Chapter 11
Uncertainty Analysis by Bayesian Inference

George Arhonditsis, Dong-Kyun Kim, Noreen Kelly, Alex Neumann, and Aisha Javed

Abstract The scientific methodology of mathematical models and their credibility to form the basis of public policy decisions have been frequently challenged. The development of novel methods for rigorously assessing the uncertainty underlying model predictions is one of the priorities of the modeling community. Striving for novel uncertainty analysis tools, we present the Bayesian calibration of process-based models as a methodological advancement that warrants consideration in ecosystem analysis and biogeochemical 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 mechanisms 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, the rigorous assessment of the expected consequences of different management actions, the optimization of the sampling design of monitoring programs, and the consistency with the scientific process of progressive learning and the policy practice of adaptive management. 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 and the Bay of Quinte—two eutrophic systems in Ontario, Canada—as case studies.

11.1 Does Uncertainty Really Matter?

In the context of environmental management, the central objectives of policy analysis and decision-making are to identify the important drivers of ecological degradation, to pinpoint the sources of controversy, and to help anticipate the unexpected. The explicit consideration of uncertainty enables one to think more

G. Arhonditsis (✉) • D.-K. Kim • N. Kelly • A. Neumann • A. Javed
University of Toronto Scarborough, Scarborough, ON, Canada
e-mail: georgea@utsc.utoronto.ca

F. Recknagel, W.K. Michener (eds.), *Ecological Informatics*,
DOI 10.1007/978-3-319-59928-1_11

carefully about these matters, to elucidate the relative role of different causal factors, and to delineate contingency plans (Dawes 1988). Environmental problems have a way of resurfacing themselves and are rarely (if ever) solved completely. Nonetheless, even if some facets may change overtime, the core problems often remain the same. Thus, having a framework that rigorously evaluates the underlying uncertainty makes it much easier to distinguish between valid assumptions and erroneous actions and, thus, maximize the efficiency of adaptive management strategies (Morgan et al. 1992).

The concepts of "uncertainty" and "risk" are understood in a variety of different ways by scientists, stakeholders, policy makers, and the public in ecology/environmental science. Uncertainty is a generic term comprising many concepts (Pappenberger and Beven 2006). No direct measurement of an empirical quantity can be absolutely exact and, therefore, uncertainty arises from *random error* in direct measurements. In addition, biases are often introduced through the measuring apparatus and/or experimental protocols. This experimental procedure typically reflects the *systematic error* associated with the difference between the true value of the quantity of interest and the value to which the mean of the measurements converges as more measurements are taken. Another source of uncertainty lies in the *subjective judgments* used to overcome knowledge gaps and lack of empirical measurements related to the major ecological mechanisms and/or variables underlying the environmental problem at hand. *Inherent randomness* is often perceived as a distinctly different type of uncertainty in that it is in principle irreducible. Nonetheless, this indeterminacy is not considered a matter of principle in environmental science, but rather the product of our incomplete knowledge of the world. It is argued that once we shed light on unknown causal variables and important ecological processes, we should be able to reduce the apparent uncertainty. In cases of environmental policy analysis, where there is no clear empirical evidence and scientific support in favor of a certain management option, significant uncertainty arises from potential *disagreements* among decision makers and stakeholders, reflecting their different perspectives and conscious (or unconscious) biases. Perhaps, the most familiar source of uncertainty is the *variability* that environmental quantities demonstrate over time and space. While these quantities can be effectively described by frequency distributions, what we typically fail to acknowledge and effectively communicate is the degree of confidence about the parameters (mean, median, standard deviation or various percentiles) of these distributions given the available information in a certain location or time period.

Along the same line of thinking, all mathematical models are simplistic representations of natural ecosystems and, therefore, their application in an environmental policy analysis context introduces the so-called *approximation uncertainty* (Arhonditsis et al. 2007). This uncertainty stems from the assumptions made and imperfect knowledge used to determine model structure and inputs (Beck 1987; Reichert and Omlin 1997). Model input error mainly stems from the uncertainty underlying the values of model parameters, initial conditions, and forcing functions as well as the realization that all models are drastic simplifications of reality that approximate the actual processes, i.e., essentially, all parameters are effective (e.g., spatially and temporally averaged) values unlikely to be represented by fixed constants (Arhonditsis et al. 2006). Model structure error arises from (1) the selection of the appropriate state variables (model endpoints) to reproduce ecosystem functioning, given the environment management problem at hand; (2) the selection of the suitable equations among a variety of mathematical formulations for describing the ecosystem processes, e.g., linear, quadratic, sigmoidal, and hyperbolic functional forms to reproduce fish predation on zooplankton (Edwards and Yool 2000); and (3) the fact that our models are based on relationships which are derived individually in controlled laboratory environments but may not collectively yield an accurate picture of the natural ecosystem dynamics (Arhonditsis et al. 2006).

The general premise for constructing mathematical models is to mirror the complexity of natural systems and account for all the ecological processes that can potentially become important in future hypothesized ecosystem states, and thus increase our predictive ability. Nonetheless, by striving for increased model complexity, and thereby (implicitly or explicitly) embracing a reductionist description of natural system dynamics, we accentuate the disparity between what we want to tease out from a mathematical model and what can realistically be observed given the available technology, staffing, and resources to study the natural system. In doing so, it often becomes impossible to impose quantitative (or even qualitative) constraints on what should be considered "acceptable" model performance (Beven 2006). This problem profoundly undermines the very basic application of mathematical models as inverse analysis tools, i.e., any information on the levels and the variability of the state (or dependent) variables is used through the model calibration exercise to infer the most likely values of independent variables (model parameters) typically representing ecological rates and functional properties of the abiotic environment and/or the biotic communities. Instead, what modelers encounter is a situation in which several distinct choices of model inputs lead to the same model output, i.e., many sets of parameters fit the data about equally well. This non-uniqueness of the model solutions is known in the modeling literature as *equifinality* (Beven 1993). In recognition of the uncertainty and equifinality problems, it is suggested that the model calibration practice should change from seeking a single "optimal" value for each model parameter, to seeking a distribution of parameter sets that all meet a pre-defined fitting criterion (Stow et al. 2007; Arhonditsis et al. 2007). These acceptable parameter sets may then provide the basis for estimating prediction error associated with the model parameters.

Model uncertainty analysis is an attempt to formulate the joint probability distribution of model inputs and then update our knowledge about this distribution after the consideration of the calibration dataset. In this regard, Bayesian inference represents a suitable means to combine existing information (prior) with current observations (likelihood) for projecting the future. Several recent studies illustrate how Bayesian inference techniques can be used to quantify the information that data contain about model inputs, to offer insights into the covariance structure among parameter estimates, and to obtain predictions along with uncertainty bounds for model outputs (Bayarri et al. 2007; Arhonditsis et al. 2007, 2008a, b).

Specifically, Bayesian calibration schemes have been introduced with simple mathematical models and statistical formulations that explicitly accommodate measurement error, parameter uncertainty, and model structure error. Nonetheless, the emergence of the holistic management paradigm has increased the demand for even more complex biogeochemical models with considerably greater uncertainty (Zhang and Arhonditsis 2008; Ramin et al. 2011; Reichert and Schuwirth 2012). In particular, there is increasing pressure for the development of integrated water quality models that effectively connect the watershed with downstream biogeochemical processes. This need stems from the emerging management questions related to contemporary climate and land use changes that should be connected with the receiving water bodies (Rode et al. 2010). In this context, significant progress has been made in regards to the computational demands and error propagation control through complex model structures (Dietzel and Reichert 2012; Kim et al. 2014).

In this chapter, we present two case studies that illustrate how the assessment of uncertainty can assist in developing integrated environmental modeling systems, overcoming the conceptual or scale misalignment between processes of interest and supporting information, and exploiting disparate sources of data that differ with regards to their quality and resolution. The two systems are the Hamilton Harbour and Bay of Quinte, Ontario, Canada. There is a great deal of modeling work that has been done toward establishing realistic eutrophication goals and impartially evaluating the likelihood of delisting the two systems as Areas of Concerns (AOCs). Existing watershed, eutrophication, and food web models shed light on different facets of the ecosystem functioning. Here, we address several critical questions that have emerged from these models: To what extent do the models coalesce with respect to their assumptions and inference drawn? What are the major sources of uncertainty that will ultimately determine the attainment of the existing delisting goals? Our aim is to highlight the major lessons learned about the watershed dynamics, the eutrophication phenomena, and the broader implications for food web integrity. We also place special emphasis on the knowledge gaps of our current understanding of the two systems. Our thesis is that the uncertainty stemming from several "ecological unknowns" can offer critical planning information to determine the optimal management actions in the two areas.

11.2 Hamilton Harbour

11.2.1 Introduction

Located at the western end of Lake Ontario, Hamilton Harbour is a large 2150 ha embayment surrounded by a watershed of approximately 500 km^2 (HH RAP 2003). The harbour has a roughly triangular shape with a length of 8 km along its main axis and a maximum width of 6 km along its eastern shoreline. It has a maximum depth of 23 m, an average depth of 13 m, a surface area of 21.5 km^2, and a volume of 2.8×10^8 m^3. The harbour exchanges water with western Lake Ontario via the Burlington Ship Canal, which is a man-made canal, 836 m long, 89 m wide and 9.5 m deep. The residence time of the harbour is significantly reduced by these exchange flows, which have a large influence on water quality and hypolimnetic dissolved oxygen concentrations (Yerubandi et al. 2016). The majority of the loads of inorganic nutrients and organic matter entering Hamilton Harbour originate from the Woodward and Skyway wastewater treatment plants (WWTPs), combined sewer overflows (CSOs), and ArcelorMittal Dofasco and Stelco steel mills (Hiriart-Baer et al. 2009). Other significant loads are delivered by three main tributaries that feed into the Harbour: Grindstone Creek, Red Hill Creek, and Spencer Creek, which reaches the harbour through a 250 ha shallow area of both marsh and open water called Cootes Paradise (HH RAP 2003). While the Redhill Creek watershed is ~80% urbanized, much of Grindstone and Spencer Creeks remain undeveloped as less than 20% of their watershed areas has been developed (HH RAP 2003). As a consequence of the excessive loading of nutrients and other pollutants, the harbour experiences serious water quality problems, such as algal blooms, low water transparency, predominance of toxic cyanobacteria, and low hypolimnetic oxygen concentrations often beginning in early summer.

Hamilton Harbour has long been considered one of the most degraded sites in the Great Lakes, and was listed as one of the 43 Areas of Concern (AOCs)[1] in the mid-1980s by the Water Quality Board of the International Joint Commission (Hall and O’Connor 2016). Since then, the Hamilton Harbour Remedial Action Plan (RAP) has assembled a variety of government, private sector, and community participants to decide on actions to restore the harbour environment. To this end, the RAP identified a number of beneficial use impairments[2] (BUIs), including the beneficial use *Eutrophication or Undesirable Algae* (HH RAP 2003). The foundation of the remedial measures and setting of water quality goals for the restoration of the harbour was based on the premise that reducing ambient phosphorus concentrations could control the chlorophyll a concentrations and water clarity. Using a framework that involved data analysis, expert judgment, and modeling along with consideration of what was deemed desirable and achievable for the harbour (Hall et al. 2006), critical thresholds for the TP concentration were set at 17 μg L^{-1}, chlorophyll *a* concentration at 10 μg L^{-1}, Secchi disc depth at 3.0 m, whilethe

[1]Great Lakes Areas of Concern are designated geographic areas within the Great Lakes Basin that show severe environmental degradation.

[2]An impairment of beneficial uses means a change in the chemical, physical or biological integrity of the Great Lakes system sufficient to cause any of the following: Restrictions on Fish and Wildlife Consumption; Tainting of Fish and Wildlife Flavor; Degraded Fish and Wildlife Populations; Fish Tumors or Other Deformities; Bird or Animal Deformities or Reproductive Problems; Degradation of Benthos; Restrictions on Dredging Activities; Eutrophication or Undesirable Algae; Restrictions on Drinking Water Consumption or Taste and Odor Problems; Beach Closings; Degradation of Aesthetics; Added Costs to Agriculture or Industry; Degradation of Phytoplankton and Zooplankton Populations; Loss of Fish and Wildlife Habitat.

maximum allowable exogenous TP loadings in the harbour were set at 142 kg day^{-1} (Charlton 2001). Reductions of external TP loading into the harbour led to water quality improvement and resurgence of aquatic macrophytes, but the system still receives substantial loads of phosphorus, ammonia, and suspended solids from the WWTPs, as well as from non-point loading sources and, therefore, only moderate improvements in TP, chlorophyll *a* and total ammonia concentrations have been observed since the mid-1990s (Hiriart-Baer et al. 2009, 2016).

Environmental modeling has been an indispensable tool of the Hamilton Harbour restoration efforts and a variety of data-oriented and process-based models are in place to determine realistic water quality goals. However, none of the existing modeling efforts in the Hamilton Harbour had rigorously assessed the effects of the uncertainty underlying model predictions (parametric and structural error, misspecified boundary conditions) on the projected system responses, nor have models to address percentile-based standards been used (Zhang and Arhonditsis 2008). Given the substantial social and economic implications of management decisions, it is important to implement modeling practices accommodating the type of probabilistic standards that seem to be more appropriate for complex environmental systems, such as the Hamilton Harbour (Ramin et al. 2011). In the following sections, we review the modeling efforts conducted to date in order to quantitatively assess the uncertainty in implementing management actions, and to highlight the applicability of percentile-based standards for setting water quality targets in the Hamilton Harbour and its watershed.

11.2.2 Eutrophication Modeling to Elucidate the Role of Lower Food Web

A series of process-based eutrophication models were built to depict the interplay among the different ecological mechanisms underlying the eutrophication problems, and to guide a water quality criteria-setting process that explicitly acknowledges the likelihood of standards violations in Hamilton Harbour (Gudimov et al. 2010, 2011; Ramin et al. 2011, 2012). As a starting point, Ramin et al. (2011) developed an ecological model that considered the interactions among eight state variables: nitrate, ammonium, phosphate, generic phytoplankton, cyanobacteria, zooplankton, organic nitrogen, and organic phosphorus. The model was based on a two-compartment vertical segmentation representing the epilimnion and hypolimnion of the harbour. The planktonic food web model was subsequently calibrated with Bayesian inference techniques founded upon a statistical formulation that explicitly accommodated measurement error, parameter uncertainty, and model structure imperfection. Concurrently with the Ramin et al. (2011) study, Gudimov et al. (2010) conducted a second (independent) modeling exercise with an upgraded model structure that utilized a three-compartment vertical segmentation representing the epilimnion, metalimnion, and hypolimnion, included three phytoplankton functional groups to more realistically depict the continuum between diatom and cyanobacteria-dominated communities, and two zooplankton functional groups to account for the role of herbivorous and omnivorous zooplankton in the system. With these approaches, both Ramin et al. (2011) and Gudimov et al. (2010) provided a good representation of the seasonal variability of the prevailing water quality conditions and accurately reproduced the major cause-effect relationships underlying the harbour dynamics. Using the upgraded model structure, Gudimov et al. (2011) revisited several of the critical assumptions made in the previous two studies, and further explored the general uncertainty involved in their assumptions of ecosystem functioning. Building from these models, Ramin et al. (2012) used Bayesian averaging techniques to synthesize the forecasts from models of differing complexity to examine the robustness of earlier predictions regarding the harbour's response to nutrient loading scenarios (see Chap. 16).

These models collectively addressed two critical questions regarding the present status and future response of the Hamilton Harbour system: Is it possible to meet the eutrophication delisting goals of the AOC, if the RAP's proposed nutrient loading reduction targets are actually implemented? How frequently would these water quality goals be violated? The adoption of a water quality criterion that permits a pre-specified level of violations in space and time offers a more realistic assessment of the anticipated water quality conditions as it accommodates both natural variability and sampling error. Overall, similar projections were achieved by Ramin et al. (2011) and by Gudimov et al. (2010), projecting that the 17 μg TP L^{-1} target would likely be met if the RAP phosphorus-loading target of 142 kg day^{-1} were achieved. However, by using a more representative summer epilimnetic TP dataset to calibrate the eutrophication model, Gudimov et al. (2011) demonstrated that the latter water quality target was too stringent, and most likely unattainable (Fig. 11.1). As corroborated by Ramin et al. (2012), a more pragmatic goal of 20 μg TP L^{-1} would permit an acceptable frequency level of violations, e.g., $<10\%$ of the weekly samples during the stratified period (Fig. 11.1).

In contrast to the TP criterion, and depending on the assumptions made about the strength of the top-down control, as well as the importance of the internal nutrient sources (e.g., phosphorus release from the sediments, nutrient mineralization), Ramin et al. (2011) and Gudimov et al. (2010) provided evidence that the mean chlorophyll *a* target was achievable, although their projections had $>50\%$ probability of exceeding the 10 μg L^{-1} threshold level, even under the most drastic external nutrient loading reduction scenarios. In a follow-up study, Gudimov et al. (2011) revisited the ecological parameterization of the previous two models in order to test whether the chlorophyll *a* criterion could be achieved with a lower frequency of violations. With this analysis, two critical "ecological unknowns" were identified to influence the model's capacity to assess compliance with the chlorophyll *a* criterion; namely, the importance of the epilimnetic nutrient regeneration mediated by the microbial food web, and the likelihood of a structural shift in the lower food web towards a zooplankton community dominated by large-sized and fast-growing herbivores (e.g., *Daphnia*) (Gudimov et al. 2011). Given these uncertainties, Ramin et al. (2012) emphasized that the criteria setting process

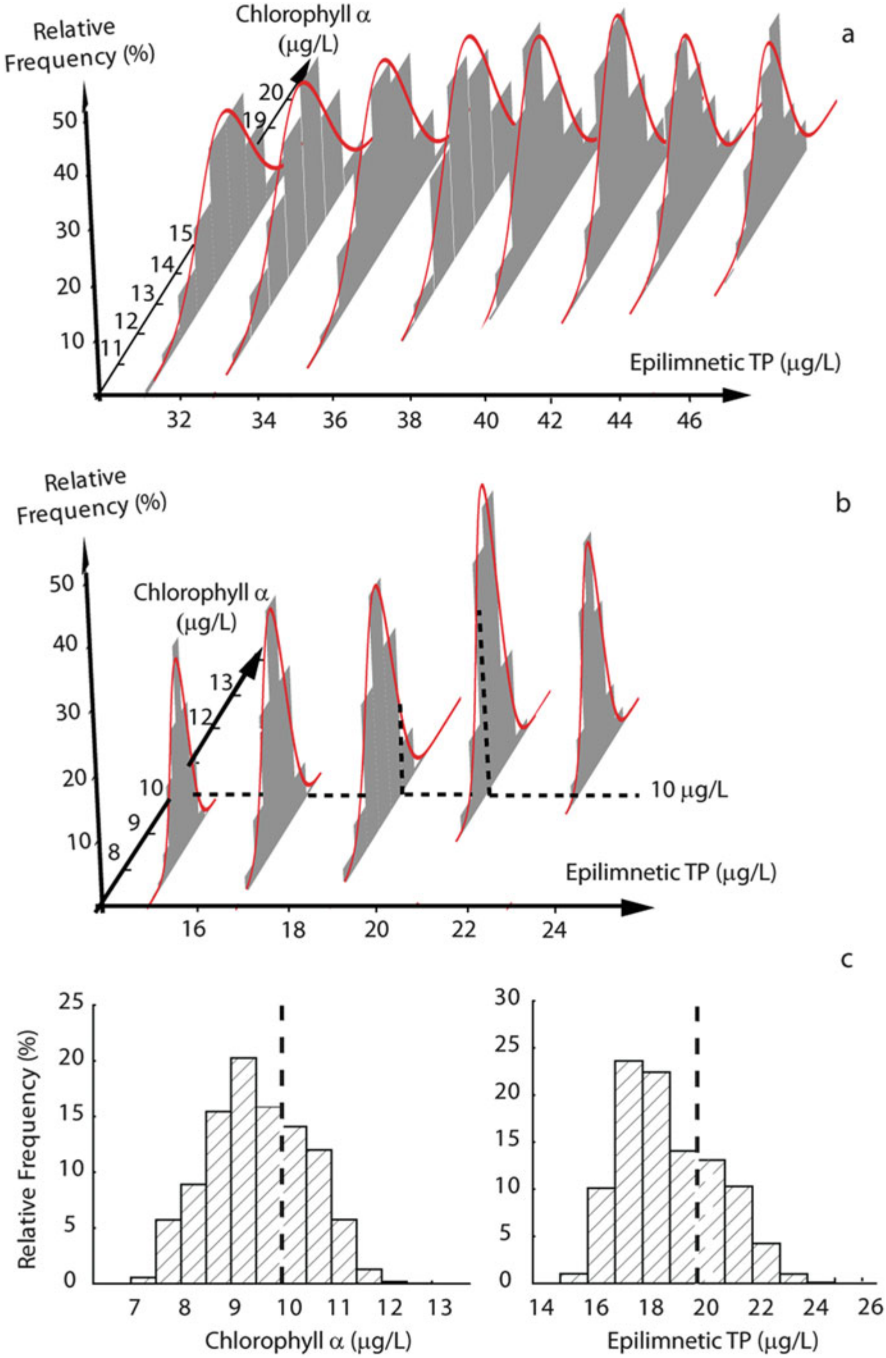


Fig. 11.1 Chlorophyll *a* predictive distributions for different levels of TP concentrations under (**a**) the present and (**b**) the Hamilton Harbour RAP loading targets (see text). Panel (**c**) illustrates the

should allow for a realistic percentage of violations of the target, such that exceedances of <10–15% of the weekly samples collected during the stratified period should still be considered as compliance, in order to explicitly accommodate the natural variability or inherent unpredictability of the system response.

In the same context, the uncertain role of planktivory and sediment diagenesis in the system emerged as two additional important ecological mechanisms for achieving the water quality targets in Hamilton Harbour. Gudimov et al. (2010) provided evidence that the anticipated structural shifts of the zooplankton community could determine the restoration rate, as well as the stability of the new trophic state in the harbour. Larger zooplankton taxa are particularly efficient in suppressing the standing phytoplankton biomass, but are also preferentially consumed by fish, and therefore the level of planktivory may shape the response rate to the nutrient loading reductions (Gudimov et al. 2010). Further, Gudimov et al. (2011) demonstrated that the epilimnetic TP concentrations were highly sensitive to the internal phosphorus loading assumptions, as a nearly two-fold increase of the sediment fluxes dramatically increased the number of violations of the TP delisting target. Thus, the internal nutrient loading from the sediments may be an important regulatory factor of the harbour.

The accuracy of the predictions made by the eutrophication model is conditional upon the credibility of the nutrient loading estimates to the harbour, which were highly uncertain and inadequately accounted for the contribution of non-point sources, episodic meteorological events (e.g., spring thaw, intense summer storms), and short-term variability at the local WWTPs (Gudimov et al. 2010, 2011). These uncertainties could potentially influence the exceedance frequency and the confidence of compliance with the water quality standards, particularly during the summer-stratified period (Gudimov et al. 2010). Given the pivotal role played by ambient phosphorus in the ecology of this system, there is a clear need to improve the tributary loading estimates in the area.

11.2.3 Nutrient Export Modeling for the Hamilton Harbour Watershed

The identification of the major nutrient source areas in the Hamilton Harbour watershed is of great management interest, as subwatersheds characterized by both high total delivery and high delivery per area are priority areas for management intervention. However, considerable knowledge gaps exist regarding the

Fig. 11.1 (continued) predictive distributions of chlorophyll *a* and epilimnetic *TP* concentrations examined to accommodate the inter- and intra-annual variability. *Vertical dashed lines* indicate the water quality targets of 10 μg·L^{-1} chl *a* and 20 μg·L^{-1} epilimnetic TP [Reproduced from Gudimov et al. (2011)]

complex interplay among hydrological factors, geological features, land uses, and spatial patterns of the built environment that modulates the attenuation rates of nutrient and contaminants. Following the development of the eutrophication models, Wellen et al. (2012, 2014a, b, c) employed two different watershed models to advance our understanding of how urban sites cycle nutrients and contaminants, so planning decisions that least impact Hamilton Harbour can be better informed.

Wellen et al. (2012, 2014a) implemented Bayesian inference techniques to parameterize the SPARROW (SPAtially Referenced Regressions On Watershed attributes) non-linear regression model in the Hamilton Harbour watershed. SPARROW is a spatially distributed, hybrid empirical/process-based model that estimates the relation between in-stream measurements of nutrient fluxes and the sources and sinks of nutrients within watersheds over annual timescales (McMahon et al. 2003). Source processes are described with export coefficients that predict TP mobilization, while the sink processes are represented by delivery factors, predicting how landscape attributes modulate the delivery of mobilized TP to streams, and attenuation coefficients, predicting the amount of the delivered TP remaining in transit per length of stream or per reservoir. With the SPARROW strategy, a two-level hierarchical structure is implemented, where watersheds are first divided into subwatersheds that each drain to a water-quality monitoring station, then each subwatershed is further divided into reach catchments draining to a particular stream segment (Schwarz et al. 2006).

Using data from Ontario's Provincial Water Quality Monitoring Network (PWQMN), Wellen et al. (2012, 2014a) offered the first estimates of export coefficients and delivery rates from the different subcatchments and generated testable hypotheses regarding the nutrient export "hot spots" in the studied watershed. The derived total phosphorus export estimates suggest that urban land uses may export more phosphorus per area than agricultural lands. This finding was 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. Wellen et al. (2014a) was able to show that subwatersheds which are both large and in close proximity to Hamilton Harbour have the highest nutrient delivery values per area, as the attenuation of their loads en route to the system is very low and the urban developments are more concentrated along the shore (Fig. 11.2).

The same modeling work has demonstrated that stream attenuation coefficients are quite variable in time (Fig. 11.3). The mechanisms that modulate the variability of nutrient attenuation across stream size are fairly well established in the literature. They generally refer to the tighter coupling of smaller streams with their streambeds, whereby biological and chemical removal processes in the sediments have greater access to nutrients in the water column (Alexander et al. 2004). The longer hydraulic residence time of smaller streams allows these processes to operate for longer times. Recent work suggests that stream stage explains the inter-annual variation of nutrient attenuation at a particular site over time, implying that the coupling between streambed and water column changes from year to year (Basu et al. 2011). Consistent with these findings, Wellen et al. (2012) showed that the inter-annual variability of the average discharge, a function of stream stage, can

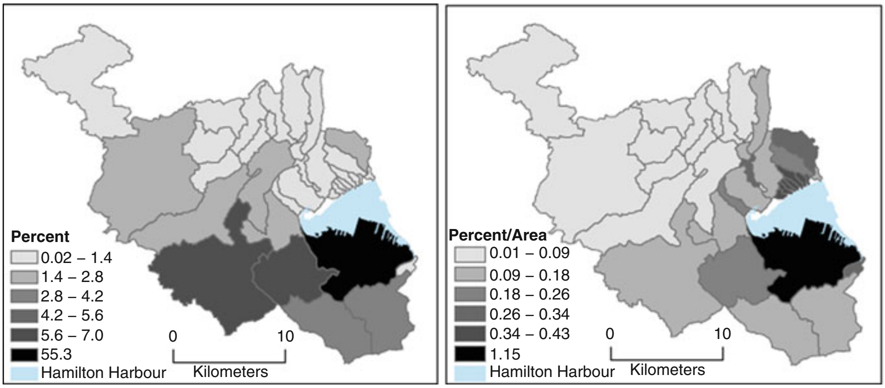


Fig. 11.2 Estimated contribution of each subwatershed to the total phosphorus loading in Hamilton Harbour. The map on the *left* expresses the load of each subwatershed as a percentage of the total phosphorus load, including the combined sewer overflows and taking into account attenuation en route to Hamilton Harbour. The map on the *right* normalizes the percentage contribution by the corresponding subwatershed areas [Reproduced from Wellen et al. (2014a)]

explain more than half of the variability of stream attenuation estimates from the SPARROW model in higher-order streams.

An interesting implication of the Hamilton Harbour's SPARROW modeling is that the year-to-year variability of the contribution of phosphorus source areas may be strongly affected by the capacity of stream reaches to attenuate nutrient loads (Fig. 11.4). Empirical studies of nutrient uptake in rivers indicate significant variability of nutrient attenuation rates at annual timescales for phosphorus (Doyle et al. 2003) and nitrogen (Claessens et al. 2009). Donner et al. (2004) found that nutrient attenuation rates varied nearly two-fold between wet and dry years in the Mississippi River, with wet years exhibiting lower attenuation. Basu et al. (2011) also showed an inverse relationship between stream stage and nutrient attenuation that was consistently manifested across spatial and temporal scales. This finding implies that fluctuations in stage (and discharge) may indeed affect the spatial location of significant nutrient source areas at various scales. While previous research has documented the variability of in-stream attenuation at annual timescales, the Hamilton Harbour modeling work allowed estimating how this variability impacts basin-scale nutrient source areas.

Wellen et al. (2014a) applied the SPARROW model to evaluate the potential improvement of parameter estimates (and the decrease of predictive uncertainty) if the precision of the currently available nutrient loading estimates in Hamilton Harbour is increased. Parameter identification was overwhelmingly improved with an increase in the spatial intensity of sampling stations, while an increase in the credibility of the measured nutrient loads significantly reduced the uncertainty of the model predictions, even when the number of stations monitored was halved (Wellen et al. 2014a). When a higher quality dataset was used to parameterize the model, the subwatersheds that displayed the greatest contraction in their 95%

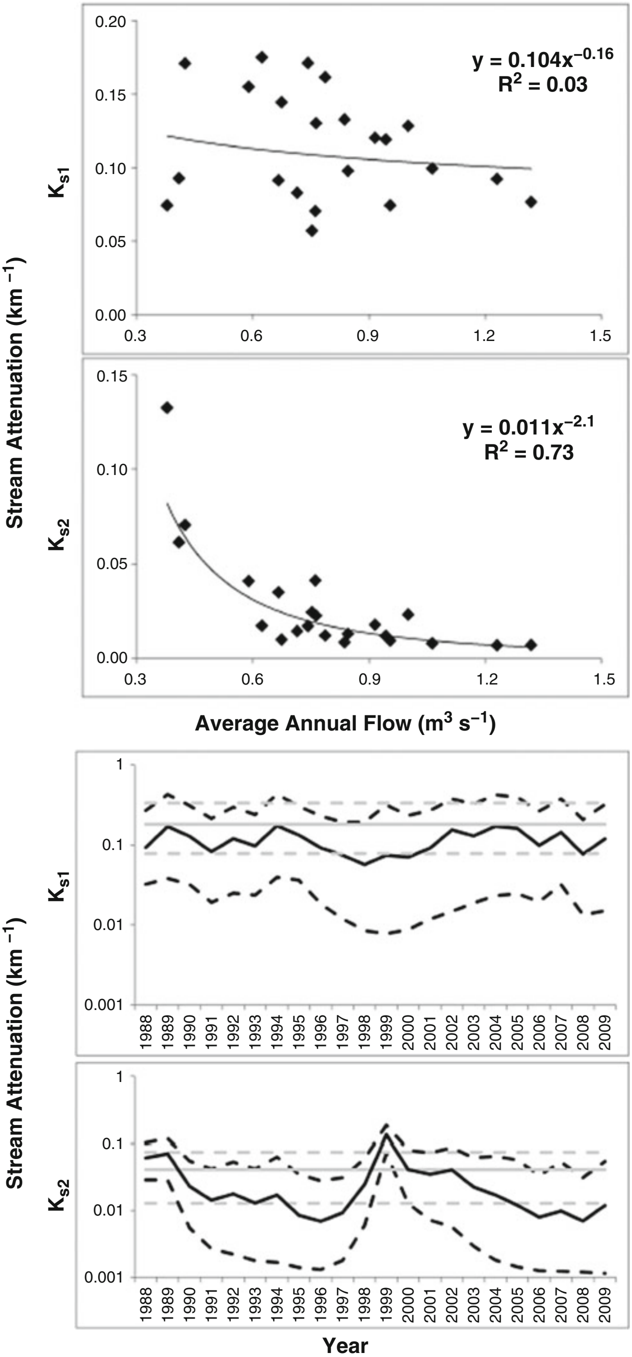


Fig. 11.3 (*Upper panels*) Scatterplots of yearly total phosphorus stream attenuation rates (k_{s1} refers to attenuation in first- and second-order streams, k_{s2} to attenuation in third- and

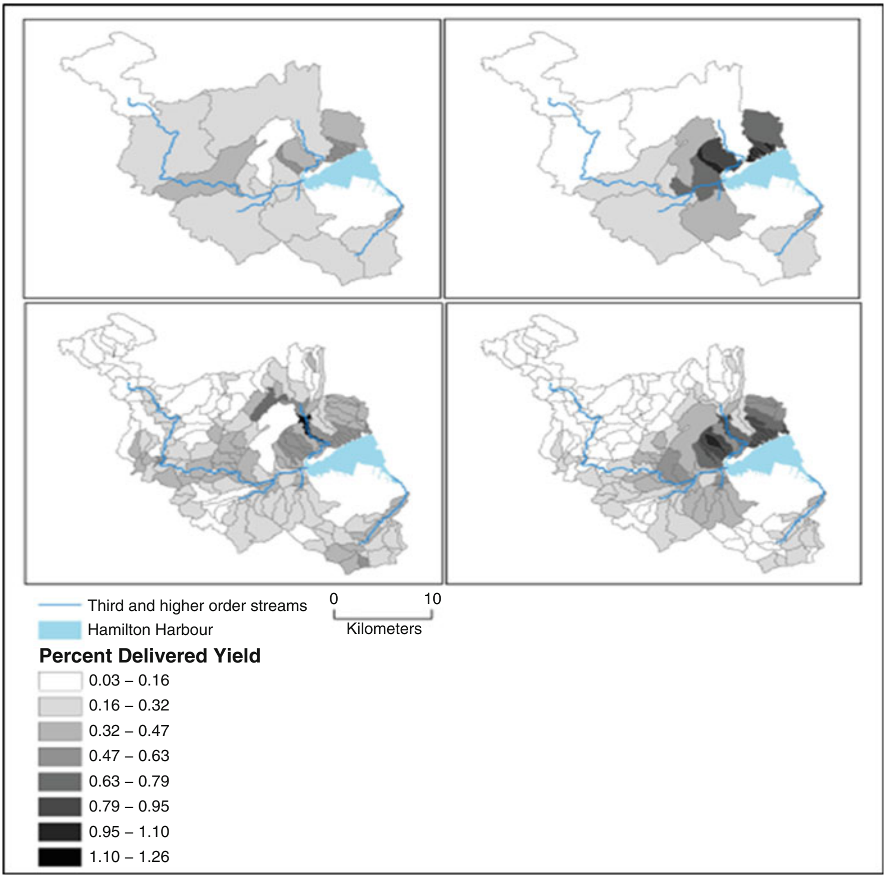


Fig. 11.4 Spatio-temporal variability of total phosphorus delivered yield at the watershed (*top panels*) and reach (*bottom panels*) scales. (*Left panels*) The percent contribution of total load into the Hamilton Harbour per square kilometer for 2006, the year with the lowest value of k_{s2}. (*Right panels*) The percent contribution of total load to the Harbour per square kilometer for 1999, the year with the highest value of k_{s2} [Reproduced from Wellen et al. (2012)]

Fig. 11.3 (continued) higher-order streams) against annual average streamflow. (*Bottom panels*) Time series plots of the two attenuation coefficients over a 22-year study period (1988–2009). *Dashed black lines* indicate upper and lower limits of the 95% credible intervals (In Bayesian statistics, a **credible interval** is an interval in the domain of a posterior probability distribution used for interval estimation. Credible intervals are analogous to confidence intervals in frequentist statistics, but differ on a philosophical basis; Bayesian intervals treat their bounds as fixed and the estimated parameter as a random variable, whereas frequentist confidence intervals treat their bounds as random variables and the parameter as a fixed value.); *solid black lines* indicate the medians of the posterior distributions of the two coefficients. *Grey lines* depict the attenuation rate values typically reported in the literature [Reproduced from Wellen et al. (2012)]

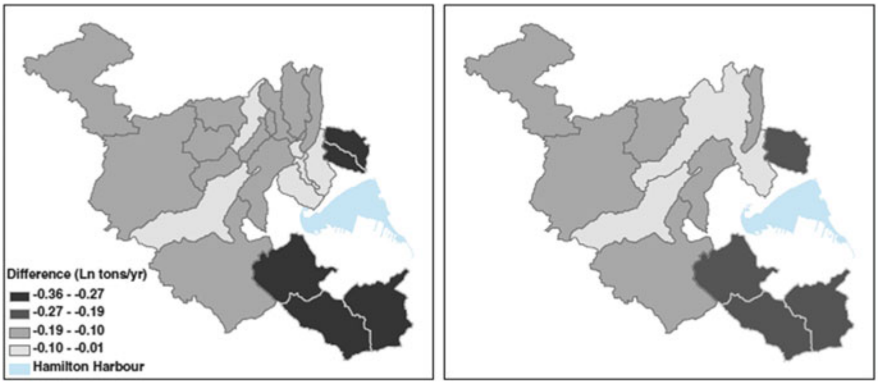
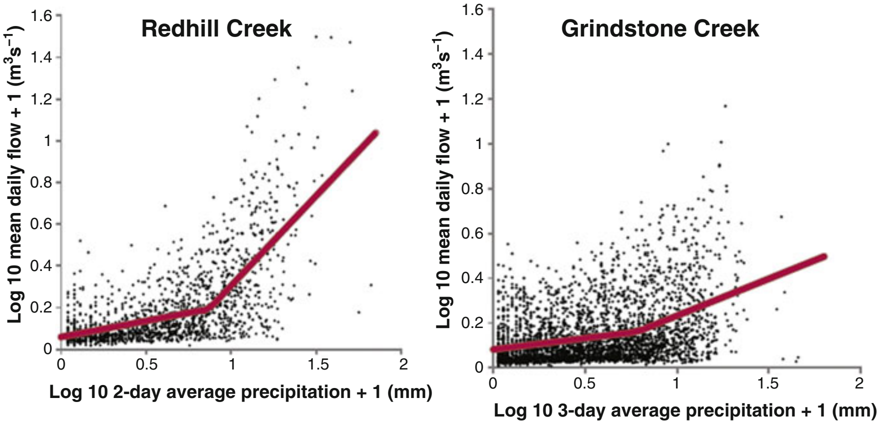


Fig. 11.5 Value of information of additional monitoring in the Hamilton Harbour watershed. Maps show the difference between the width of the 95% credible intervals of the posterior loading estimates derived from the high and the current precision scenarios for sampling with all 24 stations originally used to calibrate the SPARROW model (*right*) and sampling with a subset of 12 stations (*left*) [Reproduced from Wellen et al. (2014a)]

credible intervals were the headwater streams as well as locations closest to the harbour characterized by high delivery rates and urban land uses (Fig. 11.5). Using the uncertainty patterns provided by the SPARROW model predictions, Wellen et al. (2014a) proposed that additional water-quality data-collection efforts in the watershed should be focused on "hot spots" sites characterized by: (1) a mid-range likelihood of impairment (*i.e.* the probability of exceeding a threshold level lying within the 25–75% range); (2) model predictions of unacceptably high variance; (3) locations where data uncertainty drives the model residuals; and/or (4) locations where modeled loads showed the greatest reduction in the width of their 95% credible intervals when higher quality dataset are obtained.

Even though the SPARROW modeling exercise has gained considerable insights, the annual resolution of the latter model, along with the fact that the PWQMN program collects monthly samples primarily during baseflow conditions, impedes the accurate characterization of TP dynamics during high flow conditions. In particular, examination of the daily flows of Redhill and Grindstone Creeks supports the idea of a single threshold separating two states of response of the two Creeks to precipitation (Wellen et al. 2014b). Figure 11.6 shows scatterplots of $\log_{10}$ transformed daily flows and averages of the previous 2 or 3 days of precipitation along with the fitted piecewise regressions. These periods were chosen to implicitly include the effect of antecedent moisture. The data used are from the period 1988–2009, representing the months from May through November. Redhill Creek's threshold was estimated at a 2-day average of 7.7 mm, and would be reached by one day with 15.2 mm of precipitation or 2 days of 7.7 mm. Grindstone Creek's threshold was estimated to be a 3-day average of 5.0 mm. It was hypothesized that the watershed response to precipitation occurs in distinct states, such that precipitation depth above these thresholds triggers an



Lower Slope = 0.174 (0.02)
Upper Slope = 1.031 (0.06)
Lower Intercept = 0.20 (0.01)
Changepoint = 0.94 (0.03) = 7.7 (0.6) mm

Lower Slope = 0.10 (0.01)
Upper Slope = 0.33 (0.03)
Lower Intercept = 0.16 (0.01)
Changepoint = 0.78 (0.05) = 5.0 (0.6) mm

Fig. 11.6 Piecewise regression graphs relating the 2- or 3-day average precipitation to the daily streamflow measured from 1988 to 2009. Only data from the months May–November are plotted. Statistics below graphs show the means and, *in parentheses*, standard deviations of the parameters of the regressions [Reproduced from Wellen et al. (2014b)]

extreme state, which is characterized by a qualitatively different response of the watershed to precipitation.

To solidify this working hypothesis, Long et al. (2014, 2015) collected 87 24-h level-weighted composite samples from a variety of catchment states (rain, snowmelt, baseflow) from all four major tributaries to Hamilton Harbour between July 2010 and May 2012. The key findings from this research were as follows: (1) daily TP loads varied by three orders of magnitude between wet and dry conditions, with storm events and spring freshets driving peak daily loads in urban and agricultural watersheds, respectively; (2) areal TP loads were significantly higher from the urban relative to the agricultural watersheds; and (3) the characterization of TP concentrations during high flow conditions was essential in establishing accurate concentration versus flow relationships and subsequently nutrient load estimates. The brief but intense events that occurred less than 10% of the time were found to be responsible for 50–90% of TP loads delivered from local tributaries.

Capitalizing upon this high-resolution dataset, a SWAT model was used to simulate the water cycle and sediment export in the area (Wellen et al. 2014b, c). Surface runoff is the primary pathway through which many pollutants (including phosphorus) enter waterways, and so identifying sources of surface runoff can aid in locating possible pollutant source areas (McDowell and Srinivasan 2009). In Fig. 11.7, estimates of surface runoff generation are presented for the different land uses in Redhill and Grindstone Creeks across three formulations (i.e., different statistical configurations of the Bayesian calibration framework; see Wellen et al. 2014b). Runoff generated during the entire year was distinguished from runoff

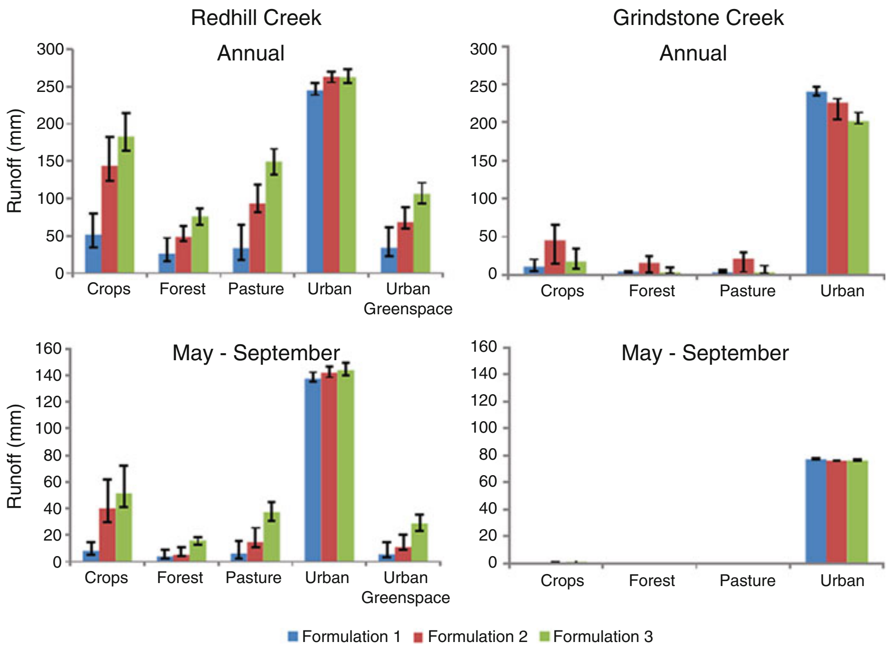


Fig. 11.7 SWAT predictions of surface runoff depths generated in Redhill and Grindstone Creeks by different land uses. Formulations 1–3 correspond to different statistical configurations of the Bayesian calibration framework as presented in Wellen et al. (2014b). The *error bars* indicate 95% credible intervals of the predictions

generated during the growing season (May–September), as this is the period when the receiving water body is most sensitive to eutrophication. In both Creeks, urban land use generated the greatest depth of runoff; 245–262 mm for Redhill Creek and 202–240 mm for Grindstone Creek. For Redhill Creek, this compares to 51–183 mm for crops, 26–76 mm for forest, 34–149 mm for pasture, and 34–106 mm for urban green space. For Grindstone Creek, the urban runoff estimate compares to 11–45 mm for crops, 3–16 mm for forest, and 3–21 mm for pasture. During the growing season, this disparity became more acute, particularly in Grindstone Creek. Between May and September, runoff generation in Redhill Creek ranged from 8–51 mm for crops, 4–16 mm for forest, 6–37 mm for pasture, and 6–29 mm for urban green space. For Grindstone Creek, this compares to 1 mm for crops, <1 mm for forest, and <1 mm for pasture. Urban areas effectively by-pass catchment storage, as nearly all the precipitation falling on them becomes surface runoff and reaches the stream in less than one day, leaving little time for evapotranspiration. While the importance of urban areas as a surface runoff source increased slightly during the growing season in Redhill Creek, the model surprisingly predicts that almost no surface runoff reaches the stream from any of the pervious surfaces in Grindstone Creek from May to September. While it is likely that the contribution of runoff for Grindstone Creek is somewhat underestimated, there seem to be important differences in soil type and/or vegetation cover between the two catchments which may be responsible for generating the markedly different amounts of runoff during the growing season.

Despite the small aerial coverage of the agricultural areas in Redhill Creek (5%) and the urban areas in Grindstone Creek (9%), these areas were responsible for a disproportionate amount of overland sediment export to streams (Fig. 11.8). Cropland was estimated to contribute between 20% and 30% of Redhill Creek's

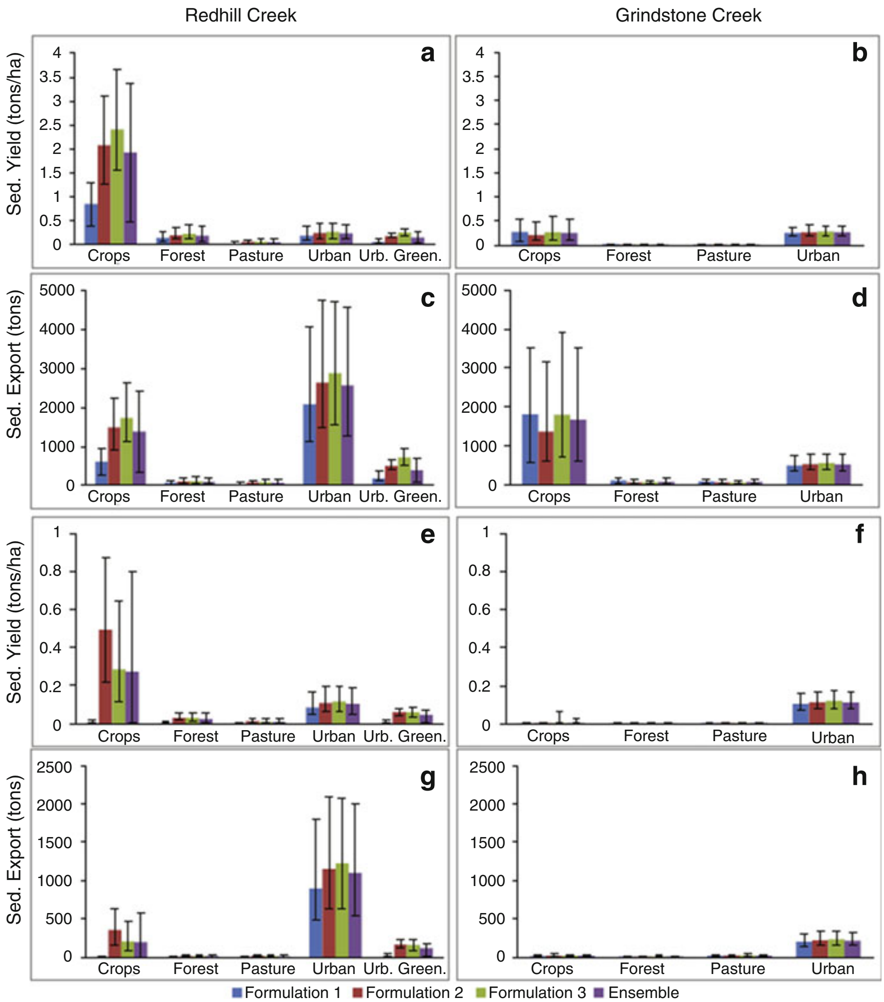


Fig. 11.8 SWAT predictions of sediment yield and export by land use for the entire annual cycle (2010–2012; a–d) and for the growing season (May–September, 2010–2012; e–h). Ensemble refers to the averaged predictions of the three statistical configurations of the Bayesian calibration framework presented in Wellen et al. (2014c). The *error bars* indicate 95% credible intervals of the predictions

total sediment export to streams (720–3299 tons), while urban areas were estimated to contribute between 17% and 36% of Grindstone Creek's total sediment export (410–1830 tons). During the growing season, urban residential areas are the main sources of sediment export to both streams, comprising 70–99% of all sediment exported to streams in Redhill Creek (217–1143 tons) and 60–81% of all estimated sediment exported to Grindstone Creek (74–214 tons).

During the calibration of the sediment routing submodel, reliable data were not available on stream bankwidth and depth. In order to draw reliable inferences on the sediment yield and streambed sediment storage status for Redhill and Grindstone Creeks at the sub-basin scale, Wellen et al. (2014c) used the entire predictive range of sediment storage for each subbasin (bed storage = upstream sediment in + erosional sediment in − downstream sediment out) (Fig. 11.9). It was assumed that if the 95% credible interval of the bed storage distribution was non-overlapping

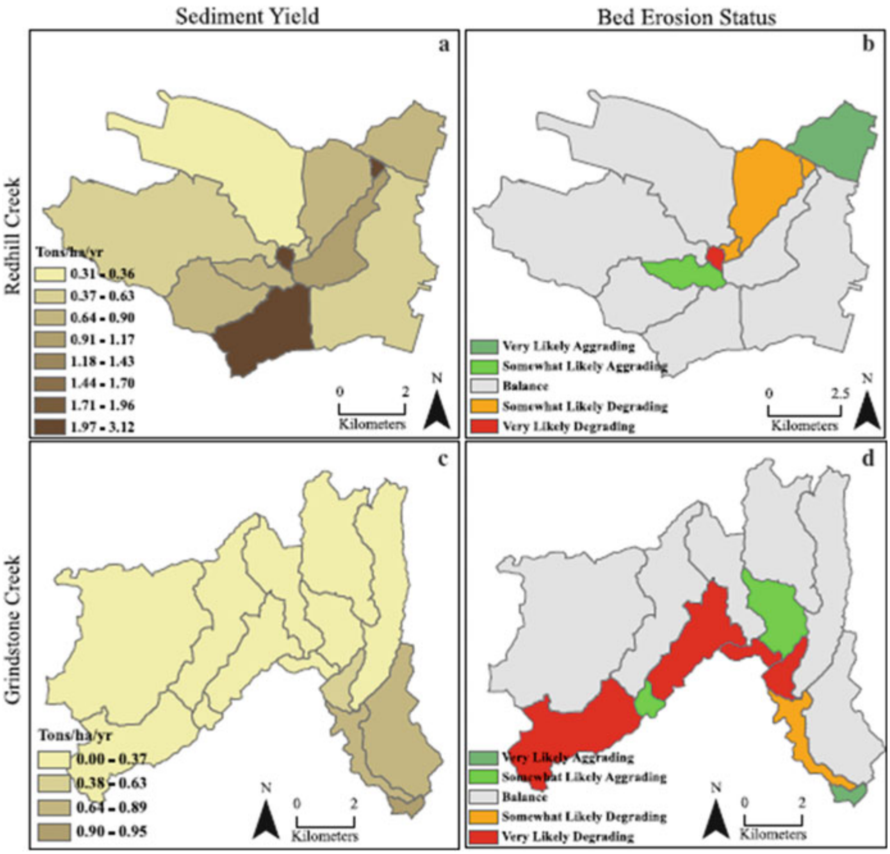


Fig. 11.9 Estimated sediment yield and bed erosion status for Redhill and Grindstone Creeks [Reproduced from Wellen et al. (2014c)]

with zero, reliable statements could be made about whether the reach was gaining or losing sediment during the period 2010–2012. If the bed storage was positive, the reach was categorized as very likely aggrading, while if the bed storage was negative, the reach was categorized as very likely degrading. If there was overlap with zero, the reach was categorized as likely aggrading or degrading, depending on which side of zero the median of the distribution laid. Some reaches categorized as balanced, as their credible intervals of absolute bed storage were less than 1 ton per year. The headwater areas of both Creeks were classified as balanced, while all the reaches losing sediment from their bed are located along the main channel. The final downstream reach was characterized as gaining sediment in both Creeks, reflecting the wider streams and gentler slopes. Notably, the sub-basin characterized as having the highest class of sediment yield in Redhill Creek's southern end was in balance, indicating that the substantial agricultural sediment mass estimated to be added to the streams in that reach was largely propagated downstream. In Grindstone Creek, there are few reaches that are storing sediment. In particular, the reaches containing most of the urban area towards the mouth of the basin are either at balance or likely degrading, implying that much of the urban sediment added to Grindstone Creek is exported downstream.

11.3 Bay of Quinte

11.3.1 Introduction

The Bay of Quinte, a Z-shaped embayment at the northern end of Lake Ontario, has experienced a long history of eutrophication problems, characterized by frequent and spatially extensive algal blooms, predominance of toxic cyanobacteria, dominance (or invasion) of undesirable fish species, and destruction of wildlife habitats (Arhonditsis et al. 2016, Shimoda et al. 2016). Because of these ecological degradation problems, the Great Lakes Water Quality Agreement between the United States and Canada established a number of objectives, guidelines, and initiatives to restore and maintain physicochemical and biological integrity. The Bay of Quinte was designated as one of the 43 Areas of Concern around the Great Lakes by the International Joint Commission (IJC) in 1986, whereby the Canadian government made a commitment to introduce a comprehensive action plan that primarily aimed to control nutrient loading from municipal sewage treatment plants. Phosphorus reduction in detergents along with upgrades at the WWTPs resulted in a dramatic reduction (>95%) of the phosphorus discharges from the 1960s, 215 kg day^{-1}, to the 2000s, <10 kg day^{-1} (Kinstler and Morley 2011).

Despite the substantial improvement of the ambient water quality conditions, high P concentrations and summer cyanobacteria blooms remain a central issue in the bay (Watson et al. 2011). Invasions of zebra (*Dreissena polymorpha*) and quagga (*Dreissena bugensis*) mussels have further complicated ecosystem structure and functioning since the mid-1990s (Dermott and Bonnell 2011). In the post-dreissenid era, total phosphorus concentrations demonstrate significant within-year variability, characterized by relatively low spring and fall levels, 10–15 μg TP L^{-1}, and high summer concentrations, $>$ 50 μg TP L^{-1} (Shimoda et al. 2016). This ambient TP variability may also stem from the biological nutrient regeneration and sediment diagenesis processes, reflecting the impact of the memory of the system (Kim et al. 2013).

Existing empirical evidence suggests that the presence of dreissenids may have led to structural changes that could ultimately be translated into an ecosystem regime shift (deYoung et al. 2008). Namely, in the Bay of Quinte, increased light penetration resulting from dreissenid filtration of suspended solids stimulated the growth of submerged macrophytes that rapidly proliferated into deeper waters (Leisti et al. 2012). Regarding the phytoplankton community, the dreissenid invasion could cause shifts of the algal assemblage stemming directly from their feeding selectivity or indirectly from an increase in water column transparency, although the role of the feedback loop associated with their nutrient recycling activity could not be ruled out (Arhonditsis et al. 2016). Specifically, the arrival of dreissenid mussels coincided with both desirable (e.g., *Aphanizomenon* and *Oscillatoria* decline) and undesirable (e.g., *Microcystis* increase) shifts in the phytoplankton community composition (Shimoda et al. 2016). The increased frequency of harmful algal blooms in the post-dreissenid period has profound ramifications for several beneficial use impairments in the Bay of Quinte, such as *Eutrophication or undesirable algae*, *Restrictions on drinking water or taste and odor problems*, and *Degradation of aesthetics*.

Environmental modeling has been an indispensable tool of the Bay of Quinte restoration efforts and a variety of data-oriented and process-based models have been used for elucidating ecosystem dynamics and evaluating the likelihood of delisting the system as an AOC. Quite recently, a network of models was developed to connect the watershed processes with the dynamics of the Bay of Quinte (Zhang et al. 2013; Arhonditsis et al. 2016; Kim et al. 2013, 2016, 2017). This integrated watershed-receiving water body modeling framework has been used to evaluate management scenarios that would lead to significant reduction of phosphorus export from the Bay of Quinte watershed and to quantify the overall uncertainty associated with the severity of the eutrophication phenomena in the area (Fig. 11.10).

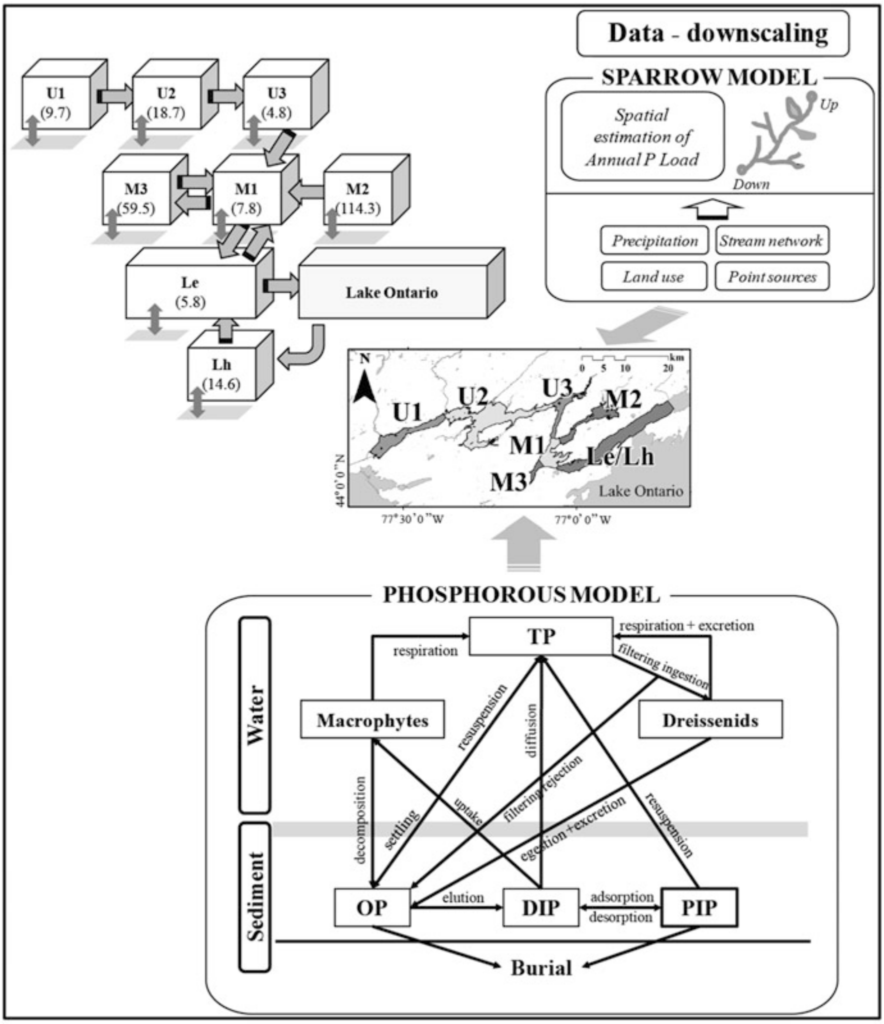


Fig. 11.10 Conceptual diagram of the integrated phosphorus-modeling framework for the Bay of Quinte. The spatial segmentation of the model for the receiving water body consists of the following compartments: (*U1*) the segment that extends from the mouth of Trent River until the city of Belleville; (*U2*) the segment that begins from the mouth of Moira River and comprises the Big Bay, Muscote Bay, and North Point Bay; and (*U3*) the area influenced by the inflows of Napanee River, extending until the outlet of Hay Bay. In the middle Bay, there are three segments corresponding to the main stem (*M1*) and the two adjacent embayments: Hay Bay (*M2*), and Picton Bay (*M3*). The lower segment of the Bay, representing the transitional area to Lake Ontario, was separated into the epilimnetic (*Le*) and hypolimnetic (*Lh*) compartments. *Numbers in parentheses* correspond to the average flushing rate of each segment [Reproduced from Arhonditsis et al. (2016)]

11.3.2 Modeling the Relationship Among Watershed Physiography, Land Use Patterns, and Phosphorus Loading

One of the emerging imperatives of eutrophication management is the advancement of our understanding of the relationships among land use, agricultural activities, hydrological processes, and water quality (Wellen et al. 2015). Prior to the watershed modeling exercise, Kim et al. (2016) implemented Self-Organizing Maps (SOM) to gain insights into the physiographical features and land-use patterns in the Bay of Quinte watershed, and to subsequently associate them with the phosphorus non-point source loading. In this application, eighteen classification variables were used, such as the landscape slope, saturated soil hydraulic conductivity, soil bulk density, and areal fractions for different land use types (lakes, ponds, alvars, bogs, coniferous swamps, deciduous swamps, fens, marshes, deciduous forests, coniferous forests, cutovers, mining areas, urban lands, pastures, and croplands) in 73 gauged and 137 ungauged subwatersheds. Thus, a total of 210 spatial units were distributed on 2-dimensional hexagonal maps, and then clustered in different groups according to their similarities.

Based on the spatial heterogeneity of these classification variables, SOM delineated six spatial clusters in the Bay of Quinte watershed with fairly distinct land-use patterns (Fig. 11.11). Coniferous and deciduous coverage along with pastures and croplands dominate the landscape in cluster 1. Different types of wetlands, such as fen ($\approx$10%), coniferous swamp ($\approx$8%), and alvar ($\approx$0.4%) have also their highest areal fraction values in the same cluster. In cluster 2, the average landscape slope is steep and the soil bulk density is high. The areal fractions of forests as well as mining and logging sites are also high. In cluster 3, most of the subwatersheds are located in the vicinity of the Bay of Quinte, where crops occupy $\approx$75% of the area. Not surprisingly, the annual TP yield per area and average TP concentrations are the highest (528 kg km^{-2} year^{-1} and 103 μg L^{-1}) in these same regions. In cluster 4, soil hydraulic conductivity is significantly higher, deciduous swamp are more abundant relative to the rest of the watershed, cropland coverage is the second highest ($\approx$41%), and thus the net TP export is high. In cluster 5, urban land represents $\approx$74% of the land-use coverage and net TP export and yield are the second highest (3.72 tonnes year^{-1} and 209 kg km^{-2} year^{-1}), which is further accentuated by the increased point source loading (2.44 tonnes year^{-1}). In cluster 6, pasture and cropland approximately correspond to 60% of the area, and these subwatersheds are mainly located adjacent to the Bay of Quinte.

Nutrient loads, yields, and deliveries at landscape and regional scales were estimated using the SPARROW model (Kim et al. 2017). The goodness-of-fit between observed and predicted TP loading values from the SPARROW model was excellent in the logarithmic scale ($r^2 > 0.95$), although there were four sites with errors greater than 10 tonnes year^{-1} when the SPARROW predictions were back-transformed to the original scale. The posterior parameter values offered insights into the patterns of phosphorus export and delivery in the Bay of Quinte

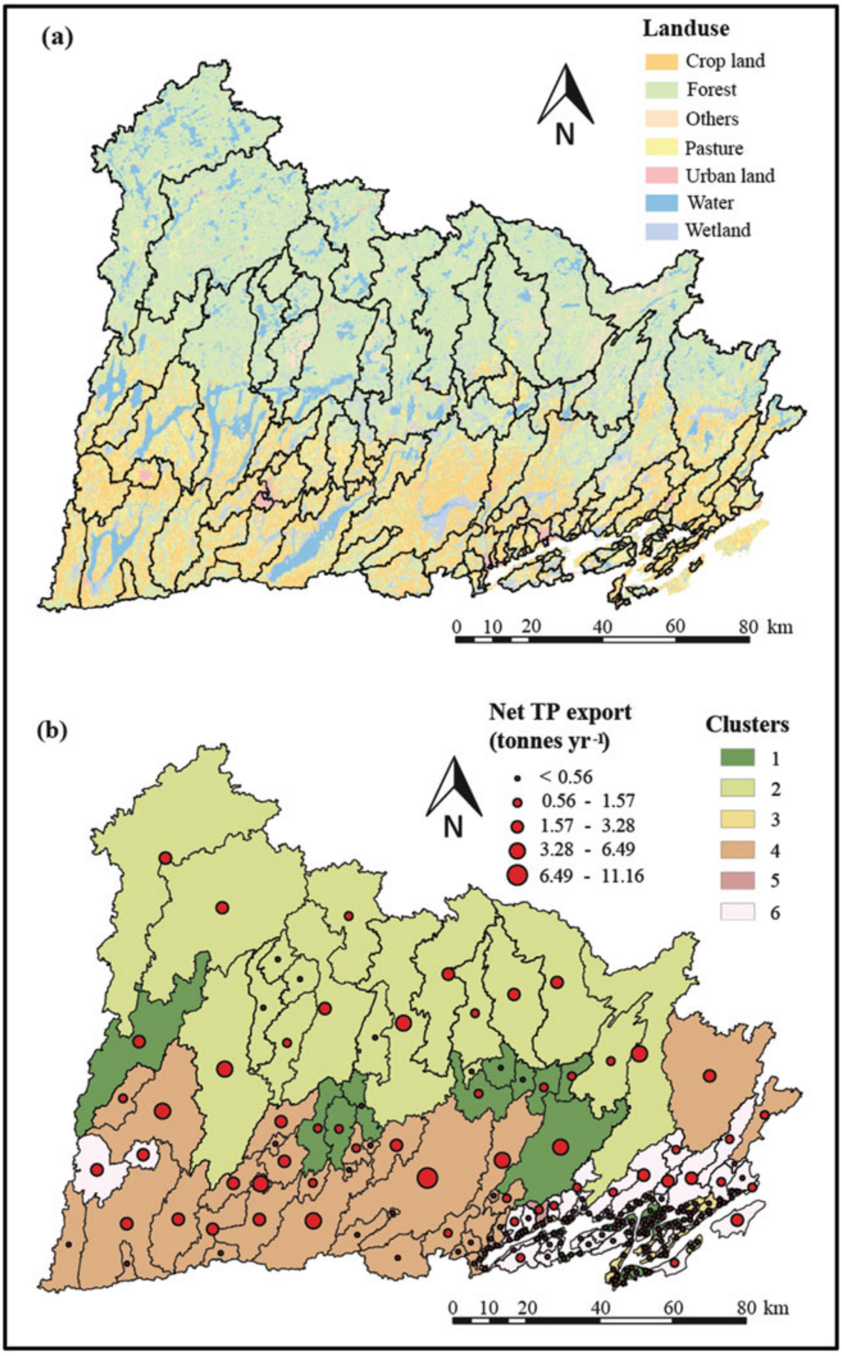


Fig. 11.11 Map of Bay of Quinte watershed: (**a**) land use types, and (**b**) classification based on artificial neural networks and associated phosphorus export per subwatershed [Reproduced from Kim et al. (2016)]

watershed. The main findings from the SPARROW modeling exercise were as follows: (1) urban areas are characterized by a fairly high areal phosphorus export with a mean estimate of 126 kg of TP per km^2 on an annual basis; (2) the contribution of phosphorus from agricultural land uses can vary considerably among the various crop types (30–127 TP kg per km^2), but is generally lower than the impact of urban sites. Similar to the Hamilton Harbour, this finding contradicts the popular notion that rates of nutrient export from urban lands are below those of agricultural lands due to lower anthropogenic nutrient subsidies, such as fertilizer implementation (Moore et al. 2004; Soldat et al. 2009). Nonetheless, other studies in the region of Southern Ontario have found urban total phosphorus export rates to be comparable (or even higher) than agricultural total phosphorus export rates (Winter and Duthie 2000); (3) the crop-specific export coefficient values were on par with those typically reported in the literature (Harmel et al. 2008); (4) the attenuation rate in low flow streams (3.7% of TP per kilometer) appears to be distinctly greater than in those with high flow (1.1% of TP per kilometer); and (5) fallow areas are responsible for approximately 70 kg of TP per km^2 on an annual basis.

In the context of watershed management, the spatial distribution of net (instead of the cumulative) TP loading that ultimately inflows into the receiving waterbody was used to identify the most influential subwatersheds (Fig. 11.12). The percentage of net loading was mostly greater in the downstream catchment of the major tributaries. By contrast, the relative contribution of the ungauged watersheds close to the bay was significantly lower primarily due to their small areal extent (Fig. 11.12a). On the other hand, the error associated with the estimates of the relative contribution of the different subwatersheds was higher in the Trent River basin (SE> 67%) than the rest of the tributaries. Interestingly, the Trent River's upper catchment also exhibited high variability in the percentage net TP loads (Fig. 11.12b). The coefficient of variation (CV) values of the relative contributions along with the net contributions normalized by the corresponding subwatershed areas were also used to delineate the hot-spots in the Bay of Quinte watershed. The highest CVs (>32%) were found in the upper catchment of Trent River (Fig. 11.12c). Counter to the error estimates, however, the ungauged watershed close to the bay was characterized by fairly high CVs (Fig. 11.12c). This trend was more pronounced when the normalized percentage TP loads were considered (Fig. 11.12d). Unlike the CV values, the normalized percentage TP loads were low in the upper catchment of Trent River, but were distinctly higher in the lower part of the watershed, especially near the bay (Fig. 11.12d). Overall, this strategy pinpointed many locations close to the water body that may be responsible for significant nutrient fluxes, due to their landscape attributes and soil characteristics (Kim et al. 2017).

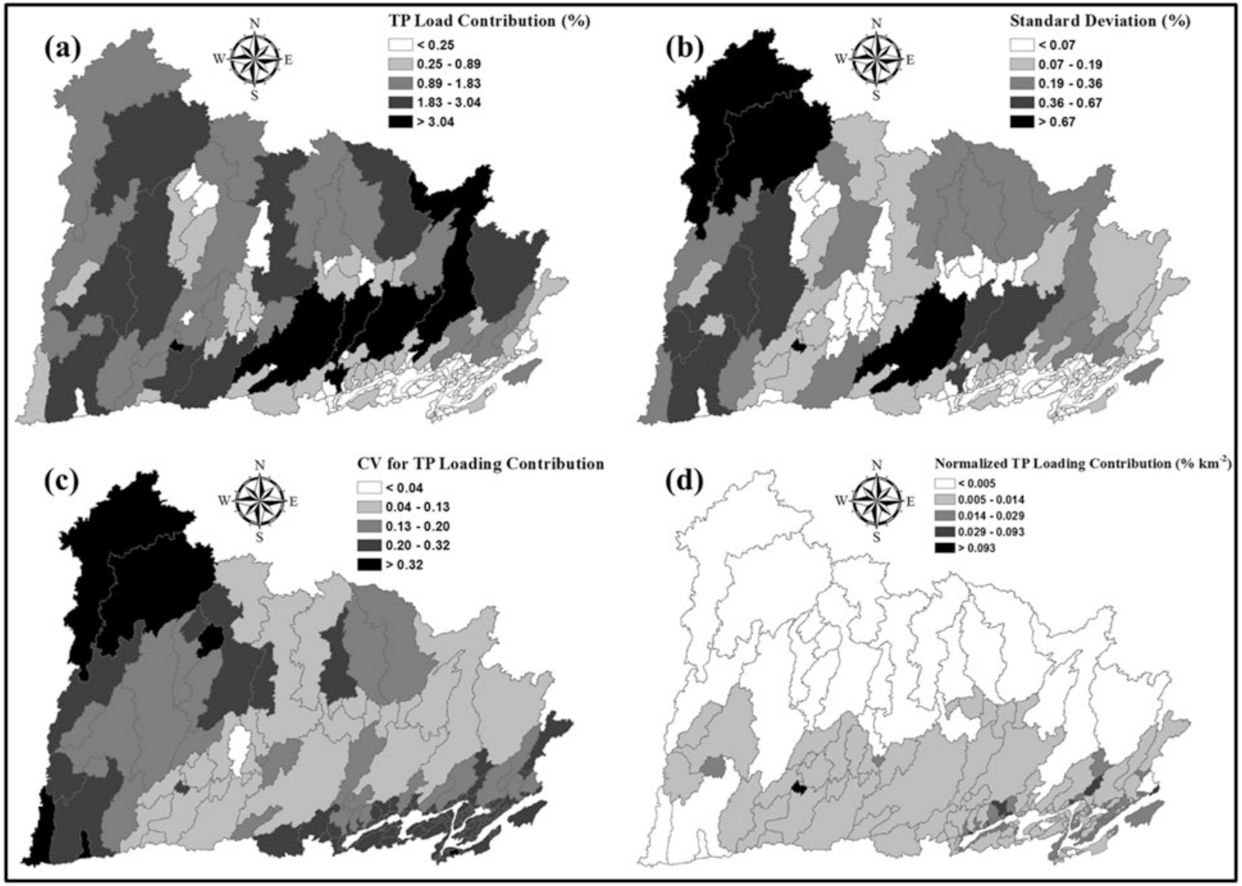


Fig. 11.12 Percentage contribution of the annual net TP loads to the Bay of Quinte: (**a**) average prediction, (**b**) standard error (SE) and (**c**) coefficient of variation of the corresponding predictions, and (**d**) average prediction normalized by the subwatershed areas [Reproduced from Kim et al. (2017)]

11.3.3 Eutrophication Risk Assessment with Process-Based Modeling and Determination of Water Quality Criteria

The basis of the eutrophication risk assessment analysis was the mechanistic model presented by Kim et al. (2013), which introduced several novel mathematical formulations regarding the representation of macrophyte dynamics; the role of dreissenids in the system; several processes related to the fate and transport of phosphorus in the sediments along with the interplay between water column and sediments, such as particulate sedimentation being dependent upon the standing algal biomass, sediment resuspension, sorption/desorption in the sediment particles, and organic matter decomposition. The model was then calibrated to match the measured TP concentrations in the upper, middle, and lower segments of the Bay during the 2002–2009 period (Kim et al. 2013; Arhonditsis et al. 2016). The model demonstrated satisfactory ability to fit the monthly TP levels in the Bay of Quinte, and was able to reproduce the end-of-summer increase of the ambient TP levels in the upper segment, even in years (e.g., 2005) when the corresponding concentrations were greater than 60 μg L^{-1}. The model also faithfully depicted the spatial gradients in the system, with distinctly higher TP levels in the upper segment relative to those experienced in the middle/lower Bay (Kim et al. 2013).

The model was then used to draw inferences on the spatial variability of the various external and internal TP flux rates in the Bay of Quinte (Fig. 11.13). The net TP contributions (sources or sinks) represent the mass of phosphorus associated with the various compartments (water column, sediments, macrophytes, dreissenids) throughout the growing season (May–October) averaged over the 2002–2009 period. In the U1 segment, the phosphorus budget is predominantly driven by the external sources (phosphorus loading: 159 kg day^{-1}) and sinks (outflows: 152 kg day^{-1}). The sediments (resuspension and diffusion from the sediments to water column minus particle settling) act as a net source of phosphorus in this segment (57.9 kg day^{-1}). Dreissenids subtract approximately 65.9 kg day^{-1} from the water column (particle filtration minus respiration) and subsequently deposit 62.4 kg day^{-1} via their excretion and particle rejection. In a similar manner, the U2 segment receives 206 kg day^{-1} from exogenous sources, including the upstream inflows, and transports downstream 190 kg day^{-1}. The net contribution of the sediments accounted for 70.4 kg day^{-1}, while dreissenids on average reduce the ambient TP levels by 112 kg day^{-1}. The main differences between the two segments in the upper Bay are the TP fluxes related to macrophyte P intake from the sediments and respiration that can reach the levels of 46.9 and 42.3 kg day^{-1} relative to the fluxes of 11.1 and 10.0 kg day^{-1} in the U1 segment. Likewise, the macrophyte intake from the sediments minus the amount of P regenerated from the decomposition of the dead plant tissues varies between 35 and 65 kg day^{-1} in segments U3 and M1, while the subsequent release of their metabolic by-products is approximately responsible for 19–26 kg day^{-1}. The settling of particulate P dominates over the resuspension and diffusion from the sediments to the water column

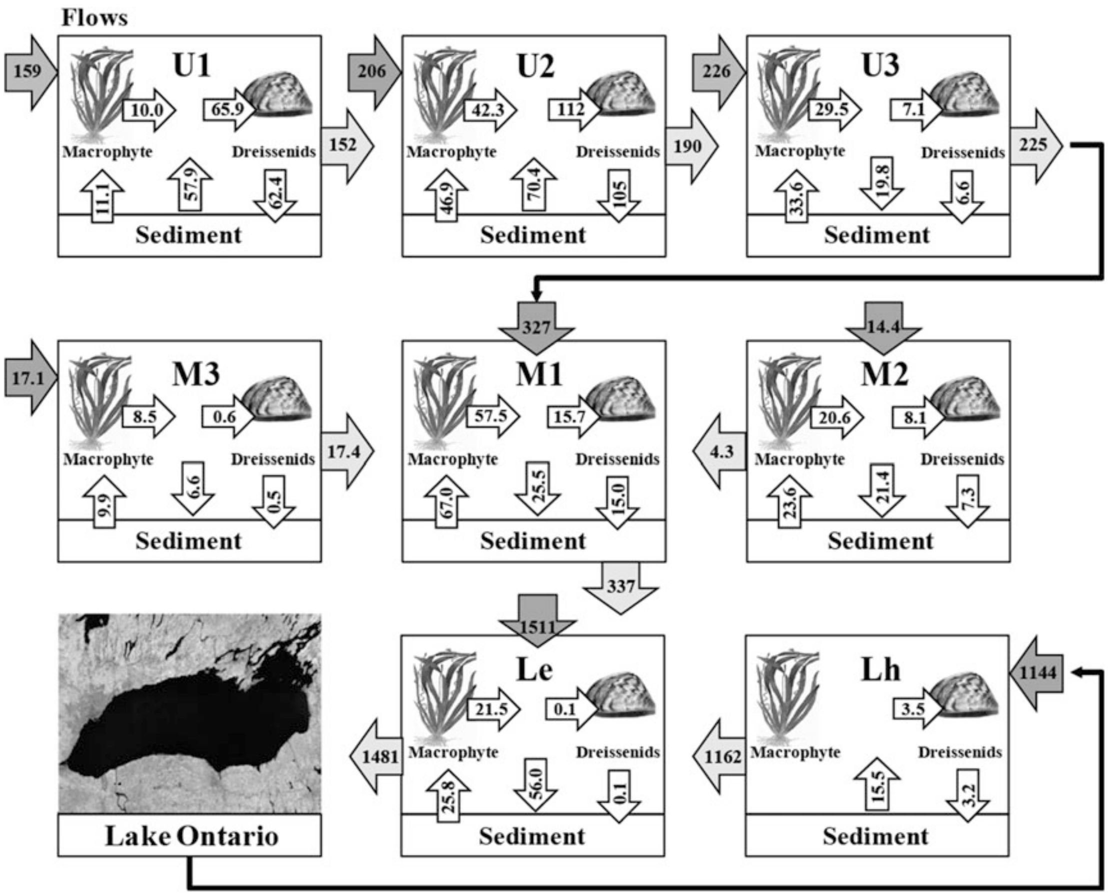


Fig. 11.13 Spatial variability of the various external and internal TP flux rates (kg day^{-1}) in the Bay of Quinte. *Arrow directions* indicate the net contribution (sources or sinks) of the various compartments (water column, sediments, macrophytes, dreissenids). *Dark gray arrows* show the TP inflows in a spatial segment, while the *light-gray arrows* depict the corresponding outflows [Reproduced from Arhonditsis et al. (2016)]

with the corresponding net fluxes ranging between 25 and 35 kg day^{-1}. In Hay Bay (M2), the fluxes mediated by the macrophytes and dreissenids primarily modulate the TP dynamics and the same pattern appears to hold true in Picton Bay (M3). In the lower Bay of Quinte (Le and Lh), the model postulates a significant pathway ($>$1100 kg P day^{-1}) through which the inflowing water masses from Lake Ontario well up from the hypolimnion to the epilimnion and are subsequently exported from the system. In the same area, the internal biotic sources (macrophytes) similarly represent an important vector of phosphorus transport.

In general, the Bay of Quinte modeling work highlights the internal recycling as one of the key drivers of phosphorus dynamics. The flow from the Trent River is the predominant driver of the dynamics in the upper segment until the main stem of the middle area. However, the sediments in the same segment release a significant amount of phosphorus and the corresponding fluxes are likely amplified by the macrophyte and dreissenid activity. From a management standpoint, the presence of a significant positive feedback loop in the upper Bay of Quinte suggests that the anticipated benefits of additional reductions of the exogenous point and non-point loading may not be realized within a reasonable time frame, i.e., 5–10 years (Kim et al. 2013). Analysis of nutrient loading scenarios showed that the restoration pace of the Bay could be slow, even if the riverine total phosphorus concentrations reach levels significantly lower than their contemporary values, $<$25 μg TP L^{-1} (Fig. 11.14; see also Kim et al. 2013; Arhonditsis et al. 2016).

Bearing in mind that the TP targeted levels merely represent a "means to an end" and not "the end itself", the actual question that the stakeholders in the area ponder is to what extent the anticipated benefits from a more efficient external phosphorus loading control could also be capitalized as a significant decrease of the algal bloom frequency? With respect to the total phytoplankton biovolume, Nicholls et al. (2002) showed that it declined after the control of phosphorus in the 1970s, but did not change significantly after the establishment of dreissenids in the system. As previously mentioned, Nicholls and Carney (2011) showed that the arrival of dreissenid mussels may be associated with positive (e.g., *Aphanizomenon* and *Anabaena* decline) effects on the integrity of the Bay of Quinte ecosystem. However, the recent increase of the cyanophyte *Microcystis* has had significant implications for the aesthetics and other beneficial uses of the Bay of Quinte, through the formation of "scums" on the water surface as well as the fact that some strains of *Microcystis* are toxin producers. These structural shifts in the phytoplankton community composition could stem directly from the feeding selectivity of dreissenids or indirectly from the improvements in the transparency of the water column (Blukacz-Richards and Koops 2012), but the role of the feedback loop associated with their nutrient recycling activity could conceivably be another important factor.

According to the predictions of a non-linear quantile regression model (Shimoda et al. 2016), the current average TP concentrations (30–40 μg L^{-1}) represent the area where the algal biovolume vs TP relationship is characterized by a steep slope and thus any further improvements in the ambient nutrient levels are likely to induce more favorable quantitative and qualitative changes in phytoplankton (Fig. 11.15). Nonetheless, existing empirical evidence from the system is indicative

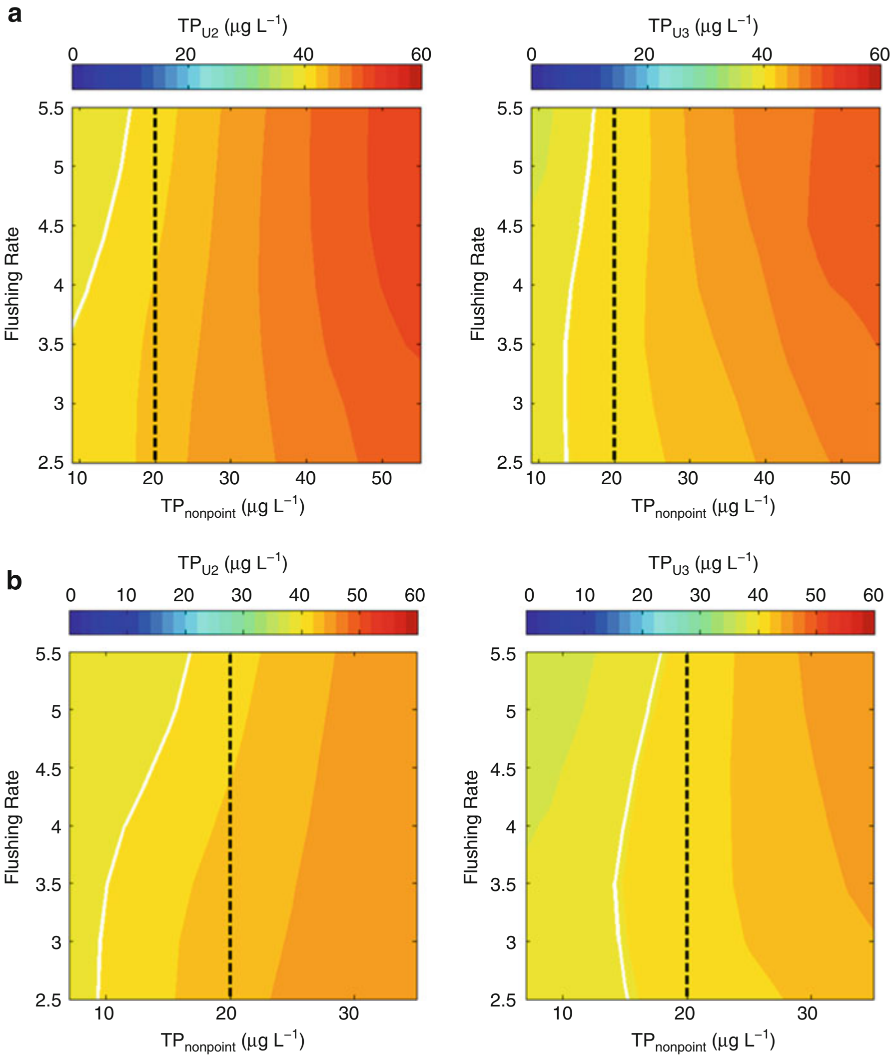


Fig. 11.14 Simulated maximum TP concentrations during the growing season (May–October) in the Bay of Quinte. *Upper panels* (**a**) refer to the predictions associated with the reference environmental conditions; and *lower panels* (**b**) represent the predictions of a TP loading reduction scenario (60% point sources, 20% non-point sources, and 50% urban storm water). The first eleven years (2002–2012) were based on real meteorological and nutrient loading conditions, while the final (12th) year was forced with a wide range of combinations of TP riverine concentrations and flows that were generated from the mean (± error) predictions of the SPARROW model. The *white contour line* corresponds to the proposed targeted level of 40 μg TP L^{-1}. The flushing rates express the frequency (number of times) of water renewal in the upper Bay during the growing season. The *black dotted line* represents a threshold level of 20 μg L^{-1} for the flow-weighted TP concentration in all the major tributaries in the upper Bay of Quinte [Reproduced from Arhonditsis et al. (2016)]

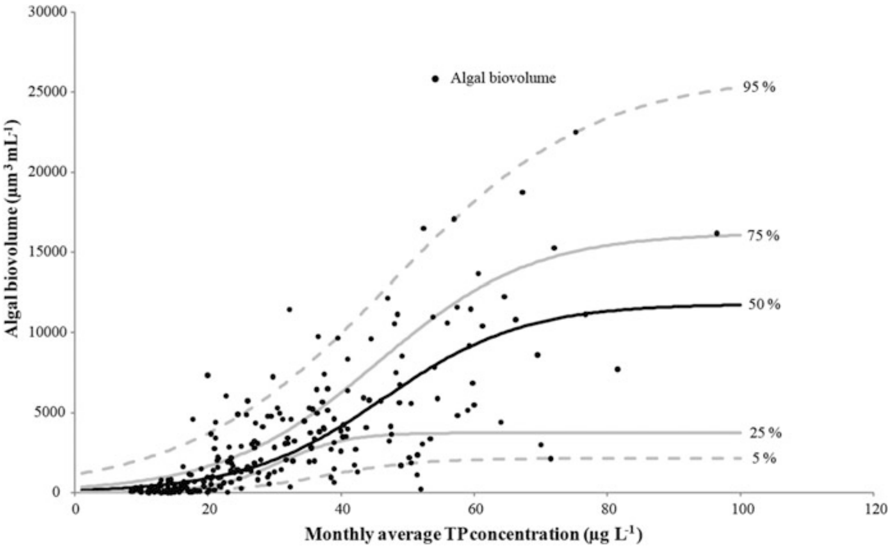


Fig. 11.15 Quantile regression model for total phytoplankton biovolume against monthly average TP concentration in the Bay of Quinte (Arhonditsis et al. 2016)

of a weak correlation between chlorophyll a and cyanobacteria toxin concentrations (Watson et al. 2011), suggesting that a complex interplay among physical, chemical, and biological factors may drive the spatiotemporal abundance and composition patterns of the algal assemblages in the Bay of Quinte (Nicholls et al. 2002). In a system like the Bay of Quinte, where both external and internal loading drives the severity of eutrophication phenomena, there will inevitably be some uncertainty in the overall eutrophication risk assessment.

There are several compelling reasons (knowledge gaps, natural variability, complex interactions among a suite of ecological mechanisms) to avoid overly confident statements about the future response of this impaired system, and thus the most prudent strategy is to explicitly recognize an acceptable level of violations of the delisting goals. Specifically, Kim et al. (2013) challenged the usefulness of the historical delisting criterion of a seasonal average TP concentration lower than 30 μg L^{-1}, as it is neither a reflection of the considerable intra-annual variability in the upper Bay nor representative of the water quality conditions in near shore areas of high public exposure (e.g., beaches). It would seem very unlikely that a single-value water quality standard monitored in a few offshore sampling stations can capture the entire range of dynamics in the system (e.g., the extremes seen in the near shore sites) or the magnitude of the end-of-summer TP peaks. Kim et al. (2013) instead advocated the pragmatic stance that the delisting objectives should revolve around extreme (and not average) values of variables of management interest and must explicitly accommodate all the sources of uncertainty (insufficient information, lack of knowledge, and natural variability) by permitting a realistic frequency of standard violations. Namely, the critical threshold level should be set at a value of 40 μg TP L^{-1}, which cannot be exceeded more than 10–15% in both time and space. Under the assumption that the TP concentrations in the Bay of Quinte follow a log-normal distribution and that TP values <15 μg L^{-1} are likely to occur only 10% of the time during the growing season, then 10–15% exceedances of the 40 μg TP L^{-1} level are approximately equivalent to a targeted seasonal average of 25–28 μg TP L^{-1}. Thus, the replacement of the historical paradigm (binary assessment) with a probabilistic approach to water quality criteria does not intend to make the delisting of AOCs easier, but rather to offer a more comprehensive method for tracking the prevailing conditions in the Bay.

11.4 Concluding Remarks

We have demonstrated some of the benefits for environmental management when identifying the uncertainties and knowledge gaps of the natural environment, differentiating between predictable and unpredictable patterns, and critically evaluating model outputs. The presentation of the model outputs as a probabilistic assessment of environmental conditions makes the model results more credible for local decision makers and stakeholders. The often-misleading deterministic statements are avoided and environmental goals are set by explicitly acknowledging an inevitable risk of not achieving 100% compliance in time and space. The acceptable level of violations is then subject to decisions that reflect different socioeconomic values and environmental priorities.

The Bayesian (iterative) nature of the presented modeling networks is conceptually similar to the policy practice of adaptive management, i.e., an iterative implementation strategy that is recommended to address the often-substantial uncertainty associated with water quality model forecasts and avoid the implementation of inefficient and flawed management plans. The use of Bayesian inference techniques is also consistent with the scientific process of progressive learning and offers a natural mechanism for sequentially updating our knowledge on model inputs and structure every time new data are collected from the system. Thus, modeling tools can be iteratively updated to accommodate the significant year-to-year variability associated with the external nutrient loading or the weather conditions, thereby serving as a reliable long-term management tool for policy analysis. Importantly, the probabilistic statements provided from the Bayesian calibration can also indicate where the limited monitoring resources should be focused (Zhang and Arhonditsis 2008). In particular, additional data collection efforts should target hot spots, where the model predictive distribution indicates a high probability of non-attaining water quality goals or, alternatively, an *unacceptably high* variance. Thus, we can assess the value of information (value of additional monitoring; "Where should additional data collection efforts be focused?") and subsequently optimize the sampling design for environmental monitoring. In other words, uncertainty does matter and its quantification is not an excuse to avoid providing answers to pressing environmental problems, but rather a prudent strategy to improve the rigor of model-based management of our natural resources!

Acknowledgements George Arhonditsis wishes to acknowledge the continuous support of his work on model uncertainty analysis from the National Sciences and Engineering Research Council of Canada (Discovery Grants). The Hamilton Harbour modeling project has received funding support from the Ontario Ministry of the Environment (Canada-Ontario Grant Agreement 120808). The Bay of Quinte modeling project was undertaken with the financial support of the Lower Trent Region Conservation Authority provided through the Bay of Quinte Remedial Action Plan Restoration Council.

References

Alexander RB, Smith RA, Schwarz GE (2004) Estimates of diffuse phosphorus sources in surface waters of the United States using a spatially referenced watershed model. Water Sci Technol 49:1–10

Arhonditsis GB, Adams-Van Harn BA, Nielsen L et al (2006) Evaluation of the current state of mechanistic aquatic biogeochemical modeling: citation analysis and future perspectives. Environ Sci Technol l40:6547–6554

Arhonditsis GB, Qian SS, Stow CA et al (2007) Eutrophication risk assessment using Bayesian calibration of process-based models: application to a mesotrophic lake. Ecol Model 208:215–229

Arhonditsis GB, Papantou D, Zhang W et al (2008a) Bayesian calibration of mechanistic aquatic biogeochemical models and benefits for environmental management. J Marine Syst 73:8–30

Arhonditsis GB, Perhar G, Zhang W et al (2008b) Addressing equifinality and uncertainty in eutrophication models. Water Resour Res 44:W01420

Arhonditsis GB, Kim D-K, Shimoda Y et al (2016) Integration of best management practices in the Bay of Quinte watershed with the phosphorus dynamics in the receiving water body: What do the models predict? Aquat Ecosyst Health Manage 19:1–18

Basu NB, Rao PSC, Thompson SE et al (2011) Spatiotemporal averaging of in-stream solute removal dynamics. Water Resour Res 47:W00J06

Bayarri MJ, Berger JO, Cafeo J et al (2007) Computer model validation with functional output. Ann Stat 35:1874–1906

Beck ME (1987) Tectonic rotations on the leading edge of South America: the Bolivian orocline revisited. Geology 15:806–808

Beven K (1993) Prophecy, reality and uncertainty in distributed hydrological modelling. Adv Water Resour 16:41–51

Beven K (2006) A manifesto for the equifinality thesis. J Hydrol 320:18–36

Blukacz-Richards EA, Koops MA (2012) An integrated approach to identifying ecosystem recovery targets: application to the Bay of Quinte. Aquat Ecosyst Health Manage 15:464–472

Charlton MN (2001) The Hamilton Harbour remedial action plan: eutrophication. Verh Internat Verein Limnol 27:4069–4072

Claessens L, Tague CL, Band LE et al (2009) Hydro-ecological linkages in urbanizing watersheds: an empirical assessment of in-stream nitrate loss and evidence of saturation kinetics. J Geophys Res Biogeosci 114:G04016

Dawes RM (1988) Rational choice in an uncertain world. Harcourt Brace Jovanovich, San Diego

Dermott R, Bonnell R (2011) Benthic fauna in the Bay of Quinte. Bay of Quinte remedial action plan: Monitoring Report #20, Kingston, ON, pp 51–71

deYoung B, Barange M, Beaugrand G et al (2008) Regime shifts in marine ecosystems: detection, prediction and management. Trends Ecol Evol 23:402–409

Dietzel A, Reichert P (2012) Calibration of computationally demanding and structurally uncertain models with an application to a lake water quality model. Environ Modell Softw 38:129–146

Donner SD, Kucharik CJ, Oppenheimer M (2004) The influence of climate on in-stream removal of nitrogen. Geophys Res Lett 31:L20509

Doyle MW, Stanley EH, Harbor JM (2003) Hydrogeomorphic controls on phosphorus retention in streams. Water Resour Res 39:1147

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

Gudimov A, Stremilov S, Ramin M, Arhonditsis GB (2010) Eutrophication risk assessment in Hamilton Harbour: system analysis and evaluation of nutrient loading scenarios. J Great Lakes Res 36:520–539

Gudimov A, Ramin M, Labencki T et al (2011) Predicting the response of Hamilton Harbour to the nutrient loading reductions: a modeling analysis of the "ecological unknowns". J Great Lakes Res 37:494–506

HH RAP (2003) Hamilton Harbour Remedial Action Plan, Report Stage 2 Update. Hamilton Harbour Technical Team. Burlington, ON

Hall JD, O'Connor K, Ranieri J (2006) Progress toward delisting a Great Lakes Area of Concern: the role of integrated research and monitoring in the Hamilton Harbour Remedial Action Plan. Environ Monit Assess 113:227–243

Hall JD, O'Connor KM (2016) Hamilton Harbour remedial action plan process: connecting science to management decisions. Aquat Ecosyst Health Manage 19:107–113

Harmel D, Qian S, Reckhow K, Casebolt P (2008) The MANAGE database: nutrient load and site characteristic updates and runoff concentration data. J Environ Qual 37:2403–2406

Hiriart-Baer VP, Milne J, Charlton MN (2009) Water quality trends in Hamilton Harbour: two decades of change in nutrients and chlorophyll a. J Great Lakes Res 35:293–301

Hiriart-Baer VP, Boyd D, Long T et al (2016) Hamilton Harbour over the last 25 years: insights from a long-term comprehensive water quality monitoring program. Aquat Ecosyst Health Manage 19:124–133

Kim D-K, Zhang W, Rao Y et al (2013) Improving the representation of internal nutrient recycling with phosphorus mass balance models: a case study in the Bay of Quinte, Ontario, Canada. Ecol Model 256:53–68

Kim D-K, Zhang W, Watson S, Arhonditsis GB (2014) A commentary on the modelling of the causal linkages among nutrient loading, harmful algal blooms, and hypoxia patterns in Lake Erie. J Great Lakes Res 40:117–129

Kim D-K, Kaluskar S, Mugalingam S, Arhonditsis GB (2016) Evaluating the relationships between watershed physiography, land use patterns, and phosphorus loading in the Bay of Quinte, Ontario, Canada. J Great Lakes Res 42:972–984

Kim D-K, Kaluskar S, Mugalingam S et al (2017) A Bayesian approach for estimating phosphorus export and delivery rates with the SPAtially Referenced Regression On Watershed attributes (SPARROW) model. Ecol Inform 37:77–91

Kinstler P, Morley A (2011) Point source phosphorus loadings 1965 to 2009. Bay of Quinte remedial action plan: monitoring report #20. Kingston, ON, pp 15–17

Leisti KE, Doka SE, Minns CK (2012) Submerged aquatic vegetation in the Bay of Quinte: Response to decreased phosphorous loading and Zebra Mussel invasion. Aquat Ecosyst Health Manage 15:442–452

Long T, Wellen C, Arhonditsis G, Boyd D (2014) Evaluation of stormwater and snowmelt inputs, land use and seasonality on nutrient dynamics in the watersheds of Hamilton Harbour, Ontario, Canada. J Great Lakes Res 40:964–979

Long T, Wellen C, Arhonditsis G et al (2015) Estimation of tributary total phosphorus loads to Hamilton Harbour, Ontario, Canada, using a series of regression equations. J Great Lakes Res 41:780–793

McDowell RW, Srinivasan MS (2009) Identifying critical source areas for water quality: 2. Validating the approach for phosphorus and sediment losses in grazed headwater catchments. J Hydrol 379:68–80
McMahon G, Alexander RB, Qian S (2003) Support of total maximum daily load programs using spatially referenced regression models. J Water Res 129:315–329
Moore RB, Johnson CM, Robinson KW, Deacon JR (2004) Estimation of total nitrogen and phosphorus in New England streams using spatially referenced regression models. US Department of the Interior, US Geological Survey, New Hampshire, p 42
Morgan MG, Henrion M, Small M (1992) Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York
Nicholls KH, Heintsch L, Carney E (2002) Univariate step-trend and multivariate assessments of the apparent effects of P loading reductions and zebra mussels on the phytoplankton of the Bay of Quinte, Lake Ontario. J Great Lakes Res 28:15–31
Nicholls KH, Carney EC (2011) The phytoplankton of the Bay of Quinte, 1972–2008: point-source phosphorus loading control, dreissenid mussel establishment, and a proposed community reference. Aquat Ecosyst Health Manage 14:33–43
Pappenberger F, Beven KJ (2006) Ignorance is bliss: or seven reasons not to use uncertainty analysis. Water Resour Res 42:W05302
Ramin M, Stremilov S, Labencki T et al (2011) Integration of numerical modeling and Bayesian analysis for setting water quality criteria in Hamilton Harbour, Ontario, Canada. Environ Modell Softw 26:337–353
Ramin M, Labencki T, Boyd D et al (2012) A Bayesian systhesis of predictions from different models for setting water quality criteria. Ecol Model 242:127–145
Reichert P, Omlin M (1997) On the usefulness of overparameterized ecological models. Ecol Model 95:289–299
Reichert P, Schuwirth N (2012) Linking statistical description of bias to multi-objective model calibration. Water Resour Res 48:W09543
Rode M, Arhonditsis G, Balin D (2010) New challenges in integrated water quality modelling. Hydrol Process 24:3447–3461
Schwarz GE, Hoos AB, Alexander RB, Smith RA (2006) The SPARROW surface water-quality model: theory, application and user documentation. U.S. Geological Survey Techniques and Methods Report, Book 6, Chapter B3; USGShttps://pubs.usgs.gov/tm/2006/tm6b3/PDF/tm6b3_part1a.pdf
Shimoda Y, Watson S, Palmer ME (2016) Delineation of the role of nutrient variability and dreissenids (Mollusca, Bivalvia) on phytoplankton dynamics in the Bay of Quinte, Ontario, Canada. Harmful Algae 55:121–136
Soldat DJ, Petrovic AM, Ketterings QM (2009) Effect of soil phosphorus levels on phosphorus runoff concentrations from turfgrass. Water Air Soil Pollut 199:33–44
Stow CA, Reckhow KH, Qian SS (2007) Approaches to evaluate water quality model parameter uncertainty for adaptive TMDL implementation. JAWRA 43:1499–1507
Watson SB, Borisko J, Lalor J (2011) Bay of Quinte harmful algal bloom programme phase I – 2009. Bay of Quinte remedial action plan: monitoring report #20. Kingston, ON, pp 27–50
Wellen C, Arhonditsis GB, Labencki T, Boyd D (2012) A Bayesian methodological framework or accommodating interannual variability of nutrient loading with the SPARROW model. Water Resour Res 48:W10505
Wellen C, Arhonditsis GB, Labencki T, Boyd D (2014a) Application of the SPARROW model in watersheds with limited information: a Bayesian assessment of the model uncertainty and the value of additional monitoring. Hydrol Process 28:1260–1283
Wellen C, Arhonditsis GB, Long T, Boyd D (2014b) Accommodating environmental thresholds and extreme events in hydrological models: a Bayesian approach. J Great Lakes Res 40:102–116

Wellen C, Arhonditsis GB, Long T, Boyd D (2014c) Quantifying the uncertainty of nonpoint source attribution in distributed water quality models: a Bayesian assessment of SWAT's sediment export predictions. J Hydrol 519:3353–3368

Wellen C, Kamran-Disfani A-R, Arhonditsis GB (2015) Evaluation of the current state of distributed watershed nutrient water quality modeling. Environ Sci Technol 49:3278–3290

Winter JG, Duthie HC (2000) Export coefficient modeling to assess phosphorus loading in an urban watershed. JAWRA 36:1053–1061

Yerubandi RR, Boegman L, Bolkhari H, Hiriart-Baer V (2016) Physical processes affecting water quality in Hamilton Harbour. Aquat Ecosyst Health Manage 19:114–123

Zhang W, Arhonditsis GB (2008) Predicting the frequency of water quality standard violations using Bayesian calibration of eutrophication models. J Great Lakes Res 34:698–720

Zhang W, Kim D-K, Rao Y et al (2013) Can simple phosphorus mass balance models guide management decision? A case study in the Bay of Quinte, Ontario, Canada. Ecol Model 257:66–79