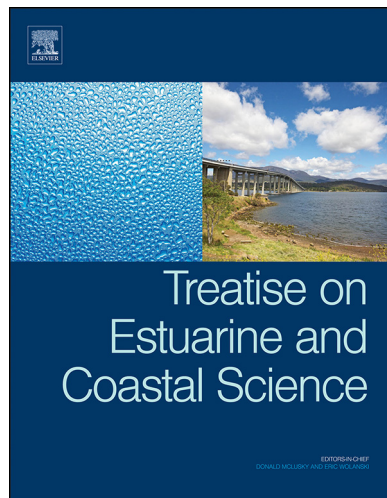


Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

This chapter was originally published in *Treatise on Estuarine and Coastal Science*, published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

Arhonditsis G, Stremilov S, Gudimov A, Ramin M, and Zhang W (2011) Integration of Bayesian Inference Techniques with Mathematical Modeling. In: Wolanski E and McLusky DS (eds.) *Treatise on Estuarine and Coastal Science*, Vol 9, pp. 173–192. Waltham: Academic Press.

© 2011 Elsevier Inc. All rights reserved.

9.10 Integration of Bayesian Inference Techniques with Mathematical Modeling

G Arhonditsis, S Stremilov, A Gudimov, and M Ramin, University of Toronto, Toronto, ON, Canada
W Zhang, Environment Canada, Burlington, ON, Canada

© 2011 Elsevier Inc. All rights reserved.

9.10.1	Evaluation of the Current State of Aquatic Biogeochemical Modeling: Where Are We?	173
9.10.2	Why Bayesian Calibration?	177
9.10.3	A Case Study: Eutrophication Risk Assessment in Hamilton Harbour	179
9.10.3.1	Model Description	180
9.10.3.1.1	Model spatial structure and forcing functions	180
9.10.3.1.2	Equations	181
9.10.3.1.3	Phytoplankton	181
9.10.3.1.4	Zooplankton	183
9.10.3.1.5	Nitrogen cycle	183
9.10.3.1.6	Phosphorus cycle	183
9.10.3.1.7	Fluxes from the sediment	183
9.10.3.2	Bayesian Framework	183
9.10.3.2.1	Statistical formulation	183
9.10.3.2.2	Prior parameter distributions	184
9.10.3.2.3	Numerical approximations for posterior distributions	184
9.10.3.3	Model Results and Prediction of the Frequency of Water-Quality Standard Violations	184
9.10.3.3.1	Exceedance frequency and confidence of compliance with water-quality standards	187
9.10.4	Conclusions and Future Perspectives	189
References		190

Abstract

The credibility of the scientific methodology of numerical models and their adequacy to form the basis of public policy decisions have been frequently challenged. The first part of this chapter aims to address the issue of model reliability by evaluating the current state of aquatic biogeochemical modeling. We provide evidence that there is still considerable controversy among modelers and the resource managers about how to develop, evaluate, and interpret mathematical models. Our arguments are that (1) models are not always developed in a consistent manner, clearly stated purpose, and predetermined acceptable model performance level, and (2) the potential users select models without properly assessing their technical value. The second part of this presentation argues that the development of novel methods for rigorously assessing the uncertainty underlying model predictions should be a top priority of the modeling community. Striving for novel uncertainty analysis tools, we introduce Bayesian calibration of process-based models as a methodological advancement that warrants consideration in aquatic ecosystem research. This modeling framework combines the advantageous features of both process-based and statistical approaches, that is, mechanistic understanding that remains within the bounds of data-based parameter estimation. The incorporation of mechanism improves the confidence in predictions made for a variety of conditions, whereas the statistical methods provide an empirical basis for parameter value selection and allow for realistic estimates of predictive uncertainty. Other advantages of the Bayesian approach include the ability to sequentially update beliefs as new knowledge is available, and the consistency with the scientific process of progressive learning and the policy practice of adaptive management. Finally, we illustrate some of the anticipated benefits from the Bayesian calibration framework, well suited for stakeholders and policy makers when making environmental management decisions, using the Hamilton Harbour – a eutrophic system in Ontario, Canada – as a case study.

9.10.1 Evaluation of the Current State of Aquatic Biogeochemical Modeling: Where Are We?

Mechanistic aquatic biogeochemical models have formed the scientific basis for environmental management decisions by providing a predictive link between management actions and ecosystem response. They have provided an important tool for elucidating the interactions between climate variability and the

carbon cycling in the oceans, and thus for assessing the pace and impacts of climate change (Doney, 1999; Franks, 2002). Acknowledging their central role in aquatic ecosystem research, several compelling questions arise, such as: What is the capacity of the current models to simulate the dynamics of coastal and estuarine ecosystems? How carefully do modelers develop their models? How rigorously do we assess what a model can or cannot predict? Arhonditsis and Brett (2004) attempted to

answer some of these questions by reviewing 153 modeling studies published in the literature between 1990 and 2002. Their hypothesis was that the sizable number of aquatic ecosystem modeling studies, which successfully passed the scrutiny of the peer-review process along with the experience gained from addressing an extent of management problems, can objectively reveal the systematic biases, methodological inconsistencies, and common misconceptions characterizing the field of aquatic ecosystem modeling. Indeed, despite the heterogeneity of the modeling studies examined with respect to model complexity, type of ecosystem modeled, spatial and temporal scales, and model development objectives, this study was able to detect statistically significant trends of the model performance and to pinpoint methodological omissions in the current modeling practice.

The first interesting finding was the absence of systematic goodness-of-fit assessment of the original models, that is, plots in which simulated values were visually compared with observed data were only presented for 16.8% of the model endpoints, and even fewer (1.3%) were the cases in which thorough statistical examination of the error was reported. In the cases in which measures of fit or comparison plots were presented, Arhonditsis and Brett (2004) independently assessed state variable performance as expressed by the relative error ($RE = \sum |\text{observed values} - \text{simulated values}| / \sum \text{observed values}$) and the coefficient of determination (r^2) (Table 1). It was found that temperature and dissolved oxygen had the lowest RE (median <10%) and the highest r^2 values (the respective medians were 0.93 and 0.70). The typical limiting nutrient forms (NO_3 , NH_4 , PO_4 , and Si) in freshwater and oceanic ecosystems along with the phytoplankton biomass had intermediate fit, with median r^2 values varying from 0.40 to 0.60 and the median RE lying around the 40% level. Zooplankton dynamics were characterized by the highest RE (70%) and the widest range of r^2 (interquartile range ~0.8) and

RE (interquartile range ~85%) values. Similarly, bacteria were also poorly modeled (median r^2 value <0.06), indicating that the performance of existing mechanistic aquatic biogeochemical models declines as we move from physical–chemical to biological components of planktonic systems. On a positive note, it was found that these results were obtained without the introduction of a major ‘calibration bias’, that is, in the process of maximizing the fit for a specific state variable (usually phytoplankton biomass), the modelers did not seem to compromise on the fit for other state variables (such as limiting nutrient concentrations or herbivorous zooplankton biomass).

Arhonditsis and Brett (2004) also assessed the effects of model complexity (expressed as the number of state variables), spatial dimension (from zero- to three-dimensional models), simulation period (from days to decades), and ecosystem type on model performance. The study reported a positive correlation between the number of state variables and the RE values for the different model outputs ($r = 0.219$, $p < 0.001$). This (counterintuitive) positive trend was even stronger when considering the RE values for phytoplankton ($r = 0.248$, $p = 0.003$) and zooplankton ($r = 0.626$, $p < 0.001$) biomass, suggesting that more complex models usually result in slightly poorer model performance. It should be noted, however, that the majority of the complex models considered in this analysis belonged to the European Regional Seas Ecosystem Model (ERSEM) family, and therefore the reported complexity–performance relationship was influenced by the development purposes, modeled environments and practices, followed by this particular family of models (Baretta et al., 1995). Similarly, a (very weak) positive correlation was found between the duration of the simulation period and the state variable RE values ($r = 0.098$, $p = 0.022$), indicating that longer simulations are also increasing model misfit. Marginally significant correlations also exist between the spatial complexity of the models and their (RE values) performance trends ($r = 0.104$, $p = 0.015$).

Table 1 Performance of the aquatic biogeochemical models for the study period 1990–2002

Percentile	Temperature	Dissolved oxygen	Nitrate	Ammonium	Phosphate	Silicate	Phytoplankton	Zooplankton	Bacteria
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 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 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 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 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 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 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

Coefficient of determination (r^2) and relative error (RE%) values for: temperature, dissolved oxygen, nitrate, ammonium, phosphate, silicate, phytoplankton, zooplankton, and bacteria. Adapted from Arhonditsis, G.B., Brett, M.T., 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Marine Ecology Progress Series* 271, 13–26.

Model error also did not vary depending on the type of ecosystem modeled (lakes/reservoirs, embayments/lagoons, coastal ecosystems/estuaries, and oceanic systems). These results provide overwhelming evidence that the ambitious efforts to increase the level of ecological information mathematically represented by the models, to increase spatial complexity, and to use longer simulation periods have not led to a systematic improvement in model performance.

Aquatic ecosystem modelers do not also seem to consistently apply conventional methodological steps during the development of their models (Arhonditsis and Brett, 2004). The large majority of the published studies in the field over the past decade did not properly assess model sensitivity to the input vectors (Figure 1(a)); aquatic ecosystem modelers are still reluctant to embrace optimization techniques during model calibration (Figure 1(b)); and to assess the ability of their models to match the observed data during the 'model training' phase (Figure 1(c)). The same study also compiled information for three types of validation: (1) predictive validation defined as an evaluation of 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 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 and derived quantities). These validation procedures were carried out (or at least explicitly reported) in 47.1% of the aquatic ecosystem modeling studies, whereas the remaining 52.9% were not predictively or structurally validated (Figure 1(d)).

Building upon the results presented in the Arhonditsis and Brett (2004) study, Arhonditsis et al. (2006) attempted a second quantitative assessment of the current state of aquatic biogeochemical modeling by focusing on the features that primarily determine our decision to utilize a specific model. The main objectives of this analysis were (1) the evaluation of how the modeling community has received the 153 aquatic biogeochemical models published from 1990 to 2002 and (2) the identification of the characteristics of a model that are

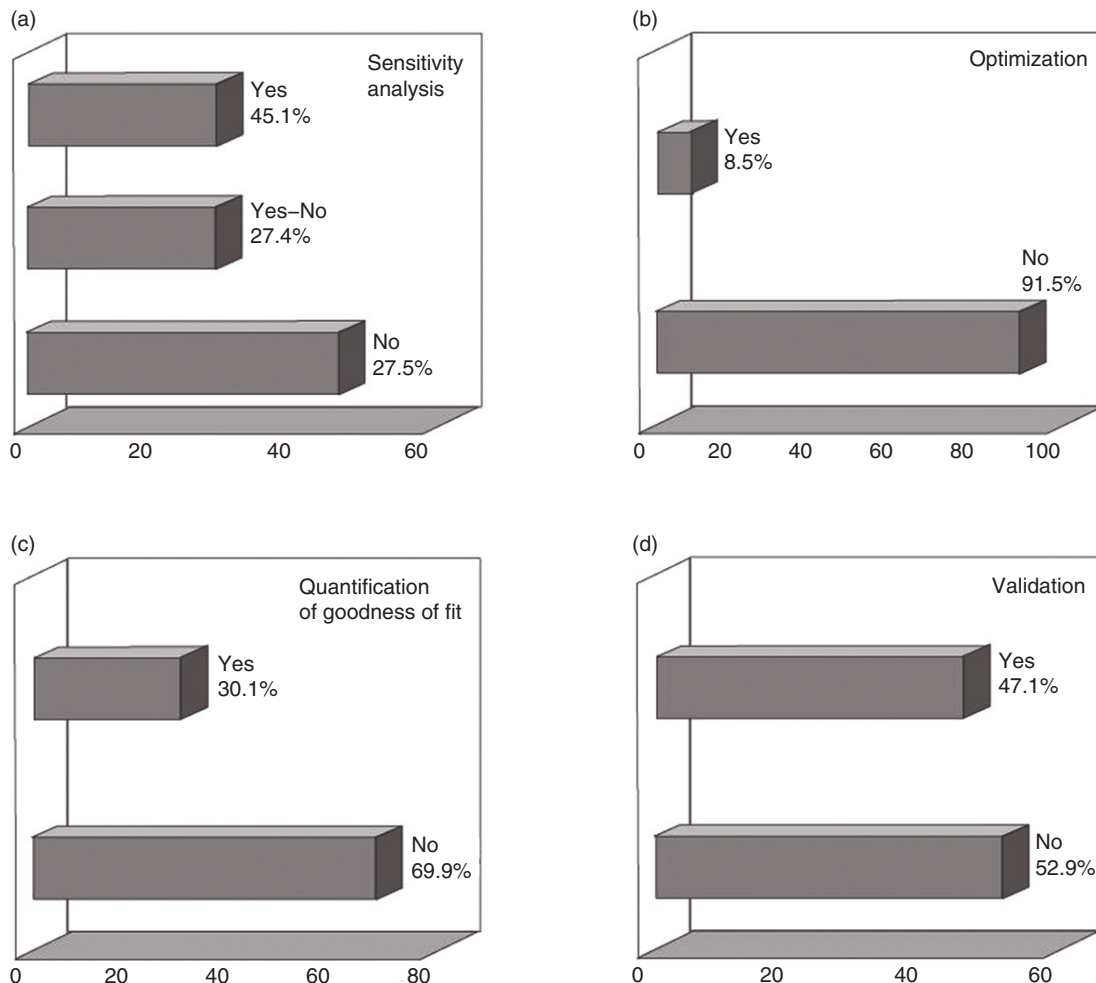


Figure 1 Proportion of aquatic biogeochemical modeling studies that (a) performed sensitivity analysis, (b) used optimization techniques for model calibration, (c) quantified the fit between model predictions and observed data, and (d) validated the models (see text for definitions of validation). The category Yes-No in (a) indicates qualitative approaches that solely tested the influence of certain model structures (e.g., alternative formulations) or parameters without providing any quantitative measures of model response to the perturbations induced to the input vector. Adapted from Arhonditsis, G.B., Brett, M.T., 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. Marine Ecology Progress Series 271, 13–26.

more attractive to the potential users and may influence the frequency of its use and subsequent citation. Citation rates are increasingly recognized as a convenient measure for assessing the importance and utility of scientific research; ideally, high-quality papers should motivate future research and should be used as a source of information by subsequent studies in the field (Adams, 2002; Leimu and Koricheva, 2005). The articles citing aquatic biogeochemical modeling studies were classified in 60 different disciplines. Oceanography was found to be the most popular thematic area of the articles that cite mechanistic

aquatic biogeochemical models and more than 27% (~1500 counts) of the total citations were related to this research topic (Figure 2(a)). The second most popular subject category was marine and freshwater biology (21.75%), followed by geosciences (10.90%) and ecology (10.33%). Disciplines more closely related to environmental management, for example, environmental sciences (8.21%), environmental engineering (2.37%), and water resources (1.81%), received a relatively low proportion of the total citations. The same study also examined the association between citation rates and several

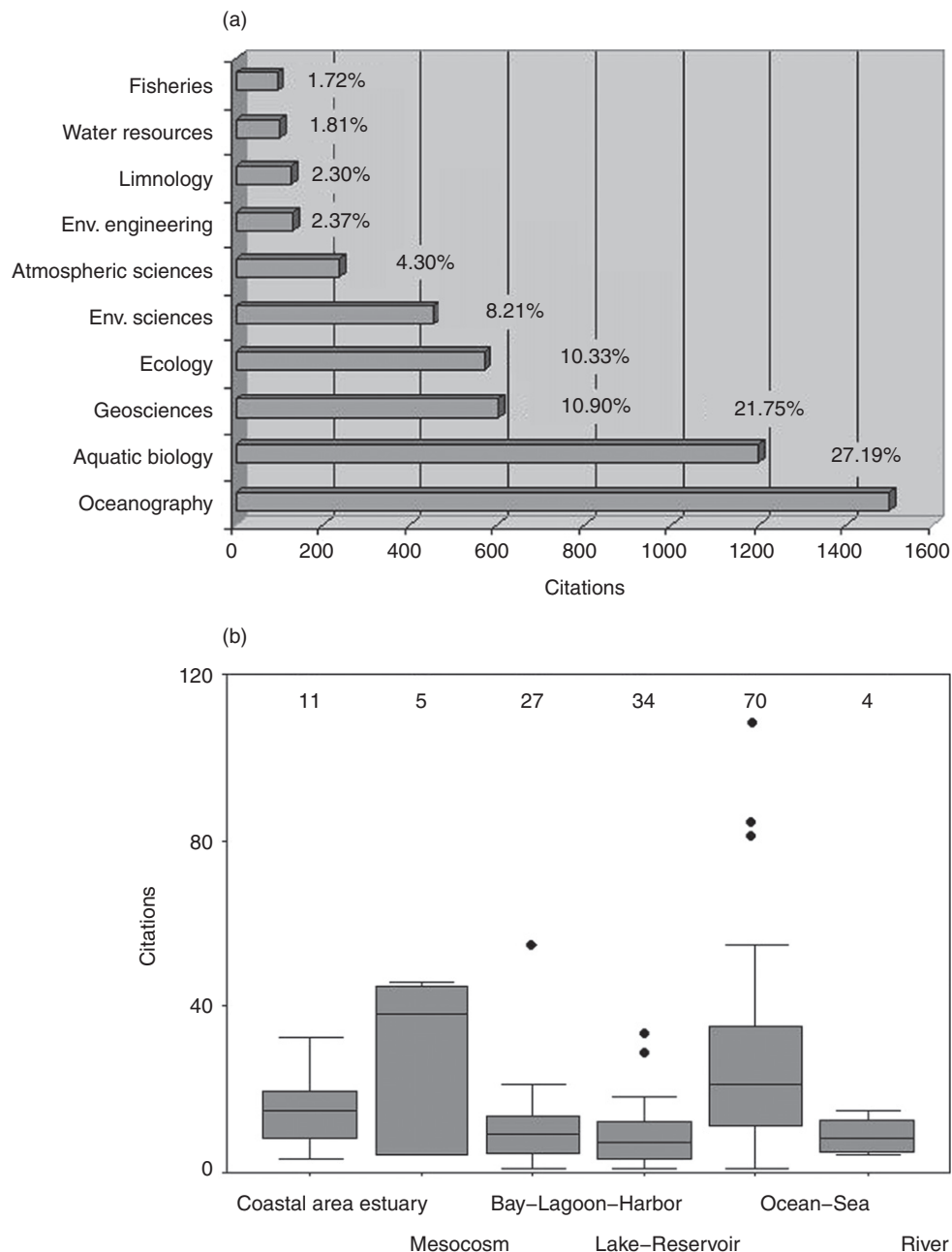


Figure 2 (a) Frequency histogram of the scientific classification (subject category) of the papers that cite mechanistic aquatic biogeochemical modeling papers. (b) Citation frequency for different types of modeled ecosystem. Adapted from Arhonditsis, G.B., Adams-VanHarn, B.A., Nielsen, L., Stow, C.A., Reckhow, K.H., 2006. Evaluation of the current state of mechanistic aquatic biogeochemical modeling: citation analysis and future perspectives. *Environmental Science and Technology* 40, 6547–6554.

characteristics of the published modeling studies: ecosystem type, methodological consistency, and model performance. Using the type of the ecosystem modeled as a criterion, the published modeling studies were classified into six categories: coastal area estuary, mesocosm, bay-lagoon-harbor, lake-reservoir, ocean-sea, and river. Ocean modeling studies have received significantly higher citations ($F=7.87$, $df=5$, $p<0.001$) among the various ecosystem types (Figure 2(b)).

Regarding the model complexity (expressed as the number of state variables) as a model feature that attracts citations, it was found that the citation rates of the individual articles were not significantly correlated with the corresponding model complexity ($r=0.111$, $p=0.183$), although there was an increasing citation trend for models with over 40 state variables (Figure 3(a)). Model performance was also not considered as a criterion for citing modeling papers; for example, citation rates and model performance for the key state variable phytoplankton were not significantly correlated ($r=0.163$, $p=0.07$; Figure 3(b)). It was also examined whether the methodological consistency of the published modeling studies is a factor that determines their citation rates. The citation counts did not differ significantly among studies whether they presented (thorough/partial) sensitivity analysis or not ($F=1.16$, $df=2$, $p=0.316$). The citation patterns of the modeling studies were not affected by whether the modelers reported assessment of the goodness of fit ($F=0.05$, $df=1$, $p=0.943$; Figure 3(c)), and by whether the original modeling studies presented structural or predictive validation ($F=2.03$, $df=1$, $p=0.156$).

Overall, after four decades of experience in the field of aquatic biogeochemical modeling, there is evidence that the current generation of models can sufficiently represent the physical and chemical components of planktonic food webs, whereas our ability to obtain accurate simulations is reduced as we move from the first to the higher trophic levels. There is also notable methodological inconsistency regarding the steps followed during the development stages of the models; that is, conventional modeling procedures, such as sensitivity analysis, validation, or even assessment of goodness of fit, have not been applied in a high proportion of the published modeling studies. The citations of the aquatic biogeochemical modeling papers are determined neither by the model complexity nor by the reported performance and methodological consistency. The type of the ecosystem being modeled is the most influential factor shaping the citation patterns of the modeling papers. Ocean modeling studies receive considerable attention and overwhelmingly dominate the total citation counts. Model application for addressing environmental management issues on a local scale faces challenges as a scientific tool. The number of published studies from lakes, reservoirs, coastal embayments, estuaries, and harbors combined was approximately equal to the number of oceanic applications, whereas the majority of these – local character – modeling studies have received fairly low citations. Arhonditsis et al. (2006) highlighted the failure to engage novelty and creativity when attempting to provide solutions to management problems as a main reason for the low citation rates, which inevitably results in unattractive modeling products that cannot export knowledge to other disciplines. Evidently, some of the current modeling practices have to change, as the need for robust modeling tools to assist with the restoration of impaired

water bodies is more pressing now than ever before; for example, the costly implementation of total maximum daily loads for pollutants during the next 5–10 years has raised the bar for innovative model developments that can accommodate rigorous error analysis (Borsuk et al., 2002).

9.10.2 Why Bayesian Calibration?

The importance of investigating the effects of uncertainty on mathematical model predictions has been extensively highlighted in the modeling literature (Omlin and Reichert, 1999; Brun et al., 2001; Reichert et al., 2002; Chen et al., 2007). The question of model credibility is important because models are used to identify polluters, to direct the use of research funds, and to determine management strategies that have considerable social and economic implications (Chapra, 1997). Erroneous model outputs and failure to account for uncertainty could produce misleading results and misallocation of limited resources during the costly implementation of alternative environmental management schemes. For better model-based decision making, the uncertainty in model projections must be reduced, or at least explicitly acknowledged, and reported in a straightforward way that can be easily used by policy planners and decision makers (Reckhow and Chapra, 1999). One weakness of the conventional calibration strategies, that is, mere adjustment of model parameters until the discrepancy between model outputs and observed data is minimized, is that although they may provide the best fit to the data available at the moment, the resulting calibration vector is specific to the given data set at hand. As new data become available, the models should be recalibrated and, in the common calibration practice, there is no formal way of considering previous results. In this sense, we do not really update previous knowledge about model input parameters, but rather we make the models data set specific (Arhonditsis et al., 2007).

Another problematic aspect of the current modeling practice is that the usual calibration methods do not address the equifinality (poor model identifiability). Equifinality is the nonuniqueness of the model solutions, whereby several distinct choices of model inputs lead to the same model outputs, that is, many sets of parameters fit the data about equally well (Beven, 1993). The main reason for the equifinality problem is that the causal mechanisms/hypotheses used for understanding how the system works internally is of substantially higher order of what can be externally observed (Beck, 1987). As a result, our ability to set quantitative (or even qualitative) constraints as to what is realistic/behavioral simulation of an ecological structure and therefore the learning capacity of a model is significantly reduced. However, having a model that realistically reflects the natural system dynamics is particularly important when the model is intended for making predictions in the extrapolation domain, that is, predicting future conditions significantly different from those used to calibrate the model (Reckhow and Chapra, 1999). For example, when a water-quality model does not operate with realistic relative/absolute magnitudes of biological rates and transport processes, even if the fit between model outputs and observations is satisfactory (“good results for the wrong reasons”), its credibility to provide predictions about how the system will respond under different external nutrient loading

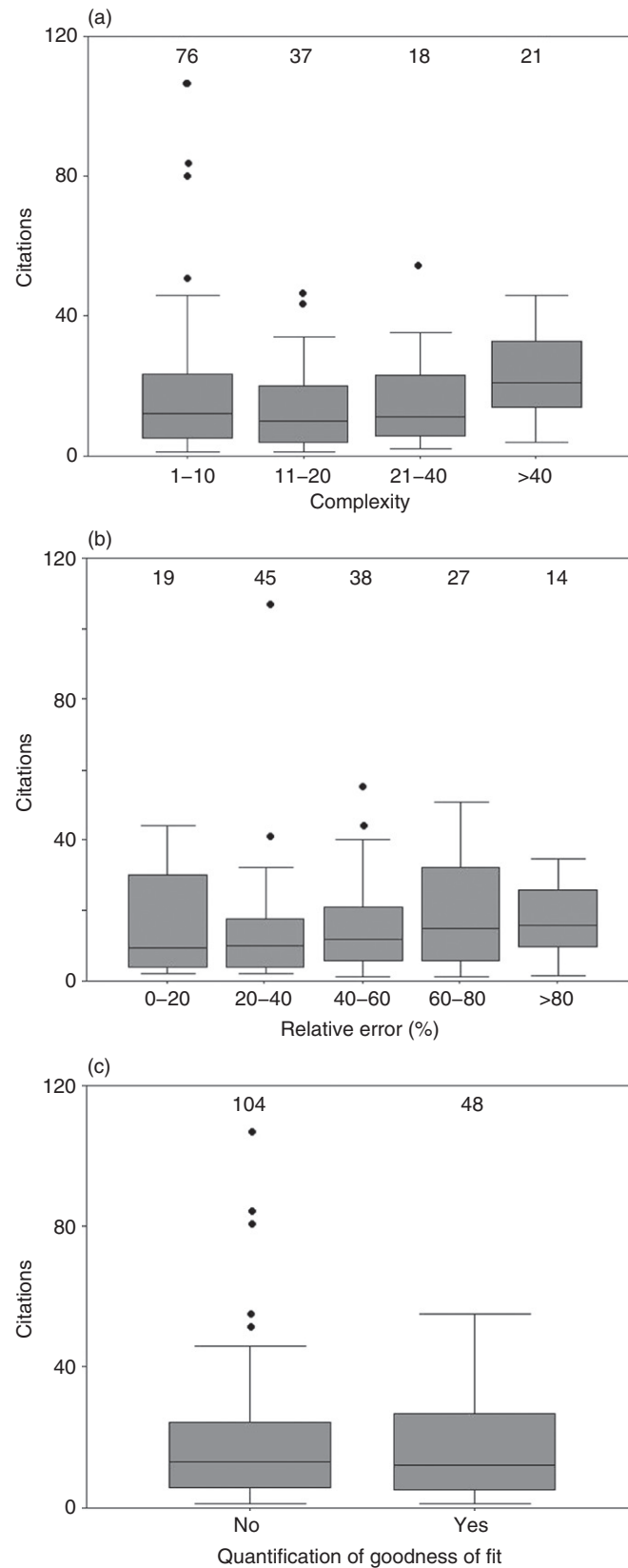


Figure 3 Citation frequency for different levels of (a) model complexity (number of state variables), (b) model performance for phytoplankton and (c) reported assessment of the goodness of fit in the original modeling study. Adapted from Arhonditsis, G.B., Adams-VanHorn, B.A., Nielsen, L., Stow, C.A., Reckhow, K.H., 2006. Evaluation of the current state of mechanistic aquatic biogeochemical modeling: citation analysis and future perspectives. *Environmental Science and Technology* 40, 6547–6554.

conditions is very limited. In this case, the application of mathematical models for extrapolative tasks is an exercise in prophecy rather than scientific action based on robust prognostic tools (Beven, 2006).

Model uncertainty analysis essentially aims to make inference about the joint probability distribution of model inputs, reflecting the amount of knowledge available for model parameters, initial conditions, forcing functions, and model structure. In this regard, Bayes' theorem provides a convenient means to combine existing information (prior) with current observations (likelihood) for projecting future ecosystem response (posterior). Hence, the Bayesian techniques are more informative than the conventional model calibration practices, and can be used to refine our knowledge of model input parameters while obtaining predictions along with uncertainty bounds for output variables (Arhonditsis et al., 2007). Despite the compelling arguments for considering Bayesian inference techniques as an integral part of the model development process, their high computational demands along with the lack of analytical expressions for the posterior distributions were until recently a major impediment for their broader application (Reichert and Omlin, 1997). Nonetheless, the advent of fast computing has allowed the development of several methods for performing Bayesian inference and the most commonly used technique is called Markov chain Monte Carlo (MCMC), a general methodology that provides a solution to the difficult problem of sampling from multidimensional distributions for the purpose of numerical integration (Gilks et al., 1996; Qian et al., 2003).

In the context of water-quality modeling, there are several recent studies illustrating how the Bayesian inference techniques can improve model forecasts and management actions over space and time. For example, Malve et al. (2005) showed how the Bayesian parameter estimation of a dynamic nonlinear model can be used to quantify the winter respiration rates (oxygen depletion per unit area of hypolimnetic surface) in a hypereutrophic shallow Finnish lake. A conceptually similar modeling approach was also used to elucidate the confounded bottom-up and top-down effects on the phytoplankton community structure of the shallow, mesotrophic Lake Pyhäjärvi (Malve et al., 2007). Arhonditsis et al. (2007, 2008a, 2008b) introduced a Bayesian calibration scheme using simple mathematical models (<10 state variables) and statistical formulations that explicitly accommodate measurement error, parameter uncertainty, and model structure error; this framework was then used to quantify the information that the data contain about model inputs, to offer insights into the covariance structure among parameter estimates, and to obtain predictions along with credible intervals for model outputs. A recent study also integrated the Bayesian calibration framework with a complex aquatic biogeochemical model simulating multiple elemental cycles and functional plankton groups to illustrate how the Bayesian parameter estimation can be used for assessing the exceedance frequency and confidence of compliance with different water-quality criteria (Zhang and Arhonditsis, 2008). In this chapter, we present another exercise in which Bayesian inference techniques are used to calibrate a fairly complex eutrophication model. Our case study is the Hamilton Harbour, an eutrophic system in the Province of Ontario, Canada, where the anticipated benefits from the Bayesian framework can be used from stakeholders and policy

makers to guide the use of millions of dollars of restoration and to dictate the best management practices.

9.10.3 A Case Study: Eutrophication Risk Assessment in Hamilton Harbour

Hamilton Harbour, a large embayment located at the Western Lake Ontario, has a long history of eutrophication problems primarily manifested as excessive algal blooms, low water transparency, predominance of toxic cyanobacteria, and low hypolimnetic oxygen concentrations during the late summer (Charlton, 1997; Hiriart-Baer et al., 2009). Since the mid-1980s, when the Harbour was identified as an area of concern (AOC) by the Water Quality Board of the International Joint Commission, the significant reductions of the exogenous nutrient loading have resulted in reduced total phosphorus (TP) concentrations and improved water clarity, which in turn has triggered aquatic macrophyte resurgence in most areas with Secchi disk depth greater than 3 m (Charlton and Le Sage, 1996). Yet, the system still receives substantial loads of phosphorus, ammonia, and suspended solids from the Burlington and Hamilton sewage treatment plants, while population growth and increasing urbanization accentuate the need for expansion of the wastewater treatment facilities (Charlton, 1997). As a result, the Hamilton Harbour Remedial Action Plan (RAP) formulated through a wide variety of government, private sector, and community participants to provide the framework for actions aimed at restoring the Harbour environment, calls for remedial measures and further reductions in sewage discharges (Charlton, 2001).

The development of appropriate nutrient load reductions and the setting of water-quality goals in the Hamilton Harbour have been founded upon an ecosystem approach that considers the complex interplay among abiotic parameters and biotic components pertinent to the beneficial uses of the Harbour (Hamilton Harbour Technical Team – Water Quality, 2007). In particular, the technical team used an analytically rigorous approach that involved data analysis, modeling, and expert judgment to determine that the respective phosphorus loading and TP concentration targets should ultimately be set at 142 kg d^{-1} and $17 \mu\text{g l}^{-1}$, whereas the environmental goals related to chlorophyll *a* concentrations ($5\text{--}10 \mu\text{g l}^{-1}$) and Secchi disk depth (3.0 m) emerged through a consensus on what was desirable and/or achievable targets for the Harbour (Charlton, 2001). The implementation of these water-quality criteria requires an operational procedure to assess the compliance of the Harbour with the desirable water-quality goals. In water resources management practice, the typical approach has been to collect a limited number of samples that are then analyzed with a statistical test to infer whether the water body is still impaired or not. Using a limited number of samples for statistical inference, however, introduces uncertainty into the assessment of compliance, and the degree of this uncertainty depends on the sampling scheme used (i.e., the quality and quantity of samples collected) as well as on the degree of confidence required from the statistical test (Borsuk et al., 2002). Beyond that, the system restoration process requires predicting future compliance, after a pollutant load reduction, usually based on a water-quality model; an endeavor that is highly uncertain mainly because the relationship between

pollutant loading and receiving water body response can never be perfectly known (Arhonditsis et al., 2006). Therefore, the practical limitations of ambient monitoring programs, the uncertainty in predicting water-quality outcomes, the intrinsic variability characterizing natural ecosystem dynamics along with the measurement error suggest that it is unrealistic to forbid any standard violations at all points in the system at all times (Barnett and O'Hagan, 1997; Reckhow et al., 2005). These shortcomings underscore the importance of adopting probabilistic methodologies in the Hamilton Harbour, such as the percentile-based approach endorsed by the US Environmental Protection Agency. The latter requires that more than 10% of samples in space and time should violate a criterion before inferring that a system does not fully support its designated use (US EPA, 1997).

Our objective herein is to illustrate how the Bayesian calibration can be used for rigorously assessing the uncertainty of eutrophication models associated with model structure and parameters. The present modeling exercise will ultimately be used to estimate the critical loads for pollutants based on acceptable probabilities of compliance with different water-quality criteria (e.g., chlorophyll *a* and TP); to elucidate the interplay among different ecological mechanisms that drive the phytoplankton variability in the system; to illustrate a methodological framework that is in alignment with the policy practice of adaptive management; and ultimately to provide a realistic platform for evaluating the potential ramifications of a variety of scenarios (e.g., urbanization/land-use changes, climate change, invasion of exotic species, and alternative restoration practices) on ecosystem functioning with special emphasis on the control of phytoplankton growth in the Harbour.

9.10.3.1 Model Description

This section provides the description of the basic conceptual design of the model. The flow diagrams of the nitrogen and phosphorus cycles used in this model are depicted in Figures 4 and 5, whereas the definitions of the model parameters are given in Table 2.

9.10.3.1.1 Model spatial structure and forcing functions

We considered a two-compartment vertical segmentation representing the epilimnion and hypolimnion of the Harbour. The depths of the two boxes varied with time and were explicitly defined based on extensive field measurements for the study period 1987–2007 (Dermott et al., 2007; Hiriart-Baer et al., 2009). During the stratified period, the epilimnion was defined as the maximum depth where the water temperature varied ≤ 1 °C relative to the temperature at 0.5 m; otherwise, we assumed a box depth of 13 m and the mass exchanges between the two compartments were computed using Fick's law (Klapwijk and Snodgrass, 1985; Hamblin and He, 2003). Other external forcing functions include the solar radiation, day length, precipitation, and evaporation based on meteorological data from Environment Canada, namely, the Canadian Daily Climate Data (1996–2002) and the Canadian Climate Normals (1971–2000). Loadings of inorganic nutrients and organic matter enter the Hamilton Harbour from the following main sources: Red Hill and Grindstone creeks, combined sewer overflows (CSOs), Dofasco and Stelco steel mills, Woodward and Skyway wastewater treatment plants (WWTPs), and Cootes Paradise. Estimates of flow and nutrient loadings are based on available data from the Water Survey of Canada and the RAP loading report (Hamilton

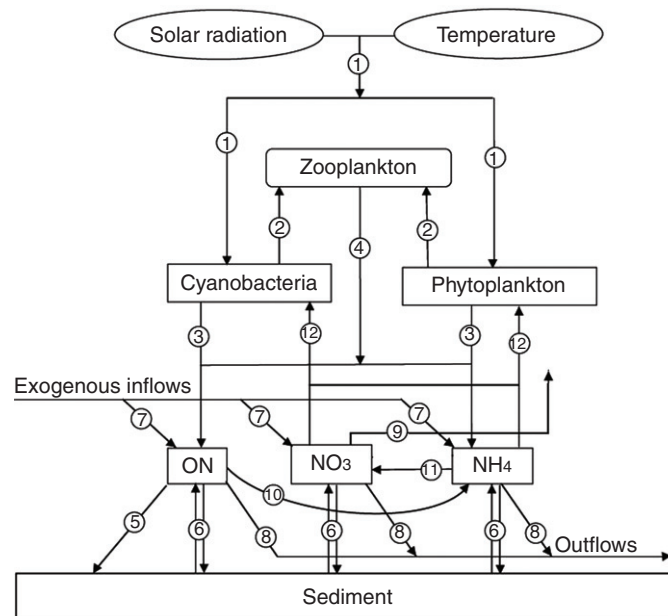


Figure 4 The nitrogen biogeochemical cycle of the model: (1) external forcing to phytoplankton growth (temperature and solar radiation); (2) zooplankton grazing; (3) phytoplankton basal metabolism excreted as NH_4 and ON; (4) zooplankton basal metabolism excreted as NH_4 and ON; (5) settling of particles; (6) water–sediment NO_3 , NH_4 , and ON exchanges; (7) exogenous inflows of NO_3 , NH_4 , and ON; (8) outflows of NO_3 , NH_4 , and ON; (9) NO_3 sinks due to denitrification; (10) ON mineralization; (11) nitrification; and (12) phytoplankton uptake.

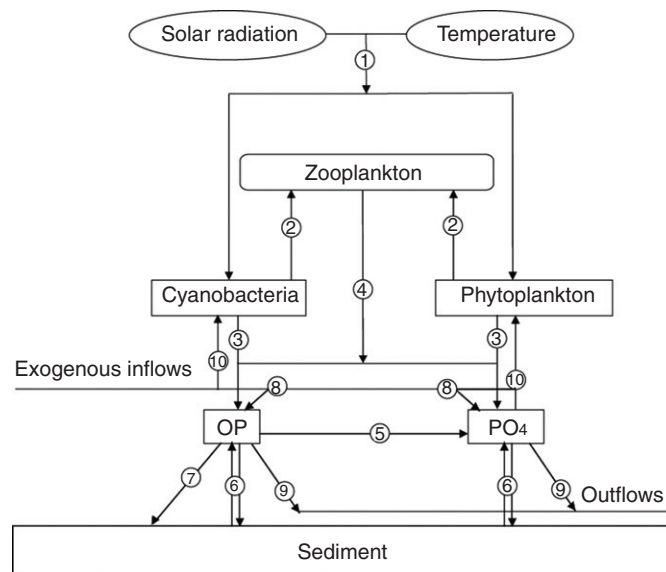


Figure 5 The phosphorus biogeochemical cycle of the model: (1) external forcing to phytoplankton growth (temperature and solar radiation); (2) zooplankton grazing; (3) phytoplankton basal metabolism excreted as PO_4 and OP; (4) zooplankton basal metabolism excreted as PO_4 and OP; (5) OP mineralization; (6) water–sediment PO_4 and OP exchanges; (7) settling of particles; (8) exogenous inflows of PO_4 and OP; (9) outflows of PO_4 and OP; and (10) phytoplankton PO_4 uptake.

Harbour Technical Team: 1996–2002 Contaminant Loadings and Concentrations to Hamilton Harbour, 2004). Similar to the Arhonditsis and Brett (2005) study, the model was run with the mean hydrological and nutrient loading annual cycle over the 1996–2002 period. The exchanges between the Hamilton Harbour and the relatively high-quality waters of Lake Ontario through the Burlington Ship Canal are another major regulatory factor of the Harbour water quality determining the dilution of the pollutant concentrations, the decrease of Harbour's residence time, and the oxygenation of the hypolimnetic waters (Barica, 1989; Hamblin and He, 2003). In particular, the winter exchanges are primarily driven by short-term oscillations due to water-level differences at the two ends of the canal, whereas the exchanges during the summer-stratified period are mediated by slowly fluctuating density gradients, that is, warm Harbour water flowing into the lake in the top layer and colder lake water flowing into the Harbour in the bottom layer (see figures 1 and 2 in Barica 1989). Existing evidence also suggests that the Hamilton Harbour–Lake Ontario interplay during the stratified conditions is much stronger and steadier than in the winter period (Hamblin and He, 2003). In this study, following the Klapwijk and Snodgrass (1985; see their figure 3) conceptual model, we assumed that 10% of the Lake Ontario inflows are directly discharged to the epilimnion, whereas 90% of the fresher oxygenated lake water replaces the hypolimnetic masses in the Harbour.

9.10.3.1.2 Equations

We developed an ecological model that considers the interactions among the eight state variables: nitrate, ammonium, phosphate, phytoplankton, cyanobacteria, zooplankton, organic nitrogen (ON), and organic phosphorus (OP).

9.10.3.1.3 Phytoplankton

The ecological submodel simulates two phytoplankton functional groups, labeled as cyanobacteria and phytoplankton, which differ with regard to their strategies for resource competition (nitrogen, phosphorus, light, and temperature) and metabolic rates as well as their morphological features (settling velocities and self-shading effects). The cyanobacteria-like group is modeled as K-strategist with low maximum growth and metabolic rate, weak P and strong N competitor, higher tolerance to low light availability, low settling velocity, and high-temperature optimum. By contrast, the more generic phytoplankton group aims to represent the rest of the phytoplankton community, having attributes of *r*-selected organisms with high maximum growth rates and higher metabolic losses, fast phosphorus and slow nitrogen kinetics, lower tolerance to low light availability, low-temperature optima, and high sinking velocities.

The governing equation for phytoplankton biomass accounts for phytoplankton production and losses due to mortality, settling, dreissenid filtration, and herbivorous zooplankton grazing. The phytoplankton growth is limited by the water temperature conditions and the nutrient and light availability. The dependence of phytoplankton growth on temperature has an optimum level (T_{opt}) and is modeled by a function similar to Gaussian probability curve (Cercio and Cole, 1994; Arhonditsis and Brett, 2005). Phosphorus dynamics within the phytoplankton cells account for luxury uptake, that is, phytoplankton nutrient uptake depends on both internal and external concentrations and is confined by maximum and minimum internal levels (Hamilton and Schladow, 1997; Arhonditsis et al., 2002). Our model explicitly considers the role of new and regenerated production using separate formulations that relate phytoplankton uptake to the ambient nitrate and ammonium concentrations (Eppley–Peterson *f*-ratio paradigm; Eppley and Peterson, 1979). Regarding the dependence of photosynthesis on solar

Table 2 Sensitivity of the posterior parameter distributions on the prior specifications

Parameters	Description	Units	Posteriors					
			Priors		Log-uniform		Lognormal 95%	
			Min.	Max.	Median	Interquartile range	Median	Interquartile range
AH _(cy)	Half-saturation constant for ammonium cyanobacteria uptake	$\mu\text{g N L}^{-1}$	30	80	76.5	5.0	120.5	21.3
AH _(phyt)	Half-saturation constant for ammonium phytoplankton uptake	$\mu\text{g N L}^{-1}$	80	150	144.1	9.8	175.9	14.4
Denitrifmax	Maximum denitrification rate	$\mu\text{g N L}^{-1} \text{d}^{-1}$	1	10	1.79	1.49	1.55	1.47
Filter _(cy)	Cyanobacteria filtering rate from dreissenids	d^{-1}	0.0045	0.0245	0.0070	0.0035	0.0070	0.0015
Filter _(phyt)	Phytoplankton filtering rate from dreissenids	d^{-1}	0.010	0.050	0.010	0.001	0.005	0.002
Growthmax _(cy)	Cyanobacteria maximum growth rate	d^{-1}	1.0	1.8	1.06	0.10	1.21	0.11
Growthmax _(phyt)	Phytoplankton maximum growth rate	d^{-1}	2.2	3.0	2.58	0.28	2.64	0.14
Ik _(cy)	Half-saturation light intensity for cyanobacteria	$\text{MJ m}^{-2} \text{d}^{-1}$	100	250	143.4	29.9	185.5	15.7
Ik _(phyt)	Half-saturation light intensity for phytoplankton	$\text{MJ m}^{-2} \text{d}^{-1}$	100	250	204.3	39.9	201.2	26.8
Kbackground	Background light extinction coefficient	m^{-1}	0.15	0.30	0.230	0.064	0.244	0.045
Kchl $a_{(cy)}$	Self-shading effect for cyanobacteria	$\text{L}(\mu\text{g chl } a\text{m})^{-1}$	0.01	0.08	0.029	0.024	0.032	0.017
Kchl $a_{(phyt)}$	Self-shading effect for phytoplankton	$\text{L}(\mu\text{g chl } a\text{m})^{-1}$	0.01	0.06	0.048	0.011	0.048	0.004
KCrefmineral	Carbon mineralization rate	d^{-1}	0.0043	0.0243	0.0059	0.0014	0.0070	0.0023
KNrefmineral	Nitrogen mineralization rate	d^{-1}	0.0043	0.0243	0.0057	0.0012	0.0051	0.0020
KPrefmineral	Phosphorus mineralization rate	d^{-1}	0.0043	0.0243	0.0062	0.0012	0.0052	0.0012
Kz	Half-saturation constant for zooplankton grazing	$\mu\text{g C L}^{-1}$	80	120	111.7	8.6	99.5	7.7
Max grazing	Zooplankton maximum grazing rate	d^{-1}	0.40	0.60	0.457	0.008	0.453	0.007
mp _(cy)	Cyanobacteria mortality rate	d^{-1}	0.01	0.05	0.019	0.005	0.015	0.003
mp _(phyt)	Phytoplankton mortality rate	d^{-1}	0.01	0.05	0.012	0.003	0.012	0.003
mz	Zooplankton mortality rate	d^{-1}	0.15	0.19	0.152	0.002	0.156	0.004
NH _(cy)	Half-saturation constant for nitrate cyanobacteria uptake	$\mu\text{g N L}^{-1}$	30	80	47.99	24.41	48.35	15.52
NH _(phyt)	Half-saturation constant for nitrate phytoplankton uptake	$\mu\text{g N L}^{-1}$	80	150	106.6	31.7	111	24.1
Nitrifmax	Maximum nitrification rate	$\mu\text{g N L}^{-1} \text{d}^{-1}$	10	30	17.19	4.26	17.89	1.06
PH _(cy)	Half-saturation constant for phosphorus cyanobacteria uptake	$\mu\text{g P L}^{-1}$	18	30	20.93	4.22	24.58	1.66
PH _(phyt)	Half-saturation constant for phosphorus phytoplankton uptake	$\mu\text{g P L}^{-1}$	5	15	12.64	3.65	13.79	1.19
Pmaxuptake _(cy)	Maximum phosphorus uptake rate for cyanobacteria	$\mu\text{g P L}^{-1} \text{d}^{-1}$	0.005	0.025	0.011	0.003	0.010	0.001
Pmaxuptake _(phyt)	Maximum phosphorus uptake rate for phytoplankton	$\mu\text{g P L}^{-1} \text{d}^{-1}$	0.01	0.05	0.020	0.007	0.022	0.001
Vsettling	Autochthonous particle settling velocity	d^{-1}	0.50	1.50	0.648	0.015	0.659	0.029
Vsettling _(biogenic)	Biogenic particle settling velocity	d^{-1}	0.05	0.25	0.068	0.028	0.076	0.016
Vsettling _(cy)	Cyanobacteria settling velocity	d^{-1}	0.01	0.05	0.037	0.017	0.042	0.011
Vsettling _(phyt)	Phytoplankton settling velocity	d^{-1}	0.12	0.25	0.185	0.028	0.196	0.022
β_N	Fraction of inert nitrogen buried into deeper sediment		0.30	0.60	0.319	0.017	0.448	0.044
β_P	Fraction of inert phosphorus buried into deeper sediment		0.80	0.95	0.897	0.013	0.868	0.015

radiation, we used Steele's equation along with Beer's law to scale photosynthetically active radiation to depth. The extinction coefficient is determined as the sum of the background light attenuation and attenuation due to chlorophyll *a* (Jassby and Platt, 1976). The phytoplankton mortality includes all internal processes that decrease algal biomass (respiration and excretion) as well as natural mortality and is assumed to increase exponentially with temperature. We also incorporated a first-order loss rate representing the filtration from the zebra and quagga mussels, which is a potentially important factor for the phytoplankton biomass levels, especially in nearshore areas (Bierman et al., 2005).

9.10.3.1.4 Zooplankton

Zooplankton grazing and losses due to natural mortality/consumption by higher predators are the two main terms in the zooplankton biomass equation. Zooplankton has three alternative food sources (the two phytoplankton groups and biogenic particulate material or detritus) grazed with preference that changes dynamically as a function of their relative proportion (Fasham et al., 1990). (It should also be noted, however, that the present model parametrization also postulates a selective zooplankton preference for phytoplankton and detritus over cyanobacteria.) Temperature-modulated zooplankton grazing was modeled using a Michaelis–Menten equation and the assimilated fraction of the grazed material 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).

9.10.3.1.5 Nitrogen cycle

There are three nitrogen forms considered in the model: nitrate (NO_3), ammonium (NH_4), and ON (Figure 4). The ammonium equation considers the phytoplankton uptake and the proportion of phytoplankton and zooplankton mortality that is returned back to the system as ammonium ions. Ammonium is also oxidized to nitrate through nitrification and the kinetics of this process is modeled as a function of the ammonium, dissolved oxygen, temperature, and light availability (Cerro and Cole, 1994; Tian et al., 2001). We used Wroblewski's model (Wroblewski, 1977) to describe ammonium inhibition of nitrate uptake. The nitrate equation also takes into account the amount of ammonium oxidized to nitrate through nitrification and the amount of nitrate lost as nitrogen gas through denitrification. The latter process is modeled as a function of dissolved oxygen, temperature, and the contemporary nitrate concentrations (Arhonditsis and Brett, 2005). The ON equation considers the contribution of phytoplankton and zooplankton mortality to the ON pool and the seasonally forced bacterial mineralization that transforms ON to ammonium. External nitrogen loads to the system and losses via the exchanges with Lake Ontario are also included.

9.10.3.1.6 Phosphorus cycle

Two state variables of the phosphorus cycle are considered in the model: phosphate (PO_4) and OP (Figure 5). The phosphate equation considers the phytoplankton uptake, the proportion of phytoplankton and zooplankton mortality/higher predation that is directly supplied into the system in inorganic form, the bacteria-mediated mineralization of OP,

and the net diffusive fluxes between epilimnion and hypolimnion. We also accounted for the phosphorus precipitation to sediment due to the iron loadings from the two steel mills, based on an empirical equation originally implemented to correct for the observed Hamilton Harbour phosphorus concentrations (Hamilton Harbour Technical Team – Water Quality, 2007). The OP equation also considers the amount of OP that is redistributed through phytoplankton and zooplankton basal metabolism. A fraction of OP settles to the sediment and another fraction is mineralized to phosphate through a first-order reaction. We also consider external phosphorus loads to the system and losses via the exchanges with Lake Ontario.

9.10.3.1.7 Fluxes from the sediment

As a simple approximation to model the role of the sediments, we followed a simple dynamic approach that relates the fluxes of nitrogen and phosphorus from the sediment with the algal and particulate matter sedimentation and burial rates while also accounting for the role of temperature (Arhonditsis and Brett, 2005). The relative magnitudes of ammonium and nitrate fluxes were also determined by nitrification and denitrification occurring at the sediment surface.

9.10.3.2 Bayesian Framework

9.10.3.2.1 Statistical formulation

Our presentation examines a statistical formulation founded upon the assumption that the eutrophication model is an imperfect simulator of the environmental system and the model error is invariant with the input conditions, that is, the difference between model and system dynamics was assumed to be constant over the annual cycle for each state variable. This formulation aims to combine field observations with simulation model outputs to update the uncertainty of model parameters, and then use the calibrated model to give predictions (along with uncertainty bounds) of the natural system dynamics. An observation i for the state variable j , y_{ij} , can be described as

$$y_{ij} = f(\theta, x_i, y_0) + \delta_j, \quad i = 1, 2, 3, \dots, n \text{ and } j = 1, \dots, m \quad [1]$$

where $f(\theta, x_i, y_0)$ denotes the eutrophication model, x_i is a vector of time-dependent control variables (e.g., boundary conditions and forcing functions) describing the environmental conditions, the vector θ is a time-independent set of the calibration model parameters, y_0 corresponds to the vector of the concentrations of the 16 state variables at the initial time point t_0 (initial conditions), and δ_j is the stochastic term that accounts for the discrepancy between the model and the natural system. Under the normality assumption, the likelihood function will be

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

where m corresponds to the number of state variables of our model for which data are available ($m = 16$); n is the number of observations in time used to calibrate the model ($n = 12$ average monthly values); $y_j = [y_{1j}, \dots, y_{mj}]^T$ and $f_j(\theta, x, y_0) = [f_{1j}(\theta, x, y_0), \dots, f_{mj}(\theta, x, y_0)]^T$ correspond to the vectors of the field observations

and model predictions for the state variable j ; $\Sigma_{\delta j} = I_n \cdot \sigma_j^2$ denotes the model structural error; and σ_j^2 represents the time-independent, variable-specific stochastic terms. In the context of the Bayesian statistical inference, the posterior density of the parameters θ and the initial conditions of the 16 state variables y_0 given the observed data y is defined as

$$p(\theta, y_0, \sigma^2 | y) = \frac{p(y | f(\theta, x, y_0, \sigma^2)) p(\theta) p(y_0) p(\sigma^2)}{\int \int p(y | f(\theta, x, y_0, \sigma^2)) p(\theta) p(y_0) p(\sigma^2) d\theta dy_0 d\sigma^2} \propto p(y | f(\theta, x, y_0, \sigma^2)) p(\theta) p(y_0) p(\sigma^2) \quad [3]$$

where $p(\theta)$ is the prior density of the model parameters θ and $p(y_0)$ is the prior density of the initial conditions of the 16 variables y_0 . 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 January monthly averages during the study period and standard deviation that was 25% of the mean value for each state variable j (Van Oijen et al., 2005); 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 , and σ^2 is

$$p(\theta, y_0, \sigma^2 | y) \propto \prod_{j=1}^m (2\pi)^{-n/2} |\Sigma_{\delta j}|^{-1/2} \times \exp \left[-\frac{1}{2} [y_j - f_j(\theta, x, y_0)]^T \Sigma_{\delta j}^{-1} [y_j - f_j(\theta, x, y_0)] \right] \times (2\pi)^{-1/2} |\Sigma_{\theta}|^{-1/2} \prod_{k=1}^l \frac{1}{\theta_k} \times \exp \left[-\frac{1}{2} [\log \theta - \theta_0]^T \Sigma_{\theta}^{-1} [\log \theta - \theta_0] \right] \times (2\pi)^{-m/2} |\Sigma_{y_0}|^{-1/2} \times \exp \left[-\frac{1}{2} [y_0 - y_{0m}]^T \Sigma_{y_0}^{-1} [y_0 - y_{0m}] \right] \times \prod_{j=1}^m \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \sigma_j^{-2(\alpha_j+1)} \exp \left(-\frac{\beta_j}{\sigma_j^2} \right) \quad [4]$$

where l is the number of the model parameters θ used for the model calibration ($l=33$); θ_0 indicates the vector of the mean values of θ in logarithmic scale; $\Sigma_{\theta} = I_l \cdot \sigma_{\theta}^T \cdot \sigma_{\theta}$ and $\sigma_{\theta} = [\sigma_{\theta 1}, \dots, \sigma_{\theta l}]^T$ corresponds to the vector of the shape parameters of the l log-normal distributions (standard deviation of $\log \theta$); the vector $y_{0m} = [y_{1,1}, \dots, y_{1,16}]^T$ corresponds to the January values of the 16 state variables; $\Sigma_{y_0} = I_m \cdot (0.25)^2 \cdot y_{0m}^T \cdot y_{0m}$; $\alpha_j (=0.001)$ and $\beta_j (=0.001)$ correspond to the shape and scale parameters of the m noninformative inverse-gamma distributions used in this analysis.

9.10.3.2.2 Prior parameter distributions

The calibration vector consists of the 33 most influential parameters as identified from a preliminary sensitivity analysis of the model. The present analysis examines two different sets of priors aiming to assess the sensitivity of the posterior patterns on the assumptions made during the prior parameter specification, which has been an historical criticism of the Bayesian inference in the literature (Dennis, 1996; Ellison, 2004). The prior parameter distributions reflected the existing knowledge (field observations, laboratory studies, literature information, and expert judgment) on the relative plausibility of their values. Specifically, the characterization of the parameter distributions was similar to the protocol used in Steinberg et al. (1997), that is, we identified the minimum

and maximum values for each parameter and then we assigned lognormal and log-uniform distributions parameterized such that 95% and 100% of the respective values were lying within the identified ranges. The group-specific parameter spaces were delineated using the framework introduced by Zhang and Arhonditsis (2008). The prior distributions of all the parameters of the model calibration vector are presented in Table 2.

9.10.3.2.3 Numerical approximations for posterior distributions

As presented in several recent studies (Arhonditsis et al., 2007, 2008a, 2008b; Zhang and Arhonditsis, 2008), sequence of realizations from the posterior distribution of the model was obtained using MCMC simulations (Gilks et al., 1996). We used the general normal-proposal Metropolis algorithm coupled with an ordered over-relaxation to control the serial correlation of the MCMC samples (Neal, 1998). In this study, we are testing two parallel chains with starting points: (1) a vector that consists of the mean values of the prior parameter distributions and (2) a vector based on a preliminary calibration of the model. The model was run for 40 000 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 is implemented in the WinBUGS Differential Interface (WBDiff) – an interface that allows numerical solution of systems of ordinary differential equations within the WinBUGS software.

9.10.3.3 Model Results and Prediction of the Frequency of Water-Quality Standard Violations

The two MCMC sequences of the model applications with the two sets of priors converged rapidly (≈ 5000 iterations) and the statistics reported were based on the last 35 000 draws by keeping every 20th iteration (thin = 20). The uncertainty underlying the values of the 33 model parameters after the MCMC sampling is depicted on the respective marginal posterior distributions (Table 2 and Figure 6). Generally, the summary statistics of the posterior parameter distributions indicate that substantial amount of knowledge was gained for the 33 parameters after updating the eutrophication model. Namely, several of the posteriors were characterized by significant shifts of their medians relative to the prior assigned values (e.g., Maxgrazing, KZ, β_p , Vsettling, and Vsettling_(phyt)), whereas the posterior inter-quartile ranges of all the parameters were significantly lower ($<50\%$) than the ranges specified prior to the calibration. Nonetheless, there were also model parameters with fairly uninformative/flat (e.g., Growthmax_(phyt), Ik_(cy), Ik_(phyt), NH_(cy), and NH_(phyt)) and bi- or multimodal posterior distributions (e.g., PH_(phyt), Pmaxuptake_(cy), Pmaxuptake_(phyt)). We also highlight the robustness of the posterior patterns on the prior parameter specifications, as the discrepancy between the posterior medians derived from the two priors was lower than 20% for the vast majority of the parameters. Notable exceptions were the two half saturation constant for ammonium uptake

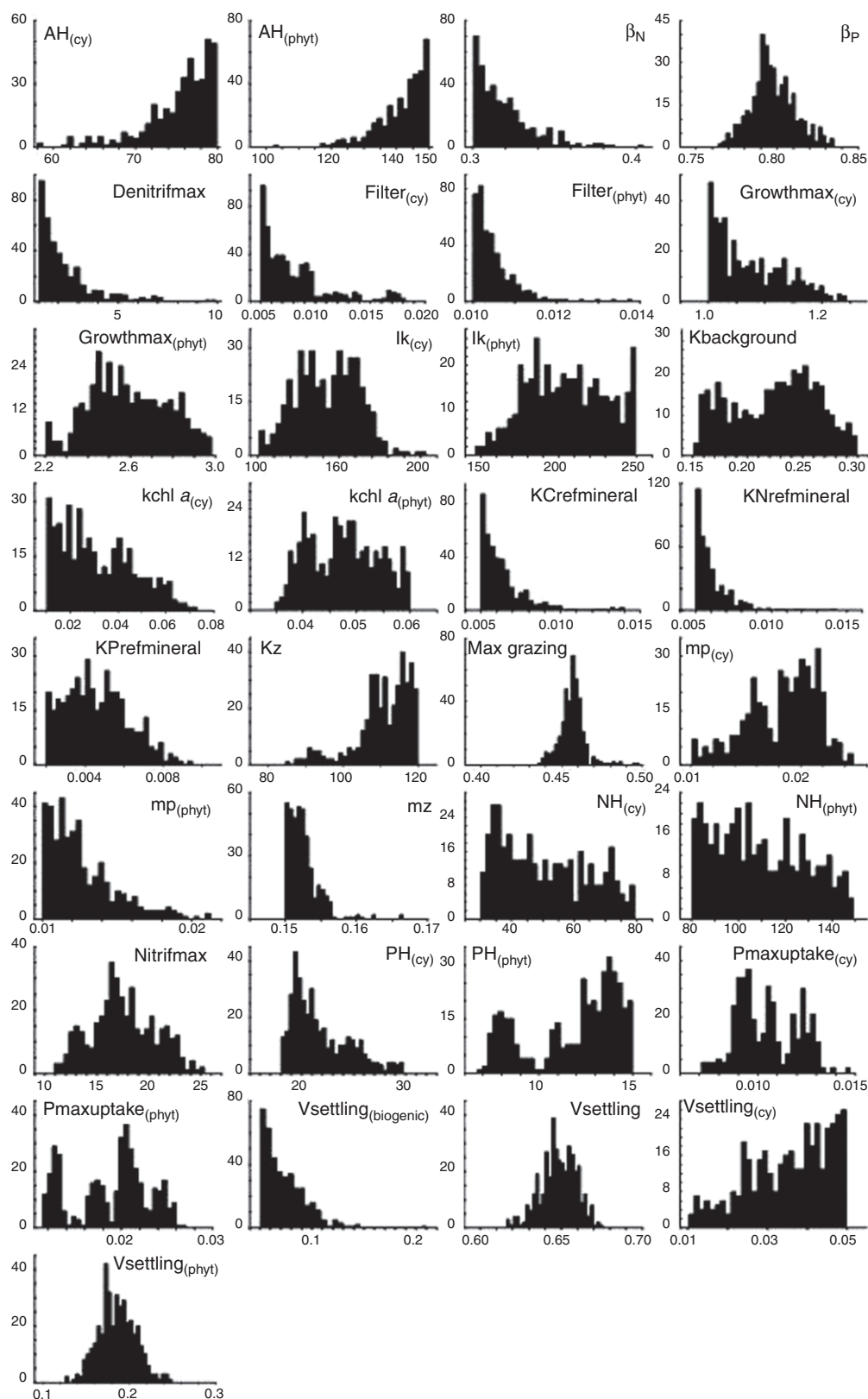


Figure 6 Posterior parameter distributions based on the assumption that the 33 parameters of the calibration vector follow a log-uniform distribution within the range provided in [Table 2](#).

($AH_{(cy)}$, $AH_{(phyt)}$), the half-saturation light intensity for cyanobacteria ($Ik_{(cy)}$), and the fraction of inert phosphorus buried into deeper sediment (β_P). The seasonally invariant (model structure) error terms (σ_j) delineate constants zone around the model predictions for the 16 state variables (Table 3). The majority of the error terms were also remarkably similar between the two experiments with the different priors. The only notable relative

(percentage) differences were the higher error terms for the epilimnetic cyanobacteria and zooplankton biomass derived from the log-uniform distributions, which were also characterized by higher coefficients of variation (>40%).

The model provided accurate representation of the system dynamics, that is, the mean predictions along with 95% credible intervals were fairly close to the observed values for phosphate, TP, ammonium, nitrate, total nitrogen, chlorophyll *a*, and total zooplankton biomass in Hamilton Harbour (Figure 7). In particular, the model accurately predicts the winter maxima ($\approx 11.5 \mu\text{g l}^{-1}$) and the summer minima ($\approx 2.2\text{--}4.7 \mu\text{g l}^{-1}$) of the epilimnetic phosphate levels as well as the hypolimnetic accumulation during the summer-stratified period ($\approx 4.2\text{--}7.3 \mu\text{g l}^{-1}$). The latter pattern is primarily driven by the interplay between the Lake Ontario inflows and the phosphorus sediment fluxes. The model underpredicts somewhat the epilimnetic TP concentrations that may stem from the assumptions made to overcome the lack of consistent data regarding the within-year variability of the exogenous loading, that is, it was assumed that the phosphorus inflow rates follow the precipitation month-to-month variability as calculated from the 1971–2000 climate normals. The model closely reproduces the winter ($\approx 4.5 \mu\text{g chl a l}^{-1}$) and the summer ($\approx 14 \mu\text{g chl a l}^{-1}$) phytoplankton levels, but seems to overpredict the spring chlorophyll *a* concentrations in that the model predicts a major spring phytoplankton bloom exceeding the level of $20 \mu\text{g chl a l}^{-1}$. This discrepancy in regard to the Harbour phenology may stem from the absence of reliable information from the system (Hiriart-Baer et al., 2009), as the sampling cruises of the monitoring programs typically do not start before the first or second week of May – a period that may

Table 3 Markov Chain Monte Carlo posterior estimates of the mean values and standard deviations of the model structure error terms

Parameters	Log-uniform		Lognormal 95%	
	Mean	Std. Dev.	Mean	Std. Dev.
$\sigma_{PO_4\text{epi}}$	0.589	0.172	0.501	0.141
$\sigma_{OP\text{epi}}$	1.170	0.268	1.112	0.281
$\sigma_{NH_4\text{epi}}$	66.78	18.76	65.74	15.79
$\sigma_{NO_3\text{epi}}$	111.8	27.40	106.4	25.87
$\sigma_{ON\text{epi}}$	7.981	1.943	7.637	2.040
$\sigma_{CYA\text{epi}}$	1.834	0.784	1.119	0.538
$\sigma_{PHYT\text{epi}}$	40.31	11.29	44.26	14.64
$\sigma_{ZOO\text{epi}}$	13.50	5.555	7.880	4.691
$\sigma_{PO_4\text{hypo}}$	0.868	0.221	0.728	0.190
$\sigma_{OP\text{hypo}}$	0.961	0.244	0.890	0.221
$\sigma_{NH_4\text{hypo}}$	50.53	18.32	46.35	11.32
$\sigma_{NO_3\text{hypo}}$	147.6	36.79	126.8	32.03
$\sigma_{ON\text{hypo}}$	11.62	2.516	11.15	2.750
$\sigma_{CYA\text{hypo}}$	3.937	1.055	3.501	0.854
$\sigma_{PHYT\text{hypo}}$	39.11	10.44	35.13	10.08
$\sigma_{ZOO\text{hypo}}$	6.084	1.378	5.980	1.420

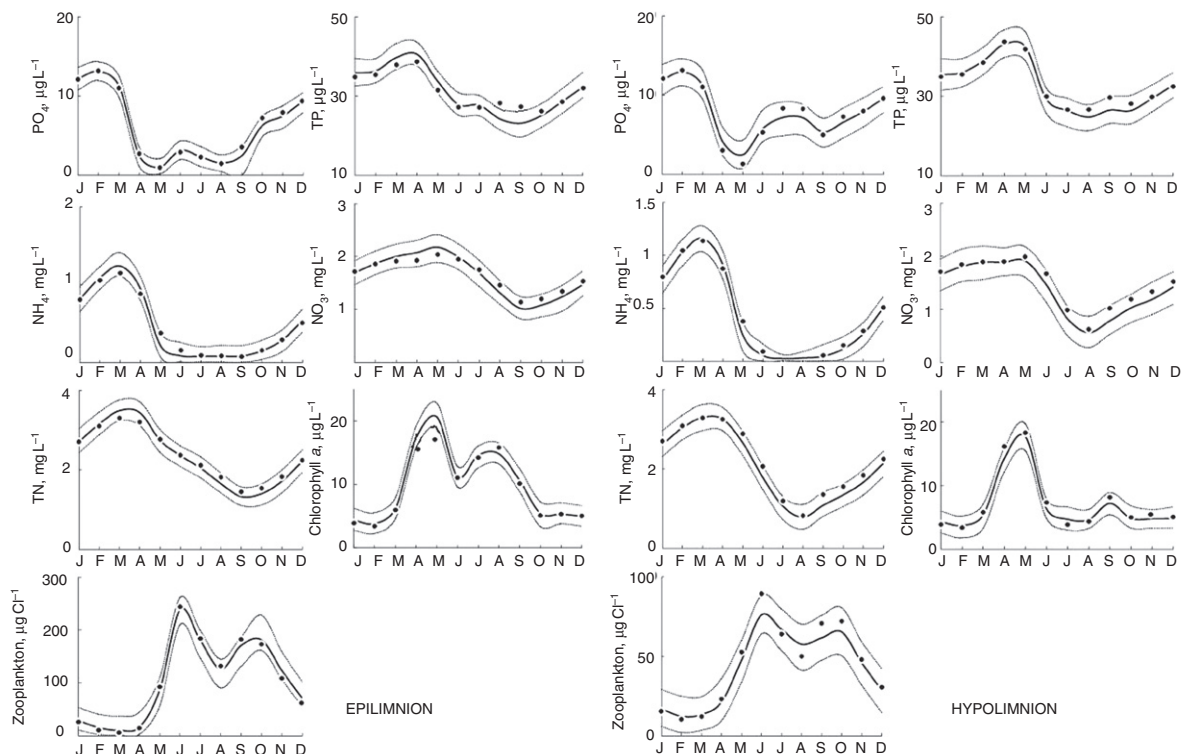


Figure 7 Comparison between the observed data (black dots) and the mean predictions (along with 95% credible intervals) for phosphate, total phosphorus, ammonium, nitrate, total nitrogen, chlorophyll *a*, and total zooplankton biomass in Hamilton Harbour.

coincide with the recession rather than the peak of the spring bloom. Our model also predicts two major peaks of the zooplankton biomass: the first peak follows the spring phytoplankton bloom ($\approx 200 \mu\text{g C l}^{-1}$), whereas the second one is predicted to occur at the end of summer–early fall ($\approx 180 \mu\text{g C l}^{-1}$). These predictions match closely with the observed patterns reported by Dermott et al. (2007; 62–63, e.g., see figures 8 and 9), if we assume an average wet-to-dry biomass ratio equal to 10 along with $0.4 \mu\text{g C per } \mu\text{g of dry zooplankton biomass}$ (Downing and Rigler, 1984).

9.10.3.3.1 Exceedance frequency and confidence of compliance with water-quality standards

The MCMC posterior samples were also used to examine the exceedance frequency and confidence of compliance with different water-quality standards under the present conditions and after 30% nutrient loading reduction. For illustration purposes, we selected two water-quality variables of management interest (i.e., chlorophyll *a* and TP concentrations), and then specified their threshold values (numerical criteria) at $15 \mu\text{g chl } a \text{ l}^{-1}$, and $25 \mu\text{g TP l}^{-1}$, respectively. For each iteration, we calculated the weekly predicted values and the

corresponding probabilities of exceeding the two water-quality criteria. The latter probabilities were calculated as follows:

$$p = P(c > c' | \theta, x, y_0, \sigma) = 1 - F\left(\frac{c' - f(\theta, x, y_0)}{\sigma}\right) \quad (5)$$

where p is the probability of the response variable exceeding a numerical criterion c' , given values of θ , x , and y_0 , σ is the model error, and $F(\cdot)$ is the value of the cumulative standard normal distribution. The weekly predicted values along with the calculated exceedance frequencies were then averaged over the summer-stratified period (June–September). The distribution of these statistics across the posterior space (1750 MCMC samples) can be used to assess the expected exceedance frequency and the confidence of compliance with the two water-quality standards given the average nutrient loading conditions of the study system, while accounting for the uncertainty that stems from the model parameters and the model error. We also note that for illustration purposes, the present analysis focuses on the mean seasonal patterns in the study system and does not consider the interannual variability, which should be accommodated to improve the realism of our risk assessment statements.

The model predicts that the expected exceedance frequencies of the $15 \mu\text{g chl } a \text{ l}^{-1}$ and $25 \mu\text{g TP l}^{-1}$ numerical criteria,

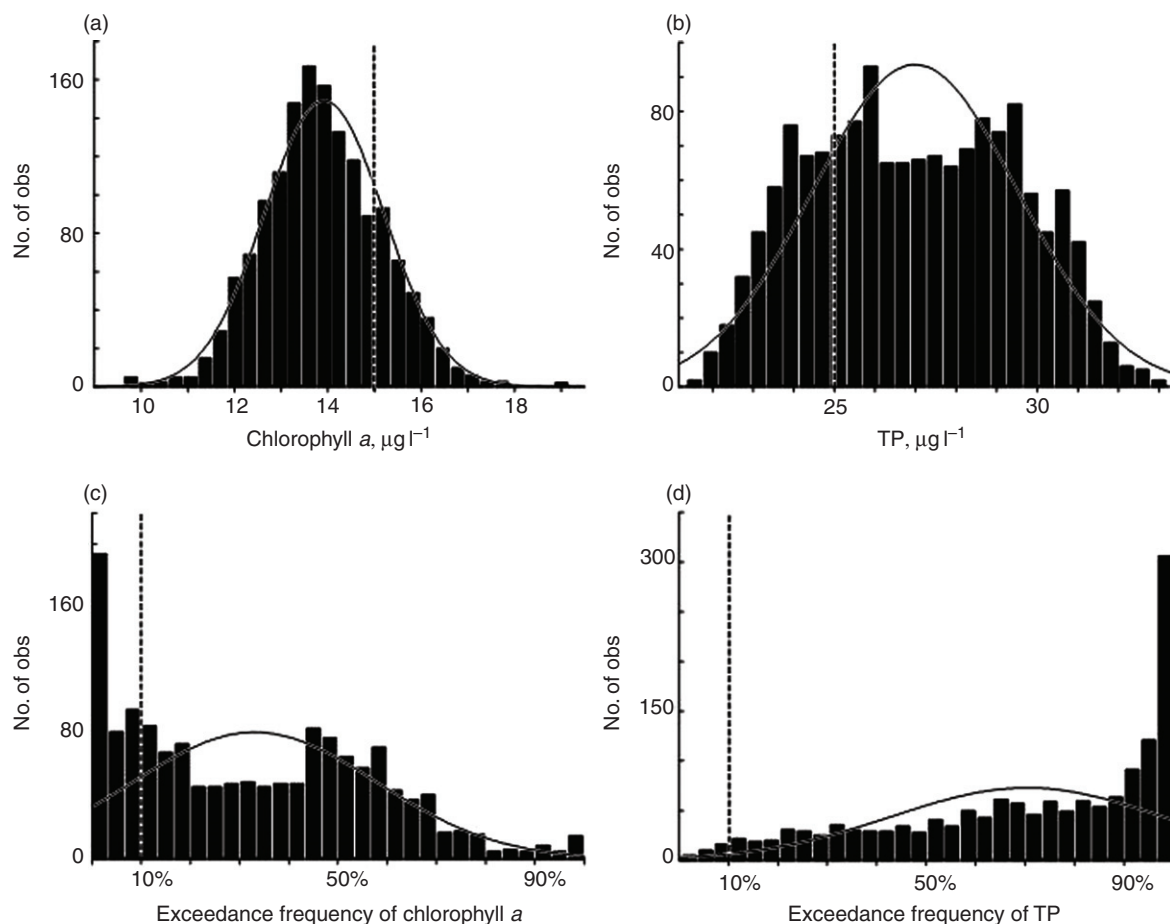


Figure 8 Marginal predictive distributions of total phosphorus and chlorophyll *a* and exceedance frequency of two water-quality standards (chlorophyll *a*, $15 \mu\text{g l}^{-1}$; total phosphorus, $25 \mu\text{g l}^{-1}$) during the summer-stratified period in Hamilton Harbour under the present conditions. In these distributions, the area below the 10% cutoff point is termed the confidence of compliance (CC), and represents the probability that the true exceedance frequency is below the 10% EPA guideline.

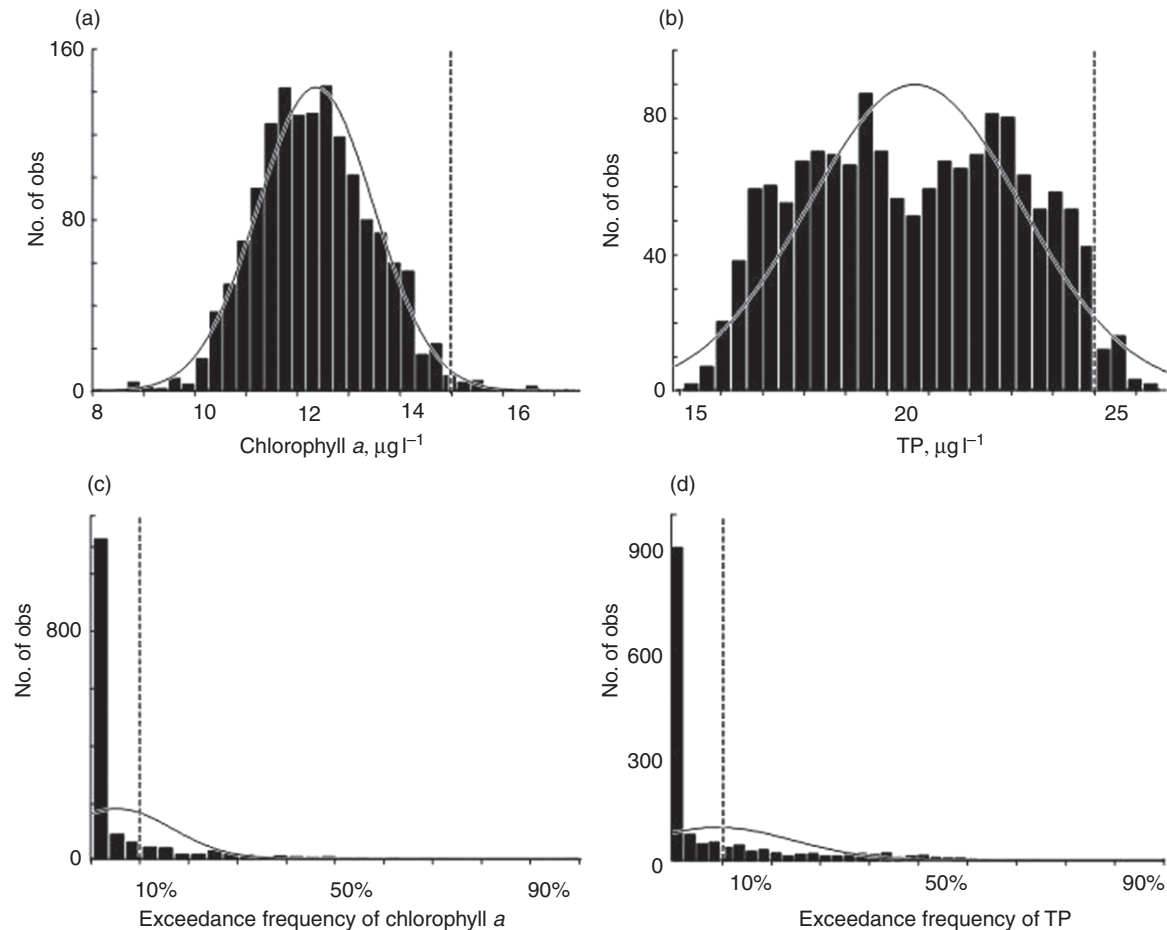


Figure 9 Marginal predictive distributions of total phosphorus and chlorophyll *a* and exceedance frequency of two water-quality standards (chlorophyll *a*, $15 \mu\text{g l}^{-1}$; total phosphorus, $25 \mu\text{g l}^{-1}$) during the summer-stratified period in Hamilton Harbour after 30% reduction of the external nutrient loading.

under the present loading conditions, are 32.9% and 70.1%, respectively (Figure 8). The corresponding confidence of compliance levels (the proportion of the exceedance frequency distribution that lies below the 10% cutoff point) were approximately 24.6% and 1.8%, and therefore it is nearly impossible to comply with the 10% Environmental Protection Agency (EPA) guideline (Figure 9). Based on the scenario of 30% reduction of the external nutrient loading, the exceedance frequencies of the chlorophyll *a* and TP threshold levels were 4.9% and 9.3%, respectively. As a result, our confidence of compliance increases to 84.5% and 72.3% for chlorophyll *a* and TP, respectively, indicating that these two water-quality criteria are likely to be met. This probabilistic assessment of the water-quality conditions should make model results more useful, because the deterministic statements are avoided and the optimal management schemes (e.g., reduction of nutrient loading) are determined by explicitly acknowledging an inevitable risk of nonattainment.

Under the present conditions, the relationship between chl *a* and TP in the surface waters was significant, although relatively weak with an r^2 value approximately equal to 0.29 (Figure 10(a)). Given that other factors beyond nutrients can limit primary production, samples were classified based on their chl *a*/particulate phosphorus (PP) ratios (Hiriart-Baer et al., 2009). Samples with chl *a*/PP ratios <0.8 were categorized as

phosphorus sufficient and those with ratios ≥ 0.8 were categorized as phosphorus limited. Indeed, the model predicts that phosphorus-limited algae occurs in the system but the frequency of the P-limited runs was lower than 1% of the total MC samples used in this analysis. Using this approach, it becomes evident that P controls growth in the Harbour only when other factors such as light availability and nitrogen supply are abundant and high growth rates are achieved. A reduction of the nutrient loading by 30% increased the slope (0.273 vs. 0.303) and the r^2 value (0.47) of the chl *a*-TP relationship, showing a more abrupt decline in chl *a* with a concomitant decrease in TP (Figure 10(b)). The number of P-limited samples also significantly increases ($\approx 20\%$), indicating that the system will be experiencing phosphorus limitation more frequently. During times of P limitation, the chl *a* versus TP relationship considerably improved, with phosphorus availability accounting for 81% of the variability in phytoplankton biomass (Figure 10(c)). Overall, these results provide evidence that the 30% nutrient loading reduction will improve the water-quality conditions in the Harbour and is likely to meet the numerical criteria of $15 \mu\text{g chl a l}^{-1}$ and $25 \mu\text{g TP l}^{-1}$. However, our analysis also suggests that the system will not achieve compliance with the current targeted environmental goals related to TP ($17 \mu\text{g l}^{-1}$) and chlorophyll *a* concentrations ($5\text{--}10 \mu\text{g l}^{-1}$), and therefore more drastic loading reductions will be necessary.

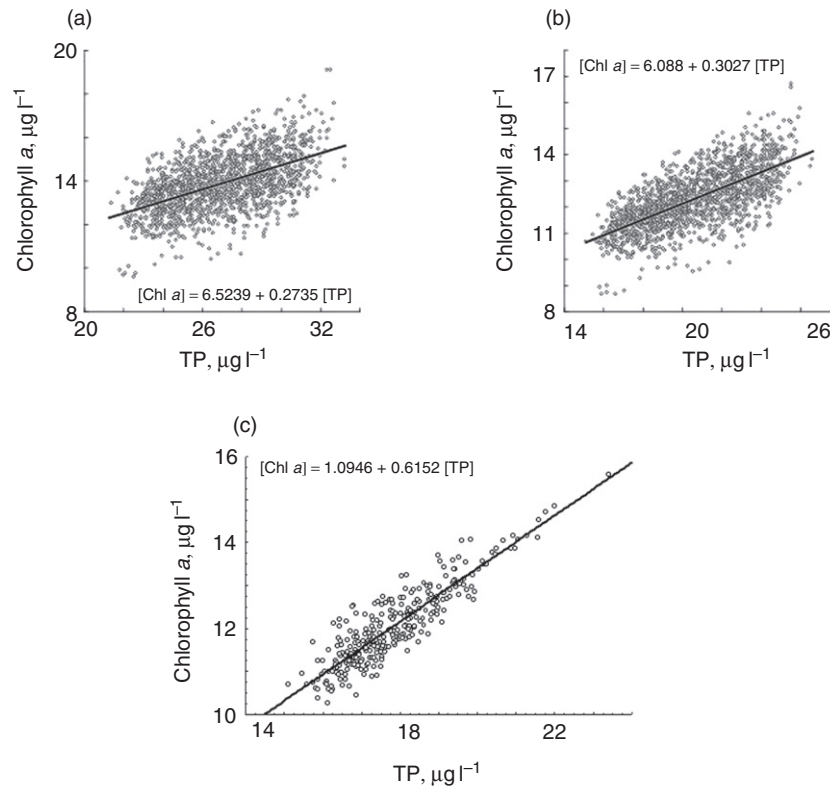


Figure 10 The predicted relationship between chlorophyll *a* and total phosphorus during the summer-stratified period in Hamilton Harbour under the present conditions (a) and after 30% reduction of the external nutrient loading (b). Panel (c) illustrates the same relationship for the data classified as phosphorus limited ($\text{chl } a:\text{PP} > 0.8$) under the loading reduction scenario.

9.10.4 Conclusions and Future Perspectives

As the articulation level of the water-quality models continues to grow, an emerging imperative is the development of novel uncertainty analysis techniques to rigorously assess the error pertaining to model structure and input parameters (Reichert and Omlin, 1997). The assessment of the uncertainty characterizing the multidimensional parameter spaces of mathematical models involves two important decisions: (1) selection of the likelihood measure to quantify model error and (2) selection of the sampling algorithms to generate a series of model realizations. In this chapter, we presented a Bayesian framework founded upon a Markov Chain Monte Carlo (MCMC) sampling algorithm and a Gaussian likelihood that enables the development of robust probabilistic analysis of error and uncertainty in model predictions by explicitly taking into account the measurement error, parameter uncertainty, and model structure imperfection. The proposed framework combines the advantageous features of both process-based and statistical approaches in that the models offer mechanistic understanding but still remain within the realm of data-based parameter estimation. The incorporation of mechanism improves the confidence in predictions made for a variety of conditions, whereas the statistical methods provide an empirical basis for parameter estimation. We have also illustrated some of the practical benefits for environmental management from the Bayesian calibration framework, such as the assessment of the exceedance frequency and confidence

of compliance with different water-quality criteria as well as probabilistic inference on cause-effect relationships pertaining to water-quality management. Furthermore, the probabilistic predictions for water-quality variables of management interest (e.g., chlorophyll *a* and dissolved oxygen) can also be used to maximize the value of information gained from environmental monitoring programs (Van Oijen et al., 2005; Zhang and Arhonditsis, 2008).

Aside from the probabilistic assessment of the water-quality conditions, another benefit of the Bayesian parameter estimation is the alignment with the policy practice of adaptive management, that is, an iterative implementation strategy that is recommended to address the often-substantial uncertainty associated with water-quality model forecasts, and to avoid the implementation of inefficient and flawed management plans (Walters, 1986). Adaptive implementation or 'learning while doing' supports initial model forecasts of management schemes with post-implementation monitoring, that is, the initial model forecast serves as the Bayesian prior, the post-implementation monitoring data serve as the sample information (the likelihood), and the resulting posterior probability (the integration of monitoring and modeling) provides the basis for revised management actions (Qian and Reckhow, 2007). The Bayesian inference and decision theory can also provide a coherent framework for decision making in problems of natural resources management (Dorazio and Johnson, 2003). Management objectives can be evaluated by integrating the probability of use attainment

for a given water-quality goal with utility functions that reflect different socioeconomic costs and benefits. The water-quality goals (resulting from specific management schemes) associated with the highest expected utility might then be chosen (Dorazio and Johnson, 2003).

A recent analysis offered a new perspective by introducing a Bayesian hierarchical framework for simultaneously calibrating aquatic biogeochemical models at multiple systems (or sites of the same system) with differences in their trophic conditions, prior precisions of model parameters, available information, measurement error, or interannual variability (Zhang and Arhonditsis, 2009). This Bayesian hierarchical proposition may be useful for a variety of applications in aquatic sciences where partial, but not complete, commonality can be assumed among the modeled units. A characteristic case is the semi-enclosed coastal embayments adjacent to the mouths of large rivers with restricted mixing with the open sea. These areas are intermediate zones in that they receive highly polluted inland waters from watersheds with significant agricultural, urban, and/or industrial activities while mixing with offshore waters of different chemical and biological characteristics (Smith et al., 1999; Diaz and Rosenberg, 2008). This type of spatial heterogeneity cannot be fully accommodated by the typical practice of developing spatially explicit mechanistic models with common parameter values over the entire systems; that is, how realistic is it to assume that the same phytoplankton growth or zooplankton grazing rates occur throughout the water body? Rather, the practical compromise between entirely site-specific and globally common parameter estimates offered by the hierarchical approach may be a conceptually more sound strategy. In a similar manner, this methodology also enables the transfer of information across systems, thereby allowing the effective modeling of systems with limited information, that is, problems of insufficient local data can be overcome by borrowing strength from well-studied sites on the basis of distributions that connect systems in space (Zhang and Arhonditsis, 2009). This outcome is also highly relevant to conservation practices of regions with high number of water resources for which complete data could never be practically collected.

Several technical issues regarding the formulation of the error structure, the selection of the parameter priors and likelihood functions, the optimal model complexity, and the computational efficiency of the Bayesian calibration scheme require particular attention and/or invite further investigation. Although the present exercise showed that the posterior patterns are fairly robust to the prior parameter specification, evidence from the literature suggests that selection of the prior parameter space or the statistical representation of the model error can significantly alter the inference (Thiemann et al., 2001; Forest et al., 2002). In this regard, robust Bayesian analysis is a promising framework to rigorously assess the conclusions drawn from typical uncertainty analysis applications based on single prior distributions and/or likelihood functions (Berger, 1994). For example, Tomassini et al. (2007) examined the robustness of the uncertainty analysis results of climate system properties using classes of parameter priors, different scaling of the observational error, and alternative likelihood functions. The posterior predictive patterns highlighted the critical role of the prior parameter distributions, and also dictated areas where future data collection

efforts should focus on to constrain climate model sensitivity. Despite its sound premise though, the broad adoption of robust Bayesian uncertainty analysis in water-quality modeling is still unclear given the computational demands that this framework entails.

Recent efforts to improve the computational efficiency of MCMC implementations of Bayesian inference for water-quality models have focused on the development of parallel algorithms (Altekar et al., 2004; Whiley and Wilson, 2004). Parallel computation for MCMC can reduce the time needed to generate a sufficient number of samples from target distributions of larger dimensions, although Whiley and Wilson (2004) assert that a good proposal distribution is of equal importance as the implementation of a parallelization scheme. Other propositions to efficiently estimate the posterior probability density function of parameters in complex high-dimensional problems involve the development of adaptive MCMC schemes that ensure ergodicity while adjusting the scale and orientation of the proposal distributions, for example, the differential evolution adaptive Metropolis (DREAM) introduced by Vrugt et al. (2008). The integration of such adaptive MCMC schemes with the generalized likelihood uncertainty estimation (GLUE) method is likely to offer flexible, easily implemented frameworks that overcome the inefficiency of the typically used Monte Carlo-based sampling to locate behavioral simulations and reduce the associated computational time (Blasone et al., 2008). An appealing alternative may be the replacement of the rigid structure of complex mathematical models with more flexible modeling tools (e.g., Bayesian networks) that have the ability to integrate quantitative descriptions of ecological processes at multiple scales and in a variety of forms (intermediate complexity mathematical models, empirical equations, and expert judgments), depending on available information (Borsuk et al., 2004).

Acknowledgments

This project has received funding support from the Ontario Ministry of the Environment (Canada-Ontario Grant Agreement 120808). Such support does not indicate endorsement by the Ministry of the contents of this material. We wish to thank Duncan Boyd and Tanya Labencki (Ontario Ministry of the Environment) for their insightful comments on an earlier version of the manuscript.

References

- Adams, D., 2002. The counting house. *Nature* 415, 726–729.
- Altekar, G., Dwarkadas S., Huelsenbeck J.P., Ronquist, F., 2004. Parallel metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* 20, 407–415.
- 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. *Journal of Marine Systems* 35, 183–205.
- 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. *Environmental Science and Technology* 40, 6547–6554.

- Arhonditsis, G.B., Brett, M.T., 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Marine Ecology Progress Series* 271, 13–26.
- Arhonditsis, G.B., Brett, M.T., 2005. Eutrophication model for Lake Washington (USA). Part I – model description and sensitivity analysis. *Ecological Modelling* 187, 140–178.
- Arhonditsis, G.B., Papantou, D., Zhang, W., Perhar, G., Massos, E., Shi, M., 2008a. Bayesian calibration of mechanistic aquatic biogeochemical models and benefits for environmental management. *Journal of Marine Systems* 73, 8–30.
- Arhonditsis, G.B., Perhar, G., Zhang, W., Massos, E., Shi, M., Das, A., 2008b. Addressing equifinality and uncertainty in eutrophication models. *Water Resources Research* 44, W01420, 15.
- Arhonditsis, G.B., Qian, S.S., Stow, C.A., Lamon, E.C., Reckhow, K.H., 2007. Eutrophication risk assessment using Bayesian calibration of process-based models: application to a mesotrophic lake. *Ecological Modelling* 208, 215–229.
- Baretta, J.W., Ebenhoh, W., Ruudij P., 1995. The European Regional Seas Ecosystem Model, a complex marine ecosystem model. *Journal of Sea Research* 33, 233–246.
- Barica, J., 1989. Unique limnological phenomena affecting water quality of Hamilton Harbour, Lake Ontario. *Journal of Great Lakes Research* 15, 519–530.
- Barnett, V., O'Hagan, A., 1997. Setting Environmental Standards: The Statistical Approach to Handling Uncertainty and Variation. Chapman and Hall, London, 112 pp.
- Beck, M.B., 1987. Water-quality modeling – a review of the analysis of uncertainty. *Water Resources Research* 23, 1393–1442.
- Berger, J., 1994. An overview of robust Bayesian analysis. *Test* 3, 5–124.
- Beven, K., 1993. Prophecy, reality and uncertainty in distributed hydrological modeling. *Advances in Water Resources* 16, 41–51.
- Beven, K.J., 2006. A manifesto for the equifinality thesis. *Journal of Hydrology* 320, 18–36.
- Bierman, V.J., Kaur, J., DePinto, J.V., Feist, T.J., Dilks, D.W., 2005. Modeling the role of zebra mussels in the proliferation of blue-green algae in Saginaw Bay, Lake Huron. *Journal of Great Lakes Research* 3, 32–55.
- Blasone, R.-S., Vrugt, J.A., Madsen, H., Rosbjerg, D., Robinson, B.A., Zyvoloski, G.A., 2008. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. *Advances in Water Resources* 31, 630–648.
- Borsuk, M.E., Stow, C.A., Reckhow, K.H., 2002. Predicting the frequency of water quality standard violations: a probabilistic approach for TMDL development. *Environmental Science and Technology* 36, 2109–2115.
- Borsuk, M.E., Stow, C.A., Reckhow, K.H., 2004. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. *Ecological Modelling* 173, 219–239.
- Brooks, S.P., Gelman, A., 1998. Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7, 434–455.
- Brun, R., Reichert, P., Kunsch, H.R., 2001. Practical identifiability analysis of large environmental simulation models. *Water Resources Research* 37, 1015–1030.
- Cerco, C.F., Cole, T.M., 1994. CE-QUAL-ICM: A Three-Dimensional Eutrophication Model, Version 1.0. User's Guide. US Army Corps of Engineers Waterways Experiments Station, Vicksburg, MS, 320 pp.
- Chapra, S.C., 1997. *Surface Water-Quality Modeling*. McGraw-Hill, New York, NY, 844 pp.
- Charlton, M.N., 1997. The sewage issue in Hamilton Harbour: implications of population growth for the remedial action plan. *Water Quality Research. Journal of Canada* 32, 407–420.
- Charlton, M.N., 2001. The Hamilton Harbour remedial action plan: eutrophication. *Verhandlungen Internationale Vereinigung für Theoretische und Angewandte Limnologie* 27, 4069–4072.
- Charlton, M.N., Le Sage, R., 1996. Water quality trends in Hamilton Harbour: 1987 to 1995. *Water Quality Research. Journal of Canada* 31, 473–484.
- Chen, C.F., Ma, H.W., Reckhow, K.H., 2007. Assessment of water quality management with a systematic qualitative uncertainty analysis. *Science of the Total Environment* 374, 13–25.
- Dennis, B., 1996. Discussion: should ecologists become Bayesians? *Ecological Applications* 6, 1095–1103.
- Dermott, R., Johannsson, O., Munawar, M., Bonnell, R., Bowen, K., Burley, M., Fitzpatrick, M., Gerlofsma, J., Niblock, H., 2007. Assessment of lower food web in Hamilton Harbour, Lake Ontario, 2002–2004. Canadian Technical Report of Fisheries and Aquatic Sciences 2729, 120.
- Diaz, R.J., Rosenberg, R., 2008. Spreading dead zones and consequences for marine ecosystems. *Science* 321, 926–929.
- Doney, S.C., 1999. Major challenges confronting marine biogeochemical modeling. *Global Biogeochemical Cycles* 13, 705–714.
- Dorazio, R.M., Johnson, F.A., 2003. Bayesian inference and decision theory – a framework for decision making in natural resource management. *Ecological Applications* 13, 556–563.
- Downing, J.A., Rigler, F.H., 1984. *A Manual on Methods for the Assessment of Second Productivity in Fresh Water*. Second ed. Blackwell Scientific, Oxford, 501 pp.
- Edwards, A.M., Yool, A., 2000. The role of higher predation in plankton population models. *Journal of Plankton Research* 22, 1085–1112.
- Ellison, A.M., 2004. Bayesian inference in ecology. *Ecological Letters* 7, 509–520.
- Eppey, 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. *Journal of Marine Research* 48, 591–639.
- Forest, C.E., Stone, P.H., Sokolov, A.P., Allen, M.R., Webster, M.D., 2002. Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science* 295, 113–117.
- Franks, P.J.S., 2002. NPZ models of plankton dynamics: their construction, coupling to physics, and application. *Journal of Oceanography* 58, 379–387.
- Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC Press, Boca Raton, FL, 512 pp.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. *Bayesian Data Analysis*. Chapman and Hall, New York, NY, p. 518.
- Hamblin, P.F., He, C., 2003. Numerical models of the exchange flows between Hamilton Harbour and Lake Ontario. *Canadian Journal of Civil Engineering* 30, 168–180.
- Hamilton Harbour Technical Team – Water Quality, 2007. Hamilton Harbour RAP water quality goals and targets review, part 1: response to the City of Hamilton's proposed wastewater system upgrades. Technical Appendix, 82.
- Hamilton, D.P., Schladow, S.G., 1997. Prediction of water quality in lakes and reservoirs. Part 1. Model description. *Ecological Modelling* 96, 91–110.
- Hiriart-Baer, V.P., Milne, J., Charlton, M.N., 2009. Long term water quality trends in Hamilton Harbour and the associated biological response. *Journal of Great Lakes Research* 35, 293–301.
- Jassby, A.D., Platt, T., 1976. Mathematical formulation of relationship between photosynthesis and light for phytoplankton. *Limnology and Oceanography* 21, 540–547.
- Klapwijk, A., Snodgrass, W.J., 1985. Model for lake–bay exchange flow. *Journal of Great Lakes Research* 11, 43–52.
- Leimu, R., Koricheva, J., 2005. What determines the citation frequency of ecological papers? *Trends in Ecology. Evolution* 20, 28–32.
- Malve, O., Laine, M., Haario, H., 2005. Estimation of winter respiration rates and prediction of oxygen regime in a lake using Bayesian inference. *Ecological Modelling* 182, 183–197.
- Malve, O., Laine, M., Haario, H., Kirkkala, T., Sarvala, J., 2007. Bayesian modeling of algal mass occurrences – using adaptive MCMC methods with a lake water quality model. *Environmental Modelling and Software* 22, 966–977.
- 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, Dordrecht, pp. 205–230.
- Omlin, M., Reichert, P., 1999. A comparison of techniques for the estimation of model prediction uncertainty. *Ecological Modelling* 115, 45–59.
- Qian, S.S., Reckhow, K.H., 2007. Combining model results and monitoring data for water quality assessment. *Environmental Science and Technology* 41, 5008–5013.
- Qian, S.S., Stow, C.A., Borsuk, M.E., 2003. On Monte Carlo methods for Bayesian inference. *Ecological Modelling* 159, 269–277.
- Reckhow, K.H., Arhonditsis, G.B., Kenney, M.A., Hauser, L., Tribo, J., Wu, C., Elcock, K.J., Steinberg, L.J., Stow, C.A., McBride, S.J., 2005. A predictive approach to nutrient criteria. *Environmental Science and Technology* 39, 2913–2919.
- Reckhow, K.H., Chapra, S.C., 1999. Modeling excessive nutrient loading in the environment. *Environmental Pollution* 100, 197–207.
- Reichert, P., Omlin, M., 1997. On the usefulness of overparameterized ecological models. *Ecological Modelling* 95, 289–299.
- Reichert, P., Schervish, M., Small, M.J., 2002. An efficient sampling technique for Bayesian inference with computationally demanding models. *Technometrics* 44, 318–327.
- Smith, V.H., Tilman, G.D., Nekola, J.C., 1999. Eutrophication: impacts of excess nutrient inputs of freshwater, marine, and terrestrial ecosystems. *Environmental Pollution* 100, 179–196.
- Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2003. WinBUGS User Manual, Version 1.4. <http://www.mrc-bsu.cam.ac.uk/bugs> (accessed July 2010).
- 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. *Ecological Modelling* 97, 35–46.
- Thiemann, M., Trosset, M., Gupta, H., Sorooshian, S., 2001. Bayesian recursive parameter estimation for hydrologic models. *Water Resource Research* 37, 2521–2535.
- Tian, R.C., Vezina, A.F., Starr, M., Saucier, F., 2001. Seasonal dynamics of coastal ecosystems and export production at high latitudes: a modeling study. *Limnology and Oceanography* 46, 1845–1859.
- Tomassini, L., Reichert, P., Knutti, R., Stocker, T.F., Borsuk, M.E., 2007. Robust Bayesian uncertainty analysis of climate system properties using Markov chain Monte Carlo methods. *Journal of Climatology* 20, 1239–1254.

- US EPA (Environmental Protection Agency), 1997. Guidelines for Preparation of the Comprehensive State Water Quality Assessments. Office of Water, US Environmental Protection Agency, Washington, DC.
- Van Oijen, M., Rougier, J., Smith, R., 2005. Bayesian calibration of process-based forest models: bridging the gap between models and data. *Tree Physiology* 25, 915–927.
- Vrugt, J.A., ter Braak C.J.F., Clark M.P., Hyman J.M., Robinson, B.A., 2008. Treatment of input uncertainty in hydrologic modeling: doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resource Research* 44, W00B09, 15.
- Walters, C.J., 1986. *Adaptive Management of Renewable Resources*. MacMillan, New York, NY.
- Whiley, M., Wilson, S.P., 2004. Parallel algorithms for Markov chain Monte Carlo methods in latent spatial Gaussian models. *Statistical Computations* 14, 171–179.
- Wroblewski, J.S., 1977. Model of phytoplankton plume formation during variable Oregon upwelling. *Journal of Marine Research* 35, 357–394.
- Zhang, W., Arhonditsis, G.B., 2008. Predicting the frequency of water quality standard violations using Bayesian calibration of eutrophication models. *Journal of Great Lakes Research* 34, 698–720.
- Zhang, W., Arhonditsis, G.B., 2009. A Bayesian hierarchical framework for calibrating aquatic biogeochemical models. *Ecological Modelling* 220, 2142–2161.

Relevant Websites

- <http://www.climate.weatheroffice.ec.gc.ca> – Environment Canada; National Climate Data and Information Archive.
- <http://www.ec.gc.ca> – Environment Canada; Water Survey of Canada.