estuary, North-Carolina. Mar. Ecol. Prog. Ser. 71, 259–274.

Cloern, J.E., 2001. Our evolving conceptual model of the coastal eutrophication problem. Mar. Ecol. Prog. Ser. 210, 223–253.

Cloern, J.E., Dufford, R., 2005. Phytoplankton community ecology: principles applied in San Francisco Bay. Mar. Ecol. Prog. Ser. 285, 11–28.

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

Downing, J.A., 1997. Marine nitrogen: phosphorus stoichiometry and the global N:P cycle. Biogeochemistry 37, 237–252.

Elmgren, R., Larsson, U., 2001. Nitrogen and the Baltic Sea: managing nitrogen in relation to phosphorus. In: The Scientific World, special ed. 1 (S2). Balkema Publishers, pp. 371–377.

Fisher, T.R., Gustafson, A.B., Sellner, K., Lacouture, R., Haas, L.W., Wetzel, R.L., Magnien, R., Everitt, D., Michaels, B., Karrh, R., 1999. Spatial and temporal variation of resource limitation in Chesapeake Bay. Mar. Biol. 133, 763–778.

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

Goldenberg, S.B., Landsea, C.W., Mestas-Nuzes, A.M., Gray, W.M., 2001. The recent increase in Atlantic hurricane activity: causes and implications. Science 293, 474–479.

Happey-Wood, C.M., 1988. Ecology of freshwater planktonic green algae. In: Sandgren, C.D. (Ed.), Growth and Reproductive Strategies of Freshwater Phytoplankton. Cambridge University Press, pp. 175–226.

Harding, L.W., 1994. Long-term trends in the distribution of phytoplankton in Chesapeake Bay: roles of light, nutrients, and streamflow. Mar. Ecol. Prog. Ser. 104, 267–291.

Hecky, R.E., Kilham, P., 1988. Nutrient limitation of phytoplankton in freshwater and marine environments—a review of recent evidence on the effects of enrichment. Limnol. Oceanogr. 33, 796–822.

Kass, R.E., Raftery, A.E., 1995. Bayes Factors. J. Am. Stat. Assoc. 90, 773–795.

Klausmeier, C.A., Litchman, E., Daufresne, T., Levin, S.A., 2004. Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton. Nature 429, 171–174.

Klaveness, D., 1988. Ecology of the cryptomonadida A first review. In: Sandgren, C.D. (Ed.), Growth and Reproductive Strategies of Freshwater Phytoplankton. Cambridge University Press, pp. 105–133.

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.

Mallin, M.A., 1994. Phytoplankton ecology of North Carolina estuaries. Estuaries 17, 561–574.

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

Malone, T.C., Conley, D.J., Fisher, T.R., Glibert, P.M., Harding, L.W., Sellner, K.G., 1996. Scales of nutrient-limited phytoplankton productivity in Chesapeake Bay. Estuaries 19, 371–385.

McNinch, J.E., Luettich, R.A., 2000. Physical processes around a cuspate foreland: implications to the evolution and long-term maintenance of a cape-associated shoal. Cont. Shelf Res. 20, 2367–2389.

Nixon, S.W., 1995. Coastal marine eutrophication—a definition, social causes and future concerns. Ophelia 41, 199–219.

Paerl, H.W., 1988. Growth and reproduction strategies of freshwater blue-green algae (cyanobacteria). In: Sandgren, C.D. (Ed.), Growth and Reproductive Strategies of Freshwater Phytoplankton. Cambridge University Press, pp. 261–315.

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

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. Environ. Sci. Technol. 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. Limnol. Oceanogr. 51, 448–462.

Peters, R.H., 1991. A Critique for Ecology. Cambridge University Press.

Piehler, M.F., Dyble, J., Moisander, P.H., Pinckney, J.L., Paerl, H.W., 2002. Effects of modified nutrient concentrations and ratios on the structure and function of the native phytoplankton community in the Neuse River Estuary, North Carolina, USA. Aquat. Ecol. 36, 371–385.

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. Estuar Coast. Shelf S 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. Can. J. Fish. Aquat. Sci. 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. Mar. Biol. 131, 371–381.

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.

Pugesek, B.H., Tomer, A., Von Eye, A. (Eds.), 2003. Structural Equation Modeling Applications in Ecological and Evolutionary Biology. Cambridge University Press, Cambridge, UK.

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. Environ. Sci. Technol. 34, 4474–4482.

Redfield, A.C., 1958. The biological control of the chemical factors in the environment. Am. Sci. 46, 205–230.

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. Mar. Ecol. Prog. Ser. 75, 133–142.

Sanudo-Wilhelmy, S.A., Tovar-Sanchez, A., Fu, F.X., Capone, D.G., Carpenter, E.J., Hutchins, D.A., 2004. The impact of surface-adsorbed phosphorus on phytoplankton Redfield stoichiometry. Nature 432, 897–901.

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

Shipley, B., 2000. Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference. Cambridge University Press.

Smayda, T.J., 2002. Adaptive ecology, growth strategies and the global bloom expansion of dinoflagellates. J. Oceanogr. 58, 281–294.

Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2003. WinBUGS User Manual, Version 1.4 [online]. Available from http://www.mrc-bsu.cam.ac.uk/bugs.

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

Twomey, L.J., Piehler, M.F., Paerl, H.W., 2005. Phytoplankton uptake of ammonium, nitrate and urea in the Neuse River Estuary, NC, USA. Hydrobiologia 533, 123–134.