



Jul 1st, 12:00 AM

Integration of Bayesian inference techniques with mathematical modelling

George B. Arhonditsis

Follow this and additional works at: <https://scholarsarchive.byu.edu/iemssconference>

Arhonditsis, George B., "Integration of Bayesian inference techniques with mathematical modelling" (2012). *International Congress on Environmental Modelling and Software*. 390.

<https://scholarsarchive.byu.edu/iemssconference/2012/Stream-B/390>

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Integration of Bayesian inference techniques with mathematical modelling

George B. Arhonditsis

Ecological Modelling Laboratory, Department of Physical & Environmental Sciences, University of Toronto, Toronto, Ontario, M1C 1A4, Canada
e-mail: georgea@utsc.utoronto.ca

Abstract: Skeptical views of the scientific value of modelling argue that there is no true model of an ecological system, but rather several adequate descriptions of different conceptual basis and structure. My study addresses this question using a complex ecosystem model, developed to guide the water quality criteria setting process in the Hamilton Harbour (Ontario, Canada), along with a simpler plankton model that considers the interplay among phosphate, detritus, and generic phytoplankton and zooplankton state variables. Predictions from the two models are combined using the respective standard error estimates as weights in a weighted model average. The two eutrophication models are used in conjunction with the SPATIally Referenced Regressions On Watershed attributes (SPARROW) watershed model. The Bayesian nature of my work is used: (i) to alleviate problems of spatiotemporal resolution mismatch between watershed and receiving waterbody models; and (ii) to overcome the conceptual or scale misalignment between processes of interest and supporting information. The lessons learned from this study will contribute towards the development of integrated modelling frameworks.

Keywords: Process-based modelling, Eutrophication, Bayesian inference, Water quality criteria, Decision making.

1. Introduction

In the context of water quality assessment, the application of process-based models typically has a deterministic character, whereby single-value predictions at each point in time and space are derived from uniquely determined model inputs. Most of the existing calibration efforts aim at reproducing the average ecological dynamics, but fail to capture the entire range of natural conditions experienced. The credibility of these practices and their adequacy in addressing environmental management problems has recently been questioned for two main reasons [Arhonditsis et al. 2007]. First, regardless of its complexity and supporting information, the application of any modeling construct involves substantial uncertainty contributed by model structure, parameters, and other associated inputs (e.g., boundary or initial conditions). Second, models parameterized to depict the average ecosystem behavior are inadequate in addressing the type of percentile-based standards needed to accommodate the natural spatiotemporal variability and may bias (underestimate) the predictions of the frequency of standard violations under various management options [Borsuk et al. 2002].

For better model-based decision analysis that can effectively support the development of environmental standards and the policy making process, the uncertainty in model predictions as well as the full range of the expected system responses must be rigorously quantified and reported in a straightforward way. 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

60 informative than the conventional model calibration practices, and can be used to
61 refine our knowledge of model input parameters while obtaining predictions along
62 with uncertainty bounds for output variables [Arhonditsis et al. 2007].
63 Despite the compelling arguments for considering Bayesian inference techniques
64 as an integral part of the model development process, their high computational
65 demands along with the lack of analytical expressions for the posterior distributions
66 was until recently a major impediment for their broader application. Nonetheless,
67 the advent of fast computing has allowed the development of several methods for
68 performing Bayesian inference and the most commonly used technique is called
69 Markov chain Monte Carlo (MCMC); a general methodology that provides a solution
70 to the difficult problem of sampling from high dimensional distributions for the
71 purpose of numerical integration. In this paper, I will discuss several promising
72 prospects of the application of Bayesian inference techniques, such as the
73 averaging of predictions from different models and the integration of watershed with
74 receiving waterbody models, which can be used from stakeholders and policy
75 makers to guide the use of millions of dollars of restoration and to dictate the Best
76 Management Practices.

77
78

79 2. Case study

80

81 Hamilton Harbour, a large embayment located at the western end of Lake Ontario,
82 has a long history of eutrophication problems primarily manifested as excessive
83 algal blooms, low water transparency, predominance of toxic cyanobacteria, and
84 low hypolimnetic oxygen concentrations during the late summer [Gudimov et al.
85 2011]. Since the mid 80s, when the Harbour was identified as one of the 43 Areas
86 of Concern (AOC) in the Great Lakes area, the Hamilton Harbour Remedial Action
87 Plan (RAP) was formulated through a variety of government, private sector, and
88 community participants to provide the framework for actions aimed at restoring the
89 Harbour environment. The foundation of the remedial measures and the setting of
90 water quality goals reflect an ecosystem-type approach that considers the complex
91 interplay between abiotic variables and biotic components pertinent to its beneficial
92 uses. The drastic nutrient loading reduction has historically played a central role in
93 the restoration efforts, although the determination of the critical levels has been a
94 thorny issue as the population growth and increasing urbanization accentuate the
95 pressure for expansion of the local wastewater treatment plants (WWTPs).

96 Recent modelling work suggests that the water quality goals for TP levels $<20 \mu\text{g L}^{-1}$,
97 chlorophyll a concentrations between $5\text{-}10 \mu\text{g L}^{-1}$, and water clarity $>3 \text{ m}$ will likely
98 be met, if the proposed phosphorus loading reductions at the level of 142 kg day^{-1}
99 are actually achieved [Ramin et al. 2011]. Yet, it was emphasized that the predictive
100 capacity of any modelling exercise in the Harbour is conditional upon the credibility
101 of the contemporary nutrient loading estimates, which are uncertain and appear to
102 inadequately account for the contribution of non-point sources, episodic
103 meteorological events (e.g., spring thaw, intense summer storms), and short-term
104 variability at the local WWTPs. The same modelling work also pinpointed two
105 important unknown factors that can potentially modulate the response of the system
106 to the exogenous nutrient loading reduction and may shape the duration of the
107 transient phase as well as the system resilience in the "post-recovery" era. First, the
108 dynamics of phosphorus in the sediment-water column interface are still poorly
109 understood, and thus the historical notion that the internal loading in the Harbour is
110 minimal may be inaccurate [Gudimov et al. 2011]. Second, we lack fundamental
111 knowledge of the regulatory factors of herbivorous zooplankton abundance and
112 composition, even though existing evidence suggests that a thriving zooplankton
113 community can be instrumental for achieving faster recovery rates in the Harbour.
114 The latter prospect highlights a central conclusion drawn from my recent work that
115 the bottom-up (i.e., nutrient loading reduction) approach historically followed in the
116 area was sufficient to bring the system in its present state, but any further
117 improvements should be sought in the context of a combined bottom-up and top-
118 down control [Ramin et al. 2011].

119
120

121 3. Integrated modelling framework

122

123

124

125

126

127

128

129

130

131

132

133

134

3.1 Watershed modelling

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

$$158 \quad Y_{i,t} = \mu_i + W_{v,t} \gamma_v + \varepsilon_{i,t} \quad \varepsilon_{i,t} \sim N(0, \sigma^2) \quad (1)$$

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

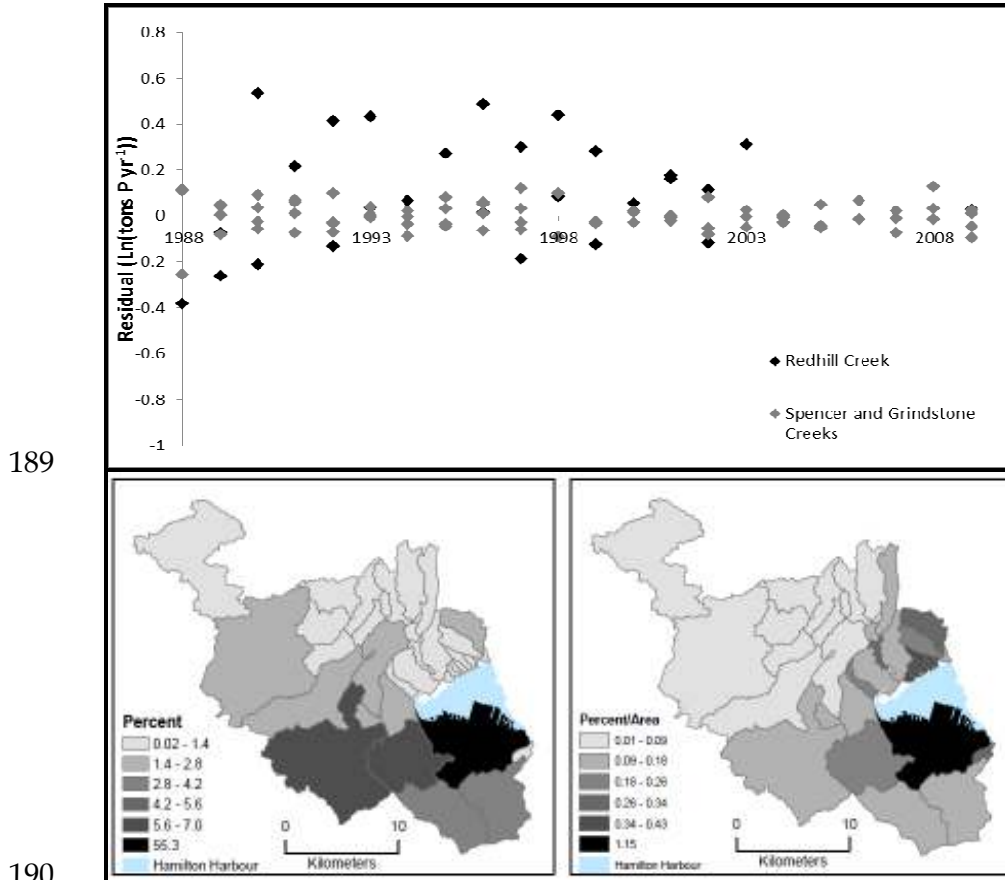
180

181

where $Y_{i,t}$ refers to the natural logarithm of the measured annual load at subwatershed monitoring station i during year t , μ_i refers to a prediction of the natural logarithm of a baseline annual load at monitoring station i estimated by the *SPARROW* equation, $W_{v,t}$ denotes a matrix of v , $1:V$, temporal forcing factors across years t , $1:T$, γ_v denotes the corresponding vector of coefficients, and $\varepsilon_{i,t}$ represents an independent spatiotemporal error. All errors are assumed independent, normally distributed, and with zero mean. The temporal variability could conceivably be accommodated by anything other than watershed landscape attributes, and the focus here is on climatic factors, namely total annual precipitation and potential evapotranspiration.

The parameterization of the *SPARROW* model was based on measured loading data from the period 1988-2007 (Fig. 1; top panel). The calibration exercise offered estimates of the export coefficients and the delivery rates from the different subcatchments and thus generated testable hypotheses regarding the nutrient export "hot spots" in the watershed. We found that sites which are both large and close to the harbour have the highest delivery values per area, as the attenuation of their loads en route to the system is very low and the urban developments in the Harbour's basin are more concentrated along the Harbour's shore (Fig. 1; bottom panel). Further, the estimates of total phosphorus export suggested that urban land uses may export more phosphorus per area than agricultural lands. This finding is somewhat contrary to the popular notion that the rates of nutrient export from urban lands are lower than those of agricultural lands due to lower nutrient subsidies. This result may be due to the very short residence time of water in urban streams and

182 the limited contact runoff has with the soil matrix, which tends to trap particulate
 183 phosphorus and chemically occlude soluble phosphorus [Wellen et al., 2012]. Soil
 184 compaction due to recent construction may cause significant declines in soil
 185 infiltration capacity and a consequent increase in the generation of runoff. The
 186 higher nutrient delivery to streams in urban areas could possibly explain higher
 187 nutrient export rates despite lower nutrient subsidies.
 188



189

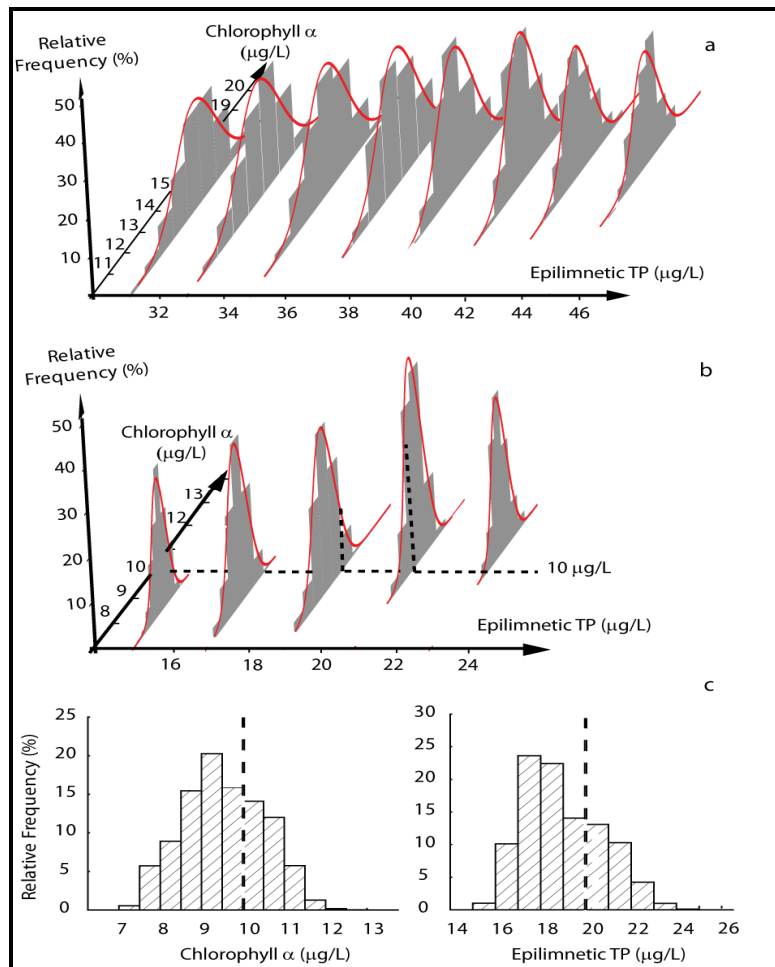
190
 191 **Figure 1:** (Top panel) Posterior median residuals of the SPARROW model
 192 predictions in two major tributaries of the Hamilton Harbour watershed. (Bottom
 193 panel) Estimated contribution of each subwatershed to the total phosphorus loading
 194 in Hamilton Harbour. The map on the left expresses the load of each subwatershed
 195 as a percentage of the total phosphorus load, including the combined sewer
 196 overflows and taking into account attenuation en route to Hamilton Harbour. The
 197 map on the right normalizes the percentage contribution by the corresponding
 198 subwatershed areas.
 199

200

201 3.2 Eutrophication modelling

202
 203 A complex eutrophication model was developed that considers the interplay among
 204 the following state variables in the epilimnion and hypolimnion of the Hamilton
 205 Harbour: nitrate (NO_3), ammonium (NH_4), phosphate (PO_4), generic phytoplankton,
 206 cyanobacteria-like phytoplankton, zooplankton, organic nitrogen (ON) and organic
 207 phosphorus (OP). The model was forced with the SPARROW outputs. To address
 208 the mismatch between the annual predictions of the watershed model and the daily
 209 resolution of the model for the receiving waterbody, we developed a Bayesian
 210 hierarchical downscaling algorithm. This approach connects the daily precipitation
 211 in the watershed with the downstream flows using logistic regression modeling and
 212 Bernoulli distribution to reproduce low and high flow regimes. A Bayesian calibration
 213 framework was then implemented, founded upon a statistical formulation that

214 explicitly accommodates measurement error, parameter uncertainty, and model
 215 structure imperfection [Ramin *et al.*, 2011].
 216 The model achieved a good representation of several key water quality variables
 217 (chlorophyll *a*, total zooplankton biomass, phosphate, and total phosphorus) and
 218 sufficiently reproduced the major cause-effect relationships underlying the Harbour
 219 dynamics. In particular, the model predicts a weakly positive *Chla-TP* relationship
 220 under the present loading conditions, while the corresponding chlorophyll *a*
 221 predictive distributions for different *TP* levels consistently exceed the targeted level
 222 of $10 \mu\text{g L}^{-1}$ (Fig. 2a). When the model is forced with the Hamilton Harbour *RAP*
 223 nutrient loading propositions, the epilimnetic *TP* concentrations dramatically
 224 decrease ($< 24 \mu\text{g L}^{-1}$), while *TP* levels lower than $20 \mu\text{g L}^{-1}$ significantly decrease
 225 the exceedance frequency of the $10 \mu\text{g L}^{-1}$ chl *a* goal (Fig. 2b). Further, the
 226 relatively discontinuous drop of the chlorophyll *a* predictive distributions around the
 227 level of $20 \mu\text{g TP L}^{-1}$ implies a severe accentuation of the phosphorus limitation
 228 of the algal growth in the system, given the posterior phytoplankton parameterization
 229 obtained. The third panel of the same figure illustrates the predictive distributions of
 230 chlorophyll *a* and epilimnetic *TP* concentrations. Generally, the modeling analysis
 231 provides evidence that the two criteria are achievable, but the water quality setting
 232 process should accommodate the natural variability by allowing for a realistic
 233 percentage of violations, e.g., exceedances of less than 10% of the weekly samples
 234 during the stratified period should still be considered as system compliance.
 235



236 **Figure 2:** Chlorophyll *a* predictive distributions for different levels of *TP*
 237 concentrations under (a) the present and (b) the Hamilton Harbour *RAP* loading
 238 targets. The third panel (c) illustrates the predictive distributions of chlorophyll *a* and
 239 epilimnetic *TP* concentrations derived from the complex eutrophication model.
 240
 241

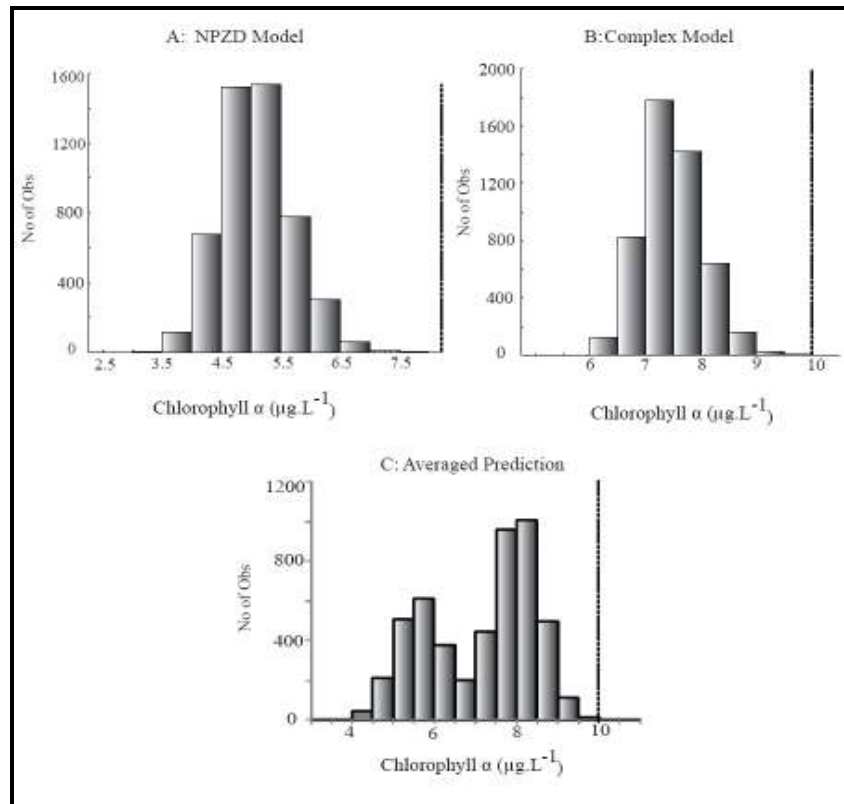
242 3.3 Bayesian Model Averaging

243

244 Recognizing that there is no true model of an ecological system, but rather several
 245 adequate descriptions of different conceptual basis and structure, Bayesian Model
 246 Averaging (BMA) is a technique designed to explicitly account for the uncertainty
 247 inherent in the model selection process [Raftery *et al.*, 2005]. By averaging over
 248 many different competing models, BMA incorporates the uncertainty about the
 249 optimal model for any given exercise into the inference drawn about parameters
 250 and prediction. Therefore, rather than picking the single “best-fit” model to predict
 251 future system responses, we can use Bayesian model averaging to provide a
 252 weighted average of the forecasts from different models. In this regard, the
 253 projections of the complex eutrophication model were tested against those from a
 254 simple model that considers the interplay among the limiting nutrient (phosphate),
 255 phytoplankton, zooplankton, and detritus (particulate phosphorus); also known as
 256 NPZD model in the literature (Ramin *et al.*, 2012).

257 The two models represent both ends of the complexity spectrum, characterized by
 258 different strengths and weaknesses. One model is a simple mathematical
 259 description of the system that accounts for the interplay between the limiting
 260 nutrient and aggregated biotic compartments such as “phytoplankton”, and
 261 “zooplankton”. This simple approach is more easily subjected to detailed uncertainty
 262 analysis and also has the advantage of fewer unconstrained parameters. The
 263 second model simulates two elemental cycles, functional phytoplankton groups,
 264 and dynamic nutrient release from the sediments. The sophisticated
 265 parameterization of the complex model provides confidence for more realistic
 266 reproduction of natural system dynamics, but the main criticism for this strategy is
 267 the inevitably poor identifiability with respect to the available data as well as the
 268 limited flexibility (high computational demands) to thoroughly examine model
 269 uncertainty to the input requirements.

270



271

272

273

274

Figure 3: Predictions of the epilimnetic summer chlorophyll a concentrations, under the proposed nutrient loading reductions by the Hamilton Harbour RAP, based on the two eutrophication models (A-B) and their averaged predictions (C).

275 The predictions from the two models were combined using the respective mean
276 model standard error estimates as weights in a weighted model average:

$$277 \quad w_{ij} = \frac{\sum_{k=1}^{MC} \frac{\sigma_{ijk}}{\bar{Y}_j}}{MC} \quad (2)$$

$$278 \quad w_{Mi} = \frac{m}{\sum_{j=1}^m w_{ij}} \quad (3)$$

$$279 \quad \overline{TP} = \sum_{i=1}^l w_{Mi} TP_{Mi} \quad \overline{chla} = \sum_{i=1}^l w_{Mi} chla_{Mi} \quad (4)$$

280 where l represents the number of models considered in this analysis ($l = 2$); m
281 corresponds to the number of state variables j of the model M_i for which data are
282 available ($m = 6$ or 11); MC is the total number of MCMC runs sampled to form the
283 model posteriors; σ_{ijk} denotes the model structural error for the state variable j of the
284 model M_i as sampled from the MCMC run k ; \bar{Y}_j represents the annual observed
285 average for the variable j , TP_{Mi} and $chla_{Mi}$ are the total phosphorus and
286 chlorophyll a predictions from the individual models weighted by the corresponding
287 weights w_{Mi} to obtain the averaged predictions \overline{TP} and \overline{chla} .

288 In particular, both models also predict that the epilimnetic chlorophyll a
289 concentrations will fall below the threshold level of $10 \mu\text{g } chla \text{ L}^{-1}$ (Fig. 3). Yet, the
290 simple model appears to support more optimistic predictions with respect to
291 phytoplankton response to the reduced ambient TP concentrations relative to the
292 complex one. Consequently, the averaged predictive distribution for chlorophyll a
293 demonstrates a distinct bimodal pattern with a primary mode at $7.5 \mu\text{g } chla \text{ L}^{-1}$,
294 reflecting the greater weight (higher performance) of the complex model, and a
295 secondary peak at $5.1 \mu\text{g } chla \text{ L}^{-1}$, associated with the simple one (Fig. 3). One of
296 the major structural differences of the two models lies in the way they handle the
297 nutrient fluxes from the sediments, i.e., a static phosphorus flux vis-à-vis a
298 mechanistic characterization that relates phosphorus release to particulate
299 sedimentation and burial rates [Ramin *et al.*, 2011]. Being part of the model
300 updating process, the simple model predicts that the sediments contribute
301 approximately $1.1 \text{ mg } P \text{ m}^2 \text{ day}^{-1}$ into the overlying water column, whereas the
302 same fluxes are raised up to $2.0 \text{ mg } P \text{ m}^2 \text{ day}^{-1}$ with the complex model. Under the
303 reduced nutrient loading scenario, the dynamic nature of the sediment response
304 with the complex model decreases the release of phosphorus at the level of 1.5 mg
305 $\text{m}^2 \text{ day}^{-1}$, which however remains well above the flux used to force the simple
306 model. This discrepancy most likely reflects one of its structural weaknesses and
307 also highlights the importance of embracing more sophisticated approaches to
308 sediment diagenesis in the Harbour. Despite all the arguments historically used to
309 downplay the relative contribution of the sediment fluxes in the system, recent
310 evidence suggests that the hypolimnetic phosphate can easily exceed the level of
311 $30 \mu\text{g } PO_4 \text{ L}^{-1}$ for extended period (3-4 weeks) during the late summer/early fall (T.
312 Labencki, unpublished data). This pattern likely suggests that the summer
313 epilimnetic environment may also be subject to intermittent nutrient pulses from the
314 hypolimnion, which in turn can have profound ramifications on the dynamics of the
315 phytoplankton community.

316 317 **4. Discussion-Future Perspectives**

318
319 Modellers must acknowledge the uncertainty pertaining to the selection of the
320 optimal model structure for a specific environmental management problem, and
321 Bayesian averaging of two or more models is a promising means for improving the
322 contemporary modelling practice. In the context of ecological process-based
323 modelling though, this approach should not be viewed solely as a framework to
324 improve our predictive devices, but rather as an opportunity to compare alternative

325 ecological structures, to challenge existing ecosystem conceptualizations, and to
326 integrate across different (and often conflicting) paradigms. Future research should
327 also focus on the refinement of the weighting schemes and other performance
328 standards to impartially synthesize the predictions of different models. Several
329 interesting statistical post-processing methods presented in the field of ensemble
330 weather forecasting will greatly benefit our attempts to develop weighting schemes
331 suitable for the synthesis of multiple ecosystem models. Some of the outstanding
332 challenges involve the development of ground rules for the features of the
333 calibration and validation domain [Anderson, 2005], the inclusion of penalties for
334 model complexity that will allow building forecasts upon parsimonious models, and
335 performance assessment that does not exclusively consider model endpoints but
336 also examines the plausibility of the underlying ecosystem structures, i.e., biological
337 rates, ecological processes or derived quantities [Arhonditsis and Brett, 2004].
338 In conclusion, Bayesian inference techniques are uniquely suitable for integrating
339 various types of models (complex dynamic models, empirical equations, expert
340 judgments) into one coherent framework, while rigorously assessing the uncertainty
341 associated with model structures, parameters and other inputs. In particular, my
342 recent research has shown that the Bayesian paradigm can effectively alleviate
343 problems of spatiotemporal resolution mismatch among different submodels of
344 integrated environmental modelling systems, overcome the conceptual or scale
345 misalignment between processes of interest and supporting information, exploit
346 disparate sources of information that differ with regards to the measurement error
347 and resolution, and accommodate tightly intertwined environmental processes
348 operating at different spatiotemporal scales

349 **ACKNOWLEDGMENTS**

350
351
352 This project has received funding support from the Ontario Ministry of the
353 Environment (Canada-Ontario Grant Agreement 120808). Such support does not
354 indicate endorsement by the Ministry of the contents of this material.

355 **REFERENCES**

- 356
357
358 Anderson, T.R., Plankton functional type modelling: running before we can walk?
359 *Journal of Plankton. Research* 27, 1073-1081, 2005.
- 360 Arhonditsis, G.B., S.S. Qian, C.A. Stow, E.C. Lamon, and K.H. Reckhow,
361 Eutrophication risk assessment using Bayesian calibration of process-based
362 models: Application to a mesotrophic lake. *Ecological Modelling* 208, 215-229,
363 2007.
- 364 Arhonditsis, G.B., and M.T.Brett, Evaluation of the current state of mechanistic
365 aquatic biogeochemical modeling. *Marine Ecology Progress Series* 271, 13-26,
366 2004.
- 367 Borsuk, M.E., Stow, C.A., and Reckhow, K.H.. Predicting the frequency of water
368 quality standard violations: a probabilistic approach for TMDL development.
369 *Environ. Sci. Technol.* 36(10):2109–2115, 2002.
- 370 Gudimov A., M. Ramin, T. Labencki, C. Wellen, M. Shelar, Y. Shimoda, D. Boyd,
371 and G.B. Arhonditsis, Predicting the response of Hamilton Harbour to the nutrient
372 loading reductions: A modeling analysis of the “ecological unknowns”. *Journal of*
373 *Great Lakes Research* 37, 494-506, 2011.
- 374 Raftery, A.E., Balabdaoui, F., Gneiting, T., and Polakowski, M., Using Bayesian
375 model averaging to calibrate forecast ensembles, Technical Report 440,
376 Department of Statistics, University of Washington, Seattle, 2003.
- 377 Ramin, M., S. Stremilov, T. Labencki, A. Gudimov, D. Boyd, and G.B. Arhonditsis,
378 Setting water quality criteria in Hamilton Harbour, Ontario, Canada: A Bayesian
379 approach. *Environmental Modelling & Software* 26, 337-353, 2011.
- 380 Ramin M., T. Labencki, D. Boyd, D. Trolle, and G.B. Arhonditsis. A Bayesian
381 synthesis of predictions from different models for setting water quality criteria.
382 *Environmental Modelling & Software, In Press.*
- 383 Wellen, C., G.B. Arhonditsis, T. Labencki, and D. Boyd, A Bayesian methodological
384 framework to accommodate inter-annual nutrient loading variability with the
385 SPARROW model. *Water Resources Research In Press.*