



# A Bayesian approach for estimating phosphorus export and delivery rates with the SPATIally Referenced Regression On Watershed attributes (SPARROW) model



Dong-Kyun Kim<sup>a</sup>, Samarth Kaluskar<sup>a</sup>, Shan Mugalingam<sup>b</sup>, Agnes Blukacz-Richards<sup>c</sup>, Tanya Long<sup>d</sup>, Andrew Morley<sup>e</sup>, George B. Arhonditsis<sup>a,\*</sup>

<sup>a</sup> Ecological Modelling Laboratory, Department of Physical & Environmental Sciences, University of Toronto, Toronto, Ontario M1C1A4, Canada

<sup>b</sup> Lower Trent Conservation, Trenton, Ontario K8V 5P4, Canada

<sup>c</sup> Environment Canada & Climate Change, 867 Lakeshore Road, Burlington, Ontario L7R 4A6, Canada

<sup>d</sup> Environmental Monitoring and Reporting Branch, Ontario Ministry of the Environment & Climate Change, Toronto, Ontario M9P 3V6, Canada

<sup>e</sup> Kingston Regional Office, Ontario Ministry of the Environment & Climate Change, Kingston, Ontario K7M 8S5, Canada

## ARTICLE INFO

### Article history:

Received 5 May 2016

Received in revised form 12 December 2016

Accepted 14 December 2016

Available online 16 December 2016

### Keywords:

Bayesian inference

Bay of Quinte

Model complexity

SPARROW

Phosphorus loading

Urbanization

## ABSTRACT

The SPATIally Referenced Regression On Watershed attributes (SPARROW) model is used to predict total phosphorus (TP) export and delivery rates from different subcatchments in the Bay of Quinte watershed. Bayesian inference techniques were used to account for the uncertainty associated with the existing knowledge from the system as well as the sampling/analytical error of the calibration data. Our analysis suggests that urban areas are characterized by a fairly high areal phosphorus export with an approximate mean estimate of 120 kg of TP per km<sup>2</sup> on an annual basis. The contribution of phosphorus from agricultural land can vary considerably among the various crop types (30–127 TP kg per km<sup>2</sup>), but is generally lower than the impact of urban areas. Crop-specific (oat, wheat, corn, alfalfa, and fallow) export coefficient values were generally on par with those typically reported in the literature. Our analysis also suggests that the attenuation rate in low flow streams (3.7% of TP per kilometer) is distinctly greater than in those with high flow (1.1% of TP per kilometer). Using posterior simulations, we obtained TP loading estimates from ungauged subwatersheds in the area that were twice as high relative to values historically used. The predictive uncertainty of phosphorus export from different sub-basins was also used to delineate “hot-spots” in the Bay of Quinte watershed that may be responsible for significant nutrient fluxes, due to their landscape attributes and soil characteristics. Our predictions can be used as pointers for maximizing the value of information of additional monitoring by determining locations where data collection efforts should focus on. The key findings of the present modelling study will be ultimately linked with process-based models developed for the receiving waterbody to shed light on the causal connections among phosphorus loading, sediment–water column interactions, and plankton community response.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

One of the emerging imperatives of eutrophication management is the advancement of our understanding of the relationships among land-use activities, hydrological processes, and water quality (Wellen et al., 2015). The profound implications of excessive nutrient enrichment for the functioning of the receiving waterbodies highlight the importance of obtaining robust non-point source nutrient load estimates and rigorously examining optimization scenarios of the agricultural cycle (tilling, planting, fertilization, irrigation, pesticide application,

harvesting, grazing) and associated management practices (buffer strips, contour tillage, tile drainage, grassed waterways). Watershed modelling is considered a useful analysis tool to quantify exogenous phosphorus loading associated with watershed attributes, including land uses and morphological characteristics (Singh and Frevert, 2006). Many sophisticated watershed models—the Hydrological Simulation Program Fortran (HSPF, Bicknell et al., 1996), Dynamic Watershed Simulation Model (DWSM, Borah et al., 2002), and Soil and Water Assessment Tool (SWAT, Arnold et al., 1998)—have been used to simulate fate and transport of a variety of pollutants at different spatiotemporal scales. Generally, HSPF is promising for long-term, continuous simulations, while DWSM is better at reproducing short-term, episodic precipitation events (Borah et al., 2002). SWAT is a physically process-

\* Corresponding author.

E-mail address: [georgea@utsc.utoronto.ca](mailto:georgea@utsc.utoronto.ca) (G.B. Arhonditsis).

based model with detailed crop growth and management modules, and as such it is suitable for projecting the long-term impact of land-use management and climate change (Neitsch et al., 2011; Wellen et al., 2014a, 2014b). These models have an established capacity to reproduce hydrology, nutrient delivery and in-stream routing, but are often overly complex and data demanding to be applied in any but the most intensively monitored catchments (Cohn et al., 1989; Runkel et al., 2004).

As a pragmatic alternative to complex mechanistic models, simple empirical watershed models have been developed at large scales, where a priori knowledge about the dominant biogeochemical process rates may not be available (Alexander et al., 2004). These simple modelling constructs can offer first approximations of annual estimates of nutrient loads, yields, and deliveries at landscape and regional scales. A characteristic example of the latter strategy is the SPATIally Referenced Regressions on Watershed attributes (SPARROW); a statistically parameterized non-linear model with contaminant supply and process components, including non-conservative transport, mass-balance constraints, and surficial flow paths defined by topography, streams, and reservoirs (Smith et al., 1997). SPARROW effectively dissects the watershed functioning into three major processes: (i) pollutant export from different land uses, (ii) land-to-stream delivery of pollutants, and (iii) attenuation and loss in streams and lakes (Schwarz et al., 2006). In terms of scalability, the SPARROW model has been applied to both small- and large-scale watersheds (Alexander et al., 2002; Alexander et al., 2004; Hantush, 2005). Watersheds are first divided into subwatersheds, each of which drains to a water quality monitoring station. Each subwatershed is then disaggregated into reach catchments drained by a particular stream segment and the attributes of each reach catchment are used as predictor variables of in-stream water quality. Maximum likelihood and bootstrapping methods have been used to obtain SPARROW parameter estimates and associated standard errors (Alexander et al., 2002; McMahon et al., 2003; Robertson and Saad, 2011). Nonetheless, despite their sound conceptual foundation and structural simplicity, there is still considerable uncertainty, knowledge gaps, and inherent limitations with SPARROW applications (Qian et al., 2005). In this regard, the implementation of Bayesian inference techniques can be advantageous in characterizing the spatial structure of SPARROW model residuals due to autocorrelated forcing factors, e.g., climate and soils, and year-to-year variability (Qian et al., 2005; Wellen et al., 2012).

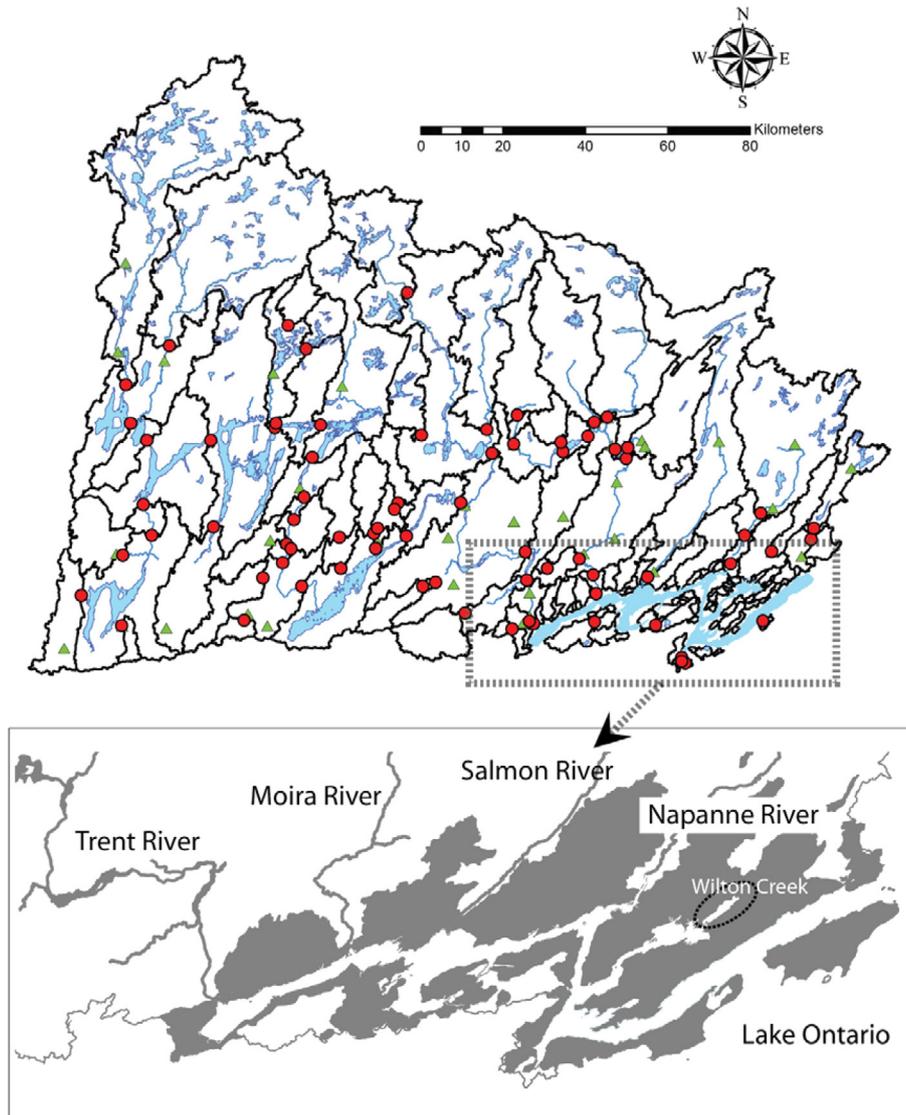
Along the same line of reasoning, Wellen et al. (2014a) focused on addressing three fairly core issues of the SPARROW modelling practice, such as: (i) the use of prior knowledge on parameter values in assisting model calibration; (ii) the error/uncertainty associated with model calibration data; and (iii) the implications of the covariance of model parameters on the inference drawn and the prediction patterns derived. Regarding the former issue, the Bayesian techniques confer a major advantage through their capacity to restrict model parameters to realistic ranges based on information either from available data or existing knowledge (Wellen et al., 2014a). Evidence about the covariance problem was provided by Qian et al. (2005), who showed that three of the SPARROW parameters were highly correlated and concentrated around a narrow “banana-shaped” region of the examined parameter space. The questionable quality of the calibration datasets along with the challenges to support predictions in areas that have modestly been monitored is another topic not heavily addressed in the literature. Both are issues of great practical importance, given that most watersheds of management interest are understudied and existing mean annual load estimates are often obtained by rating curves that are characterized by substantial uncertainty (Cohn et al., 1989; Cohn et al., 1992; Alexander et al., 2002; Alexander et al., 2004; Moatar and Meybeck, 2005). In this context, Wellen et al. (2014a) developed a series of statistical (measurement error) formulations that explicitly considered the analytic uncertainty, sampling error, inter-annual variability as well as the uncertainties stemming from the use of non-contemporaneous measurements of flow and concentration in several stream reaches.

Building upon these methodological advancements, the broader objective of the present study is to advance our understanding of how agricultural and urban sites cycle nutrients and contaminants in the Bay of Quinte watershed, Ontario Canada. The Bay of Quinte is an embayment at the northeastern end of Lake Ontario with a long history of eutrophication, characterized by frequent and spatially extensive algal blooms, and predominance of toxic cyanobacteria (Nicholls et al., 2002; Shimoda et al., 2016). Our intent is to evaluate the accuracy of the current nutrient loading estimates, quantify their uncertainty, and identify hot spots of nutrient export in the watershed. Through the SPARROW parameter estimation, our modelling work aims to provide critical planning information, so management decisions related to the Bay of Quinte can be better guided. Our modelling exercise offers estimates of export coefficients and delivery rates from different subcatchments and thus can be used to formulate (and subsequently validate) hypotheses regarding the “hot spots” of the Bay of Quinte watershed. Thus, the novelty of our study lies not only in the implementation of Bayesian inference techniques to derive rigorous nutrient loading estimates, but also in the use of model uncertainty patterns to facilitate watershed management. The key findings of the present modelling study will be ultimately linked with the process-based models developed for the receiving waterbody to elucidate the causal connections among phosphorus loading, sediment-water column interactions, and plankton community response. Empirical evidence and model predictions suggest that the interplay between inflowing nutrient loads and circulation patterns shapes the local biogeochemical processes, thereby modulating the severity of eutrophication phenomena in the system (Arhonditsis et al., 2016; Shimoda et al., 2016).

## 2. Method

### 2.1. Site description

The Bay of Quinte is a long and narrow Z-shaped embayment located at the northeastern shore of Lake Ontario (Fig. 1). The catchment area is approximately 18,604 km<sup>2</sup>. The Bay of Quinte watershed is geographically characterized by the Canadian Shield (or Precambrian Shield) in the upper basin and non-Canadian Shield (or Paleozoic limestone) in the lower basin (Minns et al., 1986). The major tributaries along the north shore of the bay are the Trent River draining 12,600 km<sup>2</sup>, the Moira River draining 2700 km<sup>2</sup>, and the Salmon and Napanee Rivers together draining 1660 km<sup>2</sup>. Trent River has a well-developed waterway, which is about 390 km long and connects Lake Ontario to Georgian Bay (Fig. 1). Given the waterway's hydrological operation, the flow regimes prevailing in Trent River are significantly different from those in the rest tributaries. Using Self-Organizing Map analysis, Kim et al. (2016) delineated six spatial clusters in the Bay of Quinte watershed, representing one forested, two agricultural, one urban, one pasture-dominated, and one transitional area. The upper basin is predominantly occupied by pristine forested areas and demonstrates lower TP riverine levels, whereas the land uses in the lower basin are associated with anthropogenic activities and consequently characterized by higher TP concentrations (Kim et al., 2016). Urbanized catchments near the Bay of Quinte include several large cities and towns, such as Peterborough (population ≈ 120,000), Trenton (≈ 20,000) and Belleville (≈ 92,000). There are also 26 sewage-treatment plants (STP) within the Bay of Quinte watershed, mainly located close to the urban development. The same locations show more dynamic flow patterns with significant within- and among-year variability and stronger relationship with precipitation. This “flashy” behaviour of urban sites is accompanied by a relatively high net TP export and yield (Kim et al., 2016). Tributaries draining agricultural catchments (e.g., wheat, oat, corn, soybean, and alfalfa) exhibit considerable variability, depending on management practices and soil properties. In particular, Kim et al. (2016) showed that soil hydraulic conductivity could significantly modulate phosphorus land-to-stream delivery, and thus determine phosphorus export from agricultural



**Fig. 1.** Map of the Bay of Quinte watershed; the red circles indicate 73 PWQMN stations (TP concentrations) and the green triangles represent 48 WSC stations (flow rates). The delineated areas with solid lines depict 210 subwatershed catchments. The grey-colored areas (at the bottom panel) represent the ungauged areas. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

areas. Pasture-based grazing systems similarly appear to play a major role in determining the fate and transport of phosphorus in the area (Kim et al., 2016).

## 2.2. Data compilation and GIS analysis

To calculate the TP loads, we compiled TP concentrations measured at 73 stations of Provincial Water Quality Monitoring Network (PWQMN) from 2002 to 2010. We also collected flow rate data recorded at 48 stations of the Water Survey of Canada over the same time period (Fig. 1). Because of the mismatch between flow and water quality monitoring locations, we developed a flow-catchment area regression model ( $r^2 > 0.99$ , Fig. S1) to project annual flow rates at the 73 PWQMN stations. There are many methods to calculating annual constituent loads when using non-continuous concentration records, but less so for non-contemporaneous records of concentration and flow. Moatar and Meybeck (2005) compared the accuracy and precision of a number of different approaches to calculate annual phosphorus loads and recommended the use of the product of means of sampled concentrations and annual discharge, similar to the approach adopted

herein. Thus, mean annual TP load were expressed as follows:

$$\ln(\text{TP Load}_i) = \ln(\text{Flow}_i) + \ln(\text{TP}_i)$$

where  $\ln(\text{TP Load}_i)$  refers to mean annual TP load in the logarithmic scale at station  $i$ ;  $\ln(\text{TP}_i)$  and  $\ln(\text{Flow}_i)$  represent mean annual TP concentrations and flow rates in logarithmic scale. Thus, mean annual TP loads on the 73 PWQMN stations could be estimated along with the corresponding standard deviation values (i.e., depicting the year-to-year loading variability). This method postulates that flow and concentration are independent and thus may result in an underestimation of annual loads (Preston et al., 1989). In this particular study, there is a moderate bias introduced as the correlation coefficients between contemporaneous measurements of flow and concentration ranged from 0.03 to 0.85, with a median of 0.41.

Point source loading estimates were based on 26 sewage treatment plants (STPs) and four sites of the Ontario Ministry of the Environment's Municipal/Industrial Strategy for Abatement (MISA) during the 1998 to 2010 period. The quantification of mean annual, point source TP loads from unserved area was based on the assumption that TP export predominantly involves septic-tank overflows, which are readily

influenced by heavy rainfall. For this reason, we estimated the number of the septic tanks based on population density, assuming that 2.5 people per household share one communal septic tank (Statistics Canada, <http://www.statcan.gc.ca>). Regarding the estimation of non-point source TP loads, we used the maps for the areal extent of specific land uses and morphological characteristics (Land Information Ontario of the Ontario Ministry of Natural Resources, <http://www.ontario.ca/environment-and-energy/land-information-ontario>). In particular for the agricultural land use, we specified the areal extent of combinations of cropland types (e.g., wheat, oat, corn, alfalfa, and fallow) and management practices (e.g., no tillage, conventional or intensive tillage, and conservation tillage) at the township dissemination level (Ontario Ministry of Agriculture and Food, <http://www.omafra.gov.on.ca>). A streamline of the Bay of Quinte watershed was extracted based on a 7.2-m digital-elevation model (DEM) of Ontario created by Natural Resource Canada (NRC). Other geographical calculations (e.g., stream lengths, drainage/subwatershed areas, lake surface areas, areal extent of the above crops and management practices) were conducted using Geographical Information Systems (GIS). With reference to calculating the areal extent of the crops and applied management practices, we had to estimate this extent using fractional coverage (due to limited information) of the crops and management practices within the subwatersheds. Each subwatershed was determined based on 73 PWQMN sites and 137 small residual drainage areas. Consequently, the former group is regarded as gauged watersheds, while the latter one as ungauged watersheds.

### 2.3. SPARROW model configuration

SPARROW is a parsimonious hybrid empirical/process-based model to estimate nutrient loads, yields, and deliveries at landscape and regional scales. The statistical basis for calibrating SPARROW models provides an objective means of empirically estimating the relation between in-stream measurements of nutrient fluxes and the sources/losses of nutrients within the watershed (Smith et al., 1997). In-stream nutrient fluxes are modelled as a non-linear function of nutrient sources (including point sources, atmospheric deposition, and agricultural and developed land use), land-delivery processes, and in-stream nutrient processing (McMahon et al., 2003). To implement the SPARROW model, the Bay of Quinte watershed was first divided into subwatersheds, each of which drains either to a water-quality-monitoring station or to a river/stream mouth. Second, each subwatershed was disaggregated into reach-specific catchments that drain to a particular stream segment (McMahon et al., 2003). We therefore created a total of 210 subwatersheds that consisted of 73 gauged and 137 ungauged ones, and used 3738 reach-specific catchments in total (Fig. 1). The governing SPARROW equation can be expressed as follows:

$$\ln(\text{Load}_i) = \left[ \ln \left\{ \sum_{n=1}^N \sum_{j=1}^{J_i} \beta_n S_{n,j} e^{(-\alpha Z_j)} H_{i,j}^S H_{i,j}^R \right\} \right] \varepsilon_i$$

where the subscripts  $i$  and  $j$  refer to subwatersheds and reach catchments, respectively;  $\ln(\text{Load}_i)$  refers to the mean annual TP load in the logarithmic scale, measured at station  $i$  in metric tons per year<sup>1</sup>;  $N$  is the total number of sources (diffuse and point sources) and  $n$  is an index for each source;  $J_i$  refers to the number of reaches in subwatershed  $i$ ;  $\beta_n$  refers to the estimated TP export coefficient for source  $n$ ;  $S_{n,j}$  refers to the quantity of source  $n$  in reach  $j$ , and therefore the product  $\beta_n S_{n,j}$  has units of metric tons per year. The parameter  $\alpha$  refers to the vector of land to water-delivery coefficients, and  $Z_j$  is a vector of the land-surface characteristics associated with drainage to reach  $j$ .  $H_{i,j}^S$  represents the fraction of TP mass originating in reach  $j$  remaining

at station  $i$  as a function of first-order loss processes in streams;  $H_{i,j}^R$  refers to the fraction of TP mass originating in reach  $j$  remaining at station  $i$  as a function of first order loss processes in lakes and reservoirs; and  $\varepsilon_i$  refers to a random multiplicative error term assumed to be independently and identically distributed across all subwatersheds.

First-order loss processes in streams include loss to sediments and biota, and are expressed as:

$$H_{i,j}^S = \exp(-K_s L_{i,j})$$

where  $K_s$  refers to the first-order loss coefficient for streams ( $\text{km}^{-1}$ ), and  $L_{i,j}$  refers to the stream length in kilometers between reach  $j$  and station  $i$ . First-order loss processes operating in lakes and reservoirs are limited to loss to sediment, which is expressed as

$$H_{i,j}^R = \prod_l \exp(-K_r q_l^{-1})$$

where  $l$  refers to any lakes or reservoirs between reach  $j$  and station  $i$ ;  $K_r$  refers to the first-order loss coefficient or settling velocity ( $\text{m year}^{-1}$ );  $q_l$  refers to the areal hydraulic loading of the lake/reservoir ( $\text{m year}^{-1}$ ).

We first examined which watershed characteristic (e.g., wetland coverage, slope of landscape, soil hydraulic conductivity, and soil bulk density) was the most closely associated with land-to-stream delivery of TP export. This exploratory analysis showed that the vertical hydraulic conductivity was the optimal surrogate variable to characterize  $Z_j$  in the Bay of Quinte watershed (Kim et al., 2016). Following Kim et al.'s (2016) SOM clustering, our watershed characterization was based on the specification of TP export ( $\beta_n$ ) by forest, urban land, pasture and cropland areas. As previously mentioned, we specified the cropland TP export coefficients according to crop types, tillage applications (or management practices), and hydrological soil groups. In doing so, we were able to obtain two SPARROW model configurations in which we consider one single crop type ( $\beta_1$ ) across all agricultural lands (Model I), and five crop types for wheat ( $\gamma_1$ ), oat ( $\gamma_2$ ), corn ( $\gamma_3$ ), alfalfa ( $\gamma_4$ ), and fallow land areas ( $\gamma_5$ ) (Model II). An additional feature of the second model was the explicit consideration of in-stream attenuation rates in low- and high-flow streams, using the flow level of  $1 \text{ m}^3 \text{ s}^{-1}$  as a cutoff point (Alexander et al., 2002). In Model II, we also conducted *post-hoc* simulations that aimed to examine the interplay among crop types, agricultural management practices, and soil characteristics. Specifically, we expressed probabilistically forty five cropland TP export coefficients associated with five crop types  $\times$  three management practices  $\times$  three hydrological soil groups (refer to as Model III hereafter). The three management practices considered were no tillage, surface tillage, and soil tillage (Fig. S2). The hydrologic soil groups comprised three categories of moderately well drained (B), imperfectly and poorly drained (C), and very poorly drained (D) soil texture, as designated by Agriculture and Agri-Food Canada (<http://www.agr.gc.ca>).

### 2.4. Bayesian estimation of SPARROW parameters

We used Bayesian inference to parameterize the SPARROW model, while accommodating the uncertainty associated with model parameters and structure as well as the sampling/analytical error of calibration data (i.e., TP loading estimates). Regarding the latter type of error, our data quality submodel stipulates that the log-transformed loadings are random variables drawn from normal distributions with mean values equal to the previously described estimates and variances representing the associated error and/or temporal variability at each site (Wellen et al., 2014a). Because this approach confounds temporal (year-to-year) variability at a site with the uncertainty of the estimation of the mean TP loads, we opted for the most conservative (largest) specification of the uncertainty as provided by the following equation:

$$\sigma_i^2 = \text{Var}(\ln(TP_i)) + \text{Var}(\ln(\text{Flow}_i)) + 2\text{Cov}(\ln(TP_i), \ln(\text{Flow}_i))$$

<sup>1</sup> The same notation  $i$  is used for both stations and subwatersheds, because each subwatershed has one station in its outlet.

where  $\sigma_i^2$  refers to the variance of the log-transformed TP loading measured at station  $i$ , and Cov refers to covariance between TP concentrations and flow rates. The statistical formulation used to parameterize the two SPARROW configurations is expressed as follows (Wellen et al., 2014a):

$$\begin{aligned}
 Y_i &\sim N(\mu_i, \sigma_i^2) \\
 \mu_i &\sim N(\ln(\text{Load}_i), \delta^2) \\
 \ln(\text{Load}_i) &= \ln \left\{ \sum_{n=1}^N \sum_{j=1}^{J_i} \beta_n S_{n,j} e^{(-\alpha z_j)} H_{i,j}^S H_{i,j}^R \right\} \\
 \delta^2 &\sim IG(0.001, 0.001)
 \end{aligned}$$

where  $Y_i$  denotes the log-transformed observed/measured TP loads at station  $i$ ,  $\sigma_i^2$  is the pre-described measurement errors,  $\mu_i$  represents a latent variable which can be referred to as the “true” loading values;  $\delta^2$  the is the model structure uncertainty determined by a non-informative inverse gamma (IG) distribution with shape and scale parameters equal to 0.001; and  $\ln(\text{Load}_i)$  corresponds to the SPARROW output.

All the prior distributions for the SPARROW model parameters are provided in Table 1, which were then updated with annual loading estimates from the gauged watershed locations. We used non-informative (flat) distribution to land-to-stream TP delivery coefficients ( $\alpha$ ), given the absence of literature information on the plausible  $\alpha$  ranges in relation to soil hydraulic conductivity. The prior distributions for the settling rate in lakes/reservoirs ( $K_r$ ) and stream attenuation rate ( $K_s$ ) were based on our previous research (Cheng et al., 2010; Wellen et al., 2014a). Assuming that the lower and upper limits of TP export rates from each of the land uses, as obtained from Beaulac and Reckhow (1982) and Harmel et al. (2008), occupied 95% of the total range in a logarithmic scale, we determined priors of TP export coefficients for cropland ( $\beta_1$ ), forest ( $\beta_2$ ), pasture ( $\beta_3$ ), and urban land ( $\beta_4$ ). The prior of the septic-tank coefficient ( $\beta_5$ ) was specified based on local empirical evidence. Our post-hoc simulations were based on probability distributions assigned to the TP export coefficients for each combination of the five crop types, three management practices, and three hydrologic soil groups as derived from the MANAGE database (Harmel et al., 2008; see also Fig. S3). In case the sample size ( $N$ ) for a particular export coefficient was limited to two records, the coefficients were determined under the assumption that those two values occupied 95% of the total range (e.g., Corn-No tillage-C and Fallow land-Surface tillage-B in Fig. S3). If there were no data for a particular combination, we used the  $\gamma_i$  specification for the corresponding crop type as global probability across all management practices and hydrological soil groups.

Bayesian inference was applied to estimate model parameters, because of its ability to include prior information in the modelling exercise and to explicitly accommodate model structural and parametric uncertainty (Gelman et al., 2004). Bayesian inference treats each parameter  $\theta$

as a random variable and uses the likelihood function to express the relative plausibility of obtaining different values of this parameter when particular data have been observed.

$$P(\theta|data) = \frac{P(\theta) \cdot P(data|\theta)}{\int_{\theta} P(\theta) \cdot P(data|\theta) d\theta}$$

whereby  $P(\theta)$  denotes the prior probability distribution of parameter  $\theta$ ,  $P(data|\theta)$  represents the likelihood of data observation given the different  $\theta$  values, and  $P(\theta|data)$  is the posterior probability that expresses the update belief on the  $\theta$  values, contingent upon the observed data. The denominator of this equation is often called the “marginal distribution of observable data”, which acts as a scaling constant that normalizes the integral of the area under the posterior probability distribution. Regarding Bayesian calibration of SPARROW model, we used Markov chain Monte Carlo (MCMC) simulations to obtain sequences of realizations from the model posterior distribution. Specifically, we used the general normal proposal Metropolis algorithm as implemented in the WinBUGS software (Lunn et al., 2000). This algorithm is based on a symmetric normal proposal distribution, whose standard deviation is adjusted over the first 4000 iterations such as the acceptance rate ranges between 20% and 40%. We ran 30,000 iterations from two chains of each model configuration, using a thin of 10 to control the serial correlation. The first 5000 samples were discarded and the posterior statistics were calculated with the remaining 5000 samples ( $= 2 \text{ chains} \times [30,000 - 5000] / 10$ ).

We evaluated the degree of change between prior and posterior parameters with the delta index; a metric of the distance between two probability distributions given by the following mathematical formula (Endres and Schindelin, 2003; Hong et al., 2005):

$$\delta_{\theta} = \sqrt{\int \left( P(\theta) \log \frac{2P(\theta)}{P(\theta) + P(\theta|data)} + P(\theta|data) \log \frac{2P(\theta|data)}{P(\theta) + P(\theta|data)} \right) d\theta}$$

where  $P(\theta)$  and  $P(\theta|data)$  represent the marginal prior and posterior distributions of a given parameter  $\theta$ , respectively. This metric is equal to zero if there is no difference between the two distributions, and has a known upper bound equal to  $\sqrt{2 \log 2}$ , if there is no overlap between the two distributions. We subsequently standardized the delta index ( $\Delta$ , normalized delta index) to be expressed as a percentage relative to this maximum value. MCMC estimates of the mean and standard deviation parameter values (i.e., TP export coefficients from different land uses, land-to-water delivery coefficients, and attenuation rates) along with the covariance structure were used to update the model (Gelman et al., 2004). Under the assumption of a

**Table 1**  
Prior parameter estimates for the two SPARROW model configurations.

Parameter	Unit	Model type	Mean	S.D.	References
Delivery coefficient ( $\alpha$ )	h cm <sup>-1</sup>	I, II	1	22,000	–
TP export coeff. of cropland ( $\beta_1$ )	tons km <sup>-2</sup> yr <sup>-1</sup>	I	0.310	0.390	Beaulac and Reckhow (1982), Harmel et al. (2008)
TP export coeff. of forest ( $\beta_2$ )	tons km <sup>-2</sup> yr <sup>-1</sup>	I, II	0.006	0.043	Beaulac and Reckhow (1982), Harmel et al. (2008)
TP export coeff. of pasture ( $\beta_3$ )			0.108	0.099	
TP export coeff. of urban area ( $\beta_4$ )			0.148	0.126	
TP export coeff. of septic tank ( $\beta_5$ )	tons tank <sup>-1</sup> yr <sup>-1</sup>		0.002	0.002	Unpublished data
TP export coeff. of wheat area ( $\gamma_1$ )	tons km <sup>-2</sup> yr <sup>-1</sup>	II	0.111	0.056	Harmel et al. (2008)
TP export coeff. of oat area ( $\gamma_2$ )			0.139	0.039	
TP export coeff. of corn area ( $\gamma_3$ )			0.162	0.167	
TP export coeff. of alfalfa area ( $\gamma_4$ )			0.053	0.025	
TP export coeff. of fallow land ( $\gamma_5$ )			0.073	0.041	
1st order settling rate in lakes ( $K_r$ )	m yr <sup>-1</sup>	I, II	12.84	4.76	Cheng et al. (2010)
1st order attenuation rate for streams ( $K_s$ )	km <sup>-1</sup>	I, II	0.04	0.17	Wellen et al. (2014a)

multivariate normal distribution, the conditional distributions of the raw (or log-transformed) parameter values are given by:

$$\hat{\theta}_{ij} = \hat{\theta}_i + \left[ \hat{\theta}_j - \hat{\theta}_j \right] \Sigma_j^{-1} \Sigma_{i,j}$$

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

where  $\hat{\theta}_{ij}$  and  $\Sigma_{ij}$  correspond to the mean value and the dispersion matrix of the parameter  $i$  conditional on the parameter vector  $j$ ; the values of the elements  $\Sigma_i$ ,  $\Sigma_{i,j}$ , and  $\Sigma_j$  correspond to the variance and covariance of the two subset of parameters; and  $\hat{\theta}_i$ ,  $\hat{\theta}_j$ ,  $\theta_j$  correspond to the posterior mean and random values of the parameters  $i$  and  $j$ , respectively. Our updated SPARROW constructs were then used to draw predictions regarding the TP export and downstream delivery rates from the ungauged subwatersheds. Finally, the post-hoc model augmentation that examined the effects of different combinations of crop types, management practices, and soil hydrological groups postulated conditional independence among the forty-five export coefficients given the posterior structure of the rest parameter vector. Relative to their literature-based prior specification, the first- and second-order moments of the land-use export coefficients for this exercise were proportionally changed based on the prior-to-posterior updating patterns of the corresponding crop type-specific ( $\gamma_{1-5}$ ) coefficients.

### 3. Results-discussion

#### 3.1. SPARROW parameterization and watershed process characterization

Our SPARROW modelling exercise aims to gain insights into the land-use export, land-to-stream delivery, lake/reservoir settling, and in-stream attenuation of phosphorus in the Bay of Quinte watershed. In this regard, we first compared the corresponding average posterior values against empirical estimates from the literature and then examined to what extent these parameter posteriors vary among the two model configurations (Table 2). In the SPARROW literature, the capacity of watershed attributes to modulate land-to-stream delivery of phosphorus was more effectively depicted by the soil drainage index (Alexander et al., 2002), soil hydrological groups (McMahon et al., 2003; Wellen et al., 2014a), and wetland coverage (Wellen et al., 2012). In the Bay of Quinte watershed, we found that the soil hydraulic conductivity was the most sensitive variable and the derived land-to-stream delivery coefficients ( $\alpha$ ) ranged from 0.172 to 0.197 h cm<sup>-1</sup>,

indicating that it takes approximately 10–12 min for TP to be vertically transmitted by 1 cm (Fig. 2). In a similar manner, Smith et al. (1997) and Alexander et al. (2004) applied soil permeability (cm h<sup>-1</sup>), using NRCS State Soil Geographic data (STATSGO, U.S. Department of Agriculture, 1994; Schwarz and Alexander, 1995), and reported land-to-stream phosphorus delivery estimates somewhat lower than ours, ranging between 0.026 and 0.073 h cm<sup>-1</sup>. On the other hand, Moore et al.'s (2004) SPARROW application in New England streams found that none of the variables used to test for phosphorus loss on the landscape (such as soil permeability, percent wetland) were significant predictors of phosphorus loads at either 85- or 95% confidence levels, and thus assumed that the land-delivery losses are factored into the source coefficients for forested, agricultural and developed land areas. It is also interesting to note that the mean soil hydraulic conductivity in the Bay of Quinte watershed is distinctly lower ( $\approx 1.95$  cm h<sup>-1</sup>; Kim et al., 2016) relative to Moore et al. (2004), STATSGO-based soil permeability in New England streams ( $\approx 11.3$  cm h<sup>-1</sup>).

Regarding the agricultural land-use ( $\beta_1$ ) coefficient, we found a posterior estimate of 34 kg P km<sup>-2</sup> yr<sup>-1</sup> which was significantly lower than the mean value of the prior distribution assigned (310 kg P km<sup>-2</sup> yr<sup>-1</sup>) from existing literature compilations (see Table 1, Beaulac and Reckhow, 1982). Based on the same generic “agricultural land” characterization, Alexander et al. (2004) reported an average TP export rate of 33 kg P km<sup>-2</sup> yr<sup>-1</sup> for the entire United States, but also demonstrated that the Mississippi regional watersheds were characterized by nearly four-fold rates (123 kg P km<sup>-2</sup> yr<sup>-1</sup>). Moore et al. (2004) reported that 108 ( $\pm 26$ ) kg of phosphorus are entering the river system in New England for each square kilometer of agricultural land upstream per year. In a more local comparison, Wellen et al. (2014a) found a similar three-times higher mean export rate ( $\approx 100$  kg P km<sup>-2</sup> yr<sup>-1</sup>) in the Hamilton Harbour watershed relative to our estimate. Given that these varying estimates collectively reflect the influence of several cropland characteristics, such as crop types, management practices, and soil hydrological groups, the specification of TP export coefficients should be more focused in order to meaningfully assist the management decisions in the Bay of Quinte watershed. In this regard, the posterior parameterization of Model II shows phosphorus export values that are on par with estimates derived from other agricultural watersheds in North America (Table 2 and Fig. 2). Specifically, we found that areas occupied by oat have the highest TP export coefficient ( $\gamma_2$ , 127 kg P km<sup>-2</sup> yr<sup>-1</sup>), followed by wheat ( $\gamma_1$ , 76 kg P km<sup>-2</sup> yr<sup>-1</sup>), fallow land ( $\gamma_5$ , 72 kg P km<sup>-2</sup> yr<sup>-1</sup>), corn ( $\gamma_3$ , 44 kg P km<sup>-2</sup> yr<sup>-1</sup>), and alfalfa ( $\gamma_4$ , 30 kg P km<sup>-2</sup> yr<sup>-1</sup>). Moreover, reviewing the literature-based

**Table 2**  
Posterior estimates and interpretation of the SPARROW model parameters.

Parameter	Unit	Model type	Mean	S.D.	Interpretation
Delivery coefficient ( $\alpha$ )	h cm <sup>-1</sup>	I	0.190	0.065	It takes approximately 11.5 min for TP to be transmitted for 1 cm.
		II	0.172	0.063	It takes approximately 10 min for TP to be transmitted for 1 cm.
TP export coefficient of cropland ( $\beta_1$ )	tons km <sup>-2</sup> yr <sup>-1</sup>	I	0.034	0.013	34 kg of TP per km <sup>2</sup> are released from cropland on an annual basis.
TP export coefficient from wheat areas ( $\gamma_1$ )		II	0.076	0.032	76 kg of TP per km <sup>2</sup> are released from wheat areas on an annual basis.
TP export coefficient from oat areas ( $\gamma_2$ )			0.127	0.065	127 kg of TP per km <sup>2</sup> are released from oat areas on an annual basis.
TP export coefficient from corn areas ( $\gamma_3$ )			0.044	0.026	44 kg of TP per km <sup>2</sup> are released from corn areas on an annual basis.
TP export coefficient from alfalfa areas ( $\gamma_4$ )			0.030	0.011	30 kg of TP per km <sup>2</sup> are released from alfalfa areas on an annual basis.
TP export coefficient from fallow lands ( $\gamma_5$ )			0.072	0.038	72 kg of TP per km <sup>2</sup> are released from fallow lands on an annual basis.
TP export coefficient of forest ( $\beta_2$ )	tons km <sup>-2</sup> yr <sup>-1</sup>	I	0.010	0.003	10 kg of TP per km <sup>2</sup> are released from forests on an annual basis.
		II	0.018	0.006	18 kg of TP per km <sup>2</sup> are released from forests on an annual basis.
TP export coefficient of pasture ( $\beta_3$ )	tons km <sup>-2</sup> yr <sup>-1</sup>	I	0.026	0.012	26 kg of TP per km <sup>2</sup> are released from pasture on an annual basis.
		II	0.040	0.020	40 kg of TP per km <sup>2</sup> are released from pasture on an annual basis.
TP export coefficient of urban area ( $\beta_4$ )	tons km <sup>-2</sup> yr <sup>-1</sup>	I	0.119	0.082	119 kg of TP per km <sup>2</sup> are released from urban area on an annual basis.
		II	0.126	0.087	126 kg of TP per km <sup>2</sup> are released from urban area on an annual basis.
TP export coefficient of septic tank ( $\beta_5$ )	tons tank <sup>-1</sup> yr <sup>-1</sup>	I	0.001	0.0004	1 kg of TP is released from septic tank on an annual basis.
		II	0.001	0.001	1 kg of TP is released from septic tank on an annual basis.
1st order settling rate in lakes ( $K_r$ )	m yr <sup>-1</sup>	I	2.952	1.332	TP settles with an average rate of 2.95 m per year in lakes.
		II	4.980	1.551	TP settles with an average rate of 4.9 m per year in lakes.
1st order attenuation rate for streams ( $K_s$ )	km <sup>-1</sup>	I	0.002	0.004	0.2% of TP is attenuated per kilometer in streams.
1st order attenuation rate for small streams ( $K_{s1}$ )		II	0.037	0.022	3.7% of TP is attenuated per kilometer in small streams.
1st order attenuation rate for large streams ( $K_{s2}$ )			0.011	0.007	1.1% of TP is attenuated per kilometer in large streams.

probability distributions assigned to different combinations of crop types, management practices (no tillage, conservational and conventional methods), and hydrological soil groups (B, C and D) (Table S1), we note that wheat, corn and fallow land displayed very wide TP export ranges, varying between 22 and 248 kg P km<sup>-2</sup> yr<sup>-1</sup>, 7–133 kg P km<sup>-2</sup> yr<sup>-1</sup>, and 28–3499 kg P km<sup>-2</sup> yr<sup>-1</sup>, respectively. By contrast, alfalfa (42–49 kg P km<sup>-2</sup> yr<sup>-1</sup>) and oat areas (170–179 kg P km<sup>-2</sup> yr<sup>-1</sup>) were characterized by low TP export variability, primarily due to the number and nature of the pertinent study sites registered in MANAGE database. Nonetheless, while we cannot rule out the likelihood of bias in some of the probability distributions used for our post-hoc simulations (see following sections), we also note that the value of information of the existing calibration dataset along with the uncertainties in delineating the areal extent of specific crop type-management practice-hydrological soil group do not allow to meaningfully characterize the corresponding forty five TP export coefficients in a statistical sense.

Regarding TP export rates pertaining to other land uses, the coefficients of forest ( $\beta_2$ ), pasture ( $\beta_3$ ), and urban areas ( $\beta_4$ ) varied among the different model configurations. The posterior mean values of  $\beta_2$  varied between 10 and 18 kg P km<sup>-2</sup> yr<sup>-1</sup> between the two models (Fig. 2), and were consistently higher than the prior mean (6 kg P km<sup>-2</sup> yr<sup>-1</sup>, Table 1) derived by the values compiled from Beaulac and Reckhow

(1982). Moore et al.'s (2004) reported coefficients for forested lands indicate that 13.4 ( $\pm 3.8$ ) kg of phosphorus enter the streams in New England for each square kilometer of forested land upstream per year. Our Model II's  $\beta_2$  posterior estimate (18 kg P km<sup>-2</sup> yr<sup>-1</sup>, Table 2) was very similar to the forest TP export (19 kg P km<sup>-2</sup> yr<sup>-1</sup>) in Alexander et al.'s (2004) US-wide SPARROW application. Regarding the pasture-related ( $\beta_3$ ) coefficient, the mean posterior estimates ranged between 26 and 40 kg P km<sup>-2</sup> yr<sup>-1</sup> (Table 2 and Fig. 2), whereas the corresponding prior was approximately two- to three-times higher (108 kg P km<sup>-2</sup> yr<sup>-1</sup>, Table 1). Even though there is limited research on TP export from pasture lands, the existing empirical evidence suggests fairly high export rates in Australian (450 kg P km<sup>-2</sup> yr<sup>-1</sup>, Nash, 2002) and North-American (2340 kg P km<sup>-2</sup> yr<sup>-1</sup>, Butler et al., 2006) watersheds. On the other hand, Alexander et al. (2004) reported a  $\beta_3$  estimate of 120 kg P km<sup>-2</sup> yr<sup>-1</sup>, which was significantly lower than the values reported by Nash (2002) and Butler et al. (2006). With respect to the pasture TP export rates, Cade-Menun et al. (2013) underscored that the associated variability may stem from the grazing intensity and precipitation/snowmelt. The same study reported that there were no significant differences between cropland and pasture for TP in runoff, but there were significant differences between specific nutrient forms, with dissolved reactive phosphorus higher in runoff from cropland

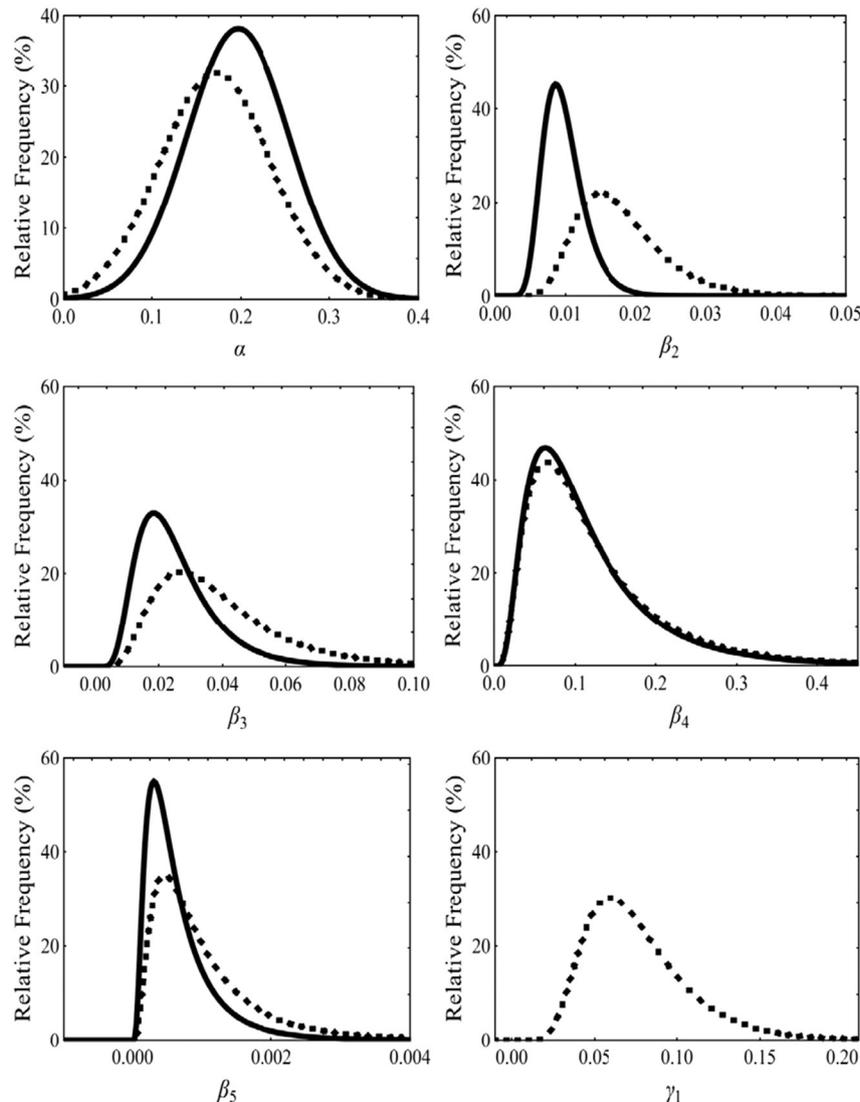


Fig. 2. Posterior distributions of SPARROW parameters: Model I (solid line) and Model II (dashed line). In the stream attenuation rate panel, black and grey dashed lines correspond to parameters assigned to small and large streams, as specified in Model II.

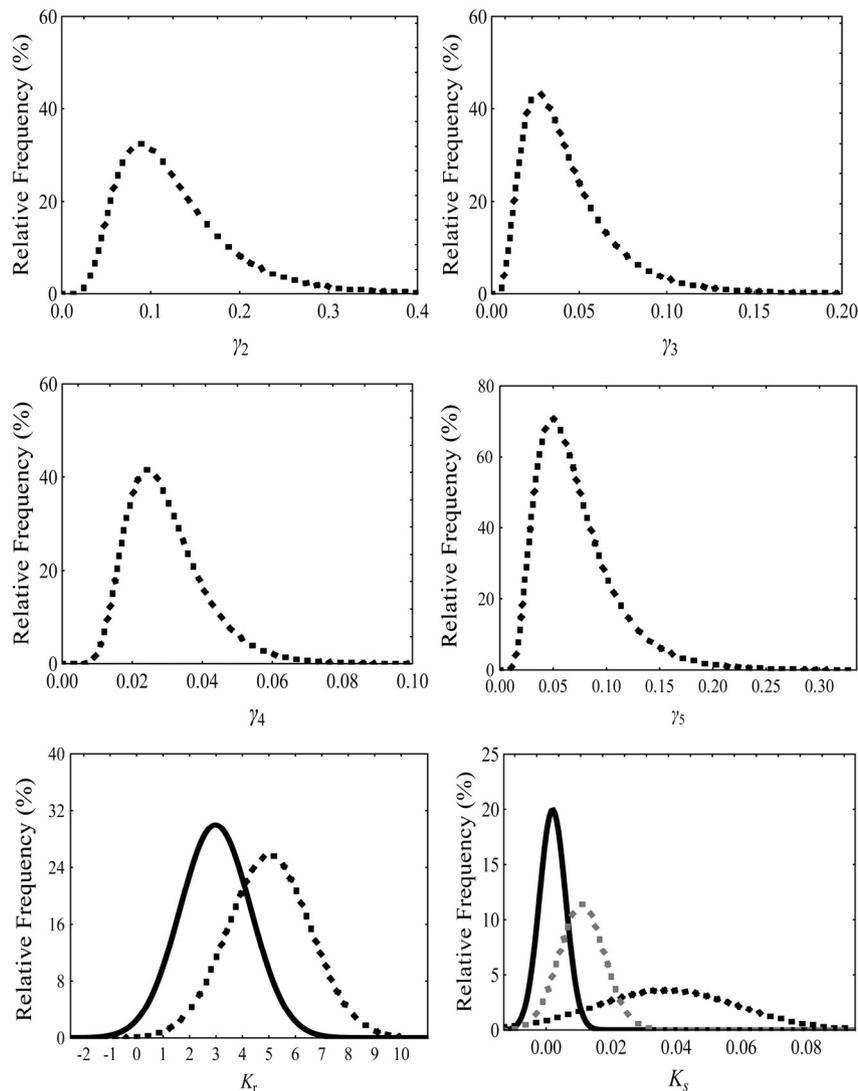


Fig. 2 (continued).

and particulate phosphorus higher in runoff from pasture (Cade-Menun et al., 2013). These results were interpreted as an evidence that there are different sources and transfer mechanisms between pasture and cropland. Pasture areas lack soil inversion and therefore the nutrients deposited by animals in dung and urine accumulate at the soil surface, which in turn makes them more susceptible to loss in runoff (Bourke et al., 2009; Owens and Shpitalo, 2009). In the present study, the relatively low  $\beta_3$  estimates are consistent with the results of a recent SOM analysis which demonstrated three-times lower TP net export (or yield) from pasture relative to cropland areas (Kim et al., 2016). In a similar manner, an older TP export survey in Southern Ontario from the early 1970s estimated a pasture TP export of  $18.1 \text{ kg P km}^{-2} \text{ yr}^{-1}$  from non-igneous/sedimentary lands which was close to our  $\beta_3$  posterior estimates (Dillon and Kirchner, 1975).

Concerning TP export from urban sites,  $\beta_4$  was highest ( $119\text{--}126 \text{ kg P km}^{-2} \text{ yr}^{-1}$ ) among all the land-use specific coefficients (Table 2 and Fig. 2). This result seems to deviate from Beaulac and Reckhow's review (1982) and Moore et al.'s (2004) New England SPARROW, in which the TP export coefficients were typically higher in croplands than in urban areas due to higher nutrient subsidies. However, Alexander et al. (2004) found three times higher TP export coefficients ( $363 \text{ kg P km}^{-2} \text{ yr}^{-1}$ ) in urban than in agricultural lands across the entire United States, and so did Wellen et al. (2014a) in the Hamilton Harbour watershed. Thus, our analysis renders support to the

hypothesis that urban TP fluxes represent the most significant non-point sources, in terms of areal export, in the Bay of Quinte watershed. Lastly, our TP export posteriors from septic tanks ( $\beta_5$ ) were consistently lower ( $\approx 1 \text{ kg P tank}^{-1} \text{ yr}^{-1}$ ) than other phosphorus sources (Table 2 and Fig. 2). Interestingly, Reckhow and Simpson (1980) presented similar septic tank TP export of  $1.0\text{--}2.2 \text{ kg P tank}^{-1} \text{ yr}^{-1}$  ( $0.4\text{--}0.9 \text{ kg P capita}^{-1} \text{ yr}^{-1}$ ). Thus, if we focus on the annual impact of TP export from septic tanks, we can infer that their total contribution may not be significant in the Bay of Quinte watershed. However, given that the septic tank density is simply estimated using demographic data in unserved areas (e.g., cottage areas) which STPs do not cover, our analysis cannot capture the seasonal variability of the associated loading resulting from the intense use of cottages in summer.

Regarding the TP settling rates in lakes/reservoirs within the Bay of Quinte watershed, our posterior estimates differed significantly from previous SPARROW studies. The two configurations provided evidence of lake/reservoir settling rates ( $K_r$ ), ranging between 2.95 and  $4.98 \text{ m yr}^{-1}$  (Table 2 and Fig. 2), which were substantially lower than our literature-based prior value of  $12.84 \pm 4.76 \text{ m yr}^{-1}$  (Table 1). Given that  $K_r$  is estimated as a function of the areal hydraulic loading of lakes/reservoirs, which in turn varies according to their morphometric characteristics (Brett and Benjamin, 2008), the considerable variability of the reported  $K_r$  values in the SPARROW literature is not surprising. For example, Alexander et al. (2004) reported a reservoir settling rate of

14.3 m yr<sup>-1</sup> across the entire United States, while Moore et al. (2004)  $K_r$  estimate was 109 m yr<sup>-1</sup> in the New England watershed. Likewise, the SPARROW application in the Hamilton Harbour watershed (Wellen et al., 2012, 2014a), provided higher posterior  $K_r$  estimates ( $\approx 12$ –13 m yr<sup>-1</sup>) than the present ones. By contrast, Robertson and Saad (2011) work with the Laurentian Great Lakes watershed, resulted in a  $K_r$  approximately equal to 4.83 m yr<sup>-1</sup>, which is very close to our Model II's estimate (4.98 m yr<sup>-1</sup>).

In-stream attenuation rates ( $K_s$ ) across our two SPARROW configurations ranged between 0.2 and 3.7% per kilometer (Table 2 and Fig. 2). Similar to our  $K_r$  findings, our mean  $K_s$  posteriors were significantly lower than the prior value assigned. Characterization of in-stream attenuation according to stream size (or class) is fairly common in the SPARROW literature. Most notably, Alexander et al. (2002) estimated a high attenuation rate of 43% km<sup>-1</sup> within small streams in the Waikato River basin in New Zealand, but calculated a negative attenuation rate of  $-0.06\%$  km<sup>-1</sup> at large streams, indicative of phosphorus export/release in the latter locations. Wellen et al. (2012) similarly estimated  $\approx 13\%$  and 3% TP loss rates per km<sup>-1</sup> in small and large streams of the Hamilton Harbour watershed, respectively. Consistent with this pattern, the corresponding  $K_s$  values with our Model II were 3.7% and 1.1% km<sup>-1</sup>, respectively. Generally, both theoretical and empirical work on stream ecology corroborates this inverse relationship between stream attenuation and stream size (Stream Solute Workshop, 1990; Alexander et al., 2000). Donner et al. (2004) also showed that in-stream removal tends to be lower under conditions of higher rainfall (or runoff) and vice versa. Along the same line of reasoning, Wellen et al. (2012) has demonstrated that stream attenuation coefficients are quite variable in time. Namely, the inter-annual variability of the average discharge, a function of stream stage, can explain more than half of the variability of SPARROW stream attenuation estimates in higher-order streams.

The mechanisms that modulate the nutrient attenuation variability across stream size are still debated in the literature. There are studies generally referring to a tighter coupling of smaller streams with their streambeds, whereby biological and chemical removal processes in the sediments have greater access to the 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 further suggests that stream stage explains the inter-annual variation of nutrient attenuation at a particular site over time, implying that the coupling between the streambed and water column changes from year-to-year (Basu et al., 2011). Others argue that the water column in headwater streams (with low flow, but high velocity due to steeper channel slopes encountered in the headwater areas) will not have tight coupling with the streambed. Even though the depth of water is relatively small, the flow conditions are quite turbulent (steeper slope and resultant high velocity), thereby making streambed erosion more likely rather than sediment settling. However, when the stream reaches the flatter downstream portion (close to the outflow point), the flow gets to be significantly high, but the velocity becomes relatively small due to flatter channel slope and larger flow section, i.e., larger depth of water and width of channel. The relatively quiescent conditions would be expected to enhance the sedimentation process. We believe that the more likely scenario here is that most of the larger settleable particles (pebbles, sand, and grit) are removed from the water column within the upstream and middle stream portions of the river, leaving only the non-settleable colloidal particles to remain in the water column, when it reaches the downstream end of the tributaries. However, the relatively quiescent flow conditions would not be enough to settle out these colloidal-type particles. Regardless of the actual mechanisms, our study overall shows that smaller streams, mainly located in the upper catchments of the Bay of Quinte watershed, are generally characterized by lower volumes of flowing water and higher in-stream attenuation, whereas stream attenuation declines in larger flow downstream locations (Fig. S4).

### 3.2. Degree-of-updating and identification patterns of SPARROW parameters

Parameter identifiability is another important issue in determining the optimal model complexity as well as its planning reliability. Factoring the “principle of parsimony” into our exercise (Spriet, 1985), we may be able to identify the optimal model complexity depending on the quantity and quality of available data (Wellen et al., 2014a). We first focused on the width of the parameter posteriors of the two SPARROW configurations. Land-to-stream delivery coefficients ( $\alpha$ ) were well identified, given the significantly lower standard deviations (SD: 0.063–0.065 h cm<sup>-1</sup>) relative to the corresponding mean values, i.e., coefficient of variation; CV = [SD/mean]  $\times 100 \approx 34$ –37% (Table 2). Likewise, the generic TP export coefficient from agricultural land,  $\beta_1$ , was characterized by a standard deviation of 13 kg P km<sup>-2</sup> yr<sup>-1</sup> and thus a 38% coefficient of variation. With Model II, the crop-specific TP-export coefficients,  $\gamma_{1-5}$ , exhibited fairly low variability: 32 kg P km<sup>-2</sup> yr<sup>-1</sup> (CV = 42%) for wheat; 65 kg P km<sup>-2</sup> yr<sup>-1</sup> (CV = 51%) for oat; 26 kg P km<sup>-2</sup> yr<sup>-1</sup> (CV = 59%) for corn; 11 kg P km<sup>-2</sup> yr<sup>-1</sup> (CV = 37%) for alfalfa; and 38 kg P km<sup>-2</sup> yr<sup>-1</sup> (CV = 53%) for fallow land. Regarding the rest land-use TP export coefficients, forest ( $\beta_2$ ), pasture ( $\beta_3$ ), and urban ( $\beta_4$ ) coefficients also showed narrow uncertainty bounds (CV: 30–33%, 46–50%, and 69–70%, respectively) between the two model configurations. Concerning the lake/reservoir settling rates ( $K_r$ ), Model II (CV = 31%) was characterized by slightly lower uncertainty relative to Model I (CV = 45%). In a similar manner, the two in-stream size-specific attenuation rates demonstrated significantly lower uncertainty with Model II (CV = 59% for  $K_{s1}$  and 64% for  $K_{s2}$ ), relative to the generic in-stream attenuation rate ( $K_s$ ) which was poorly identified (CV = 200%).

Using the normalized delta indices ( $\Delta$ ), we quantified the degree of change in the shape of prior and posterior parameter distributions (Fig. 3). The land-to-stream delivery coefficient ( $\alpha$ ) showed the highest  $\Delta$  ( $\approx 86\%$ ) with both SPARROW models, while the urban TP-export coefficient ( $\beta_4$ ) exhibited one the lowest  $\Delta$  values ( $\approx 37\%$ ). Given that  $\alpha$  is one of the well-identified parameters (Table 2), as well as the only parameter assigned a non-informative prior distribution (Table 1), the highest  $\Delta$  for this coefficient is plausible. By contrast, the lowest  $\Delta$  for  $\beta_4$ , may suggest that the phosphorus fluxes from urban areas obtained from the prior literature closely resemble to those in the Bay of Quinte watershed (Table 2). The generic agricultural coefficient ( $\beta_1$ ) showed a fairly high  $\Delta$  value  $\approx 70\%$  (Fig. 3a), reflecting the fact that not only the prior mean (310 kg P km<sup>-2</sup> yr<sup>-1</sup>, Table 1) was almost ten-times higher than the corresponding posterior mean (34 kg P km<sup>-2</sup> yr<sup>-1</sup>, Table 2), but also that the  $\beta_1$ 's CV was reduced through the prior-to-posterior cycle from  $>100\%$  to 38% (Tables 1–2). Except for corn ( $\gamma_3$ ) ( $\approx 65\%$ ) and alfalfa ( $\gamma_4$ ) ( $\approx 57\%$ ), none of the crop-specific TP export coefficients demonstrated  $\Delta$  values  $>50\%$ . In particular, the posteriors for fallow ( $\gamma_5$ ) and oat ( $\gamma_2$ ) areas were fairly similar to the distributions assigned prior to the model calibration (Table S1; Fig. 3b and Fig. S2). Among the rest land-use export coefficients ( $\beta_{2-5}$ ), the estimates of the TP fluxes from forested areas ( $\beta_2$ ) were characterized by the highest  $\Delta$  values ( $\approx 72\%$ ). Likewise, the in-stream attenuation rates demonstrated very high  $\Delta$  values ( $\approx 76\%$ ) and, notably, the corresponding value in large streams ( $K_{s1} \approx 65\%$ ) was lower than in small streams ( $K_{s2} \approx 76\%$ ). On a final note, lake/reservoir settling rates ( $K_r$ ) did not differ significantly with respect to their  $\Delta$  values ( $\approx 64\%$ ) between the two SPARROW configurations examined. Overall, the parameter posterior patterns suggest that the additional complexity of Model II not only offered additional insights into the watershed process characterization, but was also supported by the information value of the calibration TP loading dataset.

### 3.3. SPARROW goodness-of-fit and predicted TP loads in the Bay of Quinte watershed

Environmental policy making and successful management implementation require robust models for linking land-use practices to in-

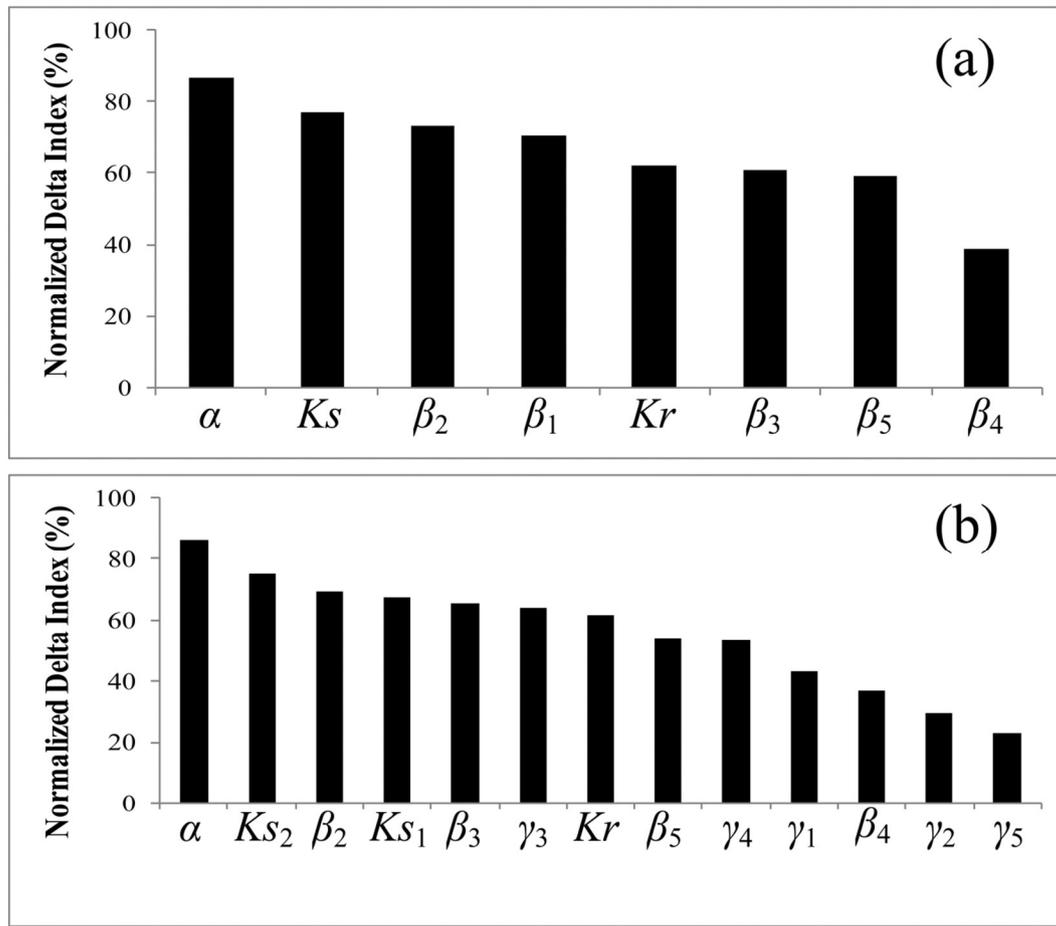


Fig. 3. Distributional changes between the prior and posterior parameters with (a) Model I and (b) Model II, using Normalized Delta Index ( $\Delta$ ).

stream nutrient concentrations and then reproducing the interplay among physical, chemical, and biological processes that controls cultural eutrophication (Arhonditsis et al., 2016). In contemporary modelling practice, model assessment is often accompanied by uncertainty estimation to establish its credibility for management purposes (Rode et al., 2010). Predictive uncertainty is closely associated with data quality/quantity, knowledge gaps, and model structure imperfection (Qian et al., 2005; Lek, 2007; Wellen et al., 2014a). In our SPARROW exercise, the uncertainty bounds of predicted TP loading gradually widened with the consideration of additional sources of error, such as the parametric uncertainty (Fig. 4a); both parametric and structural uncertainty (Fig. 4b), and parametric/structural uncertainty along with the data error

or uncertainty of the loading estimates used to calibrate the model (Fig. 4c).

We visualized the spatial TP load patterns over the entire Bay of Quinte watershed based on SPARROW Model II (Fig. 5). Observed and predicted TP loads were very close to each other (Fig. 5a and c), although the year-to-year variability of the measured loadings (Fig. 5b) was clearly greater than the predictive uncertainty of the corresponding long-term mean annual values (Fig. 5d). Mean annual TP loads were generally larger in the Trent River basin with an upstream-to-downstream gradient from low to high TP loads. The latter locations were also characterized by greater predictive uncertainty. The goodness-of-fit between observed and predicted TP loading values from the

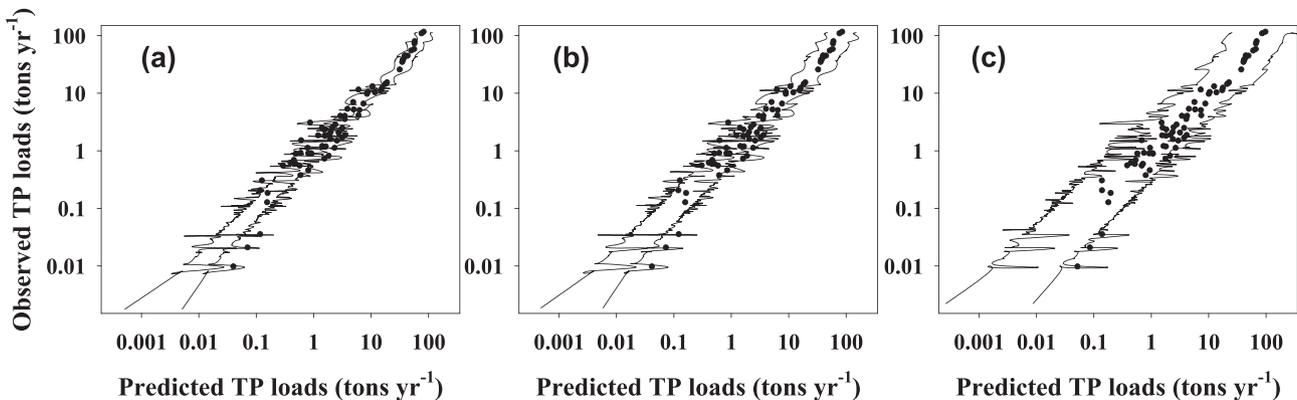
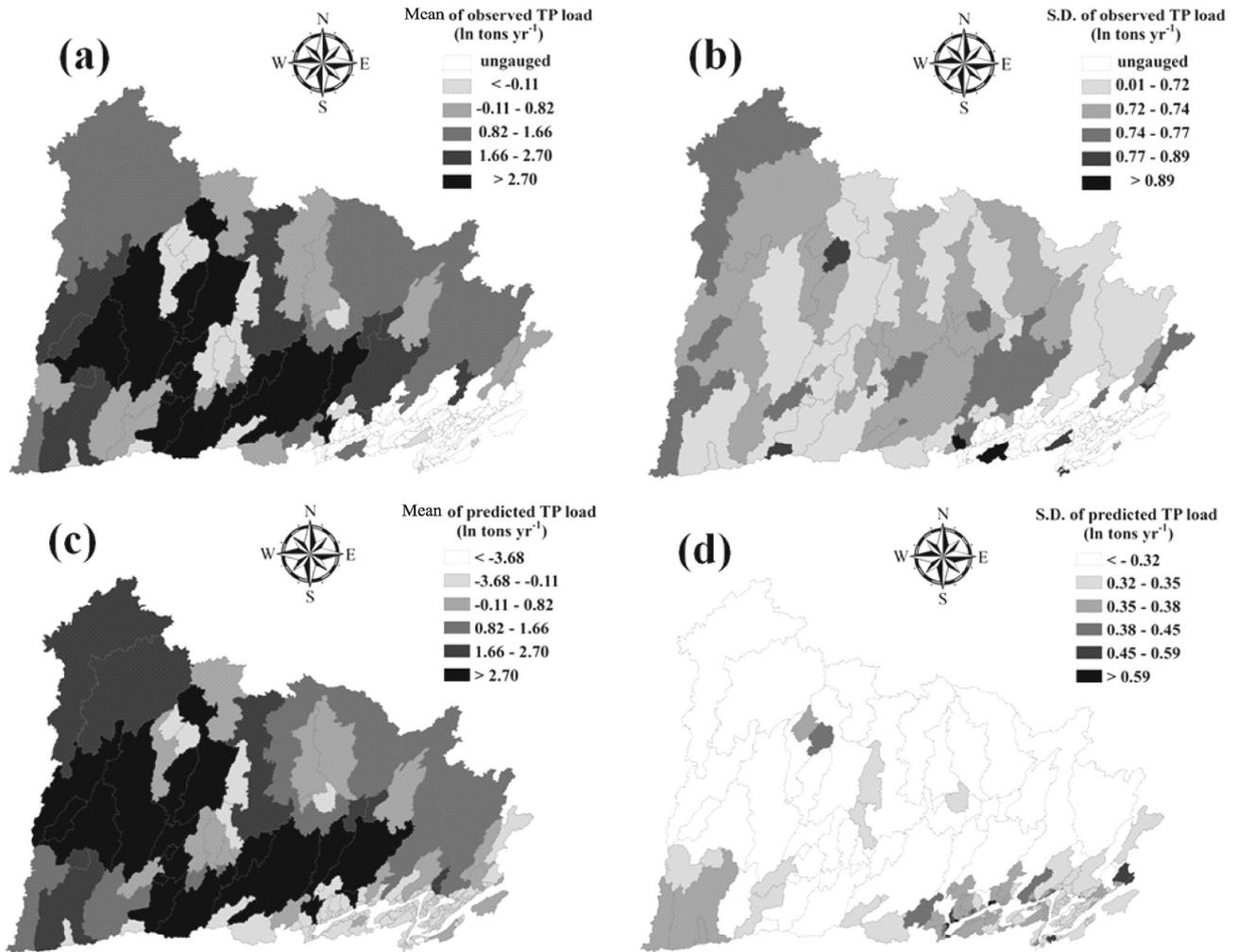


Fig. 4. Observed versus predicted TP loads with uncertainty bounds using Model II. The uncertainty was based on (a) parametric error, (b) parametric and structural error, and (c) parametric, structural, and data-measurement error.



**Fig. 5.** Mean annual TP loading estimates in the Bay of Quinte watershed: (a) mean observed TP loads, (b) standard deviation of observed TP loads, (c) mean predicted TP loads, and (d) standard deviation of predicted TP loads.

SPARROW model was excellent in the logarithmic scale ( $r^2 > 0.95$ ), although there were four sites with error  $> 10$  tonnes  $\text{yr}^{-1}$  when the SPARROW predictions were back-transformed to the original scale. The greater predictive errors were found in the lower basin, primarily near the lower Trent River where TP loads are the largest among the tributaries (Fig. S5). Another interesting finding was the substantial error found in the -predominantly agricultural- Napanee River (Fig. S5), which is characterized by the highest TP concentrations and fairly strong TP-flow relationship (Kim et al., 2013). Importantly, empirical evidence and model predictions suggest that the interplay between inflowing nutrient loads from Napanee River and local circulation patterns shapes the biogeochemical processes in the northeastern segment of the receiving waterbody, thereby modulating the severity of eutrophication phenomena (e.g., frequency of harmful algal blooms) relative to the rest of the Upper Bay of Quinte (Kim et al., 2013). Taking together the distinctly higher predictive loading uncertainty in sub-watersheds near the Bay along with the tight watershed-receiving waterbody coupling, our analysis highlights the importance of obtaining more reliable TP loading estimates from non-point sources in ungauged watershed, urban runoff, and extreme precipitation events (Long et al., 2014, 2015).

We used the spatial distribution of net (instead of the cumulative) TP loading that ultimately inflows into the receiving waterbody to identify the most influential subwatersheds (Fig. 6). The percentage of net loading was mostly greater in the downstream catchment of the major tributaries. Specifically, the lower Trent, Moira and Salmon River basins accounted for  $> 15\%$  of total annual TP loading into the

Bay of Quinte (Fig. 6a). By contrast, the relative contribution of the ungauged watersheds close to the bay was significantly lower primarily due to their small areal extent. On the other hand, the variability of the relative contribution of the different subwatersheds was higher in the Trent River basin ( $SD > 67\%$ ) than the rest of the tributaries. Interestingly, the Trent River's upper catchment also exhibited high variability of the percentage net TP loads (Fig. 6b). We attribute this pattern to the flow rates in this area (e.g., Gull River, which is the headstream of the Trent River) that are highly regulated by dam operation (Kim et al., 2016).

To delineate “hot spots” in a watershed context, McMahon et al. (2003) and Wellen et al. (2014a) introduced four criteria or measures of impairment uncertainty: (i) sites characterized by midrange likelihood of impairment (i.e., the probability of exceeding threshold nutrient values is lying within the 25–75% range); (ii) sites where model predictions have unacceptably high variance; (iii) locations where data uncertainty drives model residuals; and (iv) locations where modelled loads showed the greatest reduction in the width of their 95% credible intervals when higher quality datasets are obtained. In this study, we used the coefficient of variation values of the relative contributions along with the net contribution normalized by the corresponding subwatershed areas to identify the hot-spots in the Bay of Quinte watershed. Similar to our previous assertions, the highest CVs ( $> 32\%$ ) were found in the upper catchment of Trent River (Fig. 6c). Counter to the SD pattern, however, the ungauged watershed close to the bay were characterized by fairly high CVs (Fig. 6c). This trend was

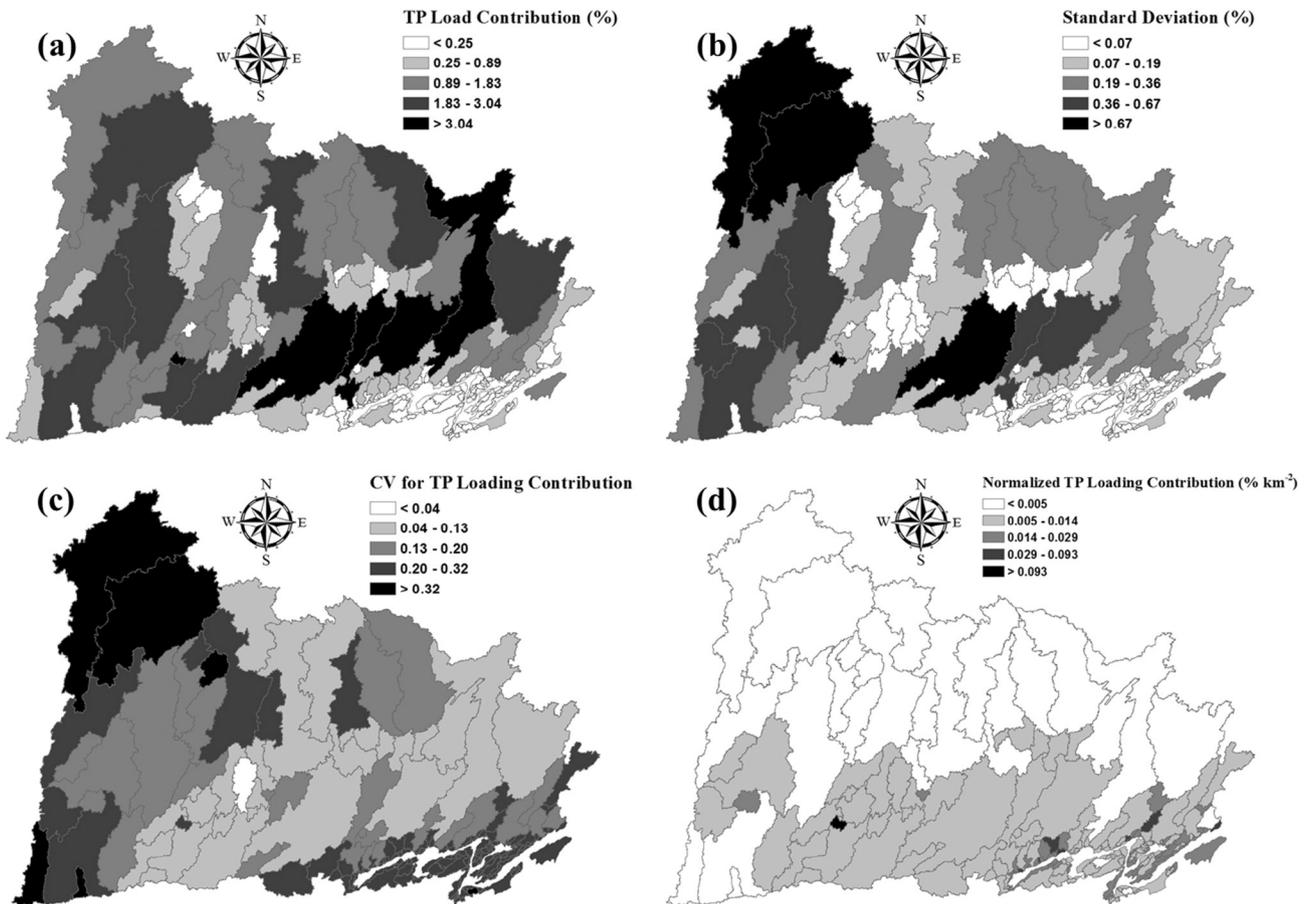


Fig. 6. Percentage contribution of the annual net TP loads to the Bay of Quinte: (a) average, (b) standard deviation, (c) coefficient of variation, and (d) areal average.

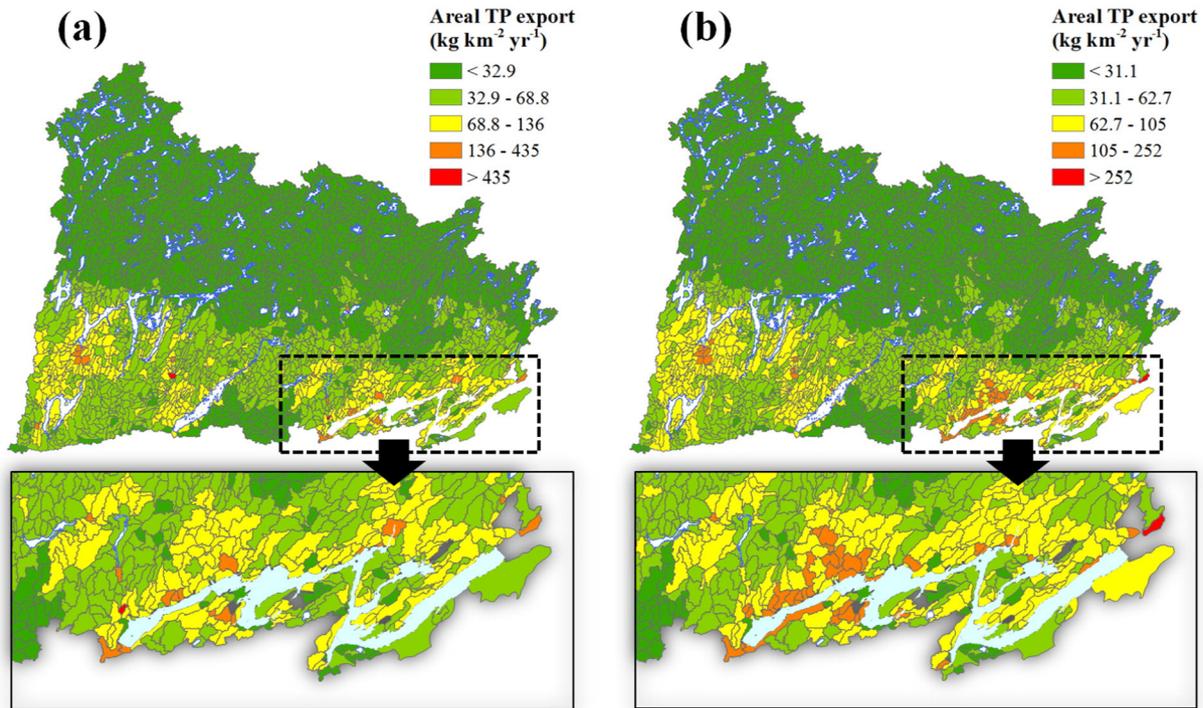
more pronounced when we considered the normalized relative TP load contributions (Fig. 6d). Unlike the CV value, the normalized percentage TP loads were low in the upper catchment of Trent River, but were distinctly higher in the lower part of the watershed, especially near the bay (Fig. 6d). Notably, the present results are very consistent with Kim et al.'s (2016) SOM analysis, who noted that the ungauged subwatersheds close to the waterbody demonstrated the highest areal TP export (i.e., TP yield) over the entire Bay of Quinte watershed.

To improve the granularity of our analysis, we further conducted post-hoc simulations with Model II that accounted for the interplay among crop-types, management practices, and soil hydrological groups (also referred to as Model III). The mean annual TP export is presented when we consider both point and non-point sources (Fig. 7a) and solely non-point sources (Fig. 7b) within the individual reaches. The distinctly higher export rates in reaches adjacent to the Bay of Quinte emerge again as a dominant pattern, with TP fluxes varying between 70 and 400 kg km<sup>-2</sup> yr<sup>-1</sup>. When designing the appropriate beneficial management practices (BMPs) in the area, this quantification of the nutrients lost from different sites with finer resolution is certainly useful, but equally important is to identify the actual mechanisms of nutrient transport. For example, summer storm events can be a significant source of runoff, with nutrient export occurring mainly in particulate form through soil erosion (Sims et al., 1998). Thus, many BMPs are intended to minimize soil disturbance, such as conservation tillage, or to trap eroded soil particles before they can enter the receiving waterbody, such as buffer strips (Kleinman et al., 2009). Nonetheless, the same methods may increase nutrient loss in dissolved form, as the lack of soil inversion can result in nutrient accumulation at the soil surface, including nutrients added through fertilizer or manure implementation (Ginting et al., 1998; Hansen et al., 2000; Tiessen et al., 2010). On the other hand, the slower snowmelt rates are not as erosive (Panuska et

al., 2008) or tend to favour the transport of materials with lower specific gravity, such as organic matter, relative to rainfall runoff (Panuska and Karthikeyan, 2010). If we also consider that frozen soils restrict particle detachment and minimize infiltration, then snowmelt runoff is expected to have lower suspended solid concentrations and higher nutrients in dissolved forms that may also be transported for longer distances (Hansen et al., 2000; Owens et al., 2011). These differences in nutrient transport processes between snowmelt and rainfall runoff can have profound implications for both seasonal and year-to-year loading variability. Shedding light on the relative importance of these mechanisms is critical for controlling nutrient inflows into the Bay of Quinte through BMP implementation.

Finally, we compared the values derived from the SPARROW posterior simulations with five and forty-five export coefficient from the croplands with those based on the Wilton Creek Equivalent<sup>2</sup> (WCE) in order to establish reliable TP loading estimates from the ungauged catchments of the watershed. The comparison is presented with respect to the loading contributions of the ungauged areas in all the seven spatial segments of the eutrophication model adopted in Kim et al. (2013). SPARROW estimates were consistently higher than the WCE ones throughout the water-body (Fig. 8 & Table S2). In total, the mean annual TP loading was more than two-fold with SPARROW ( $\approx 31.9$  tons yr<sup>-1</sup> and  $\approx 43.8$  tons yr<sup>-1</sup> in Models II and III, respectively) relative to WCE ( $\approx 15.3$  tons yr<sup>-1</sup>). Regarding the uncertainty of these contributions, the ungauged subwatersheds displayed SD values  $> 0.38$  ln tons P yr<sup>-1</sup> (1.46 tons P yr<sup>-1</sup>) using SPARROW, while the WCE had SD values of

<sup>2</sup> The Wilton Creek equivalent approach, originally introduced Minns et al. (1986), assumes that the daily load from an ungauged subwatershed equals the daily load from Wilton Creek weighted by the ratio of the area of the ungauged site to the Wilton Creek area.



**Fig. 7.** Mean annual TP export derived from (a) point and non-point sources, and (b) non-point source, derived from the post-hoc simulations with Model II that accounted for the interplay among crop-types, management practices, and soil hydrological groups (also referred to as Model III).

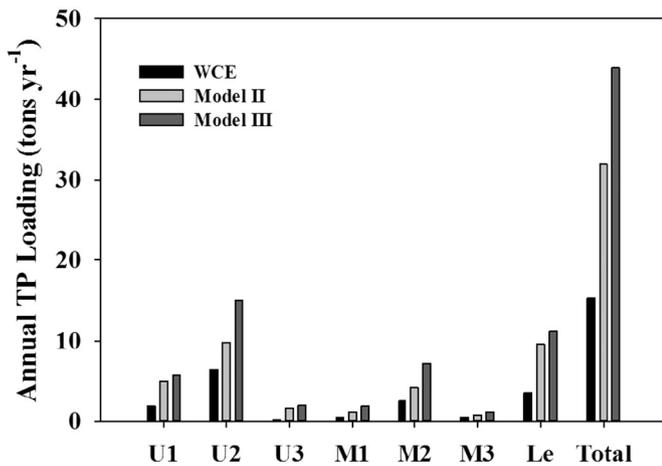
lower than  $-0.32 \ln \text{ tons P yr}^{-1}$  ( $<0.72 \text{ tons P yr}^{-1}$ ). That is, the TP loading uncertainty (or variability) in the ungauged watershed is at least twice as high as the uncertainty deriving from the WCE. Taken together the SPARROW-derived higher mean value (Fig. 8) and greater uncertainty of TP loading pertaining to the ungauged areas (Figs. 5d and 6c), we conclude that the WCE-based external TP loading used to force the existing eutrophication models in the Bay of Quinte clearly introduced an underestimation bias and downplayed the uncertainty of the corresponding predictions (Minns and Moore, 2004; Kim et al., 2013; Zhang et al., 2013); see also Fig. 9 in Kim et al. (2013).

**4. Conclusions/synthesis**

There is a considerable knowledge gap about the complex interplay among hydrological factors, morphological/geological features, and

spatial patterns that modulates the attenuation rates of nutrient and contaminants within a watershed context. In this study, our SPARROW exercise aimed at elucidating phosphorus export and delivery rates in agricultural and urban land areas, so planning decisions that least impact the Bay of Quinte can be better informed. The posterior process characterization indicated that (i) urban areas experience fairly high areal phosphorus export fluxes with an approximate annual estimate of 120 kg of TP per  $\text{km}^2$ ; (ii) the contribution of phosphorus from agricultural land uses can vary considerably among the various crop types (30–127 TP kg per  $\text{km}^2$ ), but is generally lower than the impact of urban sites. This finding appears to contradict the popular notion that nutrient export rates from urban areas are below those of agricultural lands due to lower anthropogenic nutrient subsidies, such as fertilizer implementation (Moore et al., 2004; Soldat and Petrovic, 2008; Soldat et al., 2009). In particular, other studies in Southern Ontario have found that urban total phosphorus export rates to be comparable (or even higher) than agricultural total phosphorus export rates (Winter and Duthie, 2000; Wellen et al., 2014a, 2014b), although the latter locations exhibit significant variability depending on the crop types, soil types, fertilizer applications (e.g., rate and timing) and tillage methods (e.g., conservational or conventional) (Beaulac and Reckhow, 1982; Djodjic et al., 2004; Van Es et al., 2004; Soldat and Petrovic, 2008; Tiessen et al., 2010); (iii) the crop-specific export coefficient values were on par with those typically reported in the literature (Harmel et al., 2008), but the confounding effects of the uncertainty associated with the delineation of the areal coverage of the various crop types may be responsible for some counterintuitive results (e.g., nutrient fluxes from oat farms greater than those from corn fields); (iv) fallow areas are responsible for approximately 70 kg of TP per  $\text{km}^2$  on an annual basis; and (v) 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).

In the context of watershed management, the identification of “hot-spots” (or critical areas of the catchment where practical activities could be of interest) has been based on different stipulations. For example, hot-spots along the flow paths have been defined on the basis of the



**Fig. 8.** Comparison of the mean annual TP loading estimates from the ungauged watershed using the Wilton Creek Equivalent approach and two SPARROW configurations (Models II and III).

rate of change of solute concentrations over time or space (McClain et al., 2003; Vidon et al., 2010). Others describe hot-spots with respect to processes (e.g., organic matter mineralization, denitrification) linked to biogeochemical cycles (Groffman et al., 2009a, 2009b; Andrews et al., 2011). In this study, the delineation of hot-spots revolved around the concept of predictive uncertainty of phosphorus export from different sub-basins around the Bay of Quinte watershed. This strategy pinpointed many locations close to the waterbody that may potentially be responsible for significant nutrient fluxes, due to their landscape attributes and soil characteristics. Of particular note is the fact that several of these catchments have not been monitored yet; thus, our predictions (especially, those with the improved granularity at the reach level) can be used as pointers for maximizing the value of information of additional monitoring by determining locations where data collection efforts should focus on (Arhonditsis et al., 2007).

The present study used a simple empirical modelling construct to obtain a first approximation of annual estimates of nutrient loads, yields, and delivery rates at landscape scale in the Bay of Quinte. Nonetheless, recent evidence from tributaries in Southern Ontario suggests that the study of watershed processes with a coarse spatiotemporal resolution may not be the most meaningful way to assist environmental management. Specifically, it has been shown that a significant fraction of the inflowing phosphorus is generated during a small number of brief but intense precipitation events (Long et al., 2014, 2015). Daily TP loads may vary 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 (Long et al., 2014, 2015). Thus, the characterization of TP concentrations during high flow conditions is essential in establishing accurate concentration versus flow relationships and subsequently nutrient load estimates. Along the same line of evidence, a recent examination of the daily flows in both urbanized and agricultural catchments supports the idea of a single threshold separating two states of response to precipitation (Wellen et al., 2014b). Precipitation depth above a certain threshold ( $\approx 15$  mm over a 2 or 3-day period) triggers an extreme state, which is characterized by a qualitatively different response of the watershed to precipitation (Wellen et al., 2014b). In light of our recent work establishing the significant role of event flows in the nutrient load determination, we believe that there is a need for critical planning information about the optimal restoration and/or mitigation strategies for alleviating the impact of episodic precipitation events. In this regard, several emerging issues that need to be addressed are the impact of disturbances associated with intense summer storm events on system stability; the threat to ecosystem integrity, if the frequency of such meteorological events increases as we gradually shift towards a warmer climatic regime; and the likelihood that increasing urbanization in the watershed will accentuate the impact of these episodic events.

On a final note, Long et al. (2014, 2015) offered insights into the potential impacts of climate change on nutrient export patterns from the Southern Ontario watersheds. TP loading may change from a bimodal delivery pattern (quiescent period followed by intense spring freshet) in cold winters, to an export characterized by intermittent inputs of TP during warm winters, reflective of precipitation patterns and subsequent tributary response observed during any other time of the year. While this empirical evidence is on par with recent modelling projections (Gombault et al., 2015), other hypotheses suggest that elevated winter air temperature would likely lead to increased surface permeability (as opposed to an impermeable frozen surface), resulting in a greater potential for infiltration and/or recharge and not necessarily to an increase in runoff and/or erosion events. There is also speculation that a significant winter recharge is likely to occur (in addition to the traditional spring and fall recharge episodes), along with a decrease in evaporation and/or evapotranspiration due to increased humidity and cloud cover (Valipour, 2016). In this context, one of the priorities for properly characterizing water and nutrient cycles in the area is the advancement of our understanding of the predominant processes in the

surface water-groundwater flow interface. Integration of hydrological with aquatic ecosystem processes will then allow examining Shimoda et al.'s (2011) hypothesis that the nutrient loadings imported during extreme precipitation events can induce broader changes in the system. Namely, the profound changes on the biogeochemistry and trophic functioning of the littoral zone induced by episodic pulses (summer storms) can potentially alter the contemporary algal growth and species competition patterns which in turn can be gradually propagated downstream to the Bay of Quinte. This is one of the critical pieces of information to determine the likelihood of delisting the system as an Area of Concern for the Beneficial Use Impairment “*Eutrophication or Undesirable Algae*” and, most importantly, to credibly predict its future resilience in an ever-changing environment.

## Acknowledgement

This 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. The authors are grateful to Weitao Zhang (City of Toronto), Peter Doris (Ministry of Agriculture, Food and Rural Affairs), Christine McClure (Quinte Conservation), Kristin Geater and Rimi Kalinauskas (Environment Canada) for providing us with data and invaluable feedback throughout this project.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.ecoinf.2016.12.003>.

## References

- Alexander, R.B., Smith, R.A., Schwarz, G.E., 2000. Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. *Nature* 403, 758–761.
- Alexander, R.B., Elliott, A.H., Shankar, U., McBride, G.B., 2002. Estimating the sources and transport of nutrients in the Waikato River Basin, New Zealand. *Water Resour. Res.* 38, 4-1-4-23.
- Alexander, R.B., Smith, R.A., Schwarz, G.E., 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.
- Andrews, D.M., Lin, H., Zhu, Q., Jin, L., Brantley, S.L., 2011. Hot spots and hot moments of dissolved organic carbon export and soil organic carbon storage in the Shale Hills catchment. *Vadose Zone J.* 10.
- Arhonditsis, G.B., Qian, S.S., Stow, C.A., Lamon, E.C., Reckhow, K.H., 2007. Eutrophication risk assessment using Bayesian calibration of process-based models: application to a mesotrophic lake. *Ecol. Model.* 208, 215–229.
- Arhonditsis, G.B., Kim, D.-K., Shimoda, Y., Zhang, W., Watson, S., Mugalingam, S., Dittrich, M., Geater, K., McClure, C., Keene, B., Morley, A., Richards, A., Long, T., Rao, Y.R., Kalinauskas, R., 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 Manag.* 19, 1–18.
- Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment part I: model development. *J. Am. Water Resour. Assoc.* 34, 73–89.
- Basu, N.B., Rao, P.S.C., Thompson, S.E., Loukinova, N.V., Donner, S.D., Ye, S., Sivapalan, M., 2011. Spatiotemporal averaging of in-stream solute removal dynamics. *Water Resour. Res.* 47, W00J06.
- Beaulac, M.N., Reckhow, K.H., 1982. An examination of land use – nutrient export relationships. *J. Am. Water Resour. Assoc.* 18, 1013–1024.
- Bicknell, B.R., Imhoff, J.C., Kittle, J.L., Donigan, A.S., Johanson, R.C., 1996. *Hydrological Simulation Program – Fortran: User's Manual for Release 11*. US Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA.
- Borah, D.K., Xia, R.J., Bera, M., 2002. DWSM – a Dynamics Watershed Simulation Model. In: Singh, V.P., Frevert, D. (Eds.), *Mathematical Models of Small Watershed Hydrology and Applications*, pp. 133–166.
- Bourke, D., Kurz, I., Dowding, P., O'Reilly, C., Tunney, H., Doody, D.G., O'Brien, J.E., Jeffrey, D.W., 2009. Characterisation of organic phosphorus in overland flow from grassland plots using  $^{31}\text{P}$  nuclear magnetic resonance spectroscopy. *Soil Use Manag.* 25, 234–242.
- Brett, M.T., Benjamin, M.M., 2008. A review and reassessment of lake phosphorus retention and the nutrient loading concept. *Freshw. Biol.* 53, 194–211.
- Butler, D.M., Franklin, D.H., Ranells, N.N., Poore, M.H., Green, J.T., 2006. Ground cover impacts on sediment and phosphorus export from manured riparian pasture. *J. Environ. Qual.* 35, 2178–2185.
- Cade-Menun, B.J., Bell, G., Baker-Ismail, S., Fouli, Y., Hodder, K., McMartin, D.W., Perez-Valdivia, C., Wu, K., 2013. Nutrient loss from Saskatchewan cropland and pasture in spring snowmelt runoff. *Can. J. Soil Sci.* 93, 445–458.

- Cheng, V., Arhonditsis, G.B., Brett, M.T., 2010. A reevaluation of lake-phosphorus loading models using a Bayesian hierarchical framework. *Ecol. Res.* 25, 59–76.
- Cohn, T.A., Delong, L.L., Gilroy, E.J., Hirsch, R.M., Wells, D.K., 1989. Estimating constituent loads. *Water Resour. Res.* 25, 937–942.
- Cohn, T.A., Caulder, D.L., Gilroy, E.J., Zynjuk, L.D., Summers, R.M., 1992. The validity of a simple statistical model for estimating fluvial constituent loads: an empirical study involving nutrient loads entering Chesapeake Bay. *Water Resour. Res.* 28, 2353–2363.
- Dillon, P.J., Kirchner, W.B., 1975. The effects of geology and land use on the export of phosphorus from watersheds. *Water Res.* 9, 135–148.
- Djodjic, F., Börling, K., Bergström, L., 2004. Phosphorus leaching in relation to soil type and soil phosphorus content. *J. Environ. Qual.* 33, 678–684.
- Donner, S.D., Kucharik, C.J., Oppenheimer, M., 2004. The influence of climate on in-stream removal of nitrogen. *Geophys. Res. Lett.* 31, L20509.
- Endres, D.M., Schindelin, J.E., 2003. A new metric for probability distributions. *IEEE Trans. Inf. Theory* 49, 1858–1860.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. *Bayesian Data Analysis*. Chapman & Hall/CRC, Boca Raton, Florida.
- Ginting, D., Moncrief, J.F., Gupta, S.C., Evans, S.D., 1998. Interaction between manure and tillage system on phosphorus uptake and runoff losses. *J. Environ. Qual.* 27.
- Gombault, C., Sottile, M.-F., Ngwa, F.F., Madramootoo, C.A., Michaud, A.R., Beaudin, I., Chikhaoui, M., 2015. Modelling climate change impacts on the hydrology of an agricultural watershed in southern Québec. *Can. Water Res. J.* 40, 71–86.
- Groffman, P.M., Butterbach-Bahl, K., Fulweiler, R.W., Gold, A.J., Morse, J.L., Stander, E.K., Tague, C., Tonitto, C., Vidon, P., 2009a. Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models. *Biogeochemistry* 93, 49–77.
- Groffman, P.M., Davidson, E.A., Seitzinger, S., 2009b. New approaches to modeling denitrification. *Biogeochemistry* 93, 1–5.
- Hansen, N.C., Gupta, S.C., Moncrief, J.F., 2000. Snowmelt runoff, sediment, and phosphorus losses under three different tillage systems. *Soil Tillage Res.* 57, 93–100.
- Hantush, M., 2005. TMDL Model Evaluation and Research Needs. U.S. Environmental Protection Agency, Cincinnati, OH, p. 403.
- 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.
- Hong, B., Strawderman, R.L., Swaney, D.P., Weinstein, D.A., 2005. Bayesian estimation of input parameters of a nitrogen cycle model applied to a forested reference watershed, Hubbard Brook Watershed Six. *Water Resour. Res.* 41, W03007.
- Kim, D.-K., Zhang, W., Rao, Y., Watson, S., Mugalingam, S., Labencki, T., Dittrich, M., Morley, A., Arhonditsis, G.B., 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., Kaluskar, S., Mugalingam, S., Arhonditsis, G.B., 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.
- Kleinman, P.J.A., Sharpley, A.N., Saporito, L.S., Buda, A.R., Bryant, R.B., 2009. Application of manure to no-till soils: phosphorus losses by sub-surface and surface pathways. *Nutr. Cycl. Agroecosyst.* 84, 215–227.
- Lek, S., 2007. Uncertainty in ecological models. *Ecol. Model.* 207, 1–2.
- 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., Boyd, D., Mohamed, M., O'Connor, K., 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.
- Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat. Comput.* 10, 325–337.
- McClain, E.M., Boyer, W.E., Dent, L.C., Gergel, E.S., Grimm, B.N., Groffman, M.P., Hart, C.S., Harvey, W.J., Johnston, A.C., Mayorga, E., McDowell, H.W., Pinay, G., 2003. Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystem. *Ecosystems* 6, 301–312.
- McMahon, G., Alexander, R.B., Qian, S., 2003. Support of total maximum daily load programs using spatially referenced regression models. *J. Water Resour. Plan. Manag.* 129, 315–329.
- Minns, C.K., Moore, J.E., 2004. Modelling Phosphorus Management in the Bay of Quinte, Lake Ontario in the Past, 1972 to 2001, and in the Future, Canadian Manuscript Report of Fisheries and Aquatic Sciences No. 2695. Great Lakes Laboratory for Fisheries and Aquatic Sciences, Bayfield Institute, Burlington, Ontario, p. 42.
- Minns, C.K., Hurley, D.A., Nicholls, K.H., 1986. Project Quinte: Point-source Phosphorus Control and Ecosystem Response in the Bay of Quinte, Lake Ontario. Dept. of Fisheries and Oceans, Ottawa.
- Moatar, F., Meybeck, M., 2005. Compared performances of different algorithms for estimating annual nutrient loads discharged by the eutrophic River Loire. *Hydrol. Process.* 19, 429–444.
- Moore, R.B., Johnson, C.M., Robinson, K.W., Deacon, J.R., 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.
- Nash, D., 2002. Phosphorus Transfer From Land to Water in Pasture-based Grazing Systems, Institute of Land and Food Resources. The University of Melbourne, Parkville 3010, p. 221.
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., King, K.W., 2011. *Soil and Water Assessment Tool: Theoretical Documentation, Version 2005*. Texas Water Resources Institute, Texas, USA.
- Nicholls, K.H., 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.
- Owens, L.B., Shipitalo, M.J., 2009. Runoff quality evaluations of continuous and rotational over-wintering systems for beef cows. *Agric. Ecosyst. Environ.* 129, 482–490.
- Owens, L.B., Bonta, J.V., Shipitalo, M.J., Rogers, S., 2011. Effects of winter manure application in Ohio on the quality of surface runoff. *J. Environ. Qual.* 40.
- Panuska, J.C., Karthikeyan, K.G., 2010. Phosphorus and organic matter enrichment in snowmelt and rainfall–runoff from three corn management systems. *Geoderma* 154, 253–260.
- Panuska, J.C., Karthikeyan, K., Norman, J.M., 2008. Sediment and phosphorus losses in snowmelt and rainfall runoff from three corn management systems. *Trans. ASABE* 51, 95–105.
- Preston, S.D., Bierman, V.J., Silliman, S.E., 1989. An evaluation of methods for the estimation of tributary mass loads. *Water Resour. Res.* 25, 1379–1389.
- Qian, S.S., Reckhow, K.H., Zhai, J., McMahon, G., 2005. Nonlinear regression modeling of nutrient loads in streams: a Bayesian approach. *Water Resour. Res.* 41, W07012.
- Reckhow, K.H., Simpson, J.T., 1980. A procedure using modeling and error analysis for the prediction of lake phosphorus concentration from land use information. *Can. J. Fish. Aquat. Sci.* 37, 1439–1448.
- Robertson, D.M., Saad, D.A., 2011. Nutrient inputs to the Laurentian Great Lakes by source and watershed estimated using SPARROW watershed models. *J. Am. Water Resour. Assoc.* 47, 1011–1033.
- Rode, M., Arhonditsis, G., Balin, D., Kebede, T., Krysanova, V., van Griensven, A., van der Zee, S.E.A.T.M., 2010. New challenges in integrated water quality modelling. *Hydrol. Process.* 24, 3447–3461.
- Runkel, R.L., Crawford, C.G., Cohn, T.A., 2004. Load Estimator (LOADEST): A FORTRAN Program for Estimating Constituent Loads in Streams and Rivers. U.S. Geological Survey.
- Schwarz, G.E., Alexander, R.B., 1995. State Soil Geographic (STATSGO) Data Base for the Conterminous United States. U.S. Geological Survey, Reston, VA.
- Schwarz, G.E., Hoos, A.B., Alexander, R.B., Smith, R.A., 2006. The SPARROW Surface Water-quality Model: Theory, Application, and User Documentation. US Dept. of the Interior, US Geological Survey.
- Shimoda, Y., Azim, M.E., Perhar, G., Ramin, M., Kenney, M.A., Sadreddini, S., Gudimov, A., Arhonditsis, G.B., 2011. Our current understanding of lake ecosystem response to climate change: what have we really learned from the north temperate deep lakes? *J. Great Lakes Res.* 37, 173–193.
- Shimoda, Y., Watson, S., Palmer, M.E., Koops, M., Mugalingam, S., Morley, A., Arhonditsis, G.B., 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.
- Sims, J.T., Simard, R.R., Joern, B.C., 1998. Phosphorus loss in agricultural drainage: historical perspective and current research. *J. Environ. Qual.* 27.
- Singh, V.P., Frevert, D.K., 2006. *Watershed Models*. Taylor & Francis, Boca Raton.
- Smith, R.A., Schwarz, G.E., Alexander, R.B., 1997. Regional interpretation of water-quality monitoring data. *Water Resour. Res.* 33, 2781–2798.
- Soldat, D.J., Petrovic, A.M., 2008. The fate and transport of phosphorus in turfgrass ecosystems. *Crop Sci.* 48, 2051–2065.
- Soldat, D., Petrovic, A.M., Ketterings, Q., 2009. Effect of soil phosphorus levels on phosphorus runoff concentrations from turfgrass. *Water Air Soil Pollut.* 199, 33–44.
- Spriet, J.A., 1985. Structure characterization: an overview. In: Barker, H.A., Young, P.C. (Eds.), Identification and System Parameter Estimation: Proceedings of the Seventh IFAC/IFORS Symposium. Pergamon Press, Oxford, pp. 749–756.
- Stream Solute Workshop, 1990. Concepts and methods for assessing solute dynamics in stream ecosystems. *J. N. Am. Benthol. Soc.* 95–119.
- Tiessen, K.H.D., Elliott, J.A., Yarotski, J., Lobb, D.A., Flaten, D.N., Glozier, N.E., 2010. Conventional and conservation tillage: influence on seasonal runoff, sediment, and nutrient losses in the Canadian Prairies. *J. Environ. Qual.* 39, 964–980.
- U.S. Department of Agriculture, 1994. State Soil Geographic (STATSGO) Data Base: Data Use Information. U.S. Department of Agriculture-Natural Resources Conservation Service.
- Valipour, M., 2016. Optimization of neural networks for precipitation analysis in a humid region to detect drought and wet year alarms. *Meteorol. Appl.* 23, 91–100.
- Van Es, H.M., Schindelbeck, R.R., Jokela, W.E., 2004. Effect of manure application timing, crop, and soil type on phosphorus leaching. *J. Environ. Qual.* 33, 1070–1080.
- Vidon, P., Allan, C., Burns, D., Duval, T.P., Gurwick, N., Inamdar, S., Lowrance, R., Okay, J., Scott, D., Sebestyen, S., 2010. Hot spots and hot moments in riparian zones: potential for improved water quality management. *J. Am. Water Resour. Assoc.* 46, 278–298.
- Wellen, C., Arhonditsis, G.B., Labencki, T., Boyd, D., 2012. A Bayesian methodological framework for accommodating interannual variability of nutrient loading with the SPARROW model. *Water Resour. Res.* 48, W10505.
- Wellen, C., Arhonditsis, G.B., 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, G.B., 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., Kamran-Disfani, A.-R., Arhonditsis, G.B., 2015. Evaluation of the current state of distributed watershed nutrient water quality modeling. *Environ. Sci. Technol.* 49, 3278–3290.
- Winter, J.G., Duthie, H.C., 2000. Export coefficient modeling to assess phosphorus loading in an urban watershed. *J. Am. Water Resour. Assoc.* 36, 1053–1061.
- Zhang, W., Kim, D.-K., Rao, Y., Watson, S., Mugalingam, S., Labencki, T., Dittrich, M., Morley, A., Arhonditsis, G.B., 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.

**A BAYESIAN APPROACH FOR ESTIMATING PHOSPHORUS EXPORT AND  
DELIVERY RATES WITH THE SPATIALLY REFERENCED REGRESSION ON  
WATERSHED ATTRIBUTES (SPARROW) MODEL**

**[SUPPORTING INFORMATION]**

**Dong-Kyun Kim<sup>1</sup>, Samarth Kaluskar<sup>1</sup>, Shan Mugalingam<sup>2</sup>, Agnes Richards<sup>3</sup>,  
Tanya Long<sup>4</sup>, Andrew Morley<sup>5</sup>, George B. Arhonditsis<sup>1\*</sup>**

<sup>1</sup> Ecological Modeling Laboratory

Department of Physical & Environmental Sciences, University of Toronto,  
Toronto, Ontario, Canada, M1C1A4

<sup>2</sup>Lower Trent Conservation

Trenton, Ontario, Canada, K8V 5P4

<sup>3</sup>Environment Canada, 867 Lakeshore Road, Burlington, Ontario, Canada, L7R 4A6

<sup>4</sup>Environmental Monitoring and Reporting Branch, Ontario Ministry of the Environment  
Toronto, Ontario, Canada, M9P 3V6

<sup>5</sup>Kingston Regional Office, Ontario Ministry of the Environment

Kingston, Ontario, Canada, K7M 8S5

\* Corresponding author

e-mail: [georgea@utsc.utoronto.ca](mailto:georgea@utsc.utoronto.ca), Tel.: +1 416 208 4858; Fax: +1 416 287 7279

## FIGURES LEGENDS

**Figure S1:** Relationship between catchment areas and mean annual flow rates in the Bay of Quinte watershed ( $N=48$ ).

**Figure S2:** Spatial distribution of the three management practices considered (a) no tillage, (b) surface tillage, and (c) soil tillage in the Bay of Quinte watershed.

**Figure S3:** Prior distributions of the TP export coefficients ( $\text{tons km}^{-2} \text{ yr}^{-1}$ ) for the combinations of crop types, management practices, and hydrologic soil groups.

**Figure S4:** Spatial distribution of flow (or stream) classes in the Bay of Quinte watershed.

**Figure S5:** Predictive errors of *SPARROW* (Model II) in the Bay of Quinte watershed. Green triangles indicate the locations of 26 sewage treatment plants.

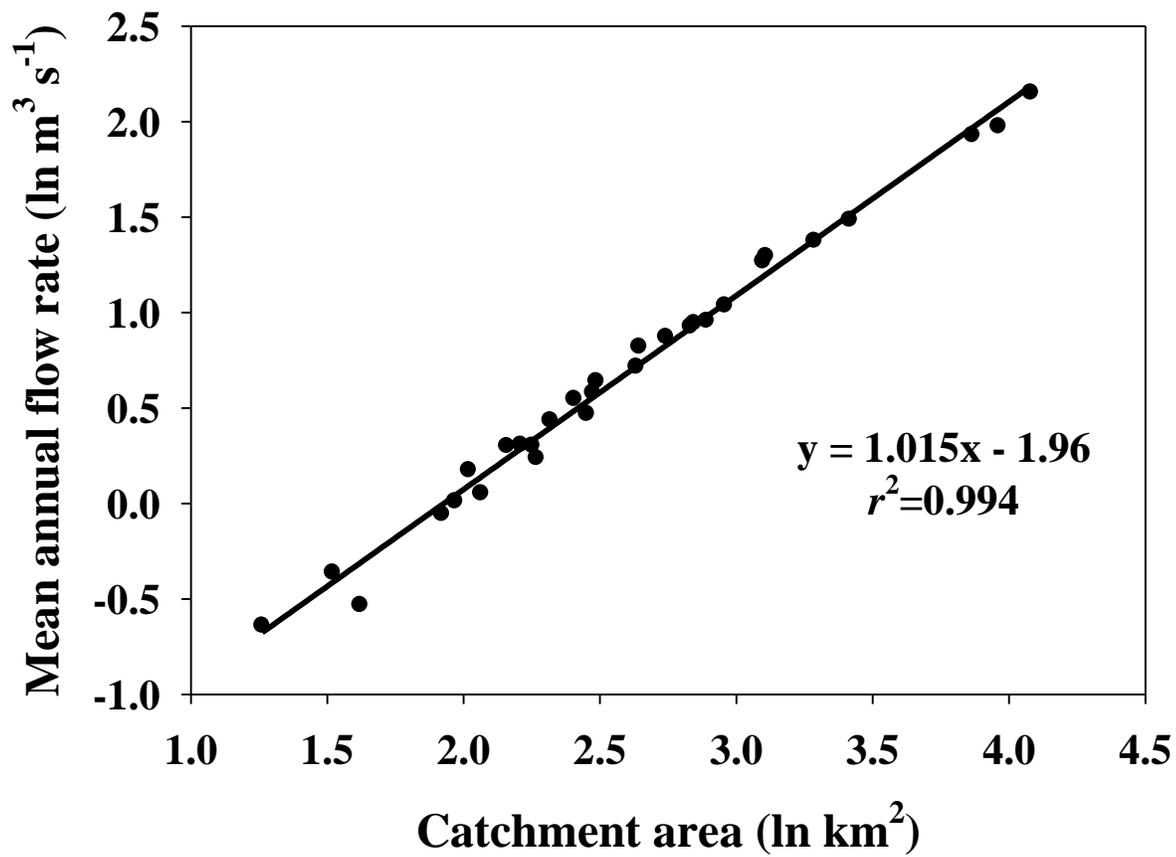
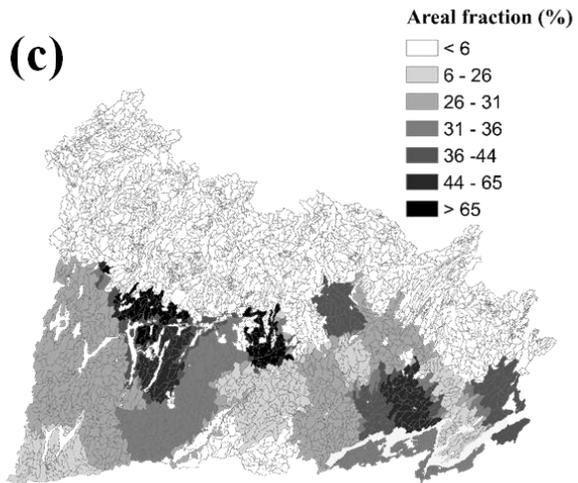
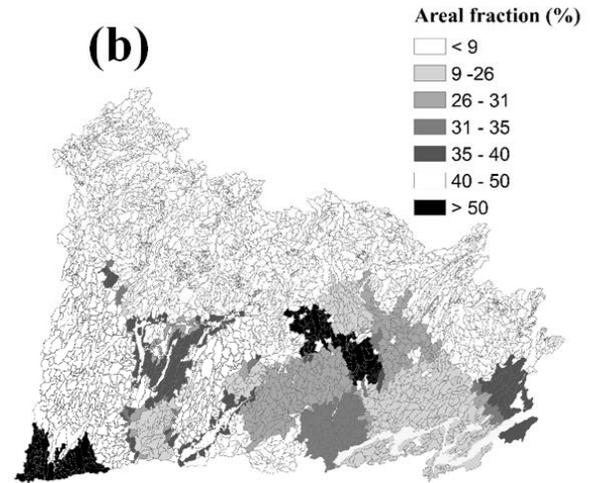
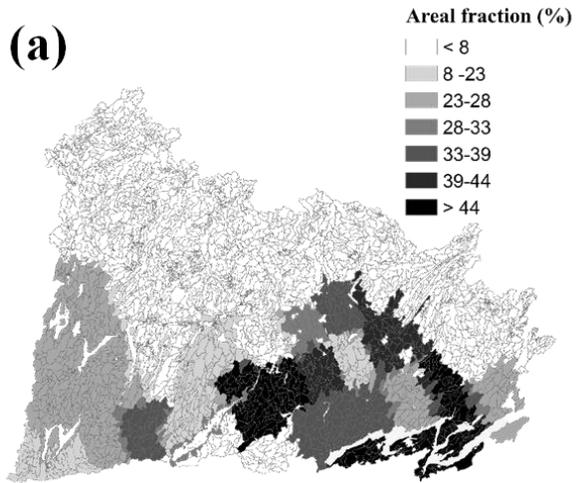
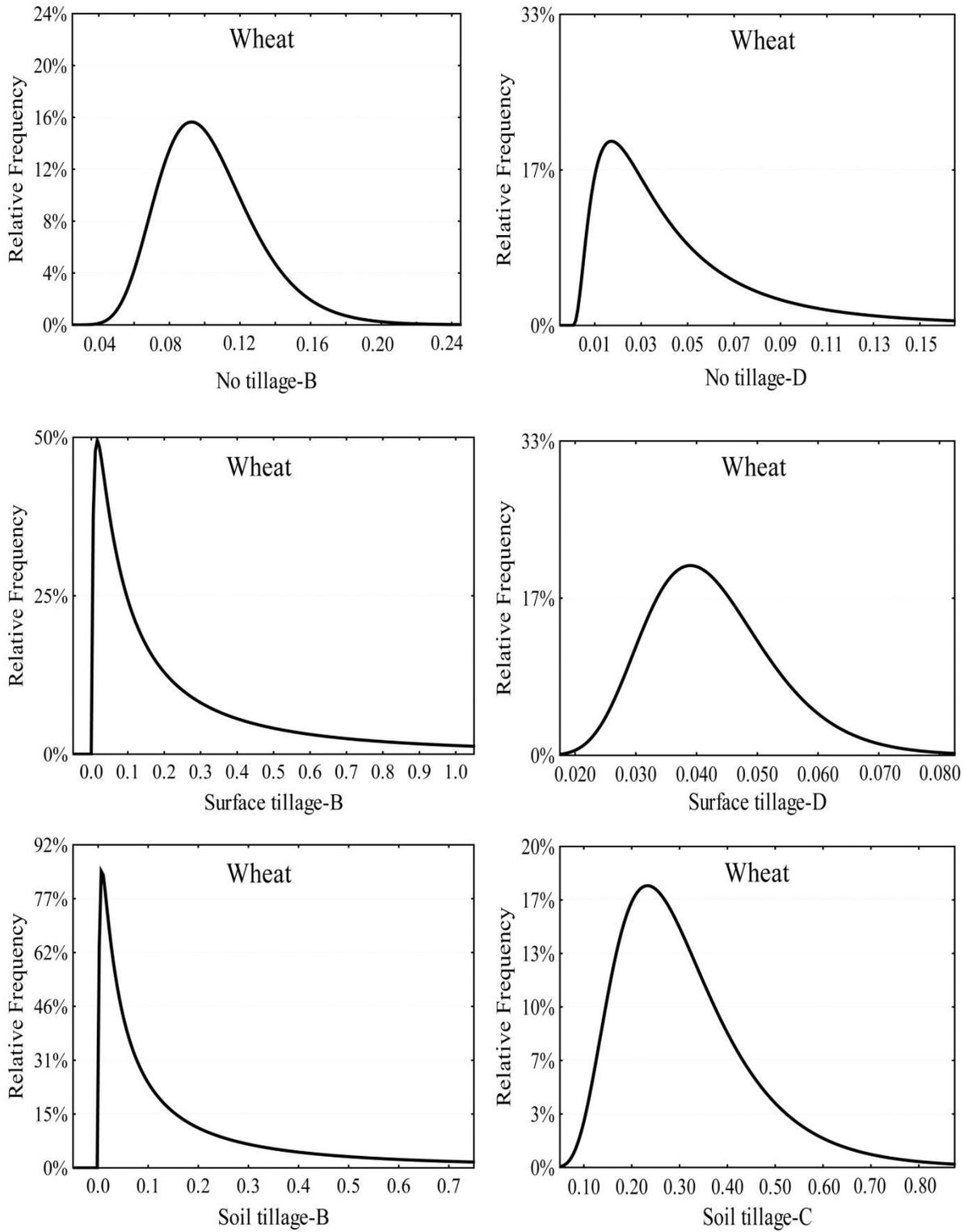


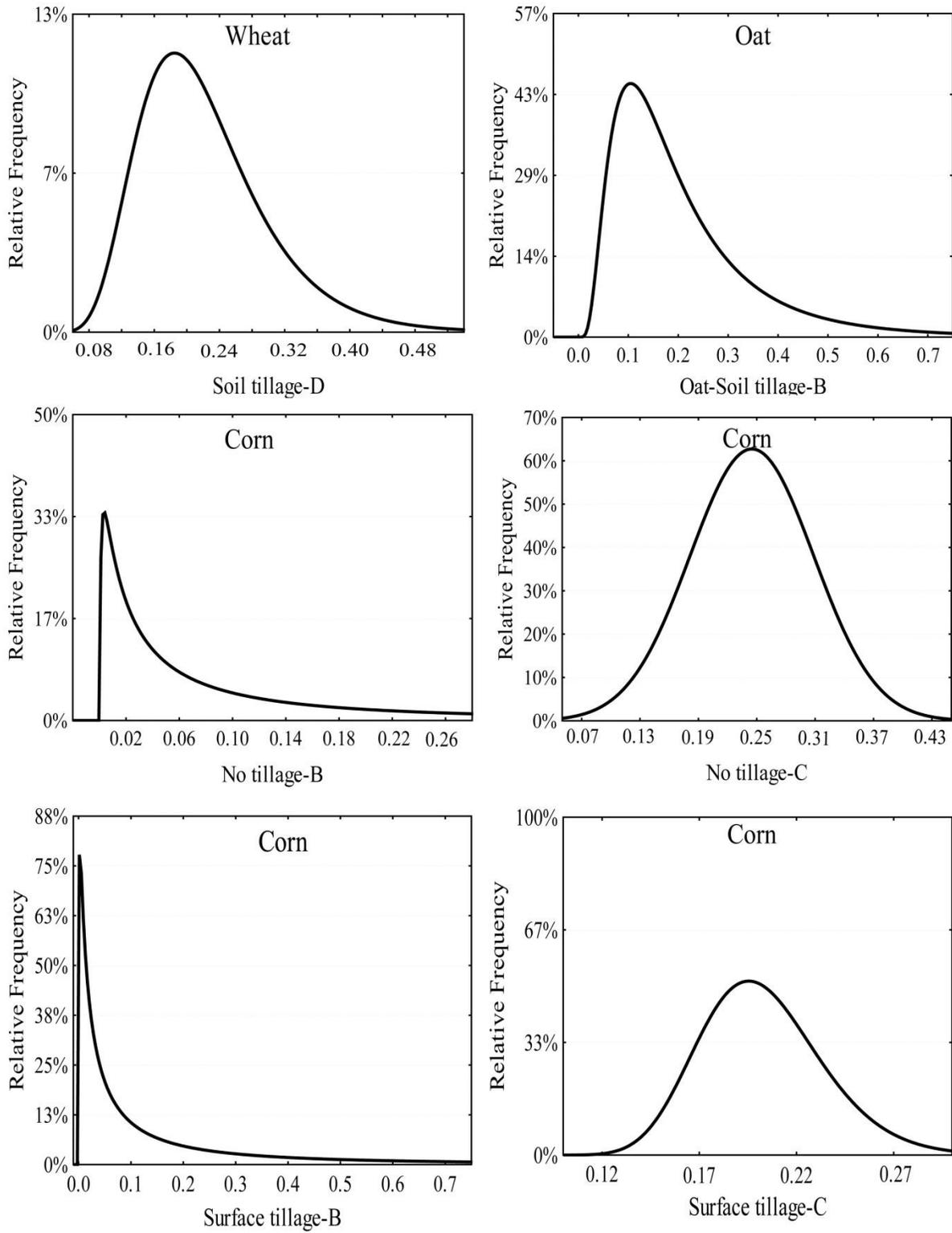
Figure S1



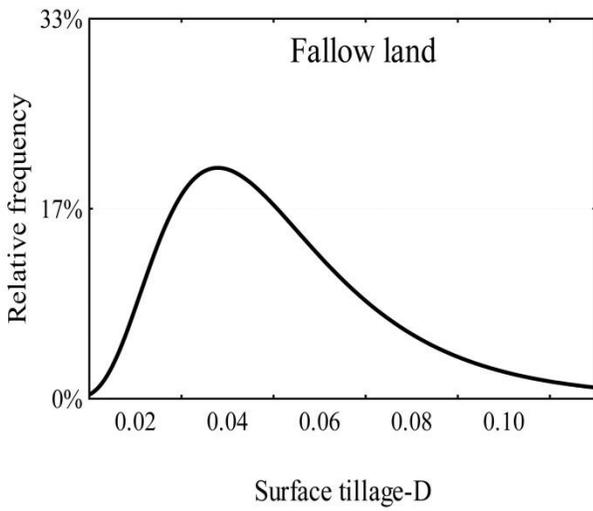
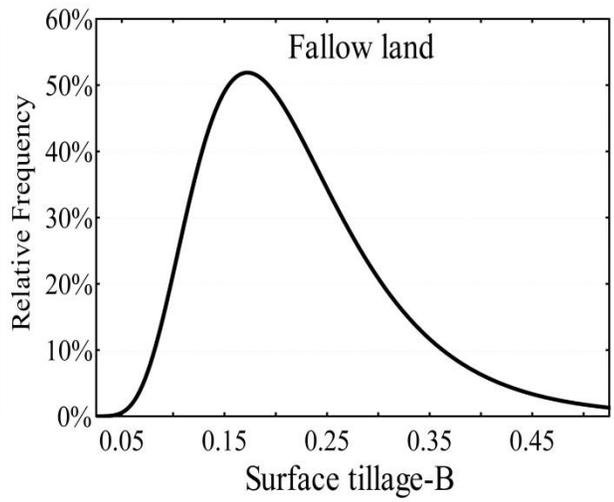
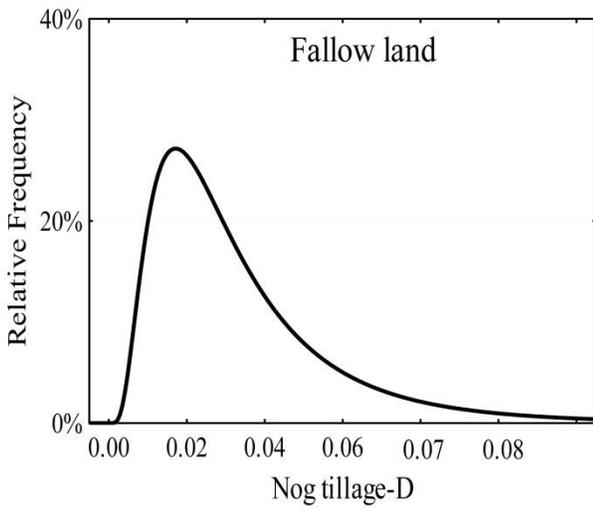
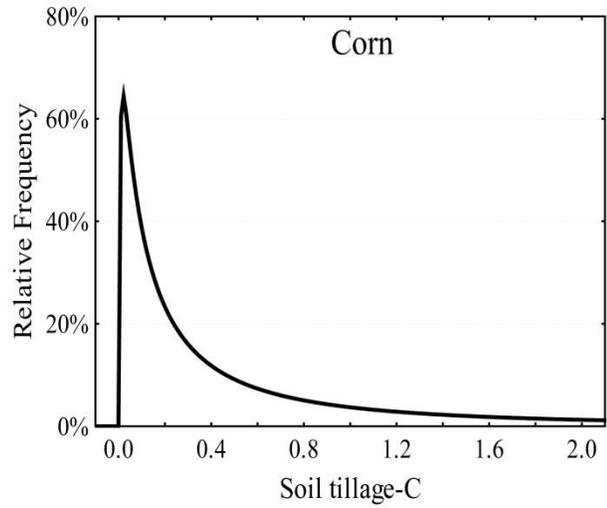
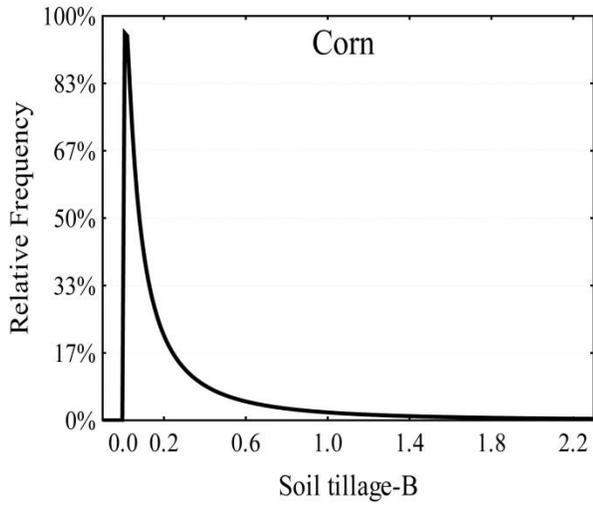
**Figure S2**



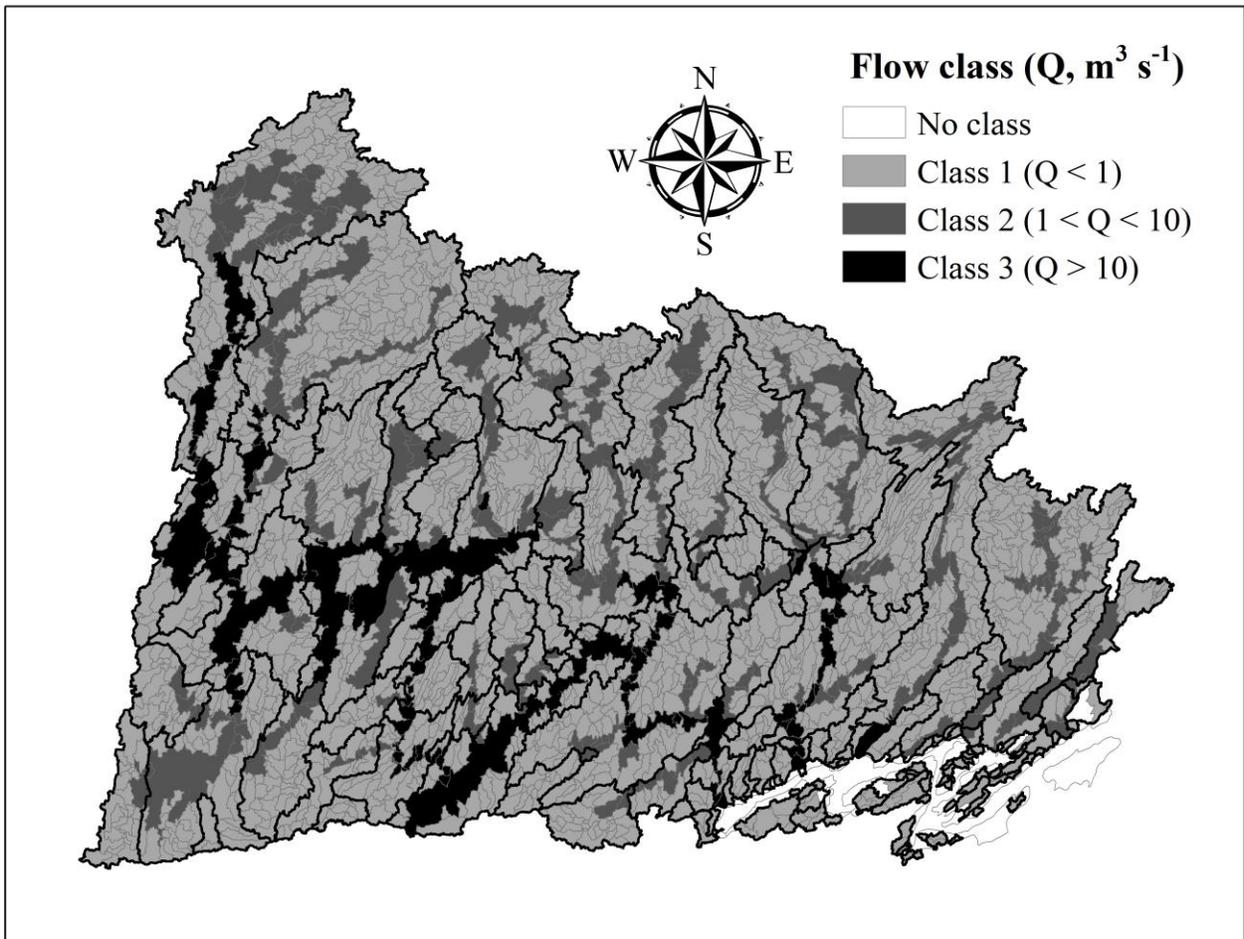
**Figure S3**



**Figure S3 (continued)**



**Figure S3 (continued)**



**Figure S4**

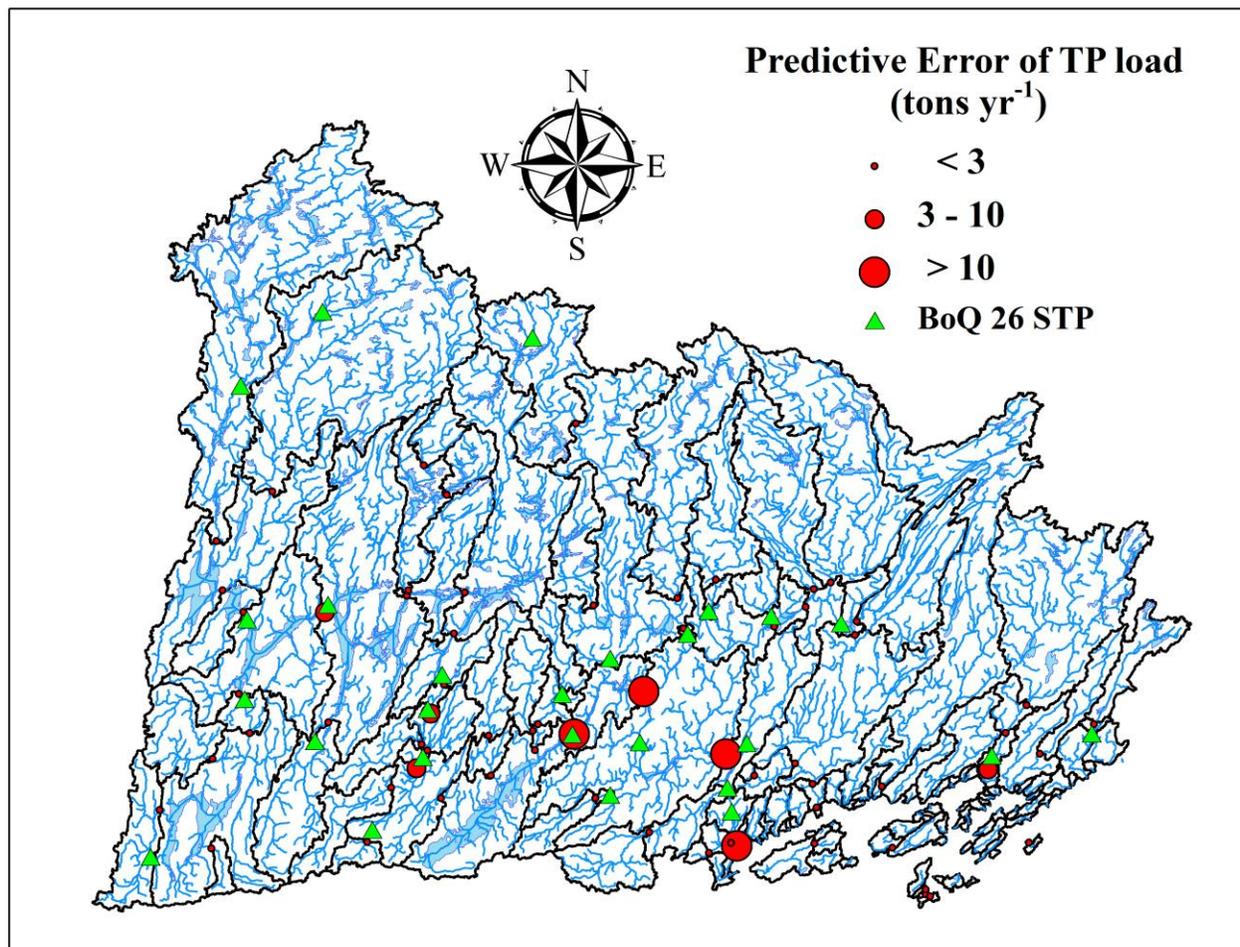


Figure S5

**Table S1:** SPARROW TP export coefficients (tons km<sup>-2</sup> year<sup>-1</sup>) used for the post-hoc simulations with Model II that accounted for the interplay among crop-types, management practices, and soil hydrological groups (also referred to as Model III).

Parameter	Mean	S.D.	Interpretation
<i>Wheat export coefficients</i>			
TP export coefficient from wheat areas with well drained soils and no tillage	0.067	0.009	67 kg of TP per km <sup>2</sup> are released from wheat areas with well drained soils and no tillage on an annual basis.
TP export coefficient from wheat areas with imperfectly and poorly drained soils and no tillage	0.111	0.028	111 kg of TP per km <sup>2</sup> are released from wheat areas with imperfectly and poorly drained soils and no tillage on an annual basis.
TP export coefficient from wheat areas with very poorly drained soils and no tillage	0.022	0.012	22 kg of TP per km <sup>2</sup> are released from wheat areas with very poorly drained soils and no tillage on an annual basis.
TP export coefficient from wheat areas with well drained soils and surface tillage	0.157	0.128	157 kg of TP per km <sup>2</sup> are released from wheat areas with well drained soils and surface tillage on an annual basis.
TP export coefficient from wheat areas with imperfectly and poorly drained soils and surface tillage	0.078	0.019	78 kg of TP per km <sup>2</sup> are released from wheat areas with imperfectly and poorly drained soils and surface tillage on an annual basis.
TP export coefficient from wheat areas with very poorly drained soils and surface tillage	0.025	0.005	25 kg of TP per km <sup>2</sup> are released from wheat areas with very poorly drained soils and surface tillage on an annual basis.
TP export coefficient from wheat areas with well drained soils and no tillage	0.126	0.533	126 kg of TP per km <sup>2</sup> are released from wheat areas with well drained soils and soil tillage on an annual basis.
TP export coefficient from wheat areas with imperfectly and poorly drained soils and soil tillage	0.248	0.080	248 kg of TP per km <sup>2</sup> are released from wheat areas with imperfectly and poorly drained soils and soil tillage on an annual basis.
TP export coefficient from wheat areas with very poorly drained soils and soil tillage	0.167	0.026	167 kg of TP per km <sup>2</sup> are released from wheat areas with very poorly drained soils and soil tillage on an annual basis.
<i>Oat export coefficients</i>			
TP export coefficient from oat areas with well drained soils and no tillage	0.170	0.154	170 kg of TP per km <sup>2</sup> are released from oat areas with well drained soils and no tillage on an annual basis.
TP export coefficient from oat areas with imperfectly and poorly drained soils and no tillage	0.172	0.157	172 kg of TP per km <sup>2</sup> are released from oat areas with imperfectly and poorly drained soils and no tillage on an annual basis.
TP export coefficient from oat areas with very poorly drained soils and no tillage	0.171	0.157	171 kg of TP per km <sup>2</sup> are released from oat areas with very poorly drained soils and no tillage on an annual basis.
TP export coefficient from oat areas with well drained soils and surface tillage	0.173	0.168	173 kg of TP per km <sup>2</sup> are released from oat areas with well drained soils and surface tillage on an annual basis.
TP export coefficient from oat areas with imperfectly and poorly drained soils and surface tillage	0.171	0.154	171 kg of TP per km <sup>2</sup> are released from oat areas with imperfectly and poorly drained soils and surface tillage on an annual basis.
TP export coefficient from oat areas with very poorly drained soils and surface tillage	0.173	0.159	173 kg of TP per km <sup>2</sup> are released from oat areas with very poorly drained soils and surface tillage on an annual basis.
TP export coefficient from oat areas with well drained soils and no tillage	0.179	0.065	179 kg of TP per km <sup>2</sup> are released from oat areas with well drained soils and soil tillage on an annual basis.
TP export coefficient from oat areas with imperfectly and poorly drained soils and soil tillage	0.029	0.008	29 kg of TP per km <sup>2</sup> are released from oat areas with imperfectly and poorly drained soils and soil tillage on an annual basis.
TP export coefficient from oat areas with very poorly drained soils and soil tillage	0.171	0.156	171 kg of TP per km <sup>2</sup> are released from oat areas with very poorly drained soils and soil tillage on an annual basis.

***Oat export coefficients***

TP export coefficient from corn areas with well drained soils and no tillage	0.007	0.009	7 kg of TP per km <sup>2</sup> are released from corn areas with well drained soils and no tillage on an annual basis.
TP export coefficient from corn areas with imperfectly and poorly drained soils and no tillage	0.113	0.007	113 kg of TP per km <sup>2</sup> are released from corn areas with imperfectly and poorly drained soils and no tillage on an annual basis.
TP export coefficient from corn areas with very poorly drained soils and no tillage	0.074	0.135	74 kg of TP per km <sup>2</sup> are released from corn areas with very poorly drained soils and no tillage on an annual basis.
TP export coefficient from corn areas with well drained soils and surface tillage	0.041	0.077	41 kg of TP per km <sup>2</sup> are released from corn areas with well drained soils and surface tillage on an annual basis.
TP export coefficient from corn areas with imperfectly and poorly drained soils and surface tillage	0.089	0.004	89 kg of TP per km <sup>2</sup> are released from corn areas with imperfectly and poorly drained soils and surface tillage on an annual basis.
TP export coefficient from corn areas with very poorly drained soils and surface tillage	0.079	0.158	79 kg of TP per km <sup>2</sup> are released from corn areas with very poorly drained soils and surface tillage on an annual basis.
TP export coefficient from corn areas with well drained soils and no tillage	0.086	0.112	86 kg of TP per km <sup>2</sup> are released from corn areas with well drained soils and soil tillage on an annual basis.
TP export coefficient from corn areas with imperfectly and poorly drained soils and soil tillage	0.133	0.104	133 kg of TP per km <sup>2</sup> are released from corn areas with imperfectly and poorly drained soils and soil tillage on an annual basis.
TP export coefficient from corn areas with very poorly drained soils and soil tillage	0.077	0.145	77 kg of TP per km <sup>2</sup> are released from corn areas with very poorly drained soils and soil tillage on an annual basis.

***Oat export coefficients***

TP export coefficient from alfalfa areas with well drained soils and no tillage	0.044	0.064	44 kg of TP per km <sup>2</sup> are released from alfalfa areas with well drained soils and no tillage on an annual basis.
TP export coefficient from alfalfa areas with imperfectly and poorly drained soils and no tillage	0.043	0.069	43 kg of TP per km <sup>2</sup> are released from alfalfa areas with imperfectly and poorly drained soils and no tillage on an annual basis.
TP export coefficient from alfalfa areas with very poorly drained soils and no tillage	0.044	0.061	44 kg of TP per km <sup>2</sup> are released from alfalfa areas with very poorly drained soils and no tillage on an annual basis.
TP export coefficient from alfalfa areas with well drained soils and surface tillage	0.042	0.061	42 kg of TP per km <sup>2</sup> are released from alfalfa areas with well drained soils and surface tillage on an annual basis.
TP export coefficient from alfalfa areas with imperfectly and poorly drained soils and surface tillage	0.043	0.062	43 kg of TP per km <sup>2</sup> are released from alfalfa areas with imperfectly and poorly drained soils and surface tillage on an annual basis.
TP export coefficient from alfalfa areas with very poorly drained soils and surface tillage	0.043	0.065	43 kg of TP per km <sup>2</sup> are released from alfalfa areas with very poorly drained soils and surface tillage on an annual basis.
TP export coefficient from alfalfa areas with well drained soils and no tillage	0.043	0.067	43 kg of TP per km <sup>2</sup> are released from alfalfa areas with well drained soils and soil tillage on an annual basis.
TP export coefficient from alfalfa areas with imperfectly and poorly drained soils and soil tillage	0.043	0.062	43 kg of TP per km <sup>2</sup> are released from alfalfa areas with imperfectly and poorly drained soils and soil tillage on an annual basis.
TP export coefficient from alfalfa areas with very poorly drained soils and soil tillage	0.049	0.064	49 kg of TP per km <sup>2</sup> are released from alfalfa areas with very poorly drained soils and soil tillage on an annual basis.

***Fallow export coefficients***

TP export coefficient from fallow lands with well drained soils and no tillage	0.163	0.037	163 kg of TP per km <sup>2</sup> are released from fallow lands with well drained soils and no tillage on an annual basis.
TP export coefficient from fallow lands with imperfectly and poorly drained soils and no tillage	0.151	0.298	151 kg of TP per km <sup>2</sup> are released from fallow lands with imperfectly and poorly drained soils and no tillage on an annual basis.

TP export coefficient from fallow lands with very poorly drained soils and no tillage	0.028	0.014	28 kg of TP per km <sup>2</sup> are released from fallow lands with very poorly drained soils and no tillage on an annual basis.
TP export coefficient from fallow lands with well drained soils and surface tillage	0.218	0.094	218 kg of TP per km <sup>2</sup> are released from fallow lands with well drained soils and surface tillage on an annual basis.
TP export coefficient from fallow lands with imperfectly and poorly drained soils and surface tillage	0.60	0.394	600 kg of TP per km <sup>2</sup> are released from fallow lands with imperfectly and poorly drained soils and surface tillage on an annual basis.
TP export coefficient from fallow lands with very poorly drained soils and surface tillage	0.050	0.015	50 kg of TP per km <sup>2</sup> are released from fallow lands with very poorly drained soils and surface tillage on an annual basis.
TP export coefficient from fallow lands with well drained soils and no tillage	3.499	0.982	3.5 tons of TP per km <sup>2</sup> are released from fallow lands with well drained soils and soil tillage on an annual basis.
TP export coefficient from fallow lands with imperfectly and poorly drained soils and soil tillage	0.156	0.444	156 kg of TP per km <sup>2</sup> are released from fallow lands with imperfectly and poorly drained soils and soil tillage on an annual basis.
TP export coefficient from fallow lands with very poorly drained soils and soil tillage	0.158	0.358	158 kg of TP per km <sup>2</sup> are released from fallow lands with very poorly drained soils and soil tillage on an annual basis.

**Table S2:** Mean annual TP loading estimates from the Wilton Creek Equivalent and two SPARROW configurations (Models II and III). Sites refer to the spatial segmentation used by Kim et al. study (2013).

Site		TP loading (tons yr <sup>-1</sup> )		
		WCE	Model II	Model III
Upper Bay	U1	1.8	4.9	5.6
	U2	6.4	9.7	14.9
	U3	0.1	1.5	1.9
Middle Bay	M1	0.4	1.1	1.8
	M2	2.4	4.1	7.1
	M3	0.4	0.7	1.1
Lower Bay	Le	3.5	9.5	11.1