Evaluation of the current state of mechanistic aquatic biogeochemical modeling

George B. Arhonditsis^{1,2,*}, Michael T. Brett¹

¹Department of Civil & Environmental Engineering, More Hall, Box 352700, University of Washington, Seattle, Washington 98195, USA

²Present address: Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina 27708, USA

ABSTRACT: The need for predictive process-oriented planktonic ecosystem models is widely recognized by the aquatic science community. We conducted a meta-analysis of recent mechanistic aquatic biogeochemical models (153 studies published from 1990 to 2002), to assess their ability to predict spatial and temporal patterns in the physical, chemical and biological dynamics of planktonic systems. The selected modeling studies covered a wide range of model complexity, ecosystem-types, spatio-temporal scales and purposes for model development. Despite the heterogeneous nature of this data set, we were able to identify model behavior trends and illuminate aspects of current modeling practice that need to be reevaluated. Temperature and dissolved oxygen had the highest coefficients of determination (respective median r² values were 0.93 and 0.70) and the lowest relative error (median RE < 10%), nutrients and phytoplankton had intermediate predictability (median r^2 values ranging from 0.40 to 0.60 and median RE ~ 40%), whereas bacteria (median $r^2 = 0.06$) and zooplankton (median RE = 70%) dynamics were poorly predicted. Longer simulation periods (i.e. months to decades) reduced model predictability, and increased model complexity did not improve fit. Aquatic biogeochemical modelers need to be more consistent in how they apply conventional methodological steps during model development (i.e. sensitivity analysis, validation), and the aquatic modeling community should adopt generally accepted standards of model performance. Recent advancements in data assimilation techniques, the combination of the present family of models with goal functions (derived from non-equilibrium thermodynamics) and the development of models with a stronger physiological basis are promising frameworks for obtaining more accurate simulations of planktonic processes.

KEY WORDS: Ecological modeling \cdot Model complexity \cdot Eutrophication \cdot Aquatic biogeochemical cycles \cdot Plankton systems

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

'... Model evaluation brings to bear all the subjective and objective elements of conscious and unconscious thought and reasoning of which the individual scientist is capable ...'

E. J. Rykiel Jr (1996) Ecological Modelling 90:239

Mechanistic biogeochemical models have been used extensively in aquatic ecosystems research (Jorgensen 1994, Franks 2002). For example, they have been applied as a management tool for predicting eutrophication

(Hamilton & Schladow 1997, Reckhow & Chapra 1999), to understand oceanic ecosystems (e.g. bloom dynamics, the global carbon cycle) and to predict biotic responses to climate change (Aumont et al. 2002, Boyd & Doney 2002, Moore et al. 2002). During the last 3 decades, the veracity of the scientific methodology of earth science models and their adequacy for forming the basis of public policy decisions has frequently been challenged (Oreskes et al. 1994). Some researchers suggest ecological models have only a heuristic value and consider their validation impossible (Starfield & Bleloch 1986,

Konikow & Bredehoeft 1992), others claim they resemble 'works of fiction' based on 'properties of convenience' (Cartwright 1983). While not all ecosystem modelers accept these viewpoints, there is no doubt that great uncertainty characterizes both the technical and philosophical aspects of the modeling endeavor. Despite the extensive literature debating these issues (Beck 1987, Rykiel 1996, Klepper 1997), there is still considerable controversy amongst model developers and the resource managers who use them about how to develop, evaluate and interpret numeric earth science models.

In this study, we evaluate the current state of mechanistic aquatic biogeochemical models across the range of temporal and spatial scales typically utilized. We address several common modeling questions by providing a quantitative assessment of model performance and by assessing how model performance depends on model development. This approach contrasts with several recent narrative discussions of modeling advances (Franks 1995, Doney 1999, Frost & Kishi 1999, Kantha & Clayson 2000, Doney et al. 2002, Franks 2002, Kawamiya 2002). The data compiled represent a heterogeneous group of modeling studies, especially with respect to complexity, spatial and temporal scales and model development objectives. Given the heterogeneous nature of these studies, our intention was not to determine which specific models 'work best', but to assess the general performance of this class of models. Any mechanistic model can obtain a good fit for the right or for the wrong reasons. However, a model that provides a poor fit to observed data clearly does not give an adequate representation of the underlying processes (Oreskes et al. 1994). Hence, with this analysis we attempted to quantify model performance and identify factors associated with variation in model performance (i.e. model complexity, simulated spatio-temporal scales).

METHODS

We attempted to assess all aquatic biochemical models published from 1990–2002. The literature was searched using the electronic databases: 'Aquatic Sciences & Fisheries Abstracts', 'BIOSIS previews', 'ISI Web of science databases' and 'ScienceDirect'. The keywords: 'eutrophication model(l)-ing', 'NPZ model(l)-ing', water quality model(l)-ing', 'phytoplankton model(l)-ing', 'freshwater model(l)-ing', 'ocean model(l)-ing' and 'biogeochemical model(l)-ing' were used as search terms. To be included in our analysis, studies had to present graphs or tables in which field data were compared to simulated data; 153 papers fit this criterion (Appendix 1). Because the large majority of these studies did not quantify goodness-of-fit, we digitized the

relevant graphs in these papers to extract the original raw data. We assessed the magnitude of digitizing error by comparing the summary statistics for those studies that reported these values to the corresponding goodness-of-fit estimates obtained for the digitized data. We found an r^2 of 0.938 (y = 0.01 + 1.02 x, n = 42) between the authors' reported and our extracted estimates of fit. Less than 5% of the modeling studies assessed included information (statistics or time-series plots) for all the state variables predicted, thus we were not able to evaluate overall model performance. We computed the relative error (RE = Σ lobserved values-simulated values|/∑observed values) and r² values as diagnostic measures of individual state-variable performance. The latter measure was selected because of its common use in modeling practice for assessing goodness-of-fit, but it should be noted that there can be instances when this statistical test provides misleading results (e.g. Mayer & Butler 1993). For example, lower r² values result when there is little temporal or spatial variability to be explained for a particular state variable.

In many cases, the spatio-temporal resolution of the simulated versus observed data comparisons required subjective judgment. For example, in many time-series plots it was not obvious if the modelers' intentions were to compare monthly, weekly or other time averages. Hence, we inferred their intentions by inspecting the time between data points and/or by looking for clues in the text of these papers and aggregated the data accordingly. The 2- or 3-dimensional comparisons (XY, XZ and XYZ graphs) between simulated and observed data were based on the use of grids, and data sampling at grid-nodes. We also tested the effects of these data aggregation schemes and model spatiotemporal resolution on performance. Evaluation of the temporal resolution was based on the assignment of nominal values from 1 (comparisons where we considered as deviance the perpendicular distances from the observed data to the simulated lines) to 5 (where we compared seasonal averages). Resolution of spatial comparisons varied from 0 (zero-dimensional or 'statevariable value vs time' graphs) to 3 (3-dimensional graphs). We computed the correlation coefficients between these values and model performance (r² and RE). We found a trend in model performance with temporal resolution (r = 0.102, p = 0.024 for state variable RE; r = -0.205, p < 0.001 for state variable r^2). In contrast, the spatial dimension of the simulated versus observed data comparisons (r = 0.102, p = 0.024 for state variable RE; r = -0.040, p = 0.342 for state variable r²) did not seem to have a clear relation to model performance. (An alternative assessment of the spatiotemporal scale impact on model performance will be discussed in the fourth subsection of 'Results'.)

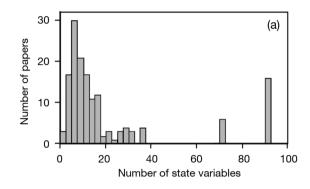
RESULTS

The mechanistic aquatic biochemical models assessed in this study had a very wide range of complexity (Fig. 1a). Simple models, similar to the Fasham et al. (1990) food-web model, consist of a small number (<10) of state variables and highly aggregated biotic compartments such as 'phytoplankton', 'bacteria' and 'zooplankton' which are modeled according to their carbon or nitrogen content. Complex models (>70 state variables) contain multiple elemental cycles (C, N, P, Si, Fe), multiple functional groups across all levels of organization, and sediment-water exchanges (Baretta et al. 1995). Numerous models have intermediate levels of complexity between these 2 extremes. During the last decade, the majority of models have been applied to oceanic systems (Fig. 1b), where they have usually been used to determine functional properties (i.e. long-term variation, nutrient supply to the subtropical gyre, limiting resources in high nutrient/ low-chlorophyll regions) or to explore the interface between the ocean's carbon cycle and global climatic responses (Dugdale et al. 1995, Hood et al. 2001, Moore et al. 2002). Models developed for smaller ecosystems (i.e. lakes, estuaries, harbors, bays and coastal areas) have usually been site-specific constructs addressing management issues such as eutrophication control (Le Pape & Menesguen 1997, Arhonditsis et al. 2002), or studying theoretical issues such as the potential productivity of upwelling systems (Ianson & Allen 2002) and the transfer of organic matter from the water column to sediments in margin systems (Tusseau-Vuillemin et al. 1998).

How consistently do modelers follow conventional modeling procedures?

The expression 'conventional modeling procedures' refers to the methodological steps typically recommended by classic modeling texts (Jorgensen 1997). We focused on 3 critical model development steps: sensitivity analysis, calibration, and validation, by exploring how these steps have been implemented in aquatic biogeochemical models published during the past decade.

Sensitivity analysis is the process by which the modeler attempts to evaluate the model sensitivity to the parameters selected, forcing functions, or state-variable submodels (Jorgensen 1997). This step in model development is essential for selecting the optimal model structure and complexity, because it indicates the accuracy required for the forcing function data and identifies the parts of the model that need to be estimated with greatest precision. We found 45.1% of the



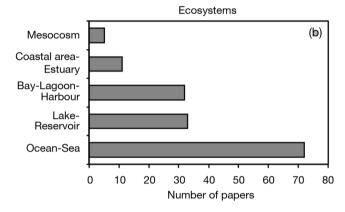


Fig. 1. Frequency histograms of (a) aquatic biogeochemical model complexity based on number of state variables and (b) types of ecosystems modeled

models published from 1990 to 2002 did not report results of sensitivity analysis (Fig. 2a). Thorough quantification of model sensitivity (Beck 1987, Klepper 1997) was only reported in 27.5% of the studies. The remaining 27.4% of the studies ('YES–NO' in Fig. 2a) examined the influence of certain model structures (e.g. alternative formulations) or parameters without providing quantitative measures of model performance.

Calibration is the procedure by which the modeler attempts to find the best fit between computed and observed data by adjusting model parameters. The large majority of the published models (91.5%) were calibrated by tuning the model within the range of literature parameter values until a satisfactory fit was obtained (Fig. 2b). The problem with this approach is that it does not ensure that the resulting parameter set is optimal and whether any lack of fit is due to an inadequacy in the model's structure or to poor parameter choice (Kawamiya 2002). This problem can be solved by application of optimization techniques designed to search the parameter space for combinations of parameters which provide the best fit through minimization of cost functions. A wide variety of techniques designed to identify both local and global minima have

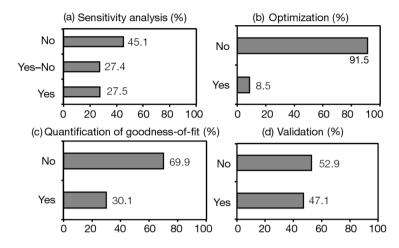


Fig. 2. Proportion of aquatic biogeochemical modeling studies that (a) performed sensitivity analyses, (b) used optimization techniques for model calibration, (c) quantified fit between simulated and field data, and (d) validated models (see first subsection of 'Results'). In (a) category 'YES-NO' indicates qualitative approaches (see first subsection of 'Results' for further explanation)

been developed during the second half of the last decade (Matear 1995, Hurtt & Armstrong 1999, Vallino 2000). Hence, the studies that implemented calibration optimization techniques still comprise only a small proportion (8.5%) of those published during the last 12 yr.

Validation is perhaps the most controversial issue in the model development process, and this controversy 'arises as much from semantic and philosophical considerations as from the selection of validation procedures' (Rykiel 1996). Generally, it is defined as a test of whether 'a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model', and it is considered an essential step for obtaining an objective measure of model reliability (Rykiel 1996, Jorgensen 1997). The present study compiled information for 4 types of validation: replicative, predictive and structural validation, and model transferability. Replicative model validation means evaluation of the agreement between predicted values and observational data from the real system during the calibration phase (Power 1993); essentially the quantification of goodness-of-fit during the formulation and estimation phases of model construction with data already acquired from the real system (Fig. 2c). Only 30.1% ('YES' in Fig. 2c) of plankton modeling studies published during the last decade statistically

evaluated the model performance, and only a small fraction (<6%) of these objectively quantified the model's performance by applying all appropriate measures of fit (i.e. RE, mean absolute error) and statistical significance (i.e. t-test, regression analysis) (Mayer & Butler 1993, Power 1993). In the remaining 69.9% of cases, the authors presented time-series plots but did not report quantitative measures of per-

Table 1. Performance of aquatic biogeochemical models for study period 1990–2002. Coefficient of determination (r²) and relative error (RE, %) values for temperature (T), dissolved oxygen (DO), nitrate, ammonium, phosphate, silicate, phytoplankton (Phyt), zooplankton (Zoop) and bacteria (Bact)

Percentile		T	DO	Nitrate	Ammonium	Phosphate	Silicate	Phyt	Zoop	Bact
10th	r^2	0.42	0.34	0.10	0.05	0.07	0.20	0.08	0.06	0.00
	RE	2 %	4 %	8%	18%	19%	18 %	20 %	17 %	21%
20th	\mathbf{r}^2	0.62	0.52	0.37	0.13	0.13	0.35	0.16	0.09	0.00
	RE	4 %	7 %	18%	30 %	26 %	30 %	26 %	31%	25 %
30th	r^2	0.81	0.58	0.47	0.18	0.20	0.46	0.30	0.12	0.01
	RE	5%	8%	26 %	34 %	32 %	32 %	32 %	44 %	33 %
40th	r^2	0.92	0.62	0.56	0.29	0.30	0.52	0.41	0.19	0.03
	RE	5%	10 %	32 %	40%	36%	34 %	37 %	52%	35 %
50th	\mathbf{r}^2	0.93	0.70	0.68	0.39	0.47	0.61	0.48	0.24	0.06
	RE	7 %	12%	36%	48%	42 %	37 %	44 %	70%	36 %
60th	r^2	0.95	0.78	0.79	0.44	0.57	0.66	0.56	0.37	0.18
	RE	7 %	14 %	44 %	55%	47 %	41%	51%	79%	37 %
70th	\mathbf{r}^2	0.96	0.86	0.84	0.57	0.69	0.69	0.63	0.69	0.21
	RE	9%	17 %	57 %	65%	55%	46%	58%	115%	42 %
80th	\mathbf{r}^2	0.97	0.88	0.91	0.78	0.80	0.80	0.76	0.90	0.24
	RE	11%	19%	68 %	77 %	69%	55%	66 %	138 %	49%
90th	\mathbf{r}^2	0.98	0.92	0.95	0.89	0.86	0.89	0.83	0.97	0.39
	RE	15 %	22 %	88%	101%	84 %	64%	79%	201 %	59%
100th	\mathbf{r}^2	0.99	1.00	1.00	0.99	0.96	0.98	0.98	1.00	0.64
	RE	25 %	31%	554 %	206 %	218%	302 %	128%	435%	66 %

formance. Fig. 2d also includes aggregated information for the other 3 types of validation: (1) predictive validation, which is defined as an evaluation of model-fit to independent data sets (model performance against data acquired from the real system after model calibration), (2) model transferability to different systems (performance of a specific model structure to different regions or ecosystem types), and (3) structural validation, which is defined as the assessment of the realistic reproduction of the operational characteristics, causal relationships, and relative magnitudes of various components of the system by the model (biological rates, derived quantities) (Power 1993). These validation procedures were carried out in 47.1% of the studies, while the remaining 52.9% were not predictively or structurally validated.

How well do aquatic biogeochemical models simulate real-world dynamics?

The 153 aguatic biogeochemical models assessed simulated the dynamics of 3386 state variables. Of these 3386 cases, the authors fully reported statistical measures of model performance (as well as time-series plots in which simulated values were compared with observed data) in 42 cases. In 569 cases, the authors presented time-series plots, but no statistical measures of performance. In 2817 cases, neither measures of statistical performance nor time series plots were reported for the state variables. In most of these cases the non-reported state variables were not the focus of these studies and/or observational data were not available for the respective comparisons. In the 569 cases where only time-series plots were presented, we digitized these graphs and used the extracted data to independently assess state-variable performance as expressed by the RE and the coefficient of determination (r²) (Fig. 3). Temperature and dissolved oxygen had the highest r² values (the respective medians were 0.93 and 0.70) and lowest RE (median < 10%). Limiting nutrients (NO₃, NH₄, PO₄ and Si) and phytoplankton biomass had intermediate fit, with median r² values ranging from 0.40 to 0.60 and the median RE around 40%. Only a fairly small number of aquatic biogeochemical studies presented model performance for bacteria (n = 14) and zooplankton (n = 30) dynamics. Both of these state variables were poorly modeled, with bacteria having the lowest median r^2 value (0.06) and zooplankton the highest RE (70%). The zooplankton predation simulations were also characterized by the widest range of r^2 (interquartile range ~ 0.8) and RE (interquartile range ~85%) values. More detailed information about the state-variable performance is provided in Table 1.

Is 'calibration bias' between state variables introduced during model calibration process?

We looked for potential 'calibration bias' introduced when calibrating aquatic biogeochemical models. Calibration bias can occur when, in the process of maximizing the fit for a specific state variable (usually phytoplankton biomass), the fit is reduced for other state variables (such as limiting nutrient concentrations or herbivorous zooplankton biomass) (Franks 1995). As previously mentioned, more than 95% of the studies we assessed did not report performance results for all state variables. In some cases this might suggest selective presentation of the state variables with better performance and/or that modelers presented the state variables they were mostly interested in. The lack of performance information for all the state variables precluded a comprehensive assessment of calibration bias in the present database. However, we sought evidence

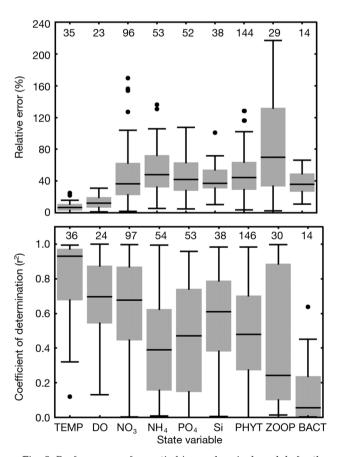


Fig. 3. Performance of aquatic biogeochemical models for the study period 1990 to 2002. Relative error (RE, %) and coefficient of determination (r^2) values for temperature (TEMP), dissolved oxygen (DO), nitrate (NO₃), ammonium (NH₄), phosphate (PO₄), silicate (Si), phytoplankton (PHYT), zooplankton (ZOOP) and bacteria (BACT). Number of studies assessed for each state variable is indicated above box-plots. Several extreme RE values were not included in these plots

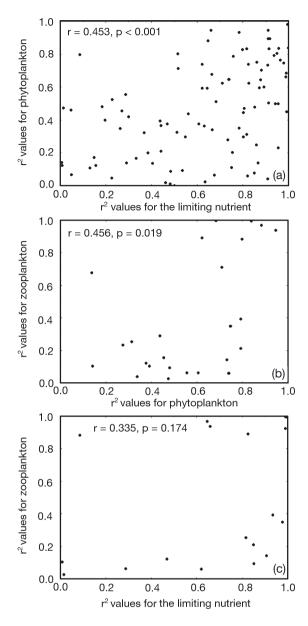


Fig. 4. Testing for 'calibration-bias' during calibration of aquatic biogeochemical models. Relationships between coefficient of determination (r²) values of (a) the limiting nutrient and phytoplankton, (b) phytoplankton and zooplankton and (c) the limiting nutrient and zooplankton

of biased calibration by comparing the fits between the key components of the aquatic biogeochemical models: limiting nutrients (inorganic nitrogen or phosphate), phytoplankton and zooplankton biomass. Despite the limitations of the underlying data, this approach provides some insight to the relative performance of the basic state variables for the aquatic biogeochemical models. The results of these comparisons did not show negative correlations between state variables (Fig. 4). In fact, significant positive correlations

were observed for the r^2 values between limiting nutrients and phytoplankton biomass (r=0.453, p<0.001), phytoplankton and zooplankton biomass (r=0.456, p=0.019), and a non-significant but still positive correlation between limiting nutrients and zooplankton biomass (r=0.335, p=0.174). Hence, the general result is that the last decade's plankton models can be differentiated into models with 'poor' to 'good' model performance with a uniformly apportioned fit between the basic structural components.

How do model complexity, spatio-temporal scale and type of modeled ecosystem impact model performance?

Determination of the most appropriate level of model complexity is an issue of particular interest to the ecological modeling community. This problem was articulated by Costanza & Sklar (1985), who examined how optimum complexity (with respect to the number of components, time and space) influenced model effectiveness (the observed variability explained). Simple models are more easily understood, have fewer unconstrained parameters and are more easily subjected to detailed sensitivity analyses, but they are often criticized as being crude oversimplifications not capable of reproducing real-world dynamics. Complicated models parameterize numerous processes and theoretically have the potential to be more accurate representations of complex natural systems. However, when model complexity is not accompanied by sufficient knowledge about the systems being modeled (i.e. in data-poor situations), complex models have been criticized for being completely artificial constructs that may include misconceptualizations of the modeled systems. The resulting uncertainty undermines their value and limits their role to that of mere heuristic tools. We therefore assessed the effect of model complexity (expressed as the number of state variables), spatial dimension (from zero to 3-dimensional) and simulation period (from days to decades) on model performance (Table 2).

A significant but quite weak negative correlation was found between the duration of the simulation period and the state variable r^2 values (r = -0.250, p < 0.001), and a very weak positive correlation was found between the simulation period and the RE (r = 0.098, p = 0.022). The same trend was identified when testing the state variables individually (r = -0.386, p < 0.001 and r = 0.239, p = 0.004 for phytoplankton r^2 and RE, respectively). In general, longer simulations resulted in slightly poorer model performance. The effects of the temporal resolution on model results were insignificant (rpartial = 0.057, p = 0.184 for the state variable RE and rpartial = -0.063, p = 0.136 for the state variable r^2) after accounting for the duration of the

Table 2. Correlation matrix between state variable performance (expressed as coefficient of determination r^2 and relative error RE) and simulation period, spatial dimension and model complexity (expressed as number state variables). Simulation period and spatial dimension: 0- to 3-dimensional modeling studies (not the reported graphs) were assigned nominal values from 0 to 3; simulation period durations of several days (<1 mo), several months (<1 yr), several years (<1 decade), and >1 decade were assigned nominal values from 1 to 4, respectively

Performance	Simulation period	Spatial dimension	Model complexity					
All state variables								
r^2	-0.250 (p < 0.001)	-0.086 (p = 0.043)	-0.108 (p = 0.011)					
RE	0.098 (p = 0.022)	0.104 (p = 0.015)	0.219 (p < 0.001)					
Phytoplankton								
r^2	-0.386 (p < 0.001)	-0.122 (p = 0.142)	-0.053 (p = 0.522)					
RE	0.239 (p = 0.004)	0.178 (p = 0.033)	0.248 (p = 0.003)					
Zooplankton	Zooplankton							
$r^{\frac{1}{2}}$	-0.408 (p = 0.025)	-0.369 (p = 0.045)	-0.205 (p = 0.276)					
RE	0.092 (p = 0.634)	0.524 (p = 0.003)	0.626 (p < 0.001)					
Nitrogen ^a								
r^2	-0.351 (p < 0.001)	-0.180 (p = 0.019)	-0.198 (p = 0.010)					
RE	0.211 (p = 0.007)	0.083 (p = 0.287)	0.160 (p = 0.039)					
Phosphorus ^a								
r ²	-0.330 (p = 0.007)	0.244 (p = 0.050)	0.086 (p = 0.495)					
RE	-0.063 (p = 0.619)	-0.072 (p = 0.570)	0.103 (p = 0.416)					
^a Includes all modeled inorganic and organic forms of this nutrient								

respective simulation periods. Hence, the previously reported significant correlations most likely reflect the influence of the simulation period, which indirectly determines the temporal resolution of the observed versus simulated data comparisons (i.e. shorter simulation periods such as less than 1 yr, were usually combined with finer temporal resolution such as days or weeks). Marginally significant correlations were found between the spatial complexity of the models and their performance trends (r = -0.086, p = 0.043 and r = 0.104, p = 0.015 for r^2 and RE, respectively). The fact that coupling physical and biological processes did not improve model predictability was not surprising, since the biogeochemical models can only be as good as the physical framework on which they are based. On the other hand, the fact that no significant degradation of model results was observed is a promising sign for future modeling development, and especially for the prospect of refining circulation models. A greater number of state variables does not seem to improve model performance, as indicated by the significant negative (but very weak) correlation between the number of state variables and the r^2 values (r = -0.108, p = 0.011), and the significant positive correlation with the RE of the model outputs (r = 0.219, p < 0.001). The RE increase was clear for phytoplankton (r = 0.248, p = 0.003) and zooplankton dynamics (r = 0.626, p < 0.001), while both nitrogen r^2 (r = -0.198, p = 0.010)

and RE (r = 0.160, p = 0.039) values were significantly degraded with increasing model complexity. This finding contradicts a common notion in modeling that the more complex the model, the more accurate it is likely to be due to a larger number of tunable parameters (Franks 1995). In theory this notion seems plausible, but in practice modelers do not appear to use all the increased degrees of freedom provided by highly parameterized models when calibrating their models. This may also be a sign of insufficient calibration effort or, alternatively, erroneous parameterization, which in turn raises questions about the current level of understanding of the relevant aquatic ecosystem processes. It should be noted however that the majority of the complex models in our dataset belong to the ERSEM (European Regional Seas Ecosystem Model) family, and the present results are—at least partially—influenced by the development purposes, modeled environments and practices followed by

this particular family of models. Finally, when using analysis of variance (ANOVA) to test for differences in model performance (both $\rm r^2$ and RE values) between system categories (i.e. mesocosm, lake-reservoir, bay-lagoon-harbor, coastal area-estuary and oceansea), we found model performance did not vary depending on the type of the ecosystem modeled. Apparently, the last decade's ambitious plankton modeling efforts to increase the level of biological detail, to increase spatial complexity and to use longer simulation periods have not led to a systematic improvement in model performance.

DISCUSSION

Present experience and future perspectives

Our intention with this meta-analysis was to carry out a systematic classification of a heterogeneous database in order to identify trends in current aquatic biogeochemical modeling practice. We showed that most plankton modelers do not evaluate sensitivity, test validity or assess the performance of their models in a consistent manner. Many aquatic biogeochemical models are still developed and applied without having a clearly stated purpose or a predetermined acceptable model performance level. Nevertheless, the need for

methodological consistency during model development has been emphatically argued during the last decade (Power 1993, Jorgensen 1994, Franks 1995, Rykiel 1996), and specification of performance criteria has been deemed an essential step for determining model acceptability. It has also been pointed out that given the subjectivity inherent in assessment of model performance, what is actually needed is a predetermined global convention (similar to the 0.05 probability level for statistical significance testing) that would be used by all modelers. The present study's quantitative assessment of aquatic biogeochemical model performance can contribute to the development of predetermined model performance criteria. As indicated by our results, in order to develop a fair test it is necessary to take into account and weigh model performance for the simulation period and probably the model complexity (number of state variables). Performance tests should not neglect biological rates (primary production or grazing rates) and derived quantities (f ratios) when testing whether the simulation realistically reproduces the functional properties of the system modeled. In cases where relevant data are not available, it should be clearly indicated that the model was not structurally validated, and therefore the match between state variables and observed data might not be the result of a correct model solution. Furthermore, during predictive validation of models it is common practice to split the data set into 2 subsets, and use the first subset for calibration and the second for testing its predictive ability. This procedure can be useful for determining whether a model was 'overfitted' to data during calibration. However, it has also been claimed that validation rigor is dependent on the difference between the calibration and validation data sets (Reckhow & Chapra 1999). For example, in studies intended to model system responses to increasing nutrient loading, the validation data set should describe enrichment conditions that significantly differ from those used during the calibration process. In cases where such data sets are not available and the 2 subsets (calibration and validation) are essentially identical, the modelers should be aware that they have not unequivocally tested the model's ability to predict new conditions (e.g. Omlin et al. 2001).

The selection of the model complexity should be driven by the system being studied and the questions being asked (Franks 1995). Some modelers have suggested, as a rule-of-thumb, starting with simple approaches and proceeding to greater complexity as warranted by data availability. However, models are quite often turned into 'mathematical exercises' by incorporating enormous complexity, which according to our results does not improve performance (Table 2). In addition, more than 95 % of the studies we assessed

did not report performance results for all state variables. Aside from the possibility of selective presentation of the state variables with better performance, this could also indicate inclusion of 'tunable' state variables in the models which are not constrained by data. If this is the case, this should be explicitly stated in the model, and the calibration process should assure that reasonable behavior for one variable is not sacrificed to achieve a good fit for others (Franks 1995).

The idea that optimal complexity and model performance assessments should be based on appropriate data sets does not imply that the observed data are a perfect criterion (Fagerstrom 1987). Observational data are just a 'snapshot' of the real system, an instantaneous record of a few components from numerous complex and interactive processes that must be carefully interpreted. Modelers should rather consider the observational data as an approximate representation of the real system that can, depending on the sampling network used, the ecosystem modeled and the questions addressed, form an objective basis for evaluating model performance. By taking into account these uncertainties, data assimilation methods attempt to provide solutions which are dynamically consistent with both the data and model outputs and can assist in model performance evaluation (Doney 1999). The choice of the misfit measure between model and observational data (cost function problem) has a significant effect on the 'best' estimate of model parameters (Evans 2003). Although this approach has been criticized for making models data interpolation algorithms, data assimilation methodologies may be one of the most promising means for solving the problem of model initialization, while also improving model parameterization and predictive ability (Gregoire et al. 2003). In addition, they can highlight missing processes, verify present model structure and guide future model reformulation (e.g. Spitz et al. 1998, Evans 1999). Thus, it is expected that modelers will embrace the assimilation techniques over the next decade beyond their current 8.5% application level, and explore their use as a part of the standard modeling development process. Currently, these approaches are more frequently combined with low-dimension ecosystem models (Evans 1999, Fennel et al. 2001). Applications to 3-dimensional models are more limited, but consist of attempts to estimate poorly quantified fluxes/biogeochemical rates and nutrient distributions as derived from large-scale physical circulation fields (Doney et al. 2002, Schlitzer 2002). Two major issues that will expedite the dissemination of the data assimilation methods on the oceanographic/limnological practice are (1) the construction of user-friendly automated systems linking the observing systems, numerical models and assimilation schemes, and (2) the adoption and refinement of approaches from other fields (meteorology) to support coupled physical-biological data assimilation.

Future research should also focus on refinements to several mechanistic aspects of current aquatic biogeochemical models (Doney 1999). We showed that zooplankton and bacteria dynamics are the most poorly simulated state variables (Fig. 3). While some studies provided fairly good simulations of phytoplankton/ zooplankton interactions (Fig. 4), most aquatic biogeochemical models were unable to obtain satisfactory performance. One reason for this might be error associated with converting the observed zooplankton data (in units of individuals per volume) to the modeled units (usually carbon or nitrogen mass per volume). Other authors have suggested that the only effective way to predict zooplankton dynamics is to fully simulate zooplankton life histories. A characteristic example is the copepod-submodel suggested by Fennel & Neumann (2001) that included eggs, nauplii, copepodites and adults as state variables. In addition, recent advances in stoichiometric nutrient recycling theory, reformulation of the zooplankton grazing term to include algal food-quality effects on assimilation efficiency, and a more detailed consideration of energy transfer through food-web processes (Touratier et al. 2001) should improve the realism of dynamic plankton simulations. Other important issues that need to be considered are the modeling of multi-nutrient interactions at the primary producer level (Flynn 2003), inclusion of the optimal nutrient/food-web framework (Reckhow & Chapra 1999), refinement of the sediment diagenesis processes (Di Toro et al. 1990), incorporation of the microbial loop and reformulation of the nutrient recycling processes (Doney et al. 2002). Concerning the latter, our meta-analysis showed that only a small number of studies explicitly modeled bacteria, which was also a poorly simulated state variable. Comparisons with the appropriate data from the real system (active instead of total bacterial abundance) and different parameterizations (first- vs higher-order kinetics) are 2 plausible strategies for improving the microbial loop simulations. Furthermore, caution should be exercised concerning the choice of the zooplankton mortality (the so-called 'closure' term). Several recent studies addressed this issue, and it was shown that the choice of this term can have a strong influence on model dynamics (e.g. Steele & Henderson 1992, Edwards & Brindley 1999, Edwards & Yool 2000). Hence, site-specific selection as derived from field experimentation seems to be the most appropriate way for choosing between a linear, quadratic, hyperbolic or sigmoid functional form (Franks 2002). Finally, recent mechanistic advancements in biogeochemical models include a more sophisticated physiological basis and

associate chemical composition changes in phytoplankton cells with ambient condition variability. For example, the model by Geider et al. (1998) accounts for acclimation of growth rate, photosynthesis, respiration, C:chlorophyll a, and the N:C of phytoplankton to irradiance. The modeling strategy introduced by Lancelot et al. (2000) simulates phytoplankton content in functional metabolites, reserve products and monomeric substrates in order to distinguish between photosynthesis and growth processes. This explicit treatment of the dynamic nature of the phytoplankton cell properties combined with models that consider the local-scale physiological structure in phytoplankton populations is also likely to improve plankton process simulations (Broekhuizen et al. 2003).

Is use of goal functions based on non-equilibrium thermodynamics the solution for more realistic models?

Aside from the previous comments on biogeochemical models, this family of models still has a fundamental weak point. Classic aquatic biogeochemical models have a rigid structure based on a fixed set of parameters, which does not allow for modifications or replacements of system components (Jorgensen 1997). However, changes in external forcing (weather conditions, nutrient loading) induce shifts in plankton community composition, and thus it would ultimately be more realistic to carry out simulations using time-varying parameters. Furthermore, it is possible that the extensive use of non-linear modeling approaches stems from attempts to get acceptable model performance using unrealistically rigid parameter sets, while linear approaches with time-varying parameters might give more reasonable results (Patten 1997). A new generation of models attempts to simulate the dynamic structure of ecosystems, using goal functions (or governing laws) that determine the self-organizing response of ecosystems to perturbations. Several goal functions derived from non-equilibrium thermodynamics have been proposed for tracking the direction of ecosystem development (Vallino 2003).

From this modeling strategy, we will use as an example the thermodynamic variable exergy, combined several times with aquatic biogeochemical models, to briefly describe the basic concept of these frameworks. Exergy is a measure of the distance from thermodynamic equilibrium (where no structure and free energy is available), which is related to the biomass and information (expressed by the number of genes) of the various system components (model state variables) (Jorgensen 1997). It is hypothesized that the way the ecosystem (model) operates is that it tends to maximize its exergy (goal function) according to the prevailing

conditions at each time step by changing the properties (model parameters) of its components (state variables). Thus, according to this scheme, the same type of models operate now with continuously changing parameter values as determined by the goal function: the optimization of the ecosystem's exergy under the variant environmental conditions. The first case studies to apply this approach gave plausible results and reproduced structural shifts in the biotic compartments, which with conventional approaches would have been feasible with only very complex models (Jorgensen 1999). This may be a promising framework for getting reliable parameter estimates and more effectively featuring ecosystem properties in models. At this point however it cannot be claimed that this structural dynamic approach gives better results than the classic aquatic biochemical models. Further research and more experience are needed before applying these models for practical management pur-

In conclusion, our meta-analysis showed that the performance of existing mechanistic aquatic biogeochemical models declines as we move from physical/chemical to biological components of planktonic systems. The same decreasing trend was identified when we compared phytoplankton with bacteria/zooplankton simulations. In addition, longer simulation periods and increased model complexity did not improve model performance. Aquatic biogeochemical modelers do not consistently apply conventional methodological steps during model development (e.g. sensitivity analysis, validation). Thus, the establishment of a systematic methodological protocol for aguatic biogeochemical model development, which is widely accepted by the modeling community, would be a very useful step. Data assimilation techniques, the development of models with a stronger physiological basis and the combination of the present family of models with goal functions that allow for time-varying parameters provide promising frameworks for improving simulations of planktonic processes. Finally, the purpose of the model development, the adopted model complexity, and the spatio-temporal scale of interest should also be considered for developing performance criteria that objectively evaluate the aquatic biogeochemical models.

Acknowledgements. We thank C. L. DeCasperi (King County, Department of Natural Resources and Parks), M. Danielsdottir (Civil and Environmental Engineering, University of Washington), K. H. Reckhow (Nicholas School of the Environment and Earth Sciences, Duke University) and 4 anonymous reviewers for very useful suggestions and comments regarding the manuscript. The study was supported by a grant from the King County, Department of Natural Resources and Parks, Wastewater Treatment Division.

LITERATURE CITED

- Arhonditsis G, Tsirtsis G, Karydis M (2002) The effects of episodic rainfall events to the dynamics of coastal marine ecosystems: applications to a semi-enclosed gulf in the Mediterranean sea. J Mar Syst 35:183–205
- Aumont O, Belviso S, Monfray P (2002) Dimethylsulfoniopropionate (DMSP) and dimethylsulfide (DMS) sea surface distributions simulated from a global three-dimensional ocean carbon cycle model. J Geophys Res C 107:3029
- Baretta JW, Ebenhoh W, Ruardij P (1995) The European-Regional Seas-Ecosystem-Model, a complex marine ecosystem model. J Sea Res 33:233–246
- Beck MB (1987) Water-quality modeling—a review of the analysis of uncertainty. Water Resour Res 23:1393–1442
- Boyd PW, Doney SC (2002) Modeling regional responses by marine pelagic ecosystems to global climate change. Geophys Res Lett 29(16):53-1-53-4
- Broekhuizen N, Oldman J, Zeldis J (2003) Sub-grid-scale differences between individuals influence simulated phytoplankton production and biomass in a shelf-sea system. Mar Ecol Prog Ser 252:61–76
- Cartwright N (1983) How the laws of physics lie. Clarendon Press, Oxford
- Costanza R, Sklar FH (1985) Articulation, accuracy and effectiveness of mathematical models—a review of freshwater wetland applications. Ecol Model 27:45–68
- Di Toro DM, Paquin PR, Subburamu K, Gruber DA (1990) Sediment oxygen demand model: methane and ammonia oxidation. J Environ Engr ASCE 116:945–986
- Doney SC (1999) Major challenges confronting marine biogeochemical modeling. Global Biogeochem Cycles 13: 705–714
- Doney SC, Kleypas JA, Sarmiento JL, Falkowski PG (2002) The US JGOFS synthesis and modeling project — an introduction. Deep-Sea Res II 49:1–20
- Dugdale RC, Wilkerson FP, Minas HJ (1995) The role of a silicate pump in driving new production. Deep-Sea Res I 42:697-719
- Edwards AM, Brindley J (1999) Zooplankton mortality and the dynamical behaviour of plankton population models. Bull Math Biol 61:303–339
- Edwards AM, Yool A (2000) The role of higher predation in plankton population models. J Plankton Res 22:1085–1112
- Evans GT (1999) The role of local models and data sets in the Joint Global Ocean Flux Study. Deep-Sea Res I 46: 1369–1389
- Evans GT (2003) Defining misfit between biogeochemical models and data sets. J Mar Syst 40:49-54
- Fagerstrom T (1987) On theory, data and mathematics in ecology. Oikos 50:258-261
- Fasham MJR Ducklow HW, McKelvie SM (1990) A nitrogenbased model of plankton dynamics in the oceanic mixed layer. J Mar Res 48:591–639
- Fennel K, Losch M, Schroter J, Wenzel M (2001) Testing a marine ecosystem model: sensitivity analysis and parameter optimization. J Mar Syst 28:45–63
- Fennel W, Neumann T (2001) Coupling biology and oceanography in models. Ambio 30:232–236
- Flynn KJ (2003) Modeling multi-nutrient interactions in phytoplankton; balancing simplicity and realism. Prog Oceanogr 56:249–279
- Franks PJS (1995) Coupled physical-biological models in oceanography. Rev Geophys 33:1177–1187
- Franks PJS (2002) NPZ models of plankton dynamics: their construction, coupling to physics, and application. J Oceanogr 58:379–387

- Frost BW, Kishi MJ (1999) Ecosystem dynamics in the eastern and western gyres of the Subarctic Pacific—a review of lower trophic level modeling. Prog Oceanogr 43:317–333
- Geider RJ, MacIntyre HL, Kana TM (1998) A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. Limnol Oceanogr 43:679–694
- Gregoire M, Brasseur P, Lermusiaux P (2003) The use of data assimilation in coupled hydrodynamic, ecological and biogeo-chemical models of the ocean. 33rd International Liège Colloquium on Ocean Dynamics. J Mar Syst 40–41: 1–3
- Hamilton DP, Schladow SG (1997) Prediction of water quality in lakes and reservoirs. 1. Model description. Ecol Model 96:91–110
- Hood RR, Bates NR, Capone DG, Olson DB (2001) Modeling the effect of nitrogen fixation on carbon and nitrogen fluxes at BATS. Deep-Sea Res II 48:1609–1648
- Hurtt GC, Armstrong RA (1999) A pelagic ecosystem model calibrated with BATS and OWSI data. Deep-Sea Res II 46: 27–61
- Ianson D, Allen SE (2002) A two-dimensional nitrogen and carbon flux model in a coastal upwelling region. Global Biogeochem Cycles 16: Art. no. 1011
- Jorgensen SE (1994) Fundamentals of ecological modeling. Elsevier Science, Amsterdam
- Jorgensen SE (1997) Integration of ecosystem theories: a pattern. Kluwer, Dordrecht
- Jorgensen SE (1999) State-of-the-art of ecological modeling with emphasis on development of structural dynamic models. Ecol Model 120:75–96
- Kantha LH, Clayson CA (2000) Small scale processes in geophysical fluid flows. Academic Press, San Diego
- Kawamiya M (2002) Numerical model approaches to address recent problems on pelagic ecosystems. J Oceanogr 58: 365–378
- Klepper O (1997) Multivariate aspects of model uncertainty analysis: tools for sensitivity analysis and calibration. Ecol Model 101:1–13
- Konikow LF, Bredehoeft JD (1992) Groundwater models cannot be validated. Adv Water Resour 15:75–83
- Lancelot C, Hannon E, Becquevort S, Veth C, De Baar HJW (2000) Modeling phytoplankton blooms and carbon export production in the Southern Ocean: dominant controls by light and iron in the Atlantic sector in austral spring 1992. Deep-Sea Res I 47:1621–1662
- Le Pape O, Menesguen A (1997) Hydrodynamic prevention of eutrophication in the Bay of Brest (France), a modeling approach. J Mar Syst 12:171–186
- Matear RJ (1995) Parameter optimization and analysis of

- ecosystem models using simulated annealing—a case study at station-P. J Mar Res 53:571–607
- Mayer DG, Butler DG (1993) Statistical validation. Ecol Model 68:21-32
- Moore JK, Doney SC, Kleypas JA, Glover DM, Fung IY (2002) An intermediate complexity marine ecosystem model for the global domain. Deep-Sea Res II 49:403–462
- Omlin M, Brun R, Reichert P (2001) Biogeochemical model of Lake Zurich: sensitivity, identifiability and uncertainty analysis. Ecol Model 141:105–123
- Oreskes N, Shrader-Frechette K, Belitz K (1994) Verification, validation, and confirmation of numerical models in the earth-sciences. Science 263:641–646
- Patten BC (1997) Synthesis of chaos and sustainability in a nonstationary linear dynamic model of the American black bear (*Ursus americanus pallas*) in the Adirondack Mountains of New York. Ecol Model 100:11–42
- Power M (1993) The predictive validation of ecological and environmental models. Ecol Model 68:33-50
- Reckhow KH, Chapra SC (1999) Modeling excessive nutrient loading in the environment. Environ Pollut 100:197–207
- Rykiel EJ (1996) Testing ecological models: the meaning of validation. Ecol Model 90:229–244
- Schlitzer R (2002) Carbon export fluxes in the Southern Ocean: results from inverse modeling and comparison with satellite-based estimates. Deep-Sea Res II 49: 1623–1644
- Spitz YH, Moisan JR, Abbott MR, Richman JG (1998) Data assimilation and a pelagic ecosystem model: parameterization using time series observations. J Mar Syst 16: 51-68
- Starfield AM, Bleloch AL (1986) Building models for conservation and wildlife management. Macmillan, New York
- Steele JH, Henderson EW (1992) The role of predation in plankton models. J Plankton Res 14:157–172
- Touratier F, Field JG, Moloney CL (2001) A stoichiometric model relating growth substrate quality (C:N:P ratios) to N:P ratios in the products of heterotrophic release and excretion. Ecol Model 139:265–291
- Tusseau-Vuillemin MH, Mortier L, Herbaut C (1998) Modeling nitrate fluxes in an open coastal environment (Gulf of Lions): transport versus biogeochemical processes. J Geophys Res C 103:7693–7708
- Vallino JJ (2000) Improving marine ecosystem models: use of data assimilation and mesocosm experiments. J Mar Res 58:117–164
- Vallino JJ (2003) Modeling microbial consortiums as distributed metabolic networks. Biol Bull (Woods Hole) 204: 174–179

Appendix 1. Aquatic biogeochemical modeling studies used for present analysis

- Aksnes DL, Ulvestad KB, Balino BM, Berntsen J, Egee JK, Svendsen E (1995) Ecological modeling in coastal waters–towards predictive physical-chemical-biological simulation models. Ophelia 41: 5–36
- Allen JI (1997) A modeling study of ecosystem dynamics and nutrient cycling in the Humber plume, UK. J Sea Res 38:333-359
- Allen JI, Blackford JC, Radford PJ (1998) An 1-D vertically resolved modeling study of the ecosystem dynamics of the middle and southern Adriatic Sea. J Mar Syst 18:265–286
- Allen JI, Howland RMH, Bloomer N, Uncles RJ (1998) Simulating the spring phytoplankton bloom in the Humber plume, UK. Mar Pollut Bull 37:295–305
- Allen JI, Blackford J, Holt J, Proctor R, Ashworth M, Siddorn J (2001) A highly spatially resolved ecosystem model for the North West European Continental Shelf. Sarsia 86:423–440
- Anderson TR, Williams PJL (1998) Modeling the seasonal cycle of dissolved organic carbon at station E-1 in the English Channel. Estuar Coast Shelf Sci 46:93–109
- Arhonditsis G, Tsirtsis G, Karydis M (2000) Quantification of the effects of non-point sources to coastal marine eutrophication: applications to a semi-enclosed gulf in the Mediterranean Sea. Ecol Model 129:209–227
- Arhonditsis G, Tsirtsis G, Karydis M (2002) The effects of episodic rainfall events to the dynamics of coastal marine ecosystems: applications to a semi-enclosed gulf in the Mediterranean Sea. J Mar Syst 35:183–205
- Asaeda T, Van Bon T (1997) Modeling the effects of macrophytes on algal blooming in eutrophic shallow lakes. Ecol Model 104: 261–287
- Asaeda T, Trung VK, Manatunge J, Van Bon T (2001) Modeling

Appendix 1 (continued)

- macrophyte-nutrient-phytoplankton interactions in shallow eutrophic lakes and the evaluation of environmental impacts. Ecol Eng 16:341–357
- Barciela RM, Garcia E, Fernandez E (1999) Modeling primary production in a coastal embayment affected by upwelling using dynamic ecosystem models and artificial neural networks. Ecol Model 120:199–211
- Baretta-Bekker JG, Baretta JW, Ebenhoh W (1997) Microbial dynamics in the marine ecosystem model ERSEM II with decoupled carbon assimilation and nutrient uptake. J Sea Res 38:195–211
- Baretta-Bekker JG, Baretta JW, Hansen AS, Riemann B (1998) An improved model of carbon and nutrient dynamics in the microbial food web in marine enclosures. Aquat Microb Ecol 14:91–108
- Bartholow J, Hanna RB, Saito L, Lieberman D, Horn M (2001) Simulated limnological effects of the Shasta Lake temperature control device. Environ Manag 27:609–626
- Benoist AP, Brinkman AG, van Diepenbeek PMJA, Waals JMJ (1998) Bekwaam, a model fit for reservoir design and management. Water Sci Technol 37:269–276
- Bierman VJ, Hinz SC, Zhu DW, Wiseman WJ, Rabalais NN, Turner RE (1994) A preliminary mass-balance model of primary productivity and dissolved-oxygen in the Mississippi river plume inner gulf shelf region. Estuaries 17:886–899
- Blackford JC (1997) An analysis of benthic biological dynamics in a North Sea ecosystem model. J Sea Res 38:213–230
- Blackford JC (2002) The influence of microphytobenthos on the Northern Adriatic ecosystem: a modeling study. Estuar Coast Shelf Sci 55:109–123
- Blackford JC, Burkill PH (2002) Planktonic community structure and carbon cycling in the Arabian Sea as a result of monsoonal forcing: the application of a generic model. J Mar Syst 36:239–267
- Bocci M, Coffaro G, Bendoricchio G (1997) Modeling biomass and nutrient dynamics in eelgrass (*Zostera marina* L.): applications to the Lagoon of Venice (Italy) and Oresund (Denmark). Ecol Model 102:67–80
- Bonnet MP, Poulin M (2002) Numerical modeling of the planktonic succession in a nutrient-rich reservoir: environmental and physiological factors leading to *Microcystis aeruginosa* dominance. Ecol Model 156:93–112
- Broekhuizen N, Heath MR, Hay SJ, Gurney WSC (1995) Modeling the dynamics of the North-Sea mesozooplankton. Neth J Sea Res 33:381–406
- Broekhuizen N, Hadfield M, Taylor AH (1998) Seasonal photoadaptation and diatom dynamics in temperate waters. Mar Ecol Prog Ser 175:227–239
- Brostrom G, Drange H (2000) On the mathematical formulation and parameter estimation of the Norwegian Sea plankton system. Sarsia 85:211–225
- Cerco CF (1995) Simulation of long-term trends in Chesapeake Bay. J Environ Eng-ASCE 121:298–310 Cerco CF, Cole T (1993) 3-dimensional eutrophication model of
- Chesapeake Bay. J Environ Eng-ASCE 119:1006–1025 Cerco CF, Meyers M (2000) Tributary refinements to Chesapeake
- Cerco CF, Meyers M (2000) Tributary refinements to Chesapeake Bay model. J Environ Eng-ASCE 126:164–174
- Cerco CF, Seitzinger SP (1997) Measured and modeled effects of benthic algae on eutrophication in Indian River Rehoboth Bay, Delaware. Estuaries 20:231–248
- Chau KW, Jin HH (2002) Two-layered, 2D unsteady eutrophication model in boundary-fitted coordinate system. Mar Pollut Bull 45: 300–310
- Chen CS, Wiesenburg DA, Xie LS (1997) Influences of river discharge on biological production in the inner shelf: a coupled biological and physical model of the Louisiana–Texas shelf. J Mar Res 55:293–320
- Chen CS, Ji RB, Schwab DJ, Beletsky D, Fahnenstiel GL and 9 others (2002) A model study of the coupled biological and physical dynamics in Lake Michigan. Ecol Model 152:145–168
- Christian JR, Verschell MA, Murtugudde R, Busalacchi AJ, McClain CR (2002) Biogeochemical modeling of the tropical Pacific Ocean. I. Seasonal and interannual variability. Deep-Sea Res II 49: 509–543
- Coffaro G, Sfriso A (1997) Simulation model of *Ulva rigida* growth in shallow water of the Lagoon of Venice. Ecol Model 102:55–66
- Connolly JP, Coffin RB (1995) Model of carbon cycling in planktonic food webs. J Environ Eng-ASCE 121:682–690
- Crispi G, Crise A, Solidoro C (2002) Coupled Mediterranean eco-

- model of the phosphorus and nitrogen cycles. J Mar Syst 33–34: 497–521
- Dadou I, Lamy F, Rabouille C, Ruiz-Pino D, Andersen V, Bianchi M, Garcon V (2001) An integrated biological pump model from the euphotic zone to the sediment: a 1-D application in the northeast tropical Atlantic. Deep-Sea Res II 48:2345–2381
- Delfuria L, Rizzoli A, Arditi R (1995) Lakemaker—a general objectoriented software tool for modeling the eutrophication process in lakes. Environ Software 10:43–64
- Doney SC, Glover DM, Najjar RG (1996) A new coupled, one-dimensional biological-physical model for the upper ocean: applications to the JGOFS Bermuda Atlantic time-series study (BATS) site. Deep-Sea Res II 43:591–624
- Dugdale RC, Barber RT, Chai F, Peng TH, Wilkerson FP (2002) Onedimensional ecosystem model of the equatorial Pacific upwelling system. Part II. Sensitivity analysis and comparison with JGOFS EqPac data. Deep-Sea Res II 49:2747–2768
- Ebenhoh W, Kohlmeier C, Radford PJ (1995) The benthic biological submodel in the European-Regional Seas Ecosystem Model. Neth J Sea Res 33:423–452
- Ebenhoh W, Baretta-Bekker JG, Baretta JW (1997) The primary production module in the marine ecosystem model ERSEM II, with emphasis on the light forcing. J Sea Res 38:173–193
- Eldridge PM, Sieracki ME (1993) Biological and hydrodynamic regulation of the microbial food-web in a periodically mixed estuary. Limnol Oceanogr 38:1666–1679
- Evans GT (1999) The role of local models and data sets in the Joint Global Ocean Flux Study. Deep-Sea Res I 46:1369–1389 Everbecq E, Gosselain V, Viroux L, Descy JP (2001) Potamon: a
- Everbecq E, Gosselain V, Viroux L, Descy JP (2001) Potamon: a dynamic model for predicting phytoplankton composition and biomass in lowland rivers. Water Res 35:901–912
- Fasham MJR, Evans GT (1995) The use of optimization techniques to model marine ecosystem dynamics at the JGOFS station at 47° N–20° W. Philos Trans R Soc Lond B 348:203–209
- Fasham MJR Ducklow HW, McKelvie SM (1990) A nitrogen-based model of plankton dynamics in the oceanic mixed layer. J Mar Res 48:591–639
- Fasham MJR, Sarmiento JL, Slater RD, Ducklow HW, Williams R (1993) Ecosystem behavior at Bermuda station-S and Ocean weather station India—a general-circulation model and observational analysis. Global Biogeochem Cycles 7:379–415
- Fennel K, Losch M, Schroter J, Wenzel M (2001) Testing a marine ecosystem model: sensitivity analysis and parameter optimization. J Mar Syst 28:45–63
- Fennel K, Spitz YH, Letelier RM, Abbott MR, Karl DM (2002) A deterministic model for N-2 fixation at stn. ALOHA in the subtropical North Pacific Ocean. Deep-Sea Res II 49:149–174
- Flindt MR, Kamp-Nielsen L (1997) Modeling of an estuarine eutrophication gradient. Ecol Model 102:143-153
- Franks PJS, Chen CS (2001) A 3-D prognostic numerical model study of the Georges bank ecosystem. Part II. Biological-physical model. Deep-Sea Res II 48:457–482
- Friedrichs MAM (2002) Assimilation of JGOFS EqPac and SeaWiFS data into a marine ecosystem model of the central equatorial Pacific Ocean. Deep-Sea Res II 49:289–319
- Gabric A, Murray N, Stone L, Kohl M (1993) Modeling the production of dimethylsulfide during a phytoplankton bloom. J Geophys Res C 98:22805–22816
- Gao HW, Feng SZ, Guan YP (1998) Modeling annual cycles of primary production in different regions of the Bohai Sea. Fish Oceanogr 7:258–264
- Garnier J, Billen G, Hannon E, Fonbonne S, Videnina Y, Soulie M (2002) Modeling the transfer and retention of nutrients in the drainage network of the Danube River. Estuar Coast Shelf Sci 54: 285–308
- Gin KYH, Guo JH, Cheong HF (1998) A size-based ecosystem model for pelagic waters. Ecol Model 112:53–72
- Guillaud JF, Andrieux F, Menesguen A (2000) Biogeochemical modeling in the Bay of Seine (France): an improvement by introducing phosphorus in nutrient cycles. J Mar Syst 25:369–386
- Hadfield MG, Sharples J (1996) Modeling mixed layer depth and plankton biomass off the west coast of South Island, New Zealand. J Mar Syst 8:1–29
- Hassan H, Hanaki K, Matsuo T (1998) A modeling approach to simulate impact of climate change in lake water quality: phytoplankton growth rate assessment. Water Sci Technol 37:177–185

Appendix 1 (continued)

- Haupt OJ, Wolf U, Von Bodungen B (1999) Modeling the pelagic nitrogen cycle and vertical particle flux in the Norwegian Sea. J Mar Syst 19:173–199
- Henderson EW, Steele JH (1995) Comparing models and observations of shelf plankton. J Plankton Res 17:1679–1692
- Hoch T, Garreau P (1998) Phytoplankton dynamics in the English Channel: a simplified 3-dimensional approach. J Mar Syst 16: 133–150
- Hood RR, Bates NR, Capone DG, Olson DB (2001) Modeling the effect of nitrogen fixation on carbon and nitrogen fluxes at BATS. Deep-Sea Res II 48:1609–1648
- Hu WP, Salomonsen J, Xu FL, Pu PM (1998) A model for the effects of water hyacinths on water quality in an experiment of physicobiological engineering in Lake Taihu, China. Ecol Model 107: 171–188
- Humborg C, Fennel K, Pastuszak M, Fennel W (2000) A box model approach for a long-term assessment of estuarine eutrophication, Szczecin Lagoon, southern Baltic. J Mar Syst 25:387–403
- Hurtt GC, Armstrong RA (1999) A pelagic ecosystem model calibrated with BATS and OWSI data. Deep-Sea Res I 46:27–61
- James RT, Bierman VJ (1995) A preliminary modeling analysis of water-quality in Lake Okeechobee, Florida—calibration results. Water Res 29:2755–2766
- James RT, Martin J, Wool T, Wang PF (1997) A sediment resuspension and water quality model of Lake Okeechobee. J Am Water Resour Assoc 33:661–680
- Jamu DM, Piedrahita RH (2002) An organic matter and nitrogen dynamics model for the ecological analysis of integrated aquaculture/agriculture systems. II. Model evaluation and application. Environ Model Software 17:583–592
- Janse JH, Aldenberg T, Kramer PRG (1992) A mathematical-model of the phosphorus cycle in Lake Loosdrecht and simulation of additional measures. Hydrobiologia 233:119–136
- Janse JH, Van Donk E, Aldenberg T (1998) A model study on the stability of the macrophyte-dominated state as affected by biological factors. Water Res 32:2696–2706
- Jayaweera M, Asaeda T (1995) Impacts of environmental scenarios on chlorophyll a in the management of shallow, eutrophic lakes following biomanipulation: an application of a numerical model. Ecol Eng 5:445–468
- Jewell PW (1995) A simple surface-water biogeochemical model. 2. Simulation of selected lacustrine and marine settings. Water Resour Res 31: 2059-2070
- Jimenez-Montealegre R, Verreth J, Steenbergen K, Moed J, Machiels M (1995) A dynamic simulation-model for the blooming of Oscillatoria-Agardhii in a monomictic lake. Ecol Model 78:17-24
- Jimenez-Montealegre R, Verdegem MCJ, van Dam A, Verreth JAJ (2002) Conceptualization and validation of a dynamic model for the simulation of nitrogen transformations and fluxes in fish ponds. Ecol Model 147: 123-152
- Karagounis I, Trosch J, Zamboni F (1993) A coupled physicalbiochemical model for forecasting water-quality—application to the northern basin of lake Lugano. Aquat Sci 55:87–102
- Kellershohn DA, Tsanis IK (1999) 3D eutrophication modeling of Hamilton Harbour: analysis of remedial options J Gt Lakes Res 25: 3–25
- Kiirikki M, Haapamaki J, Koponen J, Ruuskanen A, Sarkkula J (1998) Linking the growth of filamentous algae to the 3D-ecohydrodynamic model of the Gulf of Finland. Environ Model Software 13:503–509
- Kuhn W, Radach G (1997) A one-dimensional physical-biological model study of the pelagic nitrogen cycling during the spring bloom in the northern North Sea (FLEX '76). J Mar Res 55: 687-734
- Kuusisto M, Koponen J, Sarkkula J (1998) Modeled phytoplankton dynamics in the Gulf of Finland. Environ Model Software 13: 461–470
- Lancelot C, Hannon E, Becquevort S, Veth C, De Baar HJW (2000) Modeling phytoplankton blooms and carbon export production in the Southern Ocean: dominant controls by light and iron in the Atlantic sector in Austral spring 1992. Deep-Sea Res I 47: 1621–1662
- Lancelot C, Staneva J, Van Eeckhout D, Beckers JM, Stanev E (2002) Modeling the Danube-influenced north-western continental shelf of the Black Sea. II. Ecosystem response to changes in nutrient delivery by the Danube River after its damming in 1972. Estuar

- Coast Shelf Sci 54:473-499
- Lee JHW, Arega F (1999) Eutrophication dynamics of Tolo Harbour, Hong Kong. Mar Pollut Bull 39:187–192
- Lenhart HJ, Radach G, Ruardij P (1997) The effects of river input on the ecosystem dynamics in the continental coastal zone of the North Sea using ERSEM. J Sea Res 38:249–274
- Le Pape O, Menesguen A (1997) Hydrodynamic prevention of eutrophication in the Bay of Brest (France), a modeling approach. J Mar Syst 12:171–186
- Le Pape O, Jean F, Menesguen A (1999) Pelagic and benthic trophic chain coupling in a semi-enclosed coastal system, the Bay of Brest (France): a modeling approach. Mar Ecol Prog Ser 189:135–147
- Levy M, Memery L, Andre JM (1998) Simulation of primary production and export fluxes in the northwestern Mediterranean Sea. J Mar Res 56:197–238
- Lewis DM, Elliot JA, Lambert MF, Reynolds CS (2002) The simulation of an Australian reservoir using a phytoplankton community model: PROTECH. Ecol Model 150:107–116
- Liu KK, Chao SY, Shaw Part, Gong GC, Chen CC, Tang TY (2002) Monsoon-forced chlorophyll distribution and primary production in the South China Sea: observations and a numerical study. Deep-Sea Res I 49:1387–1412
- Lonin ŜA, Tuchkovenko YS (2001) Water quality modeling for the ecosystem of the Cienaga de Tesca coastal lagoon. Ecol Model 144:279–293
- Marmefelt E, Arheimer B, Langner J (1999) An integrated biogeochemical model system for the Baltic Sea. Hydrobiologia 393: 45–56
- Marra J, Ho C (1993) Initiation of the spring bloom in the Northeast Atlantic (47° N, 20° W) a numerical simulation. Deep-Sea Res II 40:55-73
- McGillicuddy DJ, McCarthy JJ, Robinson AR (1995) Coupled physical and biological modeling in the North Atlantic. 1. Model formulation and one-dimensional bloom processes. Deep-Sea Res I 42:1313–1357
- McClain CR, Arrigo K, Tai KS, Turk D (1996) Observations and simulations of physical and biological processes at ocean weather station P, 1951–1980. J Geophys Res C 101:3697–3713
- Menesguen A, Guillaud JF, Aminot A, Hoch T (1995) Modeling the eutrophication process in a river plume—the Seine case study (France). Ophelia 42:205–225
- Menshutkin VV, Astrakhantsev GP, Yegorova NB, Rukhovets LA, Simo TL, Petrova NA (1998) Mathematical modeling of the evolution and current conditions of the Ladoga Lake ecosystem. Ecol Model 107:1–24
- Mesple F, Casellas C, Trousellier M, Bontoux J (1995) Some difficulties in modeling chlorophyll a evolution in a high rate algal pond ecosystem. Ecol Model 78:25–36
- Moll A (1998) Regional distribution of primary production in the North Sea simulated by a three-dimensional model. J Mar Syst 16: 151–170
- Moore JK, Doney SC, Kleypas JA, Glover DM, Fung IY (2002) An intermediate complexity marine ecosystem model for the global domain. Deep-Sea Res II 49:403–462
- Muhammetoglu AB, Soyupak S (2000) A three-dimensional water quality-macrophyte interaction model for shallow lakes. Ecol Model 133:161–180
- Murray AG, Parslow JS (1999) Modeling of nutrient impacts in Port Phillip Bay—a semi-enclosed marine Australian ecosystem. Mar Freshw Res 50:597–611
- Napolitano E, Oguz T, Malanotte-Rizzoli P, Yilmaz A, Sansone E (2000) Simulations of biological production in the Rhodes and Ionian basins of the eastern Mediterranean. J Mar Syst 24:277–298
- Oguz T, Ducklow H, Malanotte-Rizzoli P, Tugrul S, Nezlin NP, Unluata U (1996) Simulation of annual plankton productivity cycle in the Black Sea by a one-dimensional physical-biological model. J Geophys Res C 101:16585–16599
- Oguz T, Ducklow HW, Purcell JE, Malanotte-Rizzoli P (2001) Modeling the response of top-down control exerted by gelatinous carnivores on the Black Sea pelagic food web. J Geophys Res C 106: 4543–4564
- Olivieri RA, Chavez FP (2000) A model of plankton dynamics for the coastal upwelling system of Monterey Bay, California. Deep-Sea Res II 47:1077–1106
- Omlin M, Reichert P, Forster R (2001) Biogeochemical model of Lake Zurich: model equations and results. Ecol Model 141:77-103

Appendix 1 (continued)

- Oschlies A, Koeve W, Garcon V (2000) An eddy-permitting coupled physical-biological model of the North Atlantic. 2. Ecosystem dynamics and comparison with satellite and JGOFS local studies data. Global Biogeochem Cycles 14:499–523
- Park SS, Lee YS (2002) A water quality modeling study of the Nakdong River, Korea. Ecol Model 152:65–75
- Patsch J, Radach G (1997) Long-term simulation of the eutrophication of the North Sea: temporal development of nutrients, chlorophyll and primary production in comparison to observations. J Sea Res 38:275–310
- Patsch J, Kuhn W, Radach G, Casiano JMS, Davila MG, Neuer S, Freudenthal T, Llinas O (2002) Interannual variability of carbon fluxes at the North Atlantic Station ESTOC. Deep-Sea Res II 49: 253–288
- Pei HP, Ma JY (2002) Study on the algal dynamic model for West Lake, Hangzhou. Ecol Model 148:67–77
- Petihakis G, Triantafyllou G, Koutsoubas D, Allen I, Dounas C (1999) Modeling the annual cycles of nutrients and phytoplankton in a Mediterranean lagoon (Gialova, Greece). Mar Environ Res 48: 37–58
- Pondaven P, Fravalo C, Ruiz-Pino D, Treguer P, Queguiner B, Jeandel C (1998) Modeling the silica pump in the permanently open ocean zone of the Southern Ocean. J Mar Syst 17:587–619
- Priyantha DGN, Asaeda T, Saitoh S, Gotoh K (1997) Modeling effects of curtain method on algal blooming in reservoirs. Ecol Model 98: 89–104
- Prunet P, Minster JF, Ruiz-Pino D, Dadou I (1996) Assimilation of surface data in a one-dimensional physical-biogeochemical model of the surface ocean. 1. Method and preliminary results. Global Biogeochem Cycles 10:111–138
- Prunet P, Minster JF, Echevin V, Dadou I (1996) Assimilation of surface data in a one-dimensional physical-biogeochemical model of the surface ocean. 2. Adjusting a simple trophic model to chlorophyll, temperature, nitrate, and pCO $_2$ data. Global Biogeochem Cycles 10:139–158
- Radach G, Lenhart HJ (1995) Nutrient dynamics in the North Sea fluxes and budgets in the water column derived from ERSEM. Neth J Sea Res 33:301–335
- Ruardij P, Vanraaphorst W (1995) Benthic nutrient regeneration in the ERSEM ecosystem model of the North Sea. Neth J Sea Res 33: 453–483
- Ruardij P, Van Haren H, Ridderinkhof H (1997) The impact of thermal stratification on phytoplankton and nutrient dynamics in shelf seas: a model study. J Sea Res 38:311–331
- Sagehashi M, Sakoda A, Suzuki M (2000) A predictive model of long-term stability after biomanipulation of shallow lakes. Water Res 34:4014–4028
- Sagehashi M, Sakoda A, Suzuki M (2001) A mathematical model of a shallow and eutrophic lake (the Keszthely Basin, Lake Balaton) and simulation of restorative manipulations. Water Res 35: 1675–1686
- Salencon MJ, Thebault JM (1996) Simulation model of a mesotrophic reservoir (Lac de Pareloup, France): MELODIA, an ecosystem reservoir management model. Ecol Model 84:163–187
- Savchuk OP (2002) Nutrient biogeochemical cycles in the Gulf of Riga: scaling up field studies with a mathematical model. J Mar Syst 32:253–280
- Schartau M, Oschlies A, Willebrand J (2001) Parameter estimates of a zero-dimensional ecosystem model applying the adjoint method. Deep-Sea Res II 48:1769–1800
- Schladow SG, Hamilton DP (1997) Prediction of water quality in lakes and reservoirs. 2. Model calibration, sensitivity analysis and application. Ecol Model 96:111–123
- Sharples J, Tett P (1994) Modeling the effect of physical variability on the midwater chlorophyll maximum. J Mar Res 52:219–238
- Shen YM, Zheng YH, Komatsu T, Kohashi N (2002) A three-dimensional numerical model of hydrodynamics and water quality in Hakata Bay. Ocean Eng 29:461–473
- Sin Y, Wetzel RL (2002) Ecosystem modeling analysis of size-structured phytoplankton dynamics in the York River estuary, Virginia (USA). II. Use of a plankton ecosystem model for investigating controlling factors on phytoplankton and nutrient dynamics. Mar Ecol Prog Ser 228:91–101
- Six KD, Maier-Reimer E (1996) Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model. Global

- Biogeochem Cycles 10:559-583
- Skliris N, Elkalay K, Goffart A, Frangoulis C, Hecq JH (2001) Onedimensional modeling of the plankton ecosystem of the northwestern Corsican coastal area in relation to meteorological constraints. J Mar Syst 27:337–362
- Soiland H, Skogen MD (2000) Validation of a three-dimensional biophysical model using nutrient observations in the North Sea. ICES J Mar Sci 57:816–823
- Soyupak S, Mukhallalati L, Yemisen D, Bayar A, Yurteri C (1997) Evaluation of eutrophication control strategies for the Keban Dam reservoir. Ecol Model 97:99–110
- Spitz YH, Moisan JR, Abbott MR, Richman JG (1998) Data assimilation and a pelagic ecosystem model: parameterization using time series observations. J Mar Syst 16:51–68
- Spitz YH, Moisan JR, Abbott MR (2001) Configuring an ecosystem model using data from the Bermuda Atlantic Time Series (BATS). Deep-Sea Res II 48:1733–1768
- Stramska M, Dickey TD (1994) Modeling phytoplankton dynamics in the Northeast Atlantic during the initiation of the spring bloom. J Geophys Res C 99:10241–10253
- Suzuki M, Sagehashi M, Sakoda A (2000) Modeling the structural dynamics of a shallow and eutrophic water ecosystem based on mesocosm observations. Ecol Model 128:221–243
- Taguchi K, Nakata K (1998) Analysis of water quality in Lake Hamana using a coupled physical and biochemical model. J Mar Syst 16:107–132
- Tamsalu R, Ennet P (1995) Ecosystem modeling in the Gulf of Finland. 2. The aquatic ecosystem model FINEST. Estuar Coast Shelf Sci 41:429–458
- Taylor AH, Harbour DS, Harris RP, Burkill PH, Edwards ES (1993) Seasonal succession in the pelagic ecosystem of the North-Atlantic and the utilization of nitrogen. J Plankton Res 15:875–891
- Tett P, Walne A (1995) Observations and simulations of hydrography, nutrients and plankton in the southern North-Sea. Ophelia 42: 371–416
- Thebault JM, Salencon MJ (1993) Simulation model of a mesotrophic reservoir (Lac-de-Pareloup, France) biological model. Ecol Model 65:1–30
- Tian RC, Vezina AF, Legendre L, Ingram RG and 7 others (2000) Effects of pelagic food-web interactions and nutrient remineralization on the biogeochemical cycling of carbon: a modeling approach. Deep-Sea Res II 47:637–662
- Tian RC, Vezina AF, Starr M, Saucier F (2001) Seasonal dynamics of coastal ecosystems and export production at high latitudes: a modeling study. Limnol Oceanogr 46:1845–1859
- Triantafyllou G, Petihakis G, Dounas C, Theodorou A (2001) Assessing marine ecosystem response to nutrient inputs. Mar Pollut Bull 43:175–186
- Tufford DL, McKellar HN (1999) Spatial and temporal hydrodynamic and water quality modeling analysis of a large reservoir on the South Carolina (USA) coastal plain. Ecol Model 114:137–173
- Tusseau MH, Lancelot C, Martin JM, Tassin B (1997) One-D coupled physical-biological model of the northwestern Mediterranean Sea. Deep-Sea Res II 44:851–880
- Varela RA, Cruzado A, Tintore J, Ladona EG (1992) Modeling the deep-chlorophyll maximum—a coupled physical-biological approach. J Mar Res 50:441–463
- Varela RA, Cruzado A, Gabaldon JE (1995) Modeling primary production in the North-Sea using the European-regional seasecosystem-model. Neth J Sea Res 33:337–361
- Vichi M, Pinardi N, Zavatarelli M, Matteucci G, Marcaccio M, Bergamini MC, Frascari F (1998) One-dimensional ecosystem model tests in the Po Prodelta area (Northern Adriatic Sea). Environ Model Software 13:471–481
- Wang PF, Martin J, Morrison G (1999) Water quality and eutrophication in Tampa Bay, Florida. Estuar Coast Shelf Sci 49:1-20
- Wirtz KW, Echkardt B (1996) Effective variables in ecosystem models with an application to phytoplankton succession. Ecol Model 92: 33–53
- Yanagi T, Inoue K, Montani S, Yamada M (1997) Ecological modeling as a tool for coastal zone management in Dokai Bay, Japan. J Mar Syst 13:123–136
- Zavatarelli M, Baretta JW, Baretta-Bekker JG, Pinardi N (2000) The dynamics of the Adriatic Sea ecosystem: an idealized model study. Deep-Sea Res I 47:937–970