Bayesian calibration of mechanistic aquatic biogeochemical models and benefits for environmental management

George B. Arhonditsis [a,*], Dimitra Papantou [b], Weitao Zhang [a], Gurbir Perhar [a], Evangelia Massos [a], Molu Shi [a]

[a] *Department of Physical & Environmental Sciences, University of Toronto, Toronto, ON, Canada M1C 1A4*
[b] *Department of Environmental Engineering, Technical University of Crete, Chania, Crete, Greece, 73100*

Abstract

Aquatic biogeochemical models have been an indispensable tool for addressing pressing environmental issues, e.g., understanding oceanic response to climate change, elucidation of the interplay between plankton dynamics and atmospheric CO_2 levels, and examination of alternative management schemes for eutrophication control. Their ability to form the scientific basis for environmental management decisions can be undermined by the underlying structural and parametric uncertainty. In this study, we outline how we can attain realistic predictive links between management actions and ecosystem response through a probabilistic framework that accommodates rigorous uncertainty analysis of a variety of error sources, i.e., measurement error, parameter uncertainty, discrepancy between model and natural system. Because model uncertainty analysis essentially aims to quantify the joint probability distribution of model parameters and to make inference about this distribution, we believe that the iterative nature of Bayes' Theorem is a logical means to incorporate existing knowledge and update the joint distribution as new information becomes available. The statistical methodology begins with the characterization of parameter uncertainty in the form of probability distributions, then water quality data are used to update the distributions, and yield posterior parameter estimates along with predictive uncertainty bounds. Our illustration is based on a six state variable (nitrate, ammonium, dissolved organic nitrogen, phytoplankton, zooplankton, and bacteria) ecological model developed for gaining insight into the mechanisms that drive plankton dynamics in a coastal embayment; the Gulf of Gera, Island of Lesvos, Greece. The lack of analytical expressions for the posterior parameter distributions was overcome using Markov chain Monte Carlo simulations; a convenient way to obtain representative samples of parameter values. The Bayesian calibration resulted in realistic reproduction of the key temporal patterns of the system, offered insights into the degree of information the data contain about model inputs, and also allowed the quantification of the dependence structure among the parameter estimates. Finally, our study uses two synthetic datasets to examine the ability of the updated model to provide estimates of predictive uncertainty for water quality variables of environmental management interest.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Bayesian calibration; Mechanistic aquatic biogeochemical models; Coastal embayments; Environmental management; Uncertainty analysis; Bayesian updating; Plankton dynamics

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

1. Introduction

A recent examination of the citation patterns of 153 aquatic biogeochemical modeling studies provided overwhelming evidence that modeling papers are cited mainly based on the questions being asked or the ecosystem being modeled (Arhonditsis et al., 2006). Models that aim to elucidate oceanic patterns are more highly cited than models developed for addressing local water quality management issues regardless of their methodological features and technical value. Using the citation frequency as a criterion for the impact of an aquatic biogeochemical modeling study, the same analysis identified several influential studies that have received an impressively high number of citations (e.g., Fasham et al., 1990; Fasham et al., 1993; Doney et al., 1996; Six and Maier-Reimer, 1996). Viewed from this perspective, the field of aquatic biogeochemical modeling follows the usual trajectory that most of the fields of knowledge follow, i.e., there are some breakthrough ideas that inspired a great deal of the research that has occurred over the past 15 years, while the rest of the studies represent incremental learning without the capacity to truly stimulate future research. However, this point does invite one to ask what it would take to prime the pump for the new breakthroughs to come along? Recognizing the role of mathematical modeling as a key research tool for understanding aquatic ecosystem dynamics, several review/synthesis papers recently debated the outstanding challenges and future directions of the field. Some researchers advocated the establishment of a systematic methodological protocol along with globally accepted performance criteria (Arhonditsis and Brett, 2004), others highlighted the importance of effectively coupling physical and biogeochemical models (Fennel and Neumann, 2001; Franks, 2002), while others identified the pressing modeling issues, technical/conceptual advances and pinpointed the areas where extra complexity should be incorporated (Doney, 1999; Anderson, 2005).

The latter topic, i.e., the demand for increasing model complexity, has been a controversial issue in the aquatic biogeochemical modeling practice, where the inevitable trade-off among complexity, generality and accuracy entails compromises that reflect different philosophical viewpoints and research priorities (Levins, 1966; Costanza and Sklar, 1985; Anderson, 2005). Recently, there is a trend towards increasing the articulation of aquatic ecosystem models with regards to the representation of specific plankton functional types (e.g., coccolithophorids, diatoms, nitrogen fixers) more closely linked to biogeochemical cycling in aquatic systems (Moore et al., 2002; Gregg et al., 2003), but the rate of this increase has also been argued in the aquatic biogeochemical modeling

literature (Le Quere, 2006; Flynn, 2006; Anderson, 2006). For example, there are views that this increase of complexity needs to be done in a gradual manner along with a "healthy dose of scepticism regarding model outcomes" (Anderson, 2005), whereas others claim that our theoretical understanding of ecosystem functioning is already far behind and the building rate of new parameterizations should be faster (Le Quere, 2006). Another interesting angle of the model complexity issue was illuminated by Flynn (2006) who pointed out that the problem is far more complex than simply the lack of sufficient data to support the more detailed simulations. Namely, the physiological basis of plankton dynamics involves an inconceivably wide array of direct and synergistic effects (trophic functionality, allelopathy, omnivory) that can never be effectively depicted in the models (Flynn, 2006). The same study even questioned our ability to simulate the variability of a single clone of a plankton species under "real world" conditions; thus the most defensible (and convenient) strategy is the description of general patterns that the aggregated models (e.g., nutrient–phytoplankton–zooplankton–detritus) can offer. While not all ecosystem modelers accept such pessimistic views, there is no doubt that the ubiquitous natural complexity imposes insurmountable barriers for attaining parsimonious ecological models (Anderson, 2006).

Implicit in the debate of increasing complexity is the importance of tightly connecting the model development process with uncertainty analysis methods that can accommodate rigorous error analysis (Pappenberger and Beven, 2006). The adoption of a reductionistic description of natural system dynamics that considers higher number of biotic subunits along with the underlying interconnections will inevitably accentuate the disparity between what ideally we want to learn and what can realistically be observed (Beck, 1987). As a result, our ability to set quantitative (or even qualitative) constraints as to what is realistic/behavioural simulation of an ecological structure will be decreased, and the resulting underidentified models will have limited predictive power and learning capacity (Flynn, 2006). The explicit recognition of the uncertainty that underlies both model structures and data also implies that the search for a single set of parameter values (global optimum) that reproduces real world patterns is not a reasonable expectation (Reichert and Omlin, 1997). Rather, the only legitimate approach is the assessment of the likelihood of different input factors (model structures/parameter sets) being acceptable simulators of the natural system, the so-called "model equifinality" (Beven and Binley, 1992). In this context, novel error analysis techniques are an essential advancement for quantifying the uncertainty in model equations

(structural uncertainty) and the effects of input uncertainties (model parameters, initial conditions, forcing functions) on model outputs. However, despite several attempts in the literature to address structural and parametric errors (Hornberger and Spear, 1981; Dilks et al., 1992; Omlin and Reichert, 1999; Brun et al., 2001), uncertainty analysis is not considered an indispensable step in the model development process; aquatic mechanistic modellers are still reluctant to assess the reliability of the critical planning information generated by the models, and the methodological consistency (whether or not the model has been subject to thorough uncertainty analysis and/or predictive/structural validation) of the original modeling papers is not considered a significant criterion for their citation (Arhonditsis and Brett, 2004; Arhonditsis et al., 2006).

The main objective of this study is to introduce a methodological framework that integrates aquatic biogeochemical modeling with Bayesian analysis. Our aim is to show how Bayesian calibration can be used to obtain insight into the degree of information the data contain about model inputs, refine our knowledge of model input parameters, and obtain predictions along with uncertainty bounds for modeled output variables. Because model uncertainty analysis basically aims to quantify the joint probability distribution of model inputs (parameters, forcing functions, and initial conditions) and to make inference about this distribution, we suggest that the iterative nature of Bayes' Theorem is a convenient means to incorporate existing knowledge and update the joint distribution as new information

becomes available. Our illustration is based on a six state variable (nitrate, ammonium, dissolved organic nitrogen, phytoplankton, zooplankton, and bacteria) model developed for gaining insight into the ecological processes that drive plankton dynamics in a coastal embayment; the Gulf of Gera, Island of Lesvos, Greece. Our study also acknowledges the lack of perfect simulators of natural system dynamics and introduces two statistical formulations that can explicitly account for the discrepancy between mathematical models and environmental systems. Finally, we use two synthetic datasets to examine the ability of the updated model to provide estimates of predictive uncertainty for water quality variables of environmental management interest.

2. Methods

2.1. Case study and model description

The case study for the examination of the Bayesian calibration framework was the Gulf of Gera, Island of Lesvos, Greece; a shallow semi-enclosed marine ecosystem that receives significant point and non-point nutrient loads from the surrounding watershed (Arhonditsis et al., 2002a). Based on a trophic classification scheme proposed for the Aegean Sea (Ignatiades et al., 1992), the Gulf of Gera can be characterized as mesotrophic with plankton dynamics driven by two main factors, i.e., the nitrogen availability and the rate of water renewal in the embayment (Arhonditsis et al., 2003a). Both field observations and simulation results
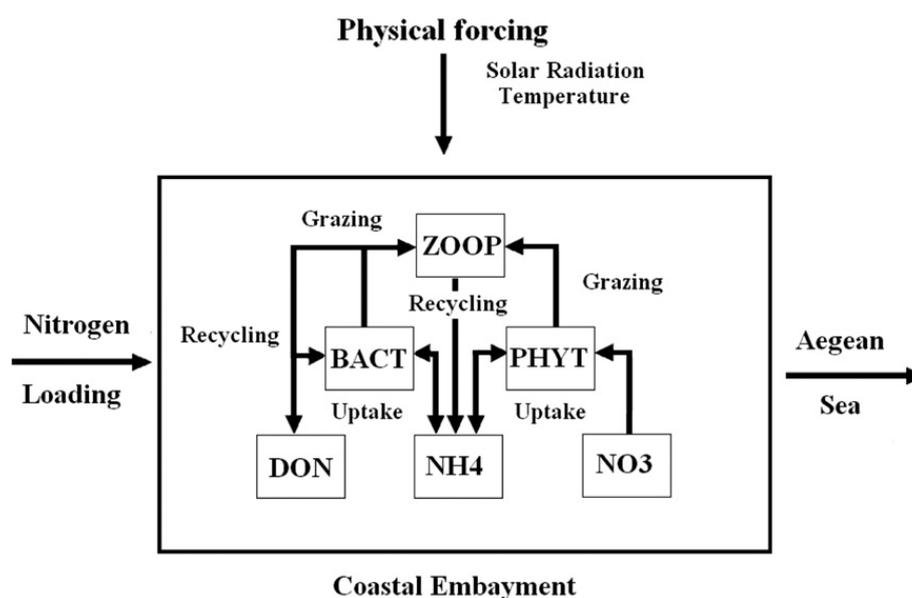

Fig. 1. The eutrophication model used for reproducing the coastal embayment dynamics. Arrows indicate flows of matter through the system. System equations and parameter definitions are provided in Tables 1 and 2.

have provided overwhelming evidence of a bimodal circulation pattern that closely determines the interplay between the abiotic environment and the biotic components of the system (Arhonditsis et al., 2000). Specifically, during the colder period of the year, the ambient temperature and the runoff inflows increase the water density in the shallow embayment and constrain the intrusion of the oligotrophic water masses from the Aegean Sea. The isolation phase of the system usually coincides with the period when the external loading is significantly increased and non-point sources account for about 40–60% of the total nutrient stock (Arhonditsis et al., 2002b). Given the latitude and local climatic conditions, the incoming solar radiation can result in an algal biomass increase up to 2–3 µg chl a/L, even in the middle of winter. When physical conditions (warm temperature, spring tides, and northern winds) are not restrictive, the water renewal rates can be less than ten days and the significant exchanges with the open sea can flush the excessive nutrient loads out of the system. Under these conditions, the embayment is nitrogen limited and the phytoplankton biomass is very low (0.5–1 µg chl a/L) (Arhonditsis et al., 2002b).

The basic conceptual design of our model builds upon the results of earlier local modeling studies and considers the basic ecological processes underlying plankton dynamics in the coastal embayment (Arhonditsis et al., 2000, 2002b). For the sake of simplicity, the spatially explicit (2-D) character of the original model was reduced to a zero-dimensional (single compartment) model that considers the interactions between the six state variables: nitrate (NO_3), ammonium (NH_4), phytoplankton (PHYT), zooplankton (ZOOP), bacteria (BACT) and dissolved organic nitrogen (DON) (Fig. 1). To accommodate the spatial (horizontal) variability observed in the coastal embayment and thus minimize the effects of the simplified model segmentation, we accordingly increased the measurement/observation error of the data used for model calibration (see Bayesian configuration of the model). The mathematical description of the ecological model and the definition of the model parameters can be found in Tables 1 and 2, respectively. The simulation model was solved numerically using the fourth-order Runge–Kutta method with a time step of one day.

2.2. Phytoplankton

The equation for algal biomass considers phytoplankton production and losses due to mortality and herbivorous zooplankton grazing. Our model explicitly considers the role of new and regenerated production using separate formulations for nitrate and ammonium

Table 1
The specific functional forms of the eutrophication model

$$\frac{dNH_4}{dt} = -(\mu_{max(phyt)}\phi_A\phi_{LT} - \alpha_{p(NH4)}m_p)\sigma_{(t)}PHYT \, N/C_{phyto}$$
$$+ \alpha_{b(NH4)}m_b\sigma_{(t)}BACT^2 N/C_{bact}$$
$$- U_2\sigma_{(t)}BACT \, N/C_{bact} + k_{min \, er}\frac{BACT}{K_R + BACT}\sigma_{(t)}DON$$
$$+ \alpha_{z(NH4)}m_z\sigma_{(tz)}ZOOP \, N/C_{zoop}$$
$$+ k\sigma_{(t)}(NH_{4(o)} - NH_4) + NH_{4EXOG}$$

$$\frac{dNO_3}{dt} = -\mu_{max(phyt)}\phi_N\phi_{LT}\sigma_{(t)}PHYT \, N/C_{phyto}$$
$$+ k\sigma_{(t)}(NO_{3(o)} - NO_3) + NO_{3EXOG}$$

$$\frac{dPHYT}{dt} = (\mu_{max(phyt)}\phi_{NA}\phi_{LT}(1 - \gamma) - m_p)\sigma_{(t)}PHYT - \sigma_{(tz)}G_p$$
$$+ k\sigma_{(t)}(PHYT_{(o)} - PHYT)$$

$$\phi_{LT} = \frac{2.718f}{k_{ext}z}(e^{-a1} - e^{-a0}) a_0 = \frac{I}{I_k} a_1 = \frac{I}{I_k}e^{-k_{ext}z}$$
$$k_{ext} = k_{chla}PHYT(chla/C) + k_{back}$$
$$\varphi_{NA} = \phi_N + \varphi_A = \frac{NO_3\exp(-\psi NH_4)}{NO_3 + NH} + \frac{NH_4}{NH_4 + AH}$$
$$\sigma_{(t)} = \frac{\left(1 - \varepsilon\cos\left(\frac{2\pi t}{365}\right)\right)}{1 + \varepsilon} \sigma_{(tz)} = \frac{\left(1 - \varepsilon\cos\left(\frac{2\pi t}{365} - 0.5\right)\right)}{1 + \varepsilon}$$

$$\frac{dBACT}{dt} = (U_1 + U_2)\sigma_{(t)}BACT - m_b\sigma_{(t)}BACT^2 - \sigma_{(tz)}G_B$$
$$+ k\sigma_{(t)}(BACT_{(o)} - BACT)$$

$$S = \min(NH_4, vDON) U_1 = \frac{\mu_{max(bact)}DON}{DH + S + DON} U_2 = \frac{\mu_{max(bact)}S}{DH + S + DON}$$

$$\frac{dZOOP}{dt} = as_p\sigma_{(tz)}G_P + as_b\sigma_{(tz)}G_B - m_z\sigma_{(tz)}ZOOP$$
$$+ k\sigma_{(t)}(ZOOP_{(o)} - ZOOP)$$

$$F = p_1PHYT + p_2BACT$$
$$G_P = \frac{g_{max}p_1PHYT}{K_z + F}ZOOP \quad p_1 = \frac{PHYT}{PHYT + BACT}$$
$$G_B = \frac{g_{max}p_2BACT}{K_z + F}ZOOP \quad p_2 = \frac{BACT}{PHYT + BACT}$$

$$\frac{dDON}{dt} = (\gamma\mu_{max(phyt)}\phi_{NA}\phi_{LT} + \alpha_{p(DON)}m_p)\sigma_{(t)}PHYT \, N/C_{phyto}$$
$$+ \alpha_{b(DON)}m_b\sigma_{(t)}BACT^2 N/C_{bact}$$
$$- U_1\sigma_{(t)}BACT \, N/C_{bact} + \alpha_{z(DON)}m_z\sigma_{(tz)}ZOOP \, N/C_{zoop}$$
$$- k_{min \, er}\frac{BACT}{K_R + BACT}\sigma_{(t)}DON$$
$$+ k\sigma_{(t)}(DON_{(o)} - DON) + DON_{EXOG}$$

Table 2
Parameter definitions of the eutrophication model

Parameter	Description	Units
γ^*	Fraction of phytoplankton production exudated as DON	
m_p^*	Phytoplankton mortality rate	day^{-1}
m_b^*	Bacterial mortality rate	$(\mu g\ C/L)^{-1} \cdot day^{-1}$
m_z^*	Zooplankton mortality rate	day^{-1}
NH*	Half saturation constant for nitrate phytoplankton uptake	$\mu g\ N/L$
AH*	Half saturation constant for ammonium phytoplankton uptake	$\mu g\ N/L$
$\mu_{max(phyt)}^*$	Maximum growth rate for phytoplankton	day^{-1}
as_p	Zooplankton assimilation efficiency for phytoplankton	0.70
as_b	Zooplankton assimilation efficiency for bacteria	0.70
DH*	Half saturation constant for bacterial uptake	$\mu g\ N/L$
I_k^*	Half saturation light intensity	$MJ/m^2 \cdot day^{-1}$
k_{back}^*	Background light extinction coefficient	m^{-1}
k_{chla}^*	Light extinction coefficient due to chlorophyll a	$L \cdot (\mu g\ chla \cdot m)^{-1}$
ψ^*	Strength of the ammonium inhibition for nitrate uptake	$(\mu g\ N/L)^{-1}$
ν	Ratio of bacterial ammonium to DON uptake	0.75
ε	Shape parameter for the trigonometric functions $\sigma_{(t)}$ and $\sigma_{(z)}$	0.50
$\mu_{max(bact)}^*$	Maximum bacterial uptake rate	day^{-1}
g_{max}^*	Zooplankton maximum grazing rate	day^{-1}
K_Z^*	Half-saturation constant for zooplankton grazing	$\mu g\ C/L$
K_R^*	Half-saturation constant for DON mineralization	$\mu g\ C/L$
k_{miner}^*	Maximum mineralization rate	day^{-1}
N/C_{phyto}	Nitrogen to carbon ratio for phytoplankton	$0.179\ \mu g\ N\ (\mu g\ C)^{-1}$
N/C_{bact}	Nitrogen to carbon ratio for bacteria	$0.222\ \mu g\ N\ (\mu g\ C)^{-1}$
N/C_{zoop}	Nitrogen to carbon ratio for zooplankton	$0.167\ \mu g\ N\ (\mu g\ C)^{-1}$
$a_{p(DON)}$	Fraction of phytoplankton mortality becoming DON	0.30
$a_{p(NH4)}$	Fraction of phytoplankton mortality becoming ammonium	0.30
$a_{b(DON)}$	Fraction of bacterial mortality becoming DON	0.30
$a_{b(NH4)}$	Fraction of bacterial mortality becoming ammonium	0.30
$a_{z(DON)}$	Fraction of zooplankton mortality becoming DON	0.30
$a_{z(NH4)}$	Fraction of zooplankton mortality becoming ammonium	0.30

The asterisks indicate parameters used during the Bayesian calibration of the model.

phytoplankton uptake (Eppley–Peterson f-ratio paradigm; Eppley and Peterson, 1979). Amongst the variety of light saturation curves (see Jassby and Platt, 1976), we used Steele's equation along with an extinction coefficient determined as the sum of the background light attenuation and attenuation due to chlorophyll a (self-shading effects). The effects of the seasonal temperature cycle on phytoplankton are described by a trigonometric function $\sigma_{(t)}$.

2.3. Zooplankton

Zooplankton grazing and losses due to natural mortality/consumption by higher predators are the main two terms in the zooplankton biomass equation. Zooplankton has two alternative food sources (phytoplankton and bacteria) grazed with preference that changes dynamically as a function of their relative proportion (Fasham et al., 1990). Zooplankton grazing was modeled using a Michaelis–Menten equation and the fraction assimilated fuels growth. In the absence of information to support more complex forms, we selected a linear closure term that represents the effects of a seasonally invariant predator biomass (see Edwards and Yool, 2000). The effects of temperature on zooplankton metabolic activities were modeled by a trigonometric function similar to the one used for phytoplankton. We also considered a lagged zooplankton growth response (≈ 30 days) during the spring warming period represented by a phase shift of -0.5 radians. To more effectively guide the Bayesian calibration of the model, we overcame the lack of consistent zooplankton biomass data by creating a semi-synthetic dataset, i.e., based on earlier simulations and existing observations from the system, the zooplankton biomass values were considered half of the contemporaneous phytoplankton biomass along with random terms sampled from normal distributions with zero mean values and standard deviations equal to 15% of the respective (generated) monthly values.

2.4. Bacteria

Bacterial growth depends on dissolved organic nitrogen and ammonium availability. Our parameterization is conceptually similar to the one introduced by Fasham et al. (1990), and considers a total bacterial nitrogenous substrate S to ensure balanced growth with a constant ratio of bacterial ammonium to DON uptake. Loss of bacterial biomass due mortality/excretion has been modeled using a quadratic function. The latter form corresponds to a loss rate dependent on the bacterial biomass itself and, aside from the metabolic

losses, may be interpreted as representing bacterivory by other consumers (e.g., heterotrophic nanoflagellates, mixotrophic flagellates, ciliates) whose biomass is proportional to that of bacteria. Zooplankton grazing is another loss term in the bacterial biomass equation.

2.5. Ammonium and nitrate

Both ammonium and nitrate equations consider phytoplankton uptake by taking into account ammonium inhibition of nitrate uptake (Wrobleski, 1977). The former equation also considers the proportion of zooplankton and bacterial excretion and mortality/predation that is returned back to the system as ammonium. Ammonium is fuelled by the bacteria-mediated DON mineralization, but is also utilized by bacteria as a source of nitrogen for cell protein synthesis (Fasham et al., 1990).

2.6. Dissolved organic nitrogen

A fraction of the primary production is exuded by phytoplankton as DON. The model also considers the contribution of phytoplankton, bacterial and zooplankton mortality to the organic nitrogen pool. Bacteria also uptake DON to obtain their carbon (i.e., DON is used as a proxy for DOC) and seasonally forced mineralization processes transform DON to ammonium.

2.7. Boundary conditions

Ammonium, nitrate and dissolved organic nitrogen loadings from the watershed were based on predictions from two different models, i.e., runoff curve number (RCN) equation and mass response functions (MRFs) appropriate for low/intermediate and high intensity rainfall events, respectively (Arhonditsis et al., 2002a). To more realistically account for the effects of the external loading conditions on the coastal embayment patterns, we also employed a stochastic treatment of the forcing functions of the model; i.e., the predicted nitrogen loads provided the mean of a Gaussian distribution with standard deviation assumed to be equal to the product of the mean relative error with the respective model predictions. Finally, a term that corresponds to the seasonally varying exchanges with the open sea was also included in the six differential equations, and the seasonal variation of the six state variables in the Aegean Sea (denoted with the subscript (O) in Table 1) was represented by trigonometric functions fitted to measured data during the study period.

3. Bayesian configuration of the model

3.1. Statistical formulations

Our Bayesian framework comprises three statistical formulations that can be distinguished by the following assumptions: i) the ecological model is a perfect simulator of the coastal embayment [Model 1], ii) the ecological model is an imperfect simulator of the coastal embayment and the model discrepancy is invariant with the input conditions (i.e., the difference between model and natural system dynamics is constant over the annual cycle for each state variable) [Model 2], and iii) the ecological model is an imperfect simulator of the coastal embayment and the model discrepancy varies with the input conditions (i.e., there is seasonally varying discrepancy between model and the environmental system for each state variable) [Model 3]. The three probabilistic approaches aim to combine field data and model outputs to update the uncertainty of model parameters, determine their covariance structure, and then use the calibrated model to give predictions (along with uncertainty bounds) of the plankton dynamics in the coastal embayment.

i) *Model 1*: The first statistical formulation is based on the assumption that our model perfectly describes the dynamics of the environmental system and the observations y for the six state variables are given by:

$$y_i = f(\theta, x_i, y_0, N_{\text{EXOG}}) + \varepsilon_i, i = 1, 2, 3, \ldots, n \tag{1}$$

where $f(\theta, x_i, y_0, N_{\text{EXOG}})$ denotes the ecological model, x_i is a vector of time dependent control variables (e.g., boundary conditions, forcing functions) describing the environmental conditions, the vector θ is a time independent set of the calibration model parameters (i.e., the 17 parameters in Table 2), y_0 corresponds to the concentrations of the six state-variables at the initial time point t_0, N_{EXOG} represents the exogenous nitrogen loadings ($NO_{3\text{EXOG}}$, $NH_{4\text{EXOG}}$, DON_{EXOG}) into the system, and ε_i denotes the observation (measurement) error that is usually assumed to be independent and identically distributed following a Gaussian distribution. The observed spatiotemporal

patterns in the Gulf of Gera provide evidence of a multiplicative measurement error (Arhonditsis et al., 2000), and thus we assumed the standard deviation to be proportional to the average monthly values for each state variable (Van Oijen et al., 2005). Specifically, we chose the monthly standard deviations to be 15% of the mean monthly values; a fraction that comprises both analytical error and spatial variability in the coastal embayment.

Based on the previous assumptions, the likelihood function that evaluates how well the simulation model is able to reproduce the observed data y at each value of θ, y_0, N_{EXOG} is given by:

$$p(y|f(\theta,x,y_0,N_{EXOG})) = \prod_{j=1}^{m} (2\pi)^{-n/2}|\Sigma_{\varepsilon j}|^{-1/2}\exp\left[-\frac{1}{2}[y_j - f_j(\theta,x,y_0,N_{EXOG})]^T \Sigma_{\varepsilon j}^{-1}[y_j - f_j(\theta,x,y_0,N_{EXOG})]\right]$$

(2)

where m and n correspond to the number of state variables ($m=6$) and the number of observations in time used to calibrate the model ($n=17$ average monthly values), respectively; $y_j=[y_{1j},..., y_{nj}]^T$ and $f_j(\theta, x, y_0)=[f_j(\theta, x_1, y_0, N_{EXOG}),..., f_j(\theta, x_m, y_0, N_{EXOG})]^T$ correspond to the vectors of the field observations and model predictions for the state variable j; and $\Sigma_{\varepsilon j}=I_n\cdot(0.15)^2\cdot y_j^T\cdot y_j$. In the context of the Bayesian statistical inference, the posterior density of the parameters θ, the initial conditions of the six state variables y_0 and the exogenous nitrogen loading N_{EXOG} given the observed data y is defined as:

$$p(\theta,y_0,N_{EXOG}|y) = \frac{p(y|f(\theta,x,y_0,N_{EXOG}))p(\theta)p(y_0)p(N_{EXOG})}{\int\int\int p(y|f(\theta,x,y_0,N_{EXOG}))p(\theta)p(y_0)p(N_{EXOG})\mathrm{d}\theta\mathrm{d}y_0\mathrm{d}N_{EXOG}}$$

$$\propto p(y|f(\theta,x,y_0,N_{EXOG}))p(\theta)p(y_0)p(N_{EXOG})$$

(3)

where $p(\theta)$ is the prior density of the model parameters θ, $p(y_0)$ is the prior density of the initial conditions of the six state variables y_0, and $p(N_{EXOG})$ is the prior density of the exogenous nitrogen loading N_{EXOG}. The formulation of the prior parameter distributions was based on the identification of the minimum and maximum values for each parameter from the pertinent literature (Fasham et al., 1990; Jorgensen et al., 1991; Arhonditsis et al., 2000, 2002b), and then we assigned lognormal distributions that 95% of their values were lying within the identified ranges (Steinberg et al., 1997) [For the sake of simplicity of this illustration, we used lognormal distributions for all the parameters but it should be noted that a whole suite of probability distributions (e.g., beta, triangular, uniform) can be considered to reflect different prior information (Hong et al., 2005)]. In a similar way to the measurement errors, the characterization of the prior density $p(y_0)$ was based on the assumption of a Gaussian distribution with a mean value derived from the observed average in the system during the first sampling date (June 11, 1996) and standard deviation that was 15% of the mean value for each state variable j; a fraction that comprises both analytical error, and spatial variability among the six sampling stations in the embayment. The prior density $p(N_{EXOG})$ of the exogenous nitrogen loading was also based on probabilistic (Gaussian) treatment of model-based estimates using standard deviations that reflected the watershed model error (Arhonditsis et al., 2002a).Thus, the resulting posterior distribution for θ, y_0 and N_{EXOG} is given by:

$$p(\theta,y_0,N_{EXOG}|y) \propto \prod_{j=1}^{m} (2\pi)^{-n/2}|\Sigma_{\varepsilon j}|^{-1/2}\exp\left[-\frac{1}{2}[y_j - f_j(\theta,x,y_0,N_{EXOG})]^T \Sigma_{\varepsilon j}^{-1}[y_j - f_j(\theta,x,y_0,N_{EXOG})]\right]$$

$$\times (2\pi)^{-l/2}|\Sigma_\theta|^{-1/2} \prod_{k=1}^{l} \frac{1}{\theta_k}\exp\left[-\frac{1}{2}[\log\theta - \theta_0]^T \Sigma_\theta^{-1}[\log\theta - \theta_0]\right]$$

$$\times (2\pi)^{-m/2}|\Sigma_{y0}|^{-1/2}\exp\left[-\frac{1}{2}[y_0 - y_{0m}]^T \Sigma_{y0}^{-1}[y_0 - y_{0m}]\right]$$

$$\times \prod_{k=1}^{o} (2\pi)^{-p/2}|\Sigma_{N_{EXOGk}}|^{-1/2}\exp\left[-\frac{1}{2}[N_{EXOGk} - N_{EXOGmk}]^T \Sigma_{N_{EXOGk}}^{-1}[N_{EXOGk} - N_{EXOGmk}]\right]$$

(4)

where l and p are the number of the model parameters θ used for the model calibration ($l=17$) and number of days of the simulation period ($p=488$); θ_0 denotes the vector of the mean values of θ (logarithmic scale); $\Sigma_\theta=I_l\cdot\sigma_\theta^T\cdot\sigma_\theta$ and $\sigma_\theta=[\sigma_{\theta 1},..., \sigma_{\theta l}]^T$ corresponds to the vector of the shape parameters of the l lognormal distributions (standard deviation of $\log\theta$); the vector $y_{0m}=[y_{11},...,y_{16}]^T$ indicates the average values of the six state variables observed in the system during the first sampling date; $\Sigma_{y0}=I_m\cdot(0.15)^2\cdot y_{0m}^T\cdot y_{0m}$; o ($=3$) corresponds to the three exogenous nitrogen loading forms (NO_3, NH_4, DON) with model-based mean values $N_{EXOGmk}=[N_{EXOGm1k},..., N_{EXOGmpk}]^T$ and

covariance matrix $\Sigma_{\text{NEXOG}k} = I_p \cdot (RE_p)^2 \cdot N^T_{\text{EXOG}mk} \cdot N_{\text{EXOG}mk}$; and RE_p represents the mean relative error of the daily exogenous nitrogen loading estimates from the watershed models (Arhonditsis et al., 2002a).

ii) *Model 2:* An advancement of the previous statistical formulation explicitly considers that the model imperfectly represents the dynamics of the coastal embayment. In this case, an observation i for the state variables j, y_{ij}, can be described as:

$$y_{ij} = f(\theta, x_i, y_0, N_{\text{EXOG}}) + \delta_j + \varepsilon_{ij}, \, i = 1, 2, 3, \dots n \text{ and } j = 1, \dots, m \tag{5}$$

where the stochastic term δ_j accounts for the discrepancy between the model $f(\theta, x, y_0, N_{\text{EXOG}})$ and the natural system, which is assumed to be invariant with the input conditions x (i.e., the difference between model and coastal embayment dynamics was assumed to be constant over the annual cycle for each state variable). With this assumption, the likelihood function will be:

$$p(y|f(\theta, x, y_0, N_{\text{EXOG}})) = \prod_{j=1}^{m} (2\pi)^{-n/2} |\Sigma_{Tj}|^{-1/2} \exp\left[-\frac{1}{2} [y_j - f_j(\theta, x, y_0, N_{\text{EXOG}})]^T \Sigma_{Tj}^{-1} [y_j - f_j(\theta, x, y_0, N_{\text{EXOG}})]\right] \tag{6}$$

$$\Sigma_{Tj} = \Sigma_{\delta j} + \Sigma_{\varepsilon j} \tag{7}$$

where $\Sigma_{\delta j} = I_n \cdot \sigma_j^2$ corresponds to the additional stochastic term of Model 2; and the prior densities $p(\sigma_j^2)$ were based on the conjugate inverse-gamma distribution (Gelman et al., 1995). Thus, the resulting posterior distribution for θ, y_0, N_{EXOG} and σ^2 is:

$$
\begin{aligned}
p(\theta, y_0, N_{\text{EXOG}}, \sigma^2 | y) \quad &\propto \prod_{j=1}^{m} (2\pi)^{-n/2} |\Sigma_{Tj}|^{-1/2} \exp\left[-\frac{1}{2} [y_j - f_j(\theta, x, y_0, N_{\text{EXOG}})]^T \Sigma_{Tj}^{-1} [y_j - f_j(\theta, x, y_0, N_{\text{EXOG}})]\right] \\
&\times (2\pi)^{-l/2} |\Sigma_\theta|^{-1/2} \prod_{k=1}^{l} \frac{1}{\theta_k} \exp\left[-\frac{1}{2} [\log\theta - \theta_0]^T \Sigma_\theta^{-1} [\log\theta - \theta_0]\right] \\
&\times (2\pi)^{-m/2} |\Sigma_{y0}|^{-1/2} \exp\left[-\frac{1}{2} [y_0 - y_{0m}]^T \Sigma_{y0}^{-1} [y_0 - y_{0m}]\right] \\
&\times \prod_{k=1}^{o} (2\pi)^{-p/2} |\Sigma_{N_{\text{EXOG}k}}|^{-1/2} \exp\left[-\frac{1}{2} [N_{\text{EXOG}k} - N_{\text{EXOG}mk}]^T \Sigma_{N_{\text{EXOG}k}}^{-1} [N_{\text{EXOG}k} - N_{\text{EXOG}mk}]\right] \\
&\times \prod_{j=1}^{m} \frac{\beta_j^{\alpha j}}{\Gamma(\alpha_j)} \sigma_j^{-2(\alpha_j + 1)} \exp\left(-\frac{\beta_j}{\sigma_j^2}\right)
\end{aligned}
\tag{8}
$$

where α_j (=0.01) and β_j (=0.01) correspond to the shape and scale parameters of the m non-informative inverse-gamma distributions used in this analysis.

iii) *Model 3:* The third statistical formulation also explicitly recognizes that the model imperfectly represents the dynamics of the environmental system but now the corresponding stochastic term varies with the input conditions x. In this case, an observation i for the state variables j, y_{ij}, can be described as:

$$y_{ij} = f(\theta, x_i, y_0, N_{\text{EXOG}}) + \delta_{ij} + \varepsilon_{ij}, \, i = 1, 2, 3, \dots n \text{ and } j = 1, \dots, m. \tag{9}$$

The modeling for all the previous terms remains unchanged. We also specify a Gaussian first order random walk model for the discrepancy term δ_{ij} to reflect that these error terms are correlated (Shaddick and Wakefield, 2002). Specifically, the vector $\delta_j = \{\delta_{1j}, \dots, \delta_{17j}\}$, $j = \text{NH}_4$, NO_3, PHYT, BACT, ZOOP, DON can be expressed as:

$$
p(\delta_{ij} | \delta_{-ij}, \omega_j^2) \sim
\begin{cases}
N(\delta_{i+1j}, \omega_j^2) & \text{for } i = 1, \\
N\left(\dfrac{\delta_{i-1j} + \delta_{i+1j}}{2}, \dfrac{\omega_j^2}{2}\right) & \text{for } i = 2, \dots, 16 \\
N(\delta_{i-1j}, \omega_j^2) & \text{for } i = 17,
\end{cases}
\tag{10}
$$

where δ_{-ij} denotes all elements of δ_j except the δ_{ij}, ω_j^2 is the conditional variance and the prior densities $p(\omega_j^2)$ were based on conjugate inverse-gamma (0.01, 0.01) distributions (Gelman et al., 1995).

3.2. Markov chain Monte Carlo

Sequence of realizations from the posterior distribution of the models was obtained using Markov chain Monte Carlo (MCMC) simulations (Gilks et al., 1998). MCMC is a general methodology that has several advantages particularly useful for environmental modeling, i.e., it can efficiently sample multi-dimensional parameter spaces and can easily handle multivariate outputs as well as large numbers of nuisance parameters (Gelman et al., 1995). In this study, we used the general normal-proposal Metropolis algorithm as implemented in the WinBUGS software (Spiegelhalter et al., 2003); this algorithm is based on a symmetric normal proposal distribution, whose standard deviation is adjusted over the first 4000 iterations such as the acceptance rate ranges between 20% and 40%. We also used an ordered over-relaxation, which generates multiple samples per iteration and then selects one that is negatively correlated with the current value of each stochastic node (Neal, 1998). The latter option resulted in an increased time per iteration but reduced within-chain correlations. The posterior simulations were based on three parallel chains with starting points: (i) a vector that consists of the mean values of the prior parameter distributions, (ii) a vector fairly similar to the one obtained from an earlier model calibration (Arhonditsis et al., 2000), and (iii) a vector that resulted from the optimization of the model with the Fletcher–Reeves conjugate-gradient method (Chapra and Canale, 1998). We used 30,000 iterations and convergence was assessed with the modified Gelman–Rubin convergence statistic (Brooks and Gelman, 1998). 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. Our framework was implemented in the WinBUGS differential Interface (WBDiff); an interface that allows numerical solution of systems of ordinary differential equations within the WinBUGS software.

3.3. Model updating

We also used the MCMC estimates of the mean and standard deviation parameter values along with the covariance structure to update the model (Gelman et al., 1995). Under the assumption of a multinormal distribution for the raw (or log transformed) parameter values, the conditional distributions are given by:

$$\hat{\theta}_{i|j} = \hat{\theta}_i + [\theta_j - \hat{\theta}_j]\Sigma_j^{-1}\Sigma_{i,j} \tag{11}$$

$$\Sigma_{i|j} = \Sigma_i - \Sigma_{j,i}\Sigma_j^{-1}\Sigma_{i,j} \quad j \in \{i+1, \dots n\} \tag{12}$$

where $\hat{\theta}_{i|j}$ and $\Sigma_{i|j}$ correspond to the mean value and the dispersion matrix of the parameter i conditional on the parameter vector j; the values of the elements Σ_i $\Sigma_{i,j}$ and Σ_j correspond to the variance and covariance of the two subset of parameters; and $\hat{\theta}_i$, $\hat{\theta}_j$, θ_j correspond to the posterior mean and random values of the parameters i and j, respectively. To examine the predictive uncertainty of the updated model, we generated two datasets from normal distributions with mean values the observed averages and standard deviations that were 15 and 25% of the mean monthly values for each state variable. The former dataset represented conditions similar to those used to calibrate the model, while the latter was characterized by an increased month-to-month variability. In a similar way to the original calibration, the measurement error was considered to be 15% of the (generated) monthly values for each state variable.

3.4. Model evaluation

Assessment of the goodness-of-fit between the model outputs and the observed data was based on the posterior predictive p-value, i.e., the Bayesian counterpart of the classical p-value. In brief, the p-value is defined as the probability that the replicated data (the posterior predictive distribution) could be more extreme than the observed data. The null hypothesis H_0 (i.e., there are no systematic discrepancies between the simulation model and the data) is rejected if the tail-area probability is close to 0.0 or 1.0; whilst the model can be regarded as plausible if the p-value is near to 0.5. The discrepancy variable chosen for carrying out the posterior predictive model checks was the x^2 test [see also Gelman et al. (1996) for a detailed description of the posterior predictive p-value]. The comparison between the two alternative

Table 3
Markov chain Monte Carlo posterior estimates of the mean values and standard deviations of the model stochastic nodes

Parameter	Prior		Model 1		Model 2		Model 3	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
γ	0.062	0.034	0.056	0.028	0.055	0.027	0.056	0.038
m_p	0.072	0.021	0.083	0.024	0.067	0.018	0.066	0.023
m_b	0.062	0.034	0.073	0.019	0.101	0.023	0.087	0.027
m_z	0.062	0.034	0.030	0.012	0.052	0.027	0.048	0.025
NH	7.452	2.720	7.591	2.798	7.473	2.740	7.532	2.764
AH	7.452	2.720	5.261	1.123	4.435	1.461	7.265	2.743
$\mu_{max(phyt)}$	1.801	0.515	4.006	0.608	2.812	0.489	3.539	0.689
DH	7.452	2.720	8.453	3.202	7.442	2.471	7.654	2.966
I_k	1.947	0.834	0.422	0.066	0.478	0.273	1.139	0.577
k_{back}	0.246	0.026	0.193	0.017	0.201	0.020	0.241	0.037
k_{chla}	0.043	0.016	0.026	0.007	0.032	0.010	0.044	0.022
ψ	1.596	0.583	1.879	0.550	1.630	0.573	1.671	0.558
$\mu_{max(bact)}$	1.801	0.515	1.167	0.260	1.502	0.316	1.356	0.369
g_{max}	0.794	0.180	0.915	0.147	0.677	0.156	0.968	0.149
K_Z	42.59	15.54	47.16	11.51	50.58	17.40	40.13	12.137
K_R	16.25	3.852	16.24	3.689	15.53	3.600	15.35	3.752
k_{miner}	0.040	0.042	0.090	0.014	0.100	0.023	0.147	0.022
σ_{NH4}					1.550	0.935		
σ_{NO3}					4.002	0.889		
σ_{DON}					22.19	5.020		
σ_{PHYT}					17.66	5.061		
σ_{zoop}					13.70	3.452		
σ_{bact}					0.237	0.258		
ω_{NH4}							1.213	0.631
ω_{NO3}							4.136	1.669
ω_{DON}							20.45	8.728
ω_{PHYT}							24.54	6.147
ω_{zoop}							9.809	2.728
ω_{bact}							0.317	0.238
$NH4_{(t0)}$	17.08	2.562	17.31	0.862	17.25	0.860	17.15	0.864
$NO3_{(t0)}$	4.252	0.638	4.294	0.213	4.285	0.216	4.300	0.217
$DON_{(t0)}$	91.05	13.66	92.12	4.647	91.95	4.640	91.35	4.564
$PHYT_{(t0)}$	9.094	1.364	9.232	0.458	9.159	0.463	9.177	0.465
$ZOOP_{(t0)}$	3.422	0.513	3.512	0.179	3.448	0.175	3.450	0.176
$BACT_{(t0)}$	16.29	2.444	16.48	0.830	16.45	0.830	16.43	0.826

models was based on the use of the Bayes factor, i.e., the posterior odds of one model over the other (assuming the prior probability on either model is 0.5). If M_1 and M_2 denote the two alternative models, the Bayes factor is:

$$B_{12} = \frac{\text{pr}(y|M_1)}{\text{pr}(y|M_2)}. \tag{13}$$

For model comparison purposes, the model likelihood ($\text{pr}(y|M_k)$; $k=1, 2$) is obtained by integrating over the unknown element (initial conditions, model parameters, error terms) space:

$$\text{pr}(y|M_k) = \int \text{pr}(y|M_k, \Theta_k)\pi(\Theta_k|M_k)\text{d}\Theta_k \tag{14}$$

where Θ_k is the unknown element vector under model M_k and $\pi(\Theta_k|M_k)$ is the prior density of Θ_k. Using the MCMC method, we can estimate $\text{pr}(y|M_k)$ from posterior samples of Θ_k. Letting $\Theta_k^{(i)}$ be samples from the posterior density $\text{pr}(\Theta_k|M_k)$, the estimated $\text{pr}(y|M_k)$ is:

$$\overline{\text{pr}(y|M_k)} = \left\{ \frac{1}{m} \sum_{i=1}^{m} \text{pr}(y|M_k, \Theta_k^{(i)})^{-1} \right\}^{-1} \tag{15}$$

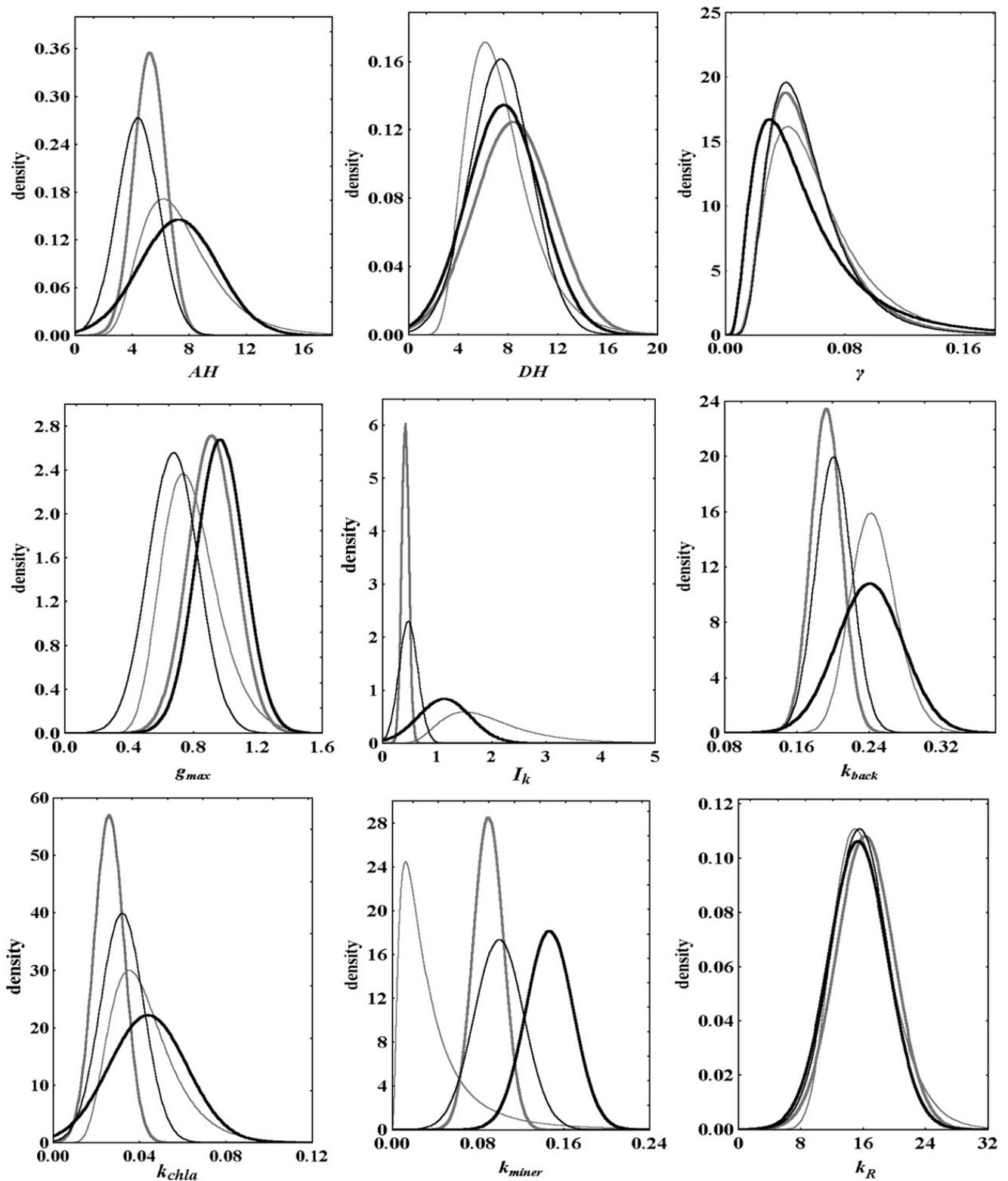the harmonic mean of the likelihood values (Kass and Raftery, 1995).

Fig. 2. Prior (thin grey lines) and posterior (Model 1: thick grey lines, Model 2: thin black lines, and Model 3: thick black lines) distributions of the eutrophication model. The posteriors depict smoothed kernel density estimates based on 12,500 MCMC samples from the three models.
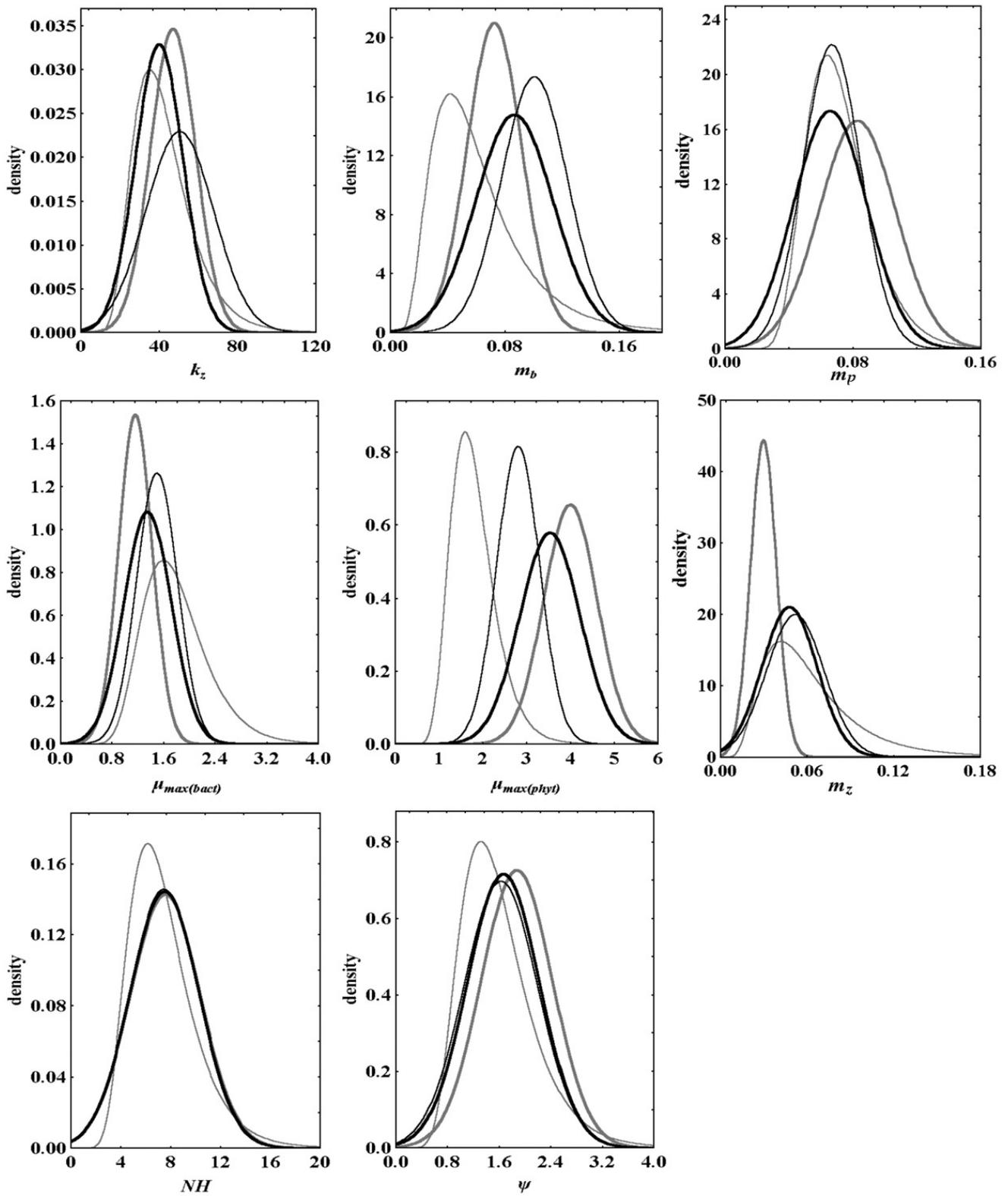
Fig. 2 (*continued*).

Table 4
Correlation matrix of the eutrophication model parameters based on MCMC posterior samples (Model 2)

	$\mu_{max(phyt)}$	$\mu_{max(bact)}$	m_p	m_b	m_z	K_R	K_Z	g_{max}	γ	DH	NH	AH	ψ	I_k	k_{back}	k_{chla}	k_{miner}
$\mu_{max(phyt)}$																	
$\mu_{max(bact)}$	−0.041																
m_p	0.119	−0.007															
m_b	−0.005	**0.983**	−0.004														
m_z	0.025	−0.039	−0.001	−0.024													
K_R	−0.033	−0.009	0.016	−0.013	−0.010												
K_Z	0.054	−0.060	0.012	−0.019	0.002	0.011											
g_{max}	−0.213	0.128	−0.017	0.070	−0.018	0.028	0.131										
γ	0.072	0.014	−0.050	0.007	−0.019	−0.026	0.001	0.012									
DH	0.007	0.044	−0.008	−0.020	0.002	−0.012	−0.025	−0.080	0.012								
NH	−0.013	0.008	−0.017	0.003	−0.013	0.014	0.005	0.004	0.006	0.019							
AH	0.097	0.174	−0.084	0.133	−0.059	0.013	−0.144	**0.286**	0.019	−0.045	0.004						
ψ	0.000	−0.009	0.015	−0.006	−0.008	0.014	0.003	−0.002	−0.008	−0.015	−0.001	−0.032					
I_k	**−0.532**	0.185	−0.005	0.128	−0.091	0.014	−0.199	**0.453**	0.050	−0.060	−0.002	**0.446**	−0.005				
k_{back}	−0.216	0.174	−0.003	0.132	−0.052	0.039	−0.130	**0.328**	0.009	−0.047	−0.001	0.241	−0.010	**0.336**			
k_{chla}	−0.064	0.079	0.000	0.064	−0.014	−0.007	−0.082	0.093	0.046	−0.019	−0.016	0.138	−0.007	0.142	0.057		
k_{miner}	**0.324**	**−0.353**	0.094	**−0.299**	0.053	**0.498**	0.140	**−0.301**	0.031	0.014	0.007	**−0.371**	0.019	**−0.449**	**−0.272**	−0.164	

Bold numbers correspond to correlation coefficients with absolute value greater than 0.250.

4. Results

The three MCMC sequences of the three models converged rapidly (≈ 5000 iterations) and the statistics reported in this study were based on the last 25,000 draws by keeping every 6th iteration (thin = 6). Based on the shifts in the most possible value and the reduction of the parameter uncertainty, we evaluated the degree of updating of model input parameters from prior to posterior. The central tendency along with the uncertainty underlying the values of the 17 model parameters after the MCMC sampling is depicted on the respective marginal posterior distributions (Table 3 and Fig. 2). The mean value of several updated distributions was shifted relative to the prior assigned values. For example, maximum phytoplankton growth and mineralization rates for Models 1, 2 and 3; zooplankton mortality, light extinction coefficient due to chlorophyll *a* and maximum bacterial uptake rate for Model 1; half saturation light intensity for Models 1 and 2; half saturation constant for ammonium phytoplankton uptake for Model 2; and bacterial mortality rate for Models 2 and 3. The standard deviations of the posterior parameter distributions were significantly reduced with the first statistical formulation (Model 1); characteristic examples were the zooplankton mortality rate, the half saturation constant for ammonium phytoplankton uptake, the half saturation light intensity, the light extinction coefficient due to chlorophyll *a*, the maximum bacterial uptake rate and the maximum mineralization rate with a relative decrease higher than 50%. On the other hand, the inclusion of the seasonally invariant stochastic term that accounts for the discrepancy between the model and the coastal embayment resulted in higher standard deviation values for most of the parameters. The same result was more evident after the addition of the seasonally variant discrepancy term (Model 3), and there were parameter distributions in which the posterior uncertainty estimates were even higher than the pre-specified ones, e.g., phytoplankton maximum growth rate, light extinction coefficient due to chlorophyll *a* and background light attenuation. Regardless of the statistical formulation used, three parameters of the calibration vector remained unaltered with respect to the first and second order moments of their posterior distributions, i.e., half saturation constant for nitrate phytoplankton uptake, strength of the ammonium inhibition for nitrate uptake, and half saturation constant for bacterial uptake. This finding probably stems from the nature of the data used to calibrate the model (bacterial biomass concentrations, relative magnitudes and temporal variability of nitrate/ammonium levels)

that did not offer insights into the role of the associated ecological processes (e.g., bacterial control on the DON mineralization rate) or supported the hypothesis underlying the prior parameter specifications (e.g., strong ammonium inhibition for nitrate uptake).

We also used the MCMC posterior samples from the Model 2 to determine the correlation structure of the seventeen model parameters (Table 4). Some of these

relationships have plausible physical explanation. For example, a higher maximum bacterial growth rate can be balanced by a higher bacterial mortality rate to accurately represent the observed bacterial biomass values in the system; a higher maximum mineralization rate can be balanced by a higher half-saturation constant for DON mineralization or by a lower bacterial mortality rate that reduces the amount of DON/ammonium being directly
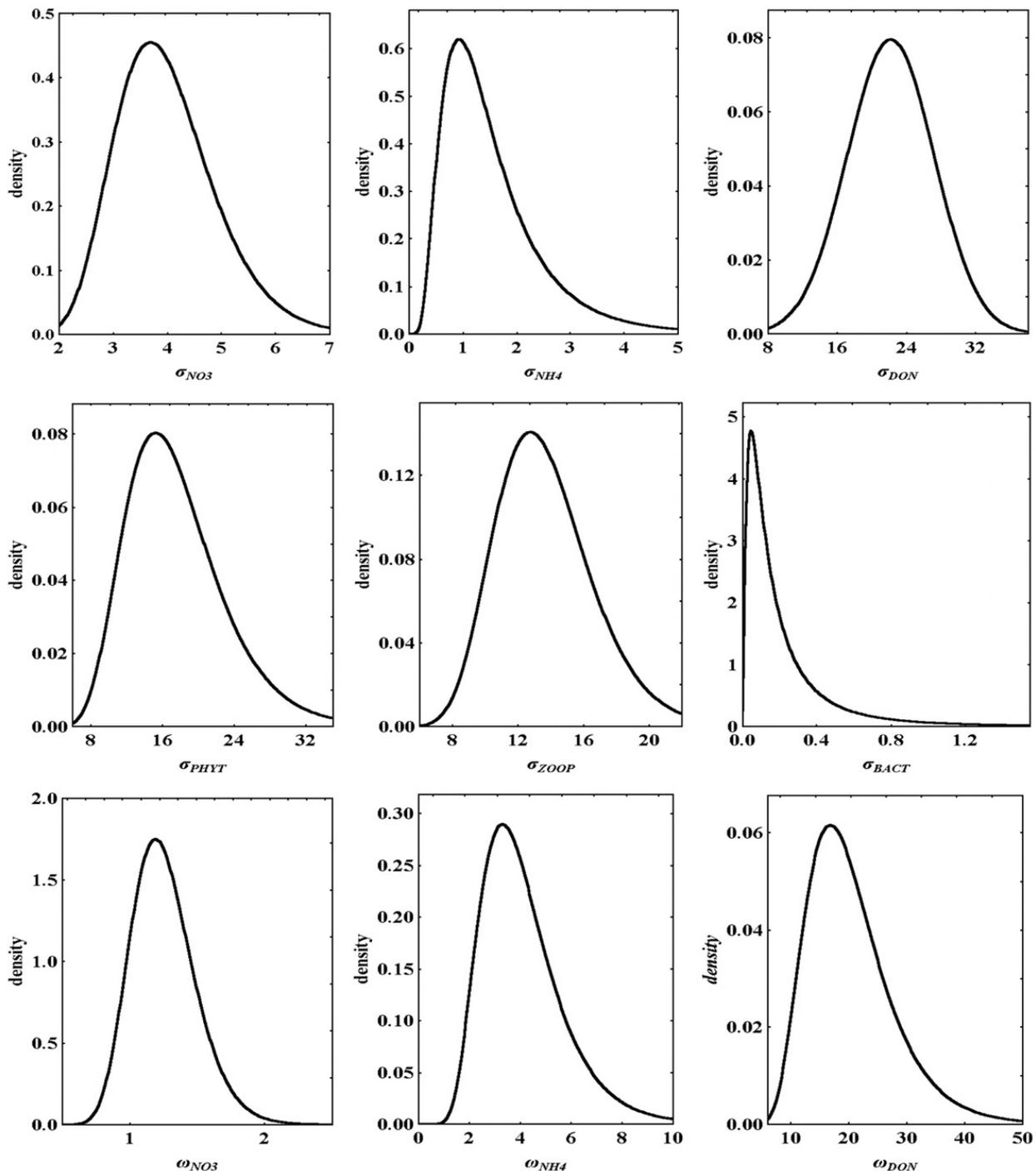

Fig. 3. Posterior distributions of the σ (seasonally invariant discrepancy between the simulation model and the natural system) and ω (conditional variance of the seasonally variant discrepancy) terms of the second and third statistical formulations, respectively.
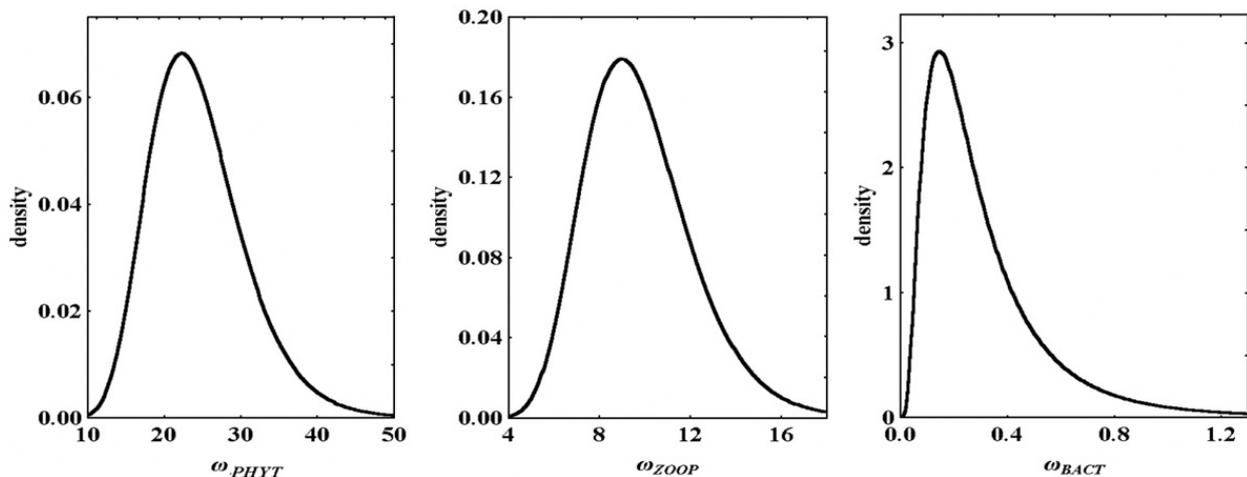
Fig. 3 (*continued*).

released in the water column; hence, the model can still fit the observed data. There were also some relationships that seem counter-intuitive and invite further explanation, such as the negative correlation between the maximum growth rate and the half-saturation constant for light intensity. Given the role of the two parameters on phytoplankton growth, a positive relationship where the two terms cancel each other out would have seemed more plausible. The negative correlation is probably driven by the two phytoplankton peaks observed in the winter months and indicates that high maximum growth rates combined with low half saturation light intensity constants (i.e., ability to attain optimal growth rates under low solar radiation availability) is the main strategy to reproduce the high winter phytoplankton biomass levels. The latter result is also reflected on the major central tendency shifts of the two parameters, i.e., the significantly higher/lower posterior mean values of the maximum growth rate and half saturation light intensity, respectively. While the winter phytoplankton community is dominated by several diatom species (e.g., *Fragilaria crotonensis*, *Fragilaria schulzi*) that can exhibit high growth rates (Arhonditsis et al., 2003b), the posterior mean values of the corresponding parameter were somewhat higher (especially for models 1 and 3) than those usually reported in the literature. To avoid parameter value shifts in areas not supported by existing evidence, one strategy is to implement a censored MCMC scheme that limits the posterior samples within the observed ranges (Spiegelhalter et al., 2003).

The posterior distributions of the error terms that correspond to the seasonally invariant discrepancy between the simulation model and the natural system are shown in Fig. 3. In the second statistical formulation, these terms reflect the model imperfection and delineate a constant zone around the model estimates for the six

state variables [i.e., the term $f(\theta, x_i, y_0, N_{EXOG})$ in Eq. (5)], and it is worth noting the relatively high coefficient of variation (CV) values of the bacteria error term (σ_{BACT}). Furthermore, to gain insight into the third statistical formulation, we plotted the model estimates vis-à-vis the discrepancy (error) terms for the nitrate, ammonium, dissolved organic nitrogen, phytoplankton, zooplankton, and bacteria biomass concentrations (Fig. 4). Despite some structural differences between earlier versions and the current model (Arhonditsis et al., 2000, 2002b), the model estimates [i.e., the term $f(\theta, x_i, y_0, N_{EXOG})$ in Eq. (9)] provided relatively similar results. Additionally, in contrast to the original calibration, the Bayesian calibration did not introduce bias (i.e., overestimation) of the bacteria biomass concentrations but it does provide evidence of more dynamic seasonal zooplankton fluctuations in the embayment [It should be noted, however, that the latter result is probably driven by the zooplankton data generated for the Bayesian calibration, whereas earlier model calibrations were based on qualitative information mainly to constrain zooplankton biomass within a plausible range.]. The discrepancy error terms [i.e., the term δ_{ij} in Eq. (5)] can be interpreted as an indicator of how well the model is matching reality. For example, the model's inability to closely reproduce the winter plankton dynamics can explain the higher December–January δ_{phyt} and δ_{zoop} values, while the higher $\delta_{August, NH4}$ also reflects the lower contemporaneous ammonium concentrations relative to the observed peaks in the system (August 1997). Finally, the posterior conditional variances (ω_j) of the seasonally variant discrepancies are also shown in Fig. 3.

The three statistical formulations were favourably supported by the data and were accepted on the basis of their posterior predictive *p*-values (0.098, 0.385 and

0.303 for Model 1, 2 and 3, respectively). The Bayes factor values $B_{21}=2.95$, $B_{31}=2.61$, $B_{23}=1.23$ did not provide strong evidence in support of one of the three alternative models but did reflect a higher performance of the Model 2 (Kass and Raftery, 1995; page 777). The comparison between the observed and posterior predictive monthly distributions for the three statistical formulations illustrates some features of the Bayesian calibration. The Model 1 consistently underestimates the nitrate levels throughout the study period, while both mean predictions and 95% credible intervals failed to

reproduce/include several observed ammonium, dissolved organic nitrogen and winter plankton biomass peaks (Fig. 5). The addition of the discrepancy terms in the third statistical formulation has significantly improved the results, although there are still observed nitrate and ammonium monthly values not included within the 95% credible intervals of the third model's predictions (Fig. 5). On the other hand, the second model provided the most accurate representation of the system dynamics, i.e., the central tendency of the majority of the predictive monthly distributions was
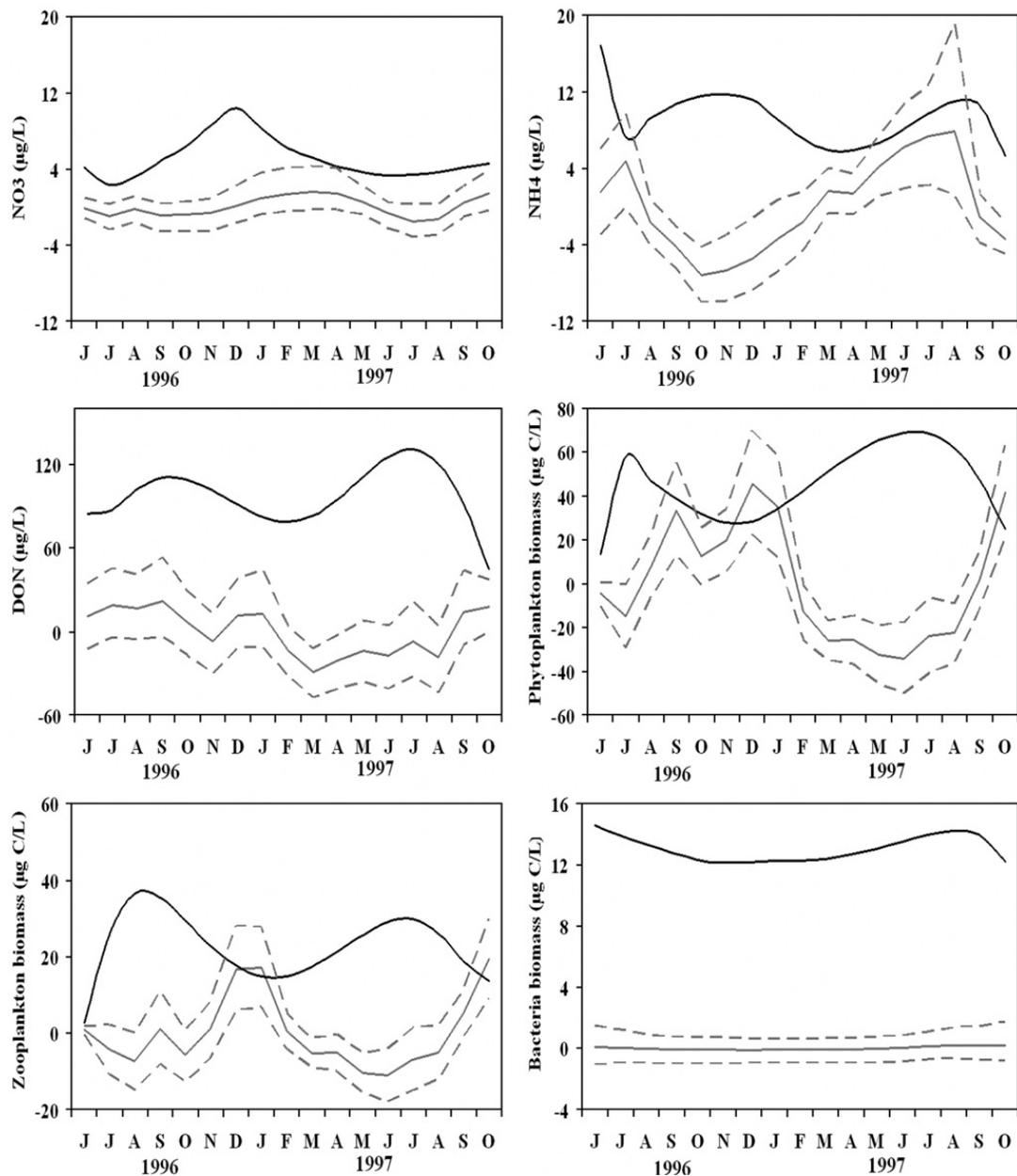

Fig. 4. Time series plots of the model estimates (black lines) and the error terms along with the 95% credible intervals (grey lines), for the ammonium, nitrate, dissolved organic nitrogen, phytoplankton, bacteria, and zooplankton biomass concentrations (Model 3). The error terms represent the discrepancy between the model and the natural system.
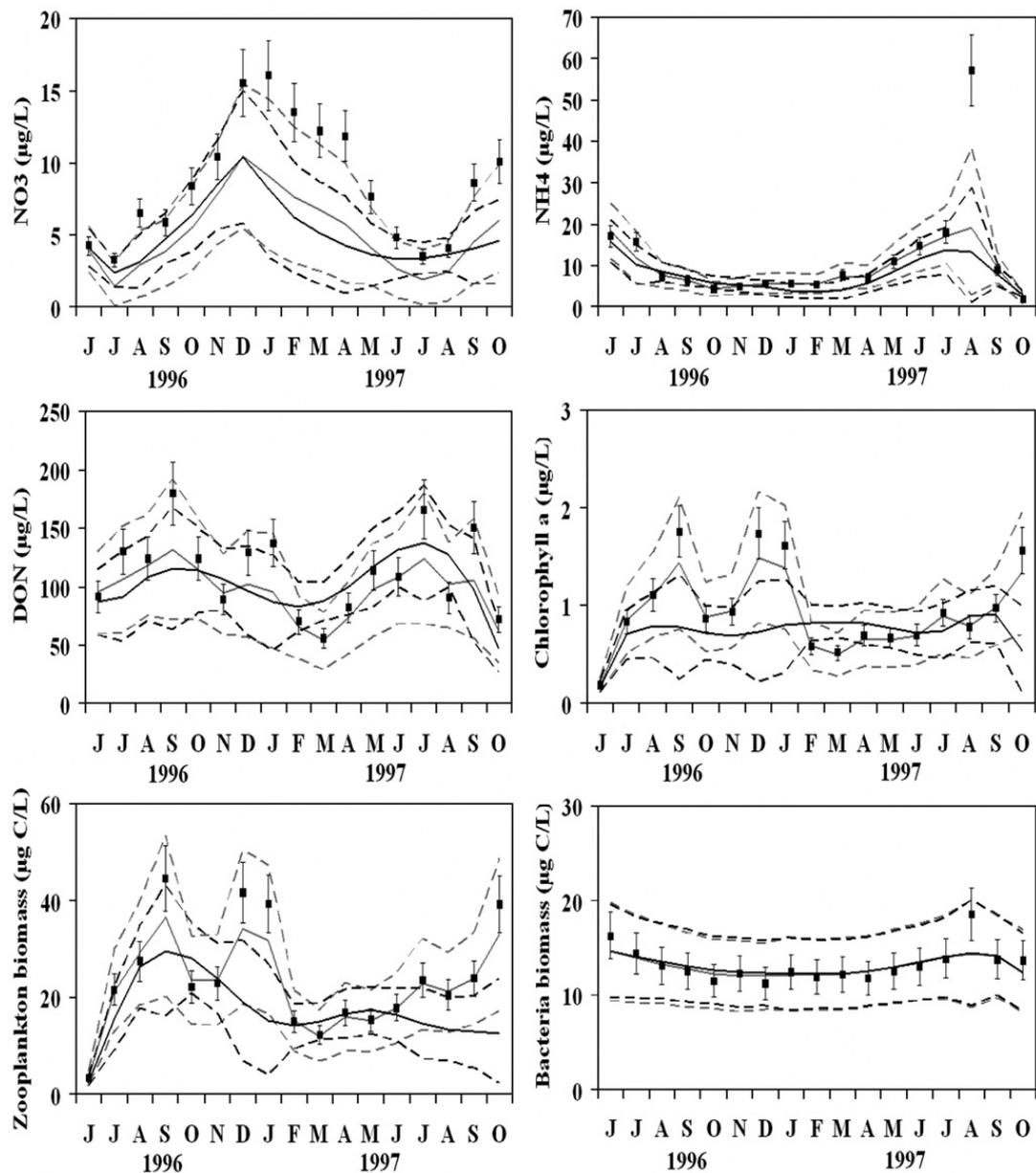
Fig. 5. Comparison between the observed and posterior predictive monthly distributions for ammonium, nitrate, dissolved organic nitrogen, chlorophyll *a* (1 g carbon=20 mg chlorophyll), bacteria, and zooplankton biomass based on 12,500 Markov chain Monte Carlo posterior samples from Models 1 (black lines) and 3 (grey lines). Continuous and dashed lines correspond to the mean predictions and the 95% credible intervals, respectively. The error bars express the measurement error/spatial variability relative to the observed data.

fairly close to the observed values of the six state variables, while the unusually high ammonium concentration observed in August 1997 was the only one not bracketed by the model uncertainty bounds (Fig. 6).

The highest performing statistical formulation (Model 2) was also used to examine the predictive ability of the updated model under two different conditions representing similar overall mean values and similar (dataset 1) or higher month-to-month (dataset 2) variability for the six state variables. As previously described, the updated parameter conditional

distributions were based on the MCMC estimates of the respective mean and standard deviation values along with their covariance structure, while the shape and scale parameters of the inverse-gamma distributions used to represent our updated beliefs for the values of the seasonally invariant discrepancy terms were estimated with the method of moments (Bernardo and Smith, 1994; pages 434). Not surprisingly, the first dataset did not alter the central tendency and standard deviation of the majority of the posterior parameter distributions (Table 5); the main exceptions were the
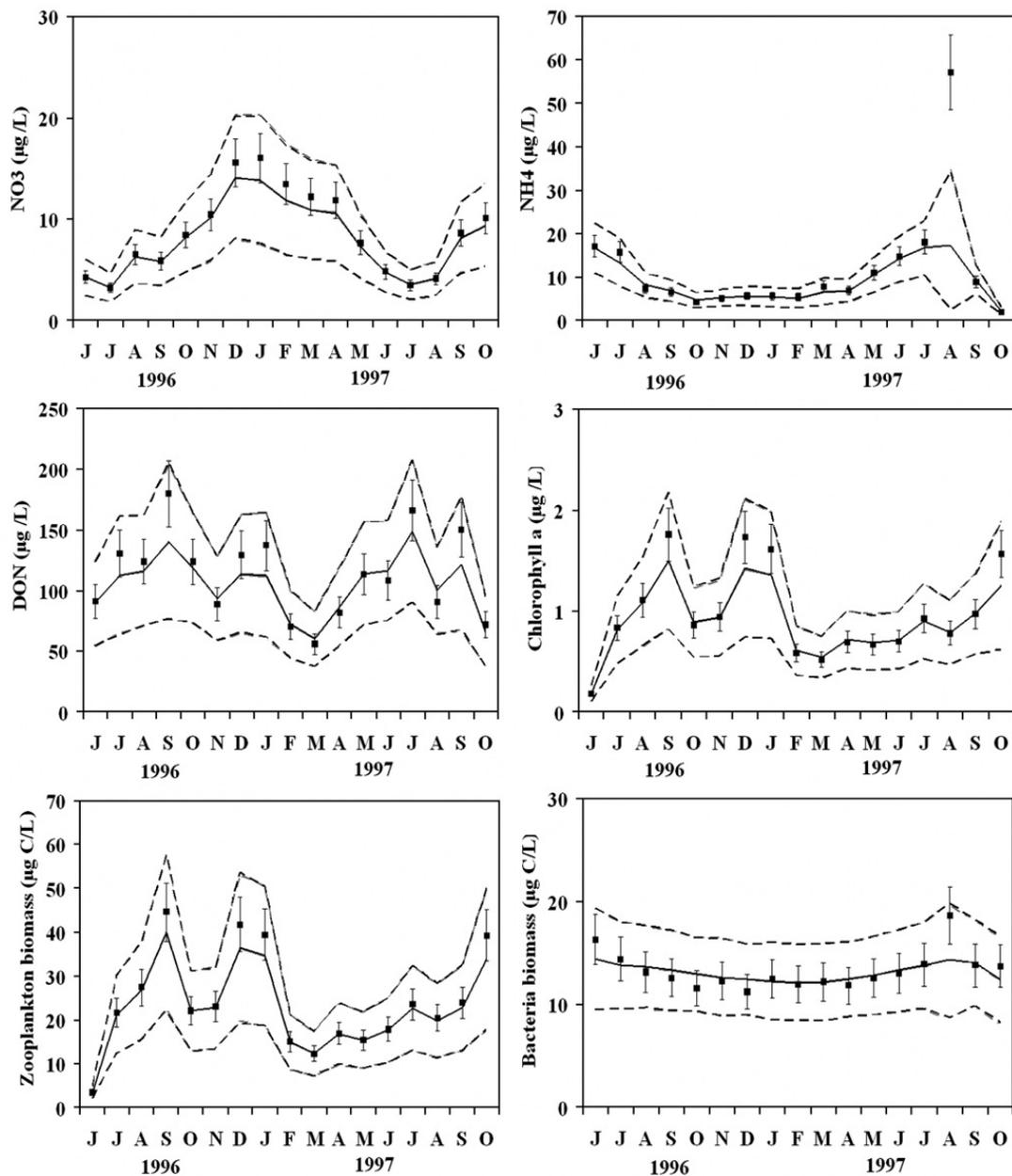
Fig. 6. Comparison between the observed and posterior predictive monthly distributions for ammonium, nitrate, dissolved organic nitrogen, chlorophyll *a*, bacteria, and zooplankton biomass based on 12,500 Markov chain Monte Carlo posterior samples from Model 2.

bacterial maximum growth and mortality rate along with the half saturation constant for light intensity in which the relative decrease of the standard deviation values was higher than 75%. Furthermore, the central tendency (≈ 20%) and standard deviation (≈ 55%) of the ammonium error term (σ_{NH4}) were also reduced mainly due to the lower August concentration used in this dataset (Fig. 7). With the second dataset, we formulated two different prior parameter distributions: i) the first set of priors was similar to the one used with the first dataset (Priors #1), and ii) the second set had parameter precisions reduced in half and aimed to provide a less constrained (flatter) prior parameter space for detecting

shifts of the joint and marginal posterior distributions (Priors #2). The first approach gave relatively similar results to those found with the first dataset, whereas the mean values of the ammonium (σ_{NH4}), dissolved organic nitrogen (σ_{DON}), and bacteria (σ_{bact}) error terms were significantly increased (Table 5). The latter finding probably indicates that the terms reflecting the mismatch between mathematical system and coastal embayment were mainly used to accommodate the pronounced month-to-month variability of the second dataset. Even higher posterior values for these discrepancy terms along with distinct changes of the moments (shifted mean values/increased standard deviation) of

Table 5
Markov chain Monte Carlo estimates of the model stochastic node mean values and standard deviations after the second updating

Parameter	Dataset 1		Dataset 2			
			Priors #1		Priors #2	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
γ	0.057	0.026	0.060	0.028	0.070	0.049
m_p	0.069	0.021	0.069	0.019	0.076	0.034
m_b	0.088	0.005	0.093	0.010	0.084	0.014
m_z	0.051	0.023	0.051	0.024	0.048	0.031
NH	7.944	2.857	7.966	2.958	8.252	4.268
AH	4.204	1.033	4.386	1.076	4.699	1.748
$\mu_{max(phyt)}$	3.031	0.353	3.063	0.370	3.236	0.628
DH	7.199	2.852	7.244	2.548	6.213	3.153
I_k	0.412	0.044	0.434	0.054	0.448	0.081
k_{back}	0.184	0.015	0.184	0.017	0.178	0.021
k_{chla}	0.030	0.008	0.030	0.008	0.029	0.011
ψ	1.775	0.610	1.763	0.601	1.924	0.924
$\mu_{max(bact)}$	1.341	0.082	1.431	0.125	1.327	0.169
g_{max}	0.581	0.111	0.690	0.135	0.671	0.169
K_Z	62.16	20.53	56.00	18.32	60.70	24.81
K_R	13.32	2.057	14.64	2.689	13.44	3.618
k_{miner}	0.110	0.015	0.114	0.020	0.110	0.024
σ_{NH4}	1.241	0.405	2.257	0.663	2.351	0.738
σ_{NO3}	3.867	0.626	4.057	0.668	4.135	0.780
σ_{DON}	22.94	5.428	30.54	8.025	32.83	9.118
σ_{chla}	16.79	3.223	18.84	3.425	19.65	4.021
σ_{zoop}	14.55	2.950	14.51	3.005	15.04	3.377
σ_{bact}	0.238	0.266	1.978	1.156	2.346	0.966
$NH4_{(t0)}$	17.18	0.853	17.31	0.875	17.30	0.881
$NO3_{(t0)}$	4.283	0.217	4.292	0.220	4.287	0.221
$DON_{(t0)}$	91.71	4.551	92.26	4.771	92.31	4.708
$PHYT_{(t0)}$	9.156	0.460	9.158	0.468	9.144	0.481
$ZOOP_{(t0)}$	3.442	0.173	3.442	0.170	3.442	0.175
$BACT_{(t0)}$	16.47	0.841	16.53	0.846	16.51	0.843

the seventeen marginal posterior parameter distributions were sampled from the wider prior parameter space (Priors #2). Moreover, the second dataset resulted in reduced CV value of the bacteria error which is probably evidence that the previously reported high value was mainly driven by the low bacteria biomass variability in the system. The performance of the updated model did not differ between the two datasets, and the results were qualitatively similar to those reported from the Bayesian calibration of the original Model 2; i.e., the mean values of the predictive monthly distributions were again close to the observed values of the six state variables, and the high late-summer ammonium concentration was the only one not included within the 95% credible intervals.

5. Conclusions

This paper addresses the urgent need for novel methodological tools that can rigorously assess the predictive uncertainty of aquatic biogeochemical models

(Flynn, 2005). The proposed framework aims to combine the advantageous features of both mechanistic and statistical approaches. Models that are based on mechanistic understanding yet remain within the bounds of data-based parameter estimation. The mechanistic foundation improves the confidence in predictions made for a variety of conditions, while the statistical methods provide an empirical basis for parameter estimation that can accommodate thorough error analysis (Borsuk et al., 2001). The Bayesian nature of our framework also uses past experience from the system along with present ecological information to project future ecosystem response. Thus, our hypothesis is that the Bayesian techniques are more informative than the conventional model calibration practices (i.e., mere adjustment of model parameters until the discrepancy between model outputs and observed data is minimized) and can be used to refine our knowledge of model input parameters, and obtain predictions along with credible intervals for modeled output variables.

We examined the efficiency of our Bayesian framework to elucidate the propagation of uncertainty arising from unknown calibration parameters, error-contaminated measurements, spatial variability, and mis-specified initial conditions, model structure or external forcing functions. The implementation of the Bayesian calibration was based on a fairly straightforward MCMC algorithm (general normal-proposal Metropolis) to efficiently sample the joint probability distribution of the multiple stochastic nodes of our plankton model. We found that 30,000 MCMC samples gave adequate summary statistics of the marginal posterior parameter distributions and the predictive model outputs. These results are in accordance with several modeling studies from a variety of disciplines that advocated the use of MCMC schemes for sampling high dimensional parameter spaces and multivariate outputs (Hegstad and More, 2001; Lee et al., 2002; Van Oijen et al., 2005). On the other hand, MCMC appropriateness for overly complex models has not been explored yet in the modeling literature, and it is argued that different Bayesian (or Bayesian-like) methodologies (GLUE, Sampling/Importance Resampling) are perhaps conceptually better suited to high dimensional input spaces with significant equifinality problems (see discussion on the paper by Kennedy and O'Hagan, 2001). Future research should involve the examination of the MCMC suitability for more complex models (≥ 15–20 state variables) extensively used in environment management.

The prior specification of the error structure can strongly influence our ability to gain insight into the
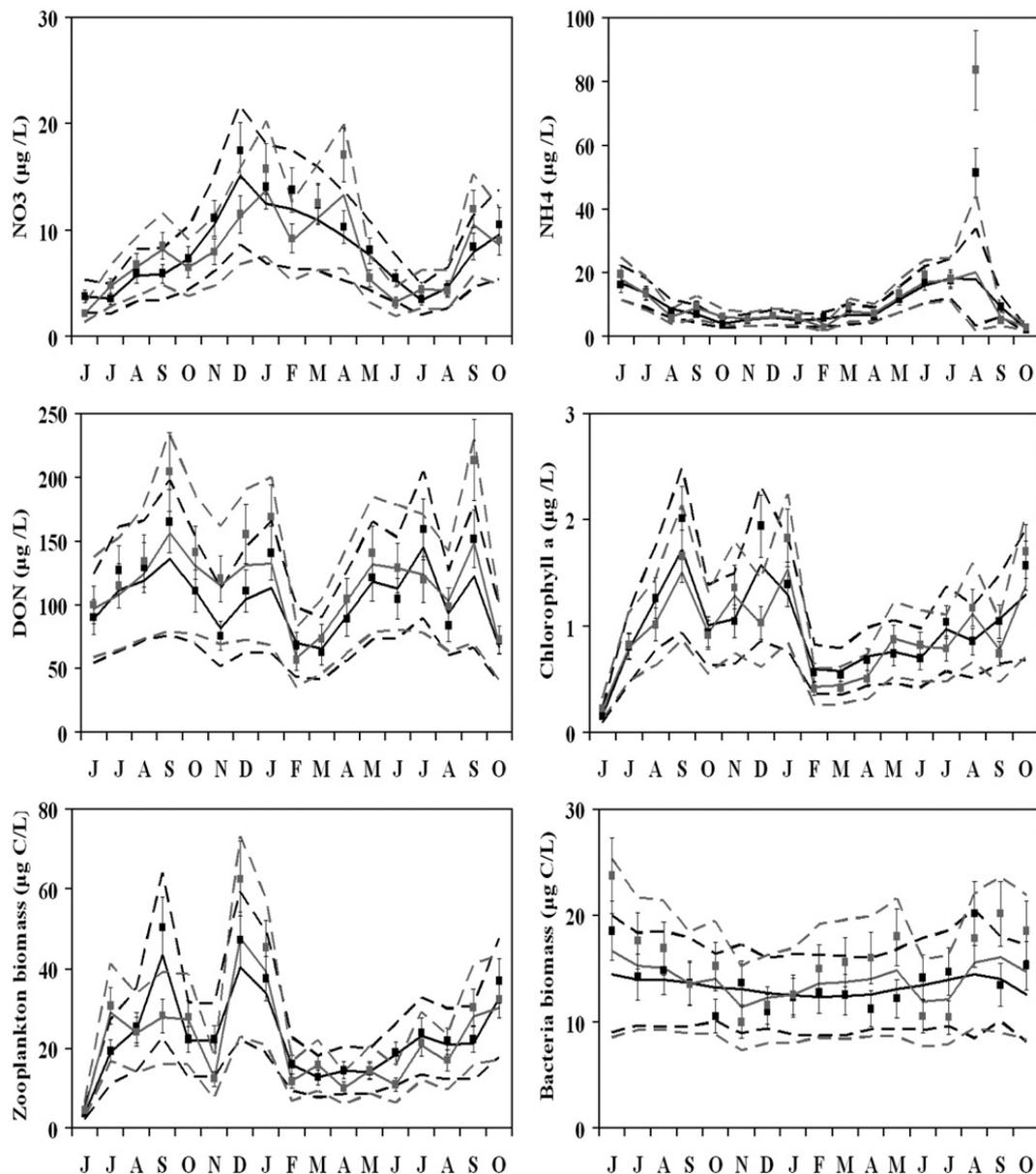
Fig. 7. Comparison between the two datasets (1: black lines and 2: grey lines) and posterior predictive monthly distributions for ammonium, nitrate, dissolved organic nitrogen, chlorophyll *a*, bacteria, and zooplankton biomass based on 12,500 Markov chain Monte Carlo posterior samples from the updated Model 2. Bayesian calibration results with the second dataset are based on the first set of prior parameter distributions (Priors #1).

degree of information the data contain about model inputs and can significantly alter the predictive outputs, i.e., mean predictions and credible intervals. Under the assumption of a "perfect" ecological model, the main sources of error considered in our analysis were the observation error and the uncertainty pertaining to exogenous nitrogen loading. The lower dimensions of the sampled space resulted in narrow peaked distributions (good information) for most of the parameters, but the problematic postulation of an ideal simulator was the main reason for the relatively poor representation of planktonic patterns in the coastal embayment. While the usual antidote for alleviating model structure imperfec-

tions is the increase of complexity (e.g., inclusion of ecological processes), we introduced an alternative approach that explicitly acknowledges the mismatch between model and natural system. This configuration can be useful in cases where the lack of information cannot reliably guide the model building process. The posterior distribution of the seasonally (in)-variant discrepancy terms is an indicator of how well the simulator is matching reality, and their inclusion has significantly improved the representation of the observed system dynamics from the predictive monthly distributions of the six state variables. In this study, the presence of the discrepancy terms also made the

interpretation of the posterior parameter distributions difficult, and the knowledge gained for the majority of the parameters was limited relative to the prior specification. In a general context, however, our experience has been that the effects of the discrepancy terms on the parameter posteriors can be quite variant depending on the prior model specification and the system being modeled; the latter assertion has also been supported by other recent studies (Higdon et al., 2004; see the results reported in their Fig. 3 and 11).

The Bayesian nature of our framework also offers a natural mechanism for sequentially updating our beliefs regarding model inputs and structure. In this study, we designed a second "training" of the model by generating two datasets representing similar average conditions with different temporal variability. Not surprisingly, aside from minor adjustments of the parameter posteriors, we found that ecological information relatively similar to the one used for the original model calibration did not cause significant alterations of the updated model. On the other hand, the consideration of data with more pronounced temporal variability mainly changed (inflated) the discrepancy term posteriors instead of the model input parameters; namely, the terms that reflect the model inadequacy and not the mathematical model itself were used to accommodate the differences in the calibration data. The latter result does not fully satisfy the basic premise of our framework to attain realistic ecological forecasts in the extrapolation domain (e.g., different nutrient loading conditions) while gaining insights into the ecosystem dynamics. Although it is probably driven by the "noisy" character of the dataset produced for this exercise, the relative importance of the discrepancy terms vis-à-vis mathematical model for extrapolative tasks invites further examination. Finally, regarding the simulation time, we found that the BUGS-language specification of the updated model took approximately half of the time required to run the original model (i.e., around 3 h on a 3.0 GHz PC machine). Hence, the commonly proposed compromise between "fidelity of the simulator" and "simulation speed" is probably not always necessary (Higdon et al., 2004); at least not for intermediate complexity models (≤ 10 state variables).

The increase of the articulation level of aquatic biogeochemical models is certainly the way forward; thorny environmental issues, such as the biotic response to climate change or the objective evaluation of management alternatives, require robust projections out of the operational model domain that cannot be accomplished without the explicit treatment of multiple biogeochemical cycles or the increase of the functional diversity of biotic communities. A major challenge of the aquatic ecosystem community is to develop a prudent strategy that can assist the evolution of the current oversimplified abstractions to more sophisticated diagnostic/prognostic modeling constructs. After several decades of modeling practice, aquatic ecosystem modelers are realizing the difficulties to forecast ecosystem behaviour; even in well-studied, data-rich systems using very sophisticated models, accurate predictions are not feasible (Arhonditsis and Brett, 2004). Differentiating the predictable from the unpredictable patterns and increasing model complexity accordingly requires careful consideration and should be tightly coupled with critical evaluation of the model outputs. In this context, our Bayesian framework is more consistent with the scientific process of progressive learning, and can be particularly useful for quantifying the uncertainty associated with model predictions.

Acknowledgments

References

Anderson, T.R., 2005. Plankton functional type modelling: running before we can walk? J. Plankton Res. 27, 1073–1081.

Anderson, T.R., 2006. Confronting complexity: reply to Le Quere and Flynn. J. Plankton Res. 28, 877–878.

Arhonditsis, G.B., Brett, M.T., 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. Mar. Ecol. Prog Ser. 271, 13–26.

Arhonditsis, G., Tsirtsis, G., Angelidis, M., Karydis, M., 2000. Quantification of the effects of non-point nutrient sources to coastal marine eutrophication: applications to a semi-enclosed gulf in the Mediterranean Sea. Ecol. Model. 129, 209–227.

Arhonditsis, G., Giourga, C., Loumou, A., Koulouri, M., 2002a. Quantitative assessment of agricultural runoff and soil erosion using mathematical modelling: applications in the Mediterranean region. Environ. Manage. 30, 434–453.

Arhonditsis, G., Tsirtsis, G., Karydis, M., 2002b. 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.

Arhonditsis, G., Eleftheriadou, M., Karydis, M., Tsirtsis, G., 2003a. Eutrophication risk assessment in coastal embayments using simple statistical models. Mar. Pollut. Bull. 46, 1174–1178.

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

Arhonditsis, G.B., Adams-VanHarn, B.A., Nielsen, L., Stow, C.A., Reckhow, K.H., 2006. Evaluation of the current state of mechanistic aquatic biogeochemical modeling: citation analysis and future perspectives. Environ. Sci. Technol. 40, 6547–6554.

Beck, M.B., 1987. Water-quality modeling — a review of the analysis of uncertainty. Water Resour. Res. 23, 1393–1442.

Bernardo, J.M., Smith, A.F.M., 1994. Bayesian Theory. John Wiley & Sons. 573. pp.

Beven, K., Binley, A., 1992. The future of distributed models — model calibration and uncertainty prediction. Hydrol. Process. 6, 279–298.

Borsuk, M.E., Higdon, D., Stow, C.A., Reckhow, K.H., 2001. A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecol. Model. 143, 165–181.

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

Brun, R., Reichert, P., Kunsch, H.R., 2001. Practical identifiability analysis of large environmental simulation models. Water Resour. Res. 37, 1015–1030.

Chapra, S.C., Canale, R.P., 1998. Numerical Methods for Engineers, 3rd Ed. McGraw-Hill. 924 pp.

Costanza, R., Sklar, F.H., 1985. Articulation, accuracy and effectiveness of mathematical models — a review of fresh-water wetland applications. Ecol. Model. 27, 45–68.

Dilks, D.W., Canale, R.P., Meier, P.G., 1992. Development of Bayesian Monte-Carlo techniques for water-quality model uncertainty. Ecol. Model. 62, 149–162.

Doney, S.C., 1999. Major challenges confronting marine biogeochemical modeling. Glob. Biogeochem. Cycles 13, 705–714.

Doney, S.C., Glover, D.M., Najjar, R.G., 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., Part 2, Top. Stud. Oceanogr. 43, 591–624.

Edwards, A.M., Yool, A., 2000. The role of higher predation in plankton population models. J. Plankton Res. 22, 1085–1112.

Eppley, R.W., Peterson, B.J., 1979. Particulate organic-matter flux and planktonic new production in the deep ocean. Nature 282, 677–680.

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

Fasham, M.J.R., Sarmiento, J.L., Slater, R.D., Ducklow, H.W., Williams, R., 1993. Ecosystem behavior at Bermuda station-S and Ocean weather station India — a general-circulation model and observational analysis. Glob. Biogeochem. Cycles 7, 379–415.

Fennel, W., Neumann, T., 2001. Coupling biology and oceanography in models. Ambio 30, 232–236.

Flynn, K.J., 2005. Castles built on sand: dysfunctionality in plankton models and the inadequacy of dialogue between biologists and modellers. J. Plankton Res. 27, 1205–1210.

Flynn, K.J., 2006. Reply to Horizons Article 'Plankton functional type modelling: running before we can walk' Anderson (2005): II. Putting trophic functionality into plankton functional types. J. Plankton Res. 28, 873–875.

Franks, P.J.S., 2002. NPZ models of plankton dynamics: their construction, coupling to physics, and application. J. Oceanogr. 58, 379–387.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian Data Analysis. Chapman and Hall, New York, p. 518.

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

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1998. Markov Chain Monte Carlo in Practice. Chapman & Hall/CRC, p. 512.

Gregg, W.W., Ginoux, P., Schopf, P.S., Casey, N.W., 2003. Phytoplankton and iron: validation of a global three-dimensional ocean biogeochemical model. Deep-Sea Res., Part 2, Top. Stud. Oceanogr. 50, 3143–3169.

Hegstad, B.K., More, H., 2001. Uncertainty in production forecasts based on well observations, seismic data, and production history. Soc. Pet. Eng. J. 6, 409–424.

Higdon, D., Kennedy, M., Cavendish, J.C., Cafeo, J.A., Ryne, R.D., 2004. Combining field data and computer simulations for calibration and prediction. SIAM J. Sci. Comput. 26, 448–466.

Hong, B.G., Strawderman, R.L., Swaney, D.P., Weinstein, D.A., 2005. Bayesian estimation of input parameters of a nitrogen cycle model applied to a forested reference watershed, Hubbard Brook Watershed Six. Water Resour. Res. 41, W03007.

Hornberger, G.M., Spear, R.C., 1981. An approach to the preliminary analysis of environmental systems. J. Environ. Manag. 12, 7–18.

Ignatiades, L., Karydis, M., Vounatsou, P., 1992. A possible method for evaluating oligotrophy and eutrophication based on nutrient concentration scales. Mar. Pollut. Bull. 24, 238–243.

Jassby, A.D., Platt, T., 1976. Mathematical formulation of relationship between photosynthesis and light for phytoplankton. Limnol. Oceanogr. 21, 540–547.

Jorgensen, S.E., Nielsen, S.E., Jorgensen, L.A., 1991. Handbook of Ecological Parameters and Ecotoxicology. Elsevier Publications. 1264 pp.

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

Kennedy, C.M., O'Hagan, A., 2001. Bayesian calibration of computer models. J. R. Stat. Soc., Ser. B Stat. Methodol. 63, 425–464.

Le Quere, C., 2006. Reply to Horizons Article 'Plankton functional type modelling: running before we can walk' Anderson (2005): I. Abrupt changes in marine ecosystems? J. Plankton Res. 28, 871–872.

Lee, H.K.H., Higdon, D.M., Bi, Z.O., Ferreira, M.A.R., West, M., 2002. Markov random field models for high-dimensional parameters in simulations of fluid flow in porous media. Technometrics 44, 230–241.

Levins, R., 1966. Strategy of model building in population biology. Am. Sci. 54, 421–431.

Moore, J.K., Doney, S.C., Kleypas, J.A., Glover, D.M., Fung, I.Y., 2002. An intermediate complexity marine ecosystem model for the global domain. Deep-Sea Res., Part 2, Top. Stud. Oceanogr. 49, 403–462.

Neal, R., 1998. Suppressing random walks in Markov chain Monte Carlo using ordered over-relaxation. In: Jordan, M.I. (Ed.), Learning in Graphical Models. Kluwer Academic Publishers, Dordrecht, pp. 205–230.

Omlin, M., Reichert, P., 1999. A comparison of techniques for the estimation of model prediction uncertainty. Ecol. Model. 115, 45–59.

Pappenberger, F., Beven, K.J., 2006. Ignorance is bliss: or seven reasons not to use uncertainty analysis. Water Resour. Res. 42, WO5302.

Reichert, P., Omlin, M., 1997. On the usefulness of overparameterized ecological models. Ecol. Model. 95, 289–299.

Shaddick, G., Wakefield, J., 2002. Modelling daily multivariate pollutant data at multiple sites. J. R. Stat. Soc., Ser. C, Appl. Stat. 51, 351–372.

Six, K.D., Maier-Reimer, E., 1996. Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model. Glob. Biogeochem. Cycles 10, 559–583.

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

Steinberg, L.J., Reckhow, K.H., Wolpert, R.L., 1997. Characterization of parameters in mechanistic models: a case study of a PCB fate and transport model. Ecol. Model. 97, 35–46.

Van Oijen, M., Rougier, J., Smith, R., 2005. Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiol. 25, 915–927.

Wrobleski, J., 1977. A model of phytoplankton plume formation during Oregon upwelling. J. Mar. Res. 35, 357–394.