Water Research 105 (2016) 527–539

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Using Bayesian hierarchical models to better understand nitrate sources and sinks in agricultural watersheds

Yongqiu Xia [a, b, *], Donald E. Weller [b], Meghan N. Williams [b], Thomas E. Jordan [b], Xiaoyuan Yan [a]

[a] Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, 210008, China
[b] Smithsonian Environmental Research Center, P.O. Box 28, 647 Contees Wharf Road, Edgewater, MD 21037-0028, USA

ARTICLE INFO

Article history:
Received 29 February 2016
Received in revised form
4 September 2016
Accepted 19 September 2016
Available online 20 September 2016

Keywords:
Bayesian hierarchical models
Export coefficient model
Nutrient sources and sinks
Uncertainty
Temporal and spatial variation

ABSTRACT

Export coefficient models (ECMs) are often used to predict nutrient sources and sinks in watersheds because ECMs can flexibly incorporate processes and have minimal data requirements. However, ECMs do not quantify uncertainties in model structure, parameters, or predictions; nor do they account for spatial and temporal variability in land characteristics, weather, and management practices. We applied Bayesian hierarchical methods to address these problems in ECMs used to predict nitrate concentration in streams. We compared four model formulations, a basic ECM and three models with additional terms to represent competing hypotheses about the sources of error in ECMs and about spatial and temporal variability of coefficients: an ADditive Error Model (ADEM), a SpatioTemporal Parameter Model (STPM), and a Dynamic Parameter Model (DPM). The DPM incorporates a first-order random walk to represent spatial correlation among parameters and a dynamic linear model to accommodate temporal correlation. We tested the modeling approach in a proof of concept using watershed characteristics and nitrate export measurements from watersheds in the Coastal Plain physiographic province of the Chesapeake Bay drainage. Among the four models, the DPM was the best–it had the lowest mean error, explained the most variability (R^2 = 0.99), had the narrowest prediction intervals, and provided the most effective tradeoff between fit complexity (its deviance information criterion, DIC, was 45.6 units lower than any other model, indicating overwhelming support for the DPM). The superiority of the DPM supports its underlying hypothesis that the main source of error in ECMs is their failure to account for parameter variability rather than structural error. Analysis of the fitted DPM coefficients for cropland export and instream retention revealed some of the factors controlling nitrate concentration: cropland nitrate exports were positively related to stream flow and watershed average slope, while instream nitrate retention was positively correlated with nitrate concentration. By quantifying spatial and temporal variability in sources and sinks, the DPM provides new information to better target management actions to the most effective times and places. Given the wide use of ECMs as research and management tools, our approach can be broadly applied in other watersheds and to other materials.

1. Introduction

The linkages between watershed nitrogen (N) loading and export are often poorly quantified because of the complexity of N pathways. Human activities have greatly increased N loading by applying fertilizer, generating human and animal waste, and burning fossil fuels (Vitousek et al., 1997). The higher N loading increases N export from terrestrial systems to streams (Jordan et al., 1997a; Gao et al., 2004); but delivery to downstream rivers, lakes, and estuaries is often less than terrestrial N export because N removal in both terrestrial and aquatic systems can temporarily or permanently reduce downstream transport of N through physical, biological, or chemical processes (Galloway et al., 2008).

Export coefficient models are useful tools for estimating loads and sources of nutrients. The basic export coefficient model assumes that different land uses are the major sources of watershed nutrient export (Norvell et al., 1979), and the model includes a

* Corresponding author. Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, 210008, China.
E-mail address: yqxia@issas.ac.cn (Y. Xia).

http://dx.doi.org/10.1016/j.watres.2016.09.033
0043-1354/

calibrated export coefficient (the rate of pollutant loading per unit of land area) for each modeled land use. Johnes (1996) developed improved models by adding coefficients to represent nutrient exports due to plant nitrogen fixation, atmospheric deposition, livestock, or people. Several models have also considered nutrient sinks. Gardner et al. (2011) modeled nitrate retention in streams using a regression relating nitrate concentration to watershed area, while Lu et al. (2013) used a first-order function relating nitrate loss to water travel time. Endreny and Wood (2003) integrated maps of watershed topography, land use, and stream channels into spatially-weighted export coefficients that predict spatial patterns of phosphorus loading in a watershed. To account for nutrient removal in streamside forests and wetlands (riparian buffers), Maillard and Santos (2008) weighted export coefficients by fixed buffer widths. More recent models have quantified buffer prevalence along topographic flow paths connecting nutrient source areas to streams, and then used that information to separate buffered from unbuffered sources (Weller et al., 2011; Weller and Baker, 2014). The ability to incorporate additional processes into basic export coefficient models has improved their applicability in watershed nutrient management.

Even with the above enhancements, export coefficient models do not account for the variability in nutrient sources and sinks that result from variations in land characteristics, weather, and management practices in both space and time (e. g., variation in soils, slopes, irrigation practices, tillage practices, rainfall volume, runoff volume, etc.) (Lu et al., 2013; Vassiljev et al., 2008). Not accounting for spatial and temporal variability limits the application of export coefficient models to management and decision making because the models do not help to identify when and where nutrient releases or nutrient retention are most important (e. g., “hot spots” and “hot moments,” McClain et al., 2003). That information is needed to make effective decisions about when and where to implement management practices.

Export coefficient models also typically do not account for model uncertainty. There are two important kinds of model uncertainty: structural and parameter uncertainty. Structural uncertainty arises from inadequately representing the true processes that control nutrient export. Much of the structural uncertainty in export coefficient models likely arises from spatial and temporal lumping (Ajami et al., 2007). Parameter uncertainty is the difference between the true parameters of a process and the parameters estimated by model calibration. Parameter uncertainty arises from using limited and uncertain calibration data, from imperfect process understanding, and from model approximation. For one export coefficient model, Khadam and Kaluarachchi (2006) found that the export coefficient of agricultural land varied as much as 72.0%, while the coefficients of urban and forest varied 59.1% and 58.2%, respectively. Structural and parameter uncertainty together yield the overall prediction uncertainty, which is often large for watershed models. Boomer et al. (2013) compared several watershed export models and showed that no model consistently matched observed discharges better than the others, and predictions differed as much as 150% for every basin considered. Prediction errors and unquantified uncertainty can foster poor understanding and ineffective decisions when export coefficient models are applied to watershed management. To understand and quantify the effects of structural and parameter uncertainties, we need methods to assess the relative probabilities that different model structures or parameter sets provide acceptable simulations of a natural system.

Bayesian methods can provide a very effective framework needed to handle spatial and temporal variability and model uncertainty, as demonstrated in models relating algal blooms and phosphate loading (Obenour et al., 2014; Wellen et al., 2012). Therefore, to better quantify nitrate sources and sinks in agricultural watersheds, we apply Bayesian hierarchical models to incorporate temporally and spatially varying processes into an export coefficient model. We examine the ability of three alternative approaches to enhance insight into model performance and parameter uncertainty. Our results document better methods to account for the major uncertainties of export coefficient models; and we apply those methods to better understand and quantify the nitrate sources and sinks that lead to observed stream nitrate levels.

2. Methods

This study builds on a published export coefficient model (the SERC model) that predicts average annual nitrate levels in streams of the Chesapeake Bay watershed (Weller et al., 2011; Weller and Baker, 2014). We modified the SERC model by adding the process of nutrient removal in streams and weekly variation in that removal. This extended SERC model (ESERC) predicts weekly rather than annual nitrate levels.

We incorporated the ESERC model into a set of four alternative models that represented different hypotheses about the nature of spatial and temporal variability in nitrate sources and sinks. The four models were the simple ESERC model (explained above) and the ADditive Error Model (ADEM), SpatioTemporal Parameter Model (STPM), and (Dynamic Parameter Model) DPM models (explained below); in which different error terms are added to the base ESERC model. The ADEM assumes that model errors arise mainly from unspecified nutrient sources or sinks (structural error) that are spatially and temporally variable, and ADEM accounts for those errors by adding temporal and spatial error terms. This simple approach has been used to model disease risk (Lagazio et al., 2001), pollution (Shaddick and Wakefield, 2002), eutrophication (Arhonditsis et al., 2008), and hydrology (Huard and Mailhot, 2006). The STPM instead assumes that model errors come mainly from using static export coefficients (parameter error), and seeks to reduce model error by allowing the model parameters (export coefficients) to vary spatially and temporally. The STPM has been used to model non-point source pollution (Lu et al., 2013; Zobrist and Reichert, 2006). The DPM also assumes that parameter error in the export coefficients is the source of model error, but the DPM differs from the STPM in its representation of how the export coefficients vary temporally and spatially. The STPM assumes that each coefficient has its own independent distribution, while the DPM instead assumes that variations in the coefficients are correlated in space and time. The DPM has also been used to characterize ecological processes (Arhonditsis et al., 2008; Sadraddini et al., 2011; Wellen et al., 2012). The following presents each of the four models in greater detail.

2.1. The SERC model

The SERC model refined the basic export coefficient model (Norvell et al., 1979) by including physiographic province as a categorical independent variable (Weller et al., 2011; Weller and Baker, 2014). This accounts for broad regional differences in export coefficients that arise from differences in topography, geology, and other geographic factors. SERC also fit separate export coefficients for all cropland and for unbuffered cropland to account for the removal of cropland nitrate in adjacent riparian buffers. To quantify unbuffered cropland, a digital elevation model was analyzed to identify the steepest-descent surface transport pathway connecting each cropland pixel in a watershed to a stream. The resulting flow paths were intersected with a land cover map to determine what fraction of the paths do not pass through a riparian buffer, yielding the fraction of unbuffered cropland pixels in the

watershed (for full details see Baker et al., 2006b; Weller et al., 2011).

The SERC model is formulated as:

$$N = \underbrace{\beta_0 + \beta_{0P} + \beta_{0a}}_{background} + \underbrace{(\beta_c + \beta_{cP}P_P + \beta_{ca}P_a)C}_{cropland} + \underbrace{(\beta_u + \beta_{uP}P_P + \beta_{ua}P_a)C_u}_{unbuffered\ cropland} + \underbrace{\beta_g G}_{grassland} + \underbrace{\beta_d D}_{developed\ land} + \varepsilon \quad (1)$$

where N is stream nitrate concentration; C, G, and D are the proportions of all cropland, grassland, and developed land in a watershed; β values are export coefficients; ε is error; and P_P and P_a are dummy variables representing the categorical variable physiographic province. The subscript u designates the proportion of unbuffered cropland C_u and its coefficientβ_u. The parameter β_c represents the rate of nitrate loss from all cropland (both buffered and unbuffered), while β_u represents an extra amount of nitrate lost only from the unbuffered cropland. Thus, the total loss from unbuffered cropland is $\beta_c + \beta_u$. These interpretations of the model parameters are described in more detail with an accompanying figure in a published paper (Fig. 2, Weller et al., 2011). When considering a single physiographic province, Eq. (1) can be simplified to:

$$N = \beta_0 + \left(\beta_c C + \beta_u C_u + \beta_d D + \beta_g G\right) + \varepsilon \quad (2)$$

2.2. The extended SERC model

The SERC model estimates nitrate delivery from the land to the stream network, but some of that delivered nitrate can be removed by in-stream processes (Gardner et al., 2011; Lu et al., 2013; Valett et al., 1996). In-stream nitrate retention can be effectively modeled using a first-order reaction (exponential decay) with loss rate k and in-stream travel time t (Lu et al., 2013; Smith et al., 1997). Here, we extend the SERC model by incorporating in-stream attenuation and temporal variation in stream nitrate concentration and in stream travel time:

$$\begin{aligned}\mu_{i,j} &= N \times \exp(-kt_{i,j}) \\ &= \beta_0 + \left(\beta_c C + \beta_u C_u + \beta_d D + \beta_g G\right)\exp(-kt_{i,j}) + \varepsilon \end{aligned} \quad (3)$$

where $\mu_{i,j}$ is the nitrate concentration at watershed monitoring station i at sampling time j. We refer to Eq. (3) as the ESERC (Extended SERC) model. Because the in-stream travel times differ among watersheds and through time, the ESERC model (unlike the SERC model) can represent temporal variation in nitrate concentration.

2.3. Bayesian hierarchical modeling and uncertainty

We used Bayesian hierarchical methods to fit models because the Bayesian framework offers many advantages. Bayesian theory provides a natural and principled way of combining prior information with data, and Bayesian models produce a probability distribution for the estimate of every parameter and prediction (Gilks, 2005). Bayesian inference also safeguards against overfitting and can provide unbiased estimates even for very small sample sizes (Congdon, 2007).

The ESERC model accommodates spatial and temporal variations in nitrate removal in streams, but not in the land sources and sinks of nitrate. To better represent variability in those sources and sinks, we developed and compared three alternative Bayesian model formulations that represent different hypotheses about the origin and structure of spatial and temporal variability. All three of

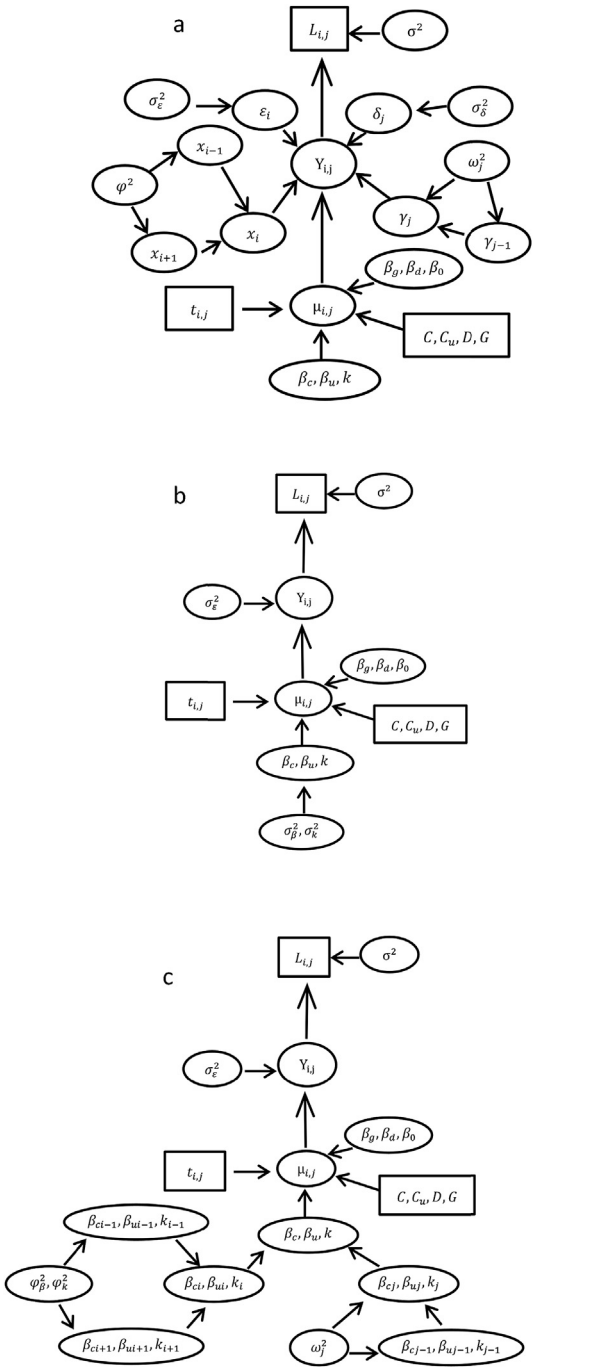

Fig. 1. Directed acyclic graphs (DAGs) describing the probabilistic relationships among variables and parameters in three alternative formulations of the Bayesian model: a) ADEM, b) STPM, and c) DPM. Ovals represent the model variables and parameters, rectangles represent the input data, and arrows represent the conditional dependencies. Symbols are described in Sections 2.1 and 2.3.

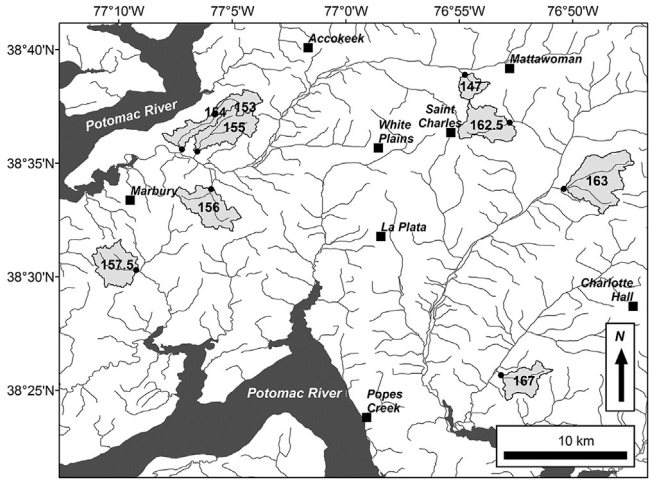

Fig. 2. Study watersheds (gray shading) and sampling stations (black circles) on streams (dark gray lines) in Charles County, Maryland about 40 km south of Washington, DC. Several towns (black squares) are show for location reference.

the models take the basic form:

$$L_{i,j} = Y_{i,j} + \varepsilon_{i,j} \tag{4}$$

where $L_{i,j}$ is the measured concentration of nitrate leaving watershed i at time j; $Y_{i,j}$ is the mean stream nitrate concentration simulated by a candidate model, and $\varepsilon_{i,j}$ is a stochastic term representing measurement error. We assume that measurement error (the difference between the observed loads and the true load) $\varepsilon_{i,j}$ is independent and normally distributed:

$$L_{i,j} \sim N\left(Y_{i,j}, \sigma^2_{i,j}\right) \tag{5}$$

where $\sigma_{i,j}$ is the standard deviation of measurement error.

A second level of uncertainty is the difference between the true load and the prediction of a candidate model:

$$Y_{i,j} = \mu_{i,j} + \varepsilon_{i,j} \tag{6}$$

where $\mu_{i,j}$ is an unbiased estimator of the "true" load by the candidate model and $\varepsilon_{i,j}$ is the associated structural or parameter error (calculated in each model as described below).

2.3.1. ADditive Error Model

The first alternative statistical formulation (Fig. 1a) assumes that the main source of error in the ESERC model is from unrepresented nitrate sources and sinks in the watershed (structural error). The ADEM equation is

$$Y_{i,j} = \mu_{i,j} + \varepsilon_{i,j} \quad \varepsilon_{i,j} = x_i + \varepsilon_i + \gamma_j + \delta_j \quad i = 1, \cdots, W; j = 1, \cdots, T \tag{7}$$

where $\mu_{i,j}$ is the ESERC prediction of nitrate concentration leaving watershed i at time j (Eq. (3)). We divided both the temporal and spatial errors into two parts: correlated error and independent error. Spatial (x_i) and temporal (γ_j) correlated error terms represent the unspecified sources or sinks; and spatial (ε_i) and temporal (δ_j) independent errors terms represent measurement errors. ε_i and δ_j are assumed to be independent and distributed normally with mean zero and precisions σ_ε^{-2} and σ_δ^{-2} ($\varepsilon_i \sim N(0, \sigma_\varepsilon^{-2})$, $\delta_j \sim N(0, \sigma_\delta^{-2})$). T is the number of water quality observations and W is the number of study watersheds. The priors for the σ terms are represented by an uninformative prior distribution, the gamma distribution with shape and scale parameters of 0.01 [Gamma (0.01, 0.01)].

Within a relatively homogenous region, the distributions of the (x_i) representing unspecified sources or sinks would be similar in different watersheds, resulting in spatial autocorrelation. We adopted a first-order random-walk effect to account for the spatial correlation among model residuals (Arhonditsis et al., 2008; Sadraddini et al., 2011; Wellen et al., 2012).

$$x_i|x_{-i} \sim \begin{cases} N\left(x_{i+1}, \varphi^2\right) & for\ i = 1 \\ N\left(\frac{x_{i-1} + x_{i+1}}{2}, \frac{\varphi^2}{2}\right) & for\ i = 2, \cdots, W-1 \\ N\left(x_{i-1}, \varphi^2\right) & for\ i = W \end{cases} \tag{8}$$

The notation x_{-i} denotes all watersheds except watershed x_i and φ^2 is the conditional variance of x_i given x_{-i}. The prior density of $x_i|x_{-i}$ is based on conjugate inverse-gamma (0.01, 0.01) distribution.

It is also likely that the unspecified sources or sinks (γ_j) are more similar between successive observations than between observations separated by longer time intervals, which would result in temporal autocorrelation. Dynamic linear modeling (DLM) can account for temporal autocorrelation (Congdon, 2007; Lamon et al., 1998b) by assuming that the level of the response at each time step is influenced by past levels (Wellen et al., 2012). We also assume the more recent past is more influential than the more distant past, and model this by including a discount factor in the DLM (Eq. (8)). The discount factor is a weight representing the fraction of information that can be obtained from the last observation period. In short, the DLM approach in this paper posits that unspecified sources or sinks vary with time, and that each time series is autocorrelated–the closer in time, the more similar are parameter values.

$$\begin{aligned} &\gamma_j \sim N\left(0, \omega_j^2\right) \\ &\omega_j^{-2} = \tau^{j-1} \times \omega_1^{-2} \quad j = 2, \cdots, T \quad \omega_1^{-2} \sim gamma(\alpha, \beta) \end{aligned} \tag{9}$$

where γ_j represent unspecified sources or sinks at time j sampled from normal distributions with zero mean and variance ω_j^2. τ is the discount factor. The value of the discount factor is selected by comparing models with identical structures but different discount factors to identify the value yielding the highest log likelihood (Lamon et al., 1998b). Discount factors control the relative weighting of recent observations versus older observations. Based on recent reports (Sadraddini et al., 2011; Wellen et al., 2012), we evaluated discounts between 0.9 and 1.0 and selected a value of $\tau = 0.95$ (Congdon, 2007; Lamon et al., 1998b). The parameter ω_1^{-2} was assumed to follow a gamma distribution with parameters α and β, where $\alpha\beta$ is the mean of the distribution and $\alpha\beta^2$ is the variance. Parameters α and β in the Gamma distributions were chosen to give the priors a high variance so that the priors are uninformative and have little effect on the posterior distributions.

2.3.2. SpatioTemporal Parameter Model

The second statistical formulation (STPM) assumes that errors in the model arise from using static export coefficients (parameter error), so STPM adds spatially and temporally variable export coefficients to account for the discrepancies between the model predictions and the observations (Fig. 1b). The STPM equation is

$$\begin{aligned} Y_{i,j} &= \mu_{i,j} + \varepsilon_{i,j} \\ \mu_{i,j} &= \beta_0 + \left(\beta_{ci,j}C + \beta_{ui,j}C_u + \beta_d D + \beta_g G\right)\exp(-k_{i,j}t_{i,j}) \\ \varepsilon_{i,j} &\sim N\left(0,\ \sigma_\varepsilon^2\right), \quad \sigma_\varepsilon \sim \text{gamma}(\alpha,\beta) \end{aligned} \tag{10}$$

where export coefficients $\beta_{ci,j}$, $\beta_{ui,j}$, and $k_{i,j}$ vary among watersheds i and times j. Each parameter for a particular watershed and time follows its own independent distribution. We did not model such variation in β_d and β_g, because these coefficients were relatively small in previous analyses (Weller et al., 2011; Weller and Baker, 2014) and in our analyses (see Results). $\varepsilon_{i,j}$ represents the stochastic error terms of the model in space and time ($\varepsilon_{i,j} \sim N(0,\ \sigma_\varepsilon^{-2})$). STPM adapts an approach used in published load apportionment models (Gardner et al., 2011; Lu et al., 2013).

2.3.3. Dynamic Parameter Model

The third alternative statistical formulation (DPM) follows STPM, but does not assume that each cropland parameter has its own independent distribution. Assuming complete independence may cause low model efficiency if watershed processes are similar among watersheds and times. Nearby watersheds could have similar parameters because they are more alike in soils, land use, or nutrient management than distant watersheds. Likewise, as observations become closer in time, model parameter values become more similar because major environmental drivers (like rainfall and temperature) that affect the parameters are more similar. Therefore, we assume that variations in the parameters of DPM are correlated in space and time (Fig. 1c). The DPM equation is

$$\begin{aligned} L_{i,j} &\sim N\left(Y_{i,j}, \sigma^2\right) \\ Y_{i,j} &= \mu_{i,j} + \varepsilon_{i,j} \\ \mu_{i,j} &= \beta_0 + \left(\beta_{ci,j}C + \beta_{ui,j}C_u + \beta_d D + \beta_g G\right)\exp(-k_{i,j}t) \end{aligned} \tag{11}$$

To effectively accommodate the correlations among parameters, we use a first-order random walk to represent spatial variability and a DLM to represent temporal variability (as described above), but replace both x_i in Eq. (8) and γ_j in Eq. (9) with θ, which can represent any of three parameters β_c, β_u, or k.

$$\begin{aligned} N_{nobuf} = &\underbrace{\beta_0 + \beta_d D + \beta_g G}_{Non-cropland\ source} + \underbrace{\beta_c C}_{Buffer\ leakage} + \underbrace{\beta_u C_u}_{Possible\ removal\ in\ restored\ buffer\ gaps} + \underbrace{\beta_u (C - C_u)}_{Removal\ in\ existing\ buffers} \\ &+ \underbrace{\left(\beta_0 + \beta_c C + \beta_u C_u + \beta_d D + \beta_g G\right)\left(1 - \exp(-k t_{i,j})\right)}_{in-stream\ removal} + \varepsilon \end{aligned} \tag{13}$$

2.4. Bayesian calibration

A Markov chain Monte Carlo (MCMC) algorithm was applied to obtain the predictive uncertainty and summaries of targeted parameters and predictions for each of the four models. We used the general Gibbs sampler algorithm implemented in the WinBUGS software (Lunn et al., 2000). We used three chains of 15,000 iterations, and samples were taken after the MCMC simulation converged to the true posterior distribution. Convergence was assessed using the modified Gelman-Rubin convergence statistic (Brooks and Gelman, 1998). The sequences converged very rapidly (≈1000 iterations). The summary statistics reported in this study were based on the last 10,000 draws by keeping every 20th iteration (thin = 20) to avoid serial correlation.

2.5. Model evaluation

To compare overall model strength among the four models and to identify the best model, we used three metrics: the coefficient of determination (R^2), root mean squared error (RMSE), and the deviance information criterion (DIC). R^2 is a common and intuitive measure of predictive skill that can be interpreted as the fraction of variance among the observations that can be explained by a model. RMSE is a frequently-used measure of the differences between predicted and observed values, and RMSE aggregates those errors across time and space into a single measure of predictive power. DIC balances the tradeoff between goodness of fit and model complexity with two components: a Bayesian term that measures goodness of fit and a penalty term for increasing model complexity.

$$DIC = \overline{D(\theta)} + p_D \tag{12}$$

where ($\overline{D(\theta)}$) is a measure of model fit that attains smaller values for better models and p_D measures model complexity as the effective number of parameters (Spiegelhalter et al., 2002). DIC is the Bayesian analog of Akaike's Information Criterion (Burnham and Anderson, 2002), and a smaller DIC value indicates a better model than a larger DIC value (Spiegelhalter et al., 2002).

2.6. Model application

We analyzed the best of the four calibrated models to predict stream nitrate levels and to quantify nitrate sources and sinks. The parameters and predictions of the model can be manipulated to quantify the sources and sinks within a watershed, including: non-cropland sources, buffer leakage, possible removal in restored buffer gaps, removal in existing buffers, and in-stream removal (Weller et al., 2011; Weller and Baker, 2014). This can be done by rearranging Eq. (3) and adding a term to represent the nitrate removed by existing buffers to produce an equation that predicts N_{nobuf}, the expected stream nitrate concentration from current land cover if there were no buffers on cropland (Weller et al., 2011) and no in-stream removal.

We applied this equation to estimate these components of the nitrate budget. Weller et al. (2011) provide more details on the derivation of this equation.

2.7. Proof of concept

2.7.1. Study watersheds

We used a subset of sampling stations from the SERC Chesapeake Bay watershed study (Jordan et al., 1997a, 1997b; Weller and Baker, 2014; Jordan et al., 1997b, c; Liu et al., 2000) to fit and compare the ESERC, ADEM, STPM and DPM models. The nine selected watersheds (numbers 147, 153, 154, 155, 156, 157.5, 162.5,

163, and 167) were in the Mattawoman Creek drainage within the Coastal Plain physiographic province (Fig. 2 in Jordan et al., 2000). The study watersheds have differing proportions of agricultural and non-agricultural land; differing stream network configurations; and no sewage outfalls, reservoirs, or lakes.

2.7.2. Water sampling

We measured water and nutrient discharges in a previous study using automated samplers to ensure adequate sampling of particulate fractions in storm flow (Jordan et al., 1997b). The samplers monitored stream depth and calculated water flow from rating curves of flow versus depth (Jordan et al., 1997b). The automated samplers used Campbell CR10 data loggers to record depth, calculate flow, and control pumps to take samples of stream water after a set amount of flow had occurred. Water samples were collected from the middle of stream. Samples were pumped more frequently at higher flow rates, up to once every 5 min during storm flow, and composited for one week to yield weekly volume-integrated water samples (Jordan et al., 1997b). The nine watersheds were sampled concurrently for 68 weeks (17 months) from April 20 1998 through August 2 1999. Nitrate was analyzed by cadmium reduction method (Jordan et al., 1997b).

2.7.3. Geographic analysis

Watershed boundaries of the study catchments were delineated within the ArcGIS 10.2 geographic information system (ESRI, 2011) by topographic analysis of digital elevation data (30 m National Elevation Dataset [NED], Gesch et al., 2002) reconciled with digital stream maps (1:24,000 National Hydrography Dataset [NHD], Simley and Carswell, 2009) by stream burning with DEM reconditioning. Full details of the watershed delineation are previously published (Baker et al., 2006b). Land cover proportions (cropland, developed land, grassland, forest, and wetland) within the study watersheds (Table 1) were obtained by intersecting the watershed boundaries with the circa 1990 National Land Cover Dataset (called NLCD 1992, USEPA, 2007; Vogelmann et al., 1998a, 1998b) within ArcGIS. Cropland in each watershed was divided into buffered and unbuffered cropland using flow path analysis of riparian buffer prevalence (Baker et al., 2006a); an analysis which quantifies buffer prevalence along the flow paths linking source areas (here croplands) to streams. Briefly, for every cropland pixel (NLCD 1992), we used the digital elevation model (30 m NED, Gesch et al., 2002) to identify the steepest descent surface pathway (O'Callaghan and Mark, 1984) connecting that cropland pixel to a stream (1:24,000 NHD, Simley and Carswell, 2009). We then determined whether or not each path went through a riparian buffer, which was defined as forest or wetland land cover (NLCD 1992) overlaying or contiguous with the stream. The fraction of flow paths that do not pass through a riparian buffer is the fraction of unbuffered cropland pixels in the watershed (Table 1). Full details of the flow path analysis as implemented within ArcGIS are previously published (Baker et al., 2006a; Weller et al., 2011).

2.7.4. In-stream travel time

The in-stream travel times for individual streams were calculated from weekly average water velocity and flow-path length in the stream. Weekly velocities (meters/second) were estimated from drainage area, dimensionless drainage area, slope, flow volume, and a dimensionless relative flow using a published empirical relationship fitted to 980 time-of-travel studies in 90 U.S. streams and rivers that represent a range of sizes, slopes, and channel geometries (Jobson, 1996):

$$\begin{aligned} V &= 0.094 + 0.0143 * D_w^{'\,0.919} * Q_w^{'-0.469} * slope^{0.159} * Q/D_w \\ D_w^{'} &= \left(D_w^{1.25} * g^{0.5}\right) \Big/ Q_w \\ Q_w^{'} &= Q/Q_w \end{aligned} \tag{14}$$

where V is weekly mean water velocity (m/s), Q is flow (m^3/s), Q_w is mean weekly flow (m^3/s), $Q_w^{'}$ is dimensionless relative discharge, D_w is drainage area (m^2), g is the acceleration of gravity (9.8 m/s^2), and $D_w^{'}$ is dimensionless relative drainage area. The travel time of pollutant loads to the downstream outlet is calculated as $T = \frac{L}{V}$ where L is the length of the stream in the watershed. As in other published models, we assumed that nitrate entering a stream reach is subject to loss along half of the stream length on average, corresponding to in-stream travel time of $T/2$ (Lu et al., 2013; Alexander et al., 2006).

2.7.5. Prior distributions of model parameters

The independent prior distributions of land use export coefficients were based on published studies in Chesapeake Bay drainage (Weller et al., 2011). The prior distribution of the stream attenuation coefficient (k) was previously compiled from the worldwide literature (Lu et al., 2013). Because that compilation included k values from streams with widely different climatic and physical characteristics, we set the prior bounds much broader than the range of literature values, so that the posterior distributions were more strongly influenced by the new data than by the priors. The same prior distributions were used for all four models (Table 2).

Table 2
Prior distributions of the parameters. The same prior distributions were used for all four models (ESERC, ADEM, STPM, and DMP).

Parameter	Description	Prior distribution
β_c	Export coefficient for all cropland	N(0.56, 3.04)
β_u	Export coefficient for unbuffered cropland	N(10.12, 4.30)
β_d	Export coefficient for developed land	N(0.02, 0.32)
β_g	Export coefficient for grassland	N(0.02 0.26)
k	River attenuation coefficient	N(0.58, 3.96)

Table 1
Total areas and percentages of land cover types in nine study watersheds. Wetlands (not shown) are less than 1% of each watershed, and the remainder is forest (not shown).

Watershed	Land area km^2	Cropland (%)	Unbuffered cropland (%)	Developed land (%)	Grassland (%)
147	2.5	11.2	5.5	31.8	11.3
153	7.24	1.2	0.1	20.0	6.5
154	3.67	1.0	0	32.4	6.8
155	7.49	0	0	8.0	4.0
156	7.82	3.6	0.8	4.9	4.1
157.5	7.21	0.7	0	0.3	4.5
162.5	7.30	2.7	1.0	54.5	10.0
163	12.51	12.3	2.4	7.7	19.5
167	5.11	18.9	5.8	0.7	26.9

Table 3
Posterior parameter means and standard deviations (SD) of the four models.

Parameters	ESERC		ADEM		STPM		DPM	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
β_c	0.38	0.17	0.45	0.44	0.56	0.14	1.13	1.65
β_u	9.08	0.43	7.83	1.19	10.12	1.87	7.69	4.32
β_d	0.22	0.07	0.41	0.19	0.47	0.06	0.52	0.07
β_g	0.11	0.16	−0.12	0.21	0.40	0.13	0.47	0.90
k	0.34	0.09	1.22	0.51	0.76	0.46	0.72	1.87

3. Results

3.1. Model comparison and assessment

The posterior parameter means and standard deviations differed among the models (Table 3). The reported values of the temporally- and spatially-varying parameters are averages of the mean and standard deviation values across all simulated watersheds and weeks, so the values in Table 3 are comparable among

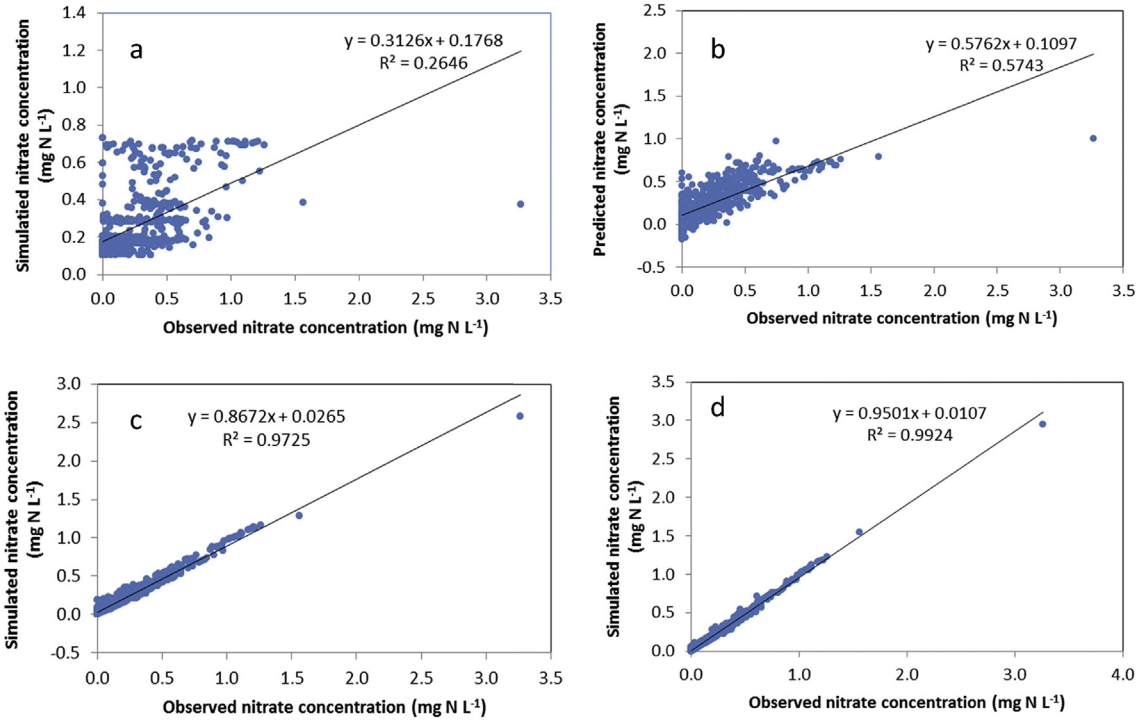

Fig. 3. Predicted versus observed nitrate concentrations for the four models: a) ESERC; b) ADEM; c) STPM; and d) DPM.

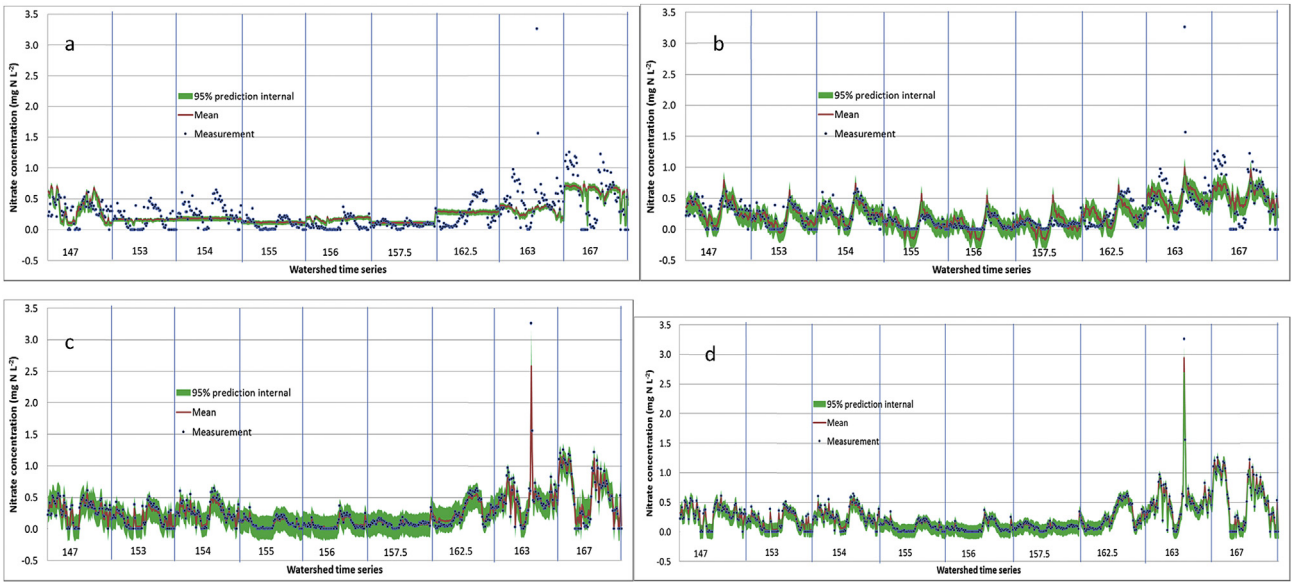

Fig. 4. Time series predictions of stream nitrate concentration from four models: a) ESERC; b) ADEM; c) STPM; and d) DPM. Each watershed time series shows 68 weeks of data from April 20, 1998 through August 2, 1999.

Table 4
Measures of skill for the four models.

Measure	ESERC	ADEM	STPM	DPM
$\overline{D}$	815	240.0	165.7	52.2
p_D	4.6	77.5	186.7	219.4
DIC	819.6	317.5	352.4	271.6
RMSE	0.245	0.185	0.057	0.028
R^2	0.265	0.574	0.972	0.992

Note: $\overline{D}$ is the posterior mean of deviance, p_D is the effective number of parameters, and DIC is the deviance information criterion.

models. For all the models, the export coefficient for unbuffered cropland was much larger than the coefficients for other land uses (Table 3). The model coefficients for all cropland are much lower than for unbuffered cropland but greater than the coefficients for developed land and grassland. The low coefficients of developed land and grassland indicate minor contributions of these land types to stream nitrate concentration. Compared to the other models, the DPM had smallest coefficient for unbuffered cropland, but had the largest coefficients for the other land types. Allowing model parameters to vary spatially and temporally substantially changed the posterior mean values compared to models with static parameters.

Measures of model accuracy indicate that the simple ESERC model is the poorest model and the DPM is the best. We assessed model accuracy using plots of observed versus predicted values (Fig. 3), the R^2 statistic, and time-series plots of the data versus the 95% prediction intervals of the models (Fig. 4). The base ESERC has a low R^2 value (26%) and the 95% prediction interval includes only a few observations (Fig. 4a). The ADEM explains more than half of the variability ($R^2 = 57\%$), and the 95% prediction interval includes most of the observations (Fig. 4b). The STPM and the DPM explained even more of the variation (R^2 values of 97% and 99%, respectively, Fig. 3); and most of the observations were within the 95% prediction intervals (Fig. 4c, d). The DPM seems superior to the STPM for three reasons: the 95% prediction interval of the DPM is much narrower (Fig. 4c, d), the plot of predictions versus observations (Fig. 3c, d) is closer to a 1:1 line (slope of 0.95 for the DPM, 0.87 for the STPM), and the intercept is smaller (0.01 for the DPM, 0.03 for the STPM).

The DPM is also the best model according to the Deviance Information Criterion (DIC, Table 4). Across the series from ESERC through ADEM to STPM, the effective number of parameters (model complexity) increased and goodness of fit improved (lower $\overline{D(\theta)}$ and higher p_D) at each step (Table 4). However, the DIC actually increased in stepping from the ADEM to the STPM. Although the STPM fits better than the ADEM (lower RMSE, higher R^2), the ADEM has a lower DIC, suggesting that it may be superior because it is more parsimonious (lower p_D). All the metrics agree that the ESERC model is the poorest model and the DPM is the best. The weight of evidence favoring the DPM is overwhelming—the DIC difference between the DIC and the next best model is 45.9, and differences greater than 10 suggest essentially no empirical support for the weaker model (Spiegelhalter et al., 2002; Burnham and Anderson, 2002). The DPM is superior despite the additional complexity introduced by allowing the export coefficients to vary in time and space (highest p_D).

3.2. Spatial and temporal variations of nitrate sources and sinks

We analyzed the results from the best of the alternative models (the DPM) to understand its implications for quantifying nitrate sources and sinks and for understanding spatial and temporal variations in sources and sinks. The model parameters (as estimated by the mean values of their posterior distributions) for in-stream attenuation (k), all cropland (β_c) and unbuffered cropland (β_u) varied greatly among watersheds and among weeks (Fig. 5). Four watersheds (153, 154, 156, 162.5, and 163) have a strong peak in the all cropland export coefficient around week number 40 in the month of January, suggesting a synchronous period of high nitrate export (Fig. 5a). On the other hand, the cropland coefficient can be negative, (especially in watersheds 147, 162.5, and 167) suggesting that cropland sometimes acts as a nitrate sink. The all cropland coefficients of two watersheds (147 and 156) are less variable than the coefficients of other watersheds. Across all the watersheds, values of the all cropland export coefficient are related to stream flow ($R^2 = 0.45$, $P < 0.01$, Fig. 6a), suggesting that almost half of the temporal variability results from variability in weekly stream flow. Much of the spatial variability (differences among watersheds) in all cropland coefficients ($R^2 = 0.73$, $P < 0.01$) and some of the variability in unbuffered cropland coefficients ($R^2 = 0.12$, $P = 0.03$) may be related to the average slope of a watershed (Fig. 7).

The export coefficients of unbuffered cropland (β_u) are larger and less variable than the all cropland coefficients (Fig. 5b). The temporal and spatial patterns of variation in the unbuffered cropland coefficients are similar to the patterns for all cropland coefficients, except in watershed 147, which has the lowest and most temporally variable β_u coefficient (Fig. 5b). Stream flow can explain 36% of temporal variability in β_u (Fig. 6b, $P < 0.01$), and the average watershed slope can explain 73% of spatial variability among watersheds (Fig. 7, $P < 0.01$).

The patterns of variability of in-stream attenuation coefficients (k) are different from the patterns for β_c and β_u (Fig. 5c). The values of k ranged from −0.11 to 1.04 across all watersheds and weeks. Negative values are possible when mineralization exceeds uptake rate so that the stream becomes a source of nitrate (Seitzinger, 1994). Among the nine watersheds, watershed 147 and 153 have the smallest in-stream attenuation coefficients, averaging 0.01 and 0.37 d^{-1}, respectively. Among watersheds, the average of weekly in-stream attenuation coefficients has a statistically significant positive correlation with the average of weekly nitrate concentrations ($R^2 = 14\%$, $P = 0.03$, Fig. 6b).

We used the DPM to predict the time series of stream nitrate concentration for each watershed, then we applied Eq. (13) to quantify the nitrate sources and sinks that lead to the predicted stream concentration at each watershed in each week (Fig. 8a). Among the nitrate sinks, in-stream removal is always small, ranging from 0.01 to 0.27 mg N L^{-1}, while removal in existing buffers ranges to higher values from 0.01 to 1.29 mg N L^{-1}. Among the nine watersheds, watersheds 163 and 167 have the greatest buffer removals (averaging 0.84 and 1.12 mg N L^{-1}, respectively), largely because they have high cropland and much of that is buffered. The potential for creating additional nitrate removal by restoring missing buffers between croplands and streams is also substantial in many watersheds, but generally less than the removal by existing buffers. Watershed 167 has greatest potential for additional removal through buffer restoration (0.50 ± 0.12 mg N L^{-1}). Cropland in watersheds 147 (0.18 ± 0.05 mg N L^{-1}), 163 (0.40 ± 0.11 mg N L^{-1}), and 167 (0.44 ± 0.18 mg N L^{-1}) releases more nitrate than do grassland and developed land. The amounts of nitrate released from grassland and developed land in watersheds 153, 154, 155, and 162.5 are much more than from cropland because of the high amounts of developed land and grassland in these watersheds.

Multiplying the predicted concentrations by the measured weekly water discharge produces estimates of stream nitrate yield (g N ha^{-1} $week^{-1}$). The dominant hydrological signal is the late winter-early spring flow peak that occurs when evapotranspiration is low. The high nitrate discharge also coincides with low temperatures so that nitrate removal processes are slower. Peaks from individual storms are superimposed on this dominant seasonal

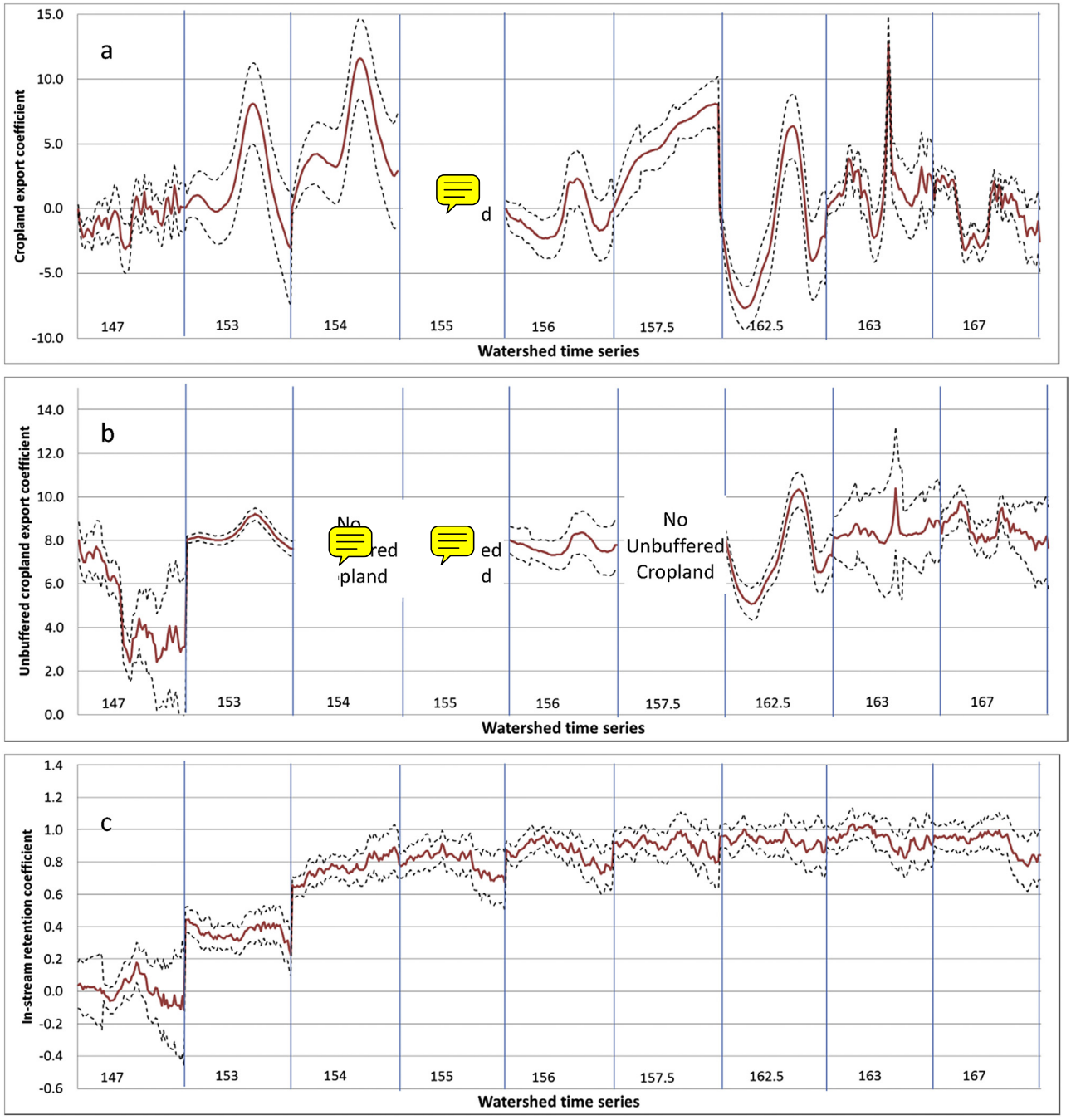

Fig. 5. Time series of the means of the posterior distributions of β_c, β_u, and k from the DPM for each study week in each watershed. Each watershed time series shows 68 weeks of data from April 20, 1998 through August 2, 1999.

pattern. In the summer, nitrate discharge into streams is low (Fig. 8b), because flow is low and riparian buffers have more time to absorb nitrate when water flow rate is lower. Moreover, high temperatures can increase denitrification rates in riparian soils and in riverbeds (Li et al., 2013). In watersheds where non-cropland dominates nitrate concentrations (watershed 153,154, 155, and 157.5), nitrate yields were also reduced by low water discharge. Although nitrate concentration was also relatively low in watershed 162.5, nitrate yield was enhanced by higher water discharge.

4. Discussion

4.1. Performance of alternate Bayesian formulations

In interpreting model results, we explicitly considered two types of model error: the observed variability that is not explained by the model and the uncertainty arising from the model parameters and the possible misspecification of the model structure (Arhonditsis et al., 2008; Stow et al., 2007). We compared alternate Bayesian formulations that can accommodate rigorous and complete error analysis. Presenting the model output as a probability distribution

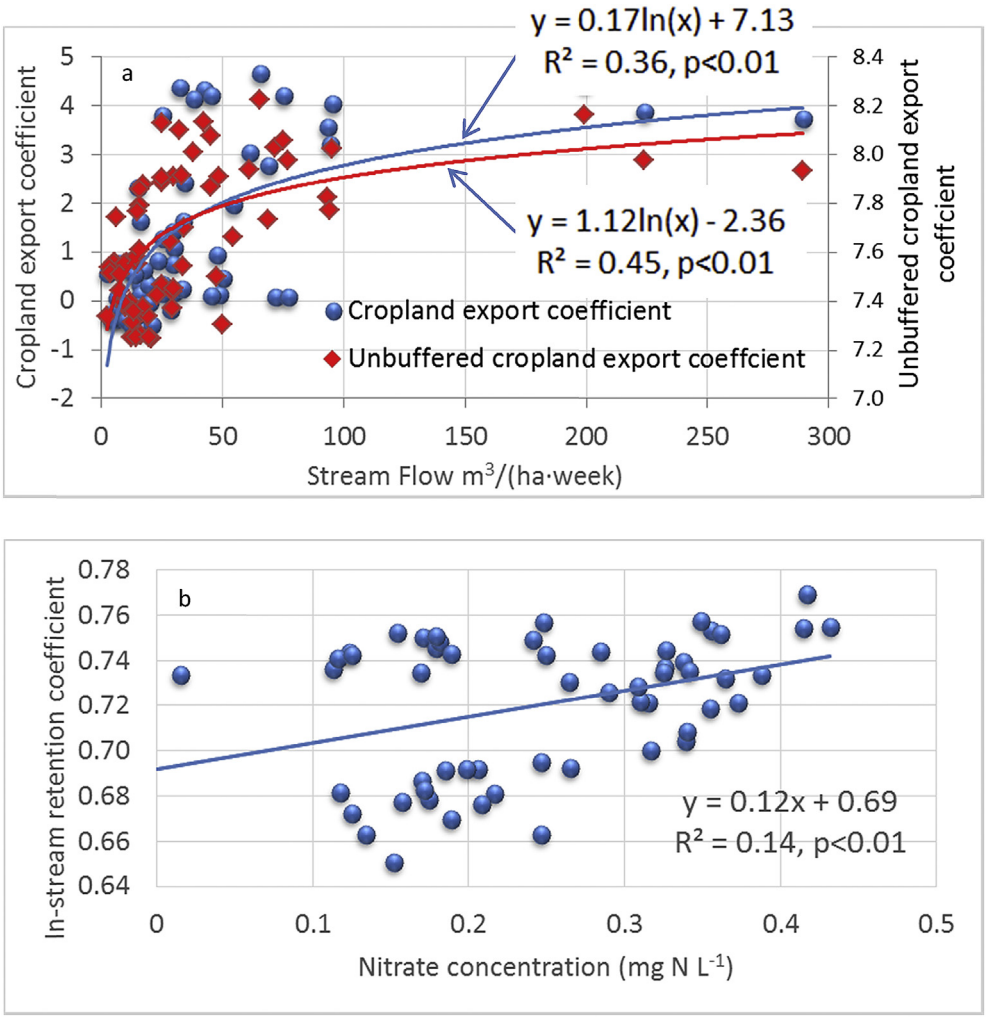

Fig. 6. Relationships of parameters β_c, β_u, and k from the DPM to stream flow and nitrate concentration.

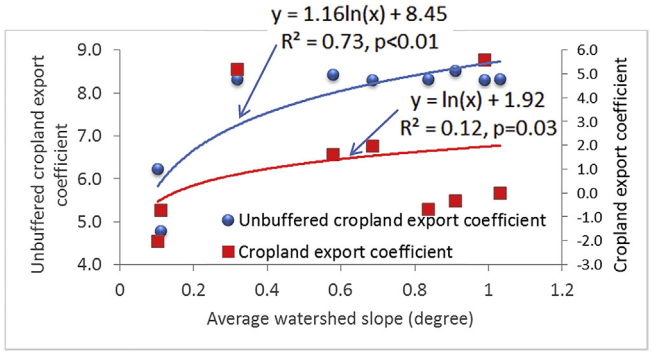

Fig. 7. Relationships of parameters β_c and β_u from the DPM to average watershed slope.

of possible nutrient loads permits the direct estimate of the probability of exceeding any particular load value. Such a probabilistic assessment of water quality conveys significantly more information than a single deterministic prediction. The probabilistic assessment also supports the percentile-based standards proposed by the U. S. Environmental Protection Agency (Office of Water, (1997)).

Different representations of the variability and uncertainty among our models led to different parameter estimates and different confidence intervals for the same underlying conceptual model (Table 3). Other Bayesian analyses have also reported such differences. For example, Arhonditsis et al. (2008) compared three ways of including multiple error sources in eutrophication models and reported that the posterior parameter values can be quite different depending on the assumptions made. Including a temporally-varying discrepancy term represented by a Gaussian first-order random-walk process gave relatively narrow credible intervals and was more informative than the other two ways of accommodating errors. Wellen et al. (2012) compared two approaches to accommodating interannual variability of phosphate loading in the SPARROW model, and they found that DLM techniques had much lower model structural error than a constant model error variance technique. On the other hand, Gronewold et al. (2009) presented three different approaches to modeling variability in two different bacteria water quality models, and found that hierarchical models including an additive noise term did not outperform simple linear models with only an additive noise term.

We used Bayesian analysis to develop probabilistic solutions for four alternative model formulations. The base ESERC model is an effective choice for modeling the effects of riparian buffers on nitrate concentrations (Weller et al., 2011; Weller and Baker, 2014). However, the ESERC model does not capture spatial variation in model parameters among watersheds or temporal variation among

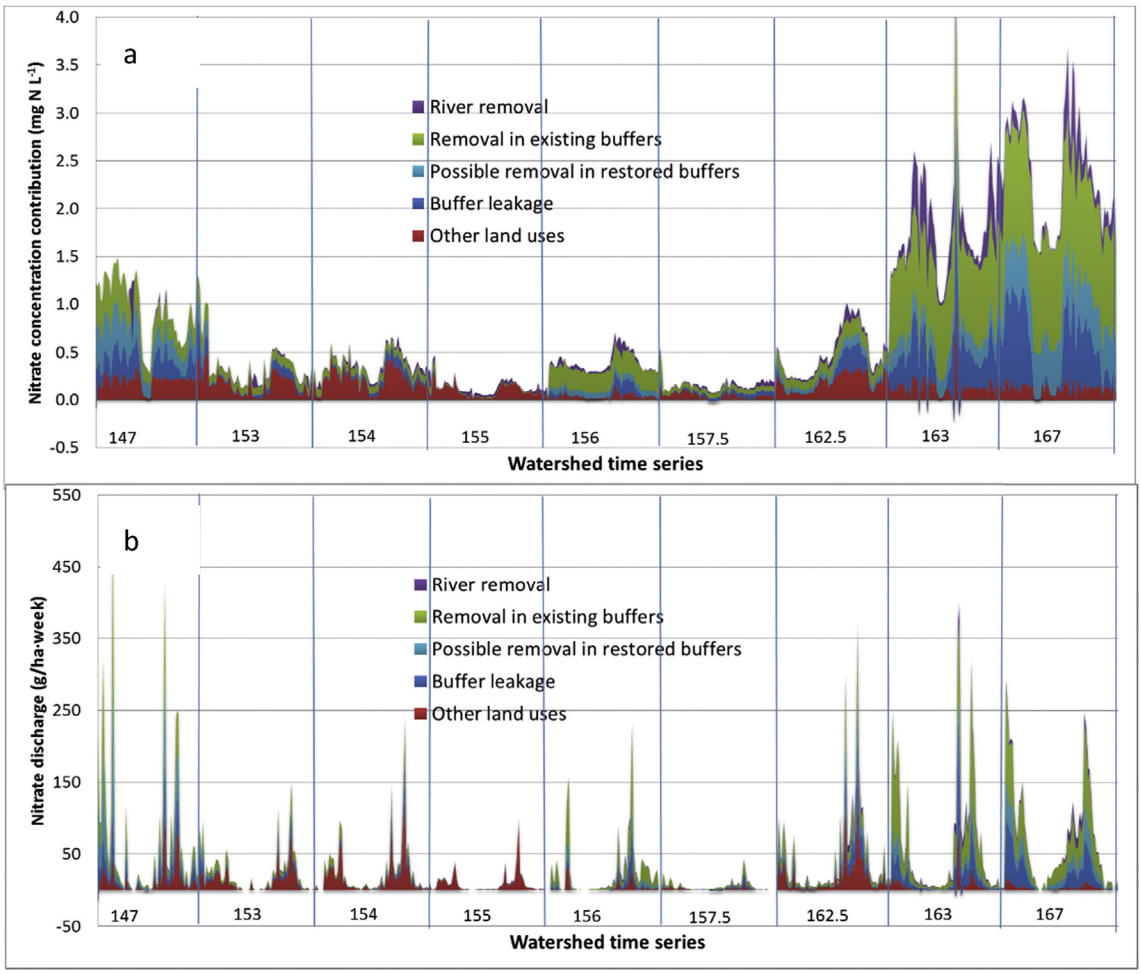

Fig. 8. Time series of the components of predicted nitrate sources and sinks lead to the stream nitrate a) concentration and b) discharge at each watershed in each week. Each watershed time series shows 68 weeks of data from April 20, 1998 through August 2, 1999.

weeks of observation. Compared to the ESERC model, the ADEM achieved higher model accuracy by adding spatial and temporal variability terms to accommodate structural error in the model. It was not the best the method in this study, probably because structural error is not the main cause of uncertainty in the ESERC model. Structural uncertainty also accounts for a relatively small component of uncertainty in some hydrological models (Huard and Mailhot, 2006). The strong performance of the STPM suggests that parameter uncertainty is the most important source of uncertainty of the ESERC model. The STPM performed well by representing spatial and temporal variations in parameters as the main source of uncertainty, as reported for other models of non-point source pollution (Lu et al., 2013; Zobrist and Reichert, 2006). The assumed correlations among parameters in the STPM foster computational efficiency, so solution times would be much less if the parameters were independently distributed (Sadraddini et al., 2011; Wellen et al., 2012).

The DPM emerged as the best model because it realistically characterized uncertainty and achieved high predictive skill. The Gaussian first-order random walk model well represents correlations among the error terms of spatial parameters, and its application is conceptually similar to a hierarchical DLM (Lamon et al., 1998a) or a Kalman filter (Meinhold and Singpurwalla, 1983). The DLM model with a discount factor explicitly recognizes temporal correlation within the time series of parameters. Gaussian first-order random walk models and data-driven priors have been used previously to characterize ecological process (Arhonditsis et al., 2008; Sadraddini et al., 2011; Wellen et al., 2012). Our study is the first study to combine the random walk and DLM to simultaneously characterize spatial and temporal parameter variability in a water quality model.

4.2. New understanding for nitrate sources and sinks

We compared our DPM results with published results from the original SERC model in the Coastal Plain province (Weller et al., 2011; Weller and Baker, 2014). We found similar strong effects of cropland and riparian buffers on stream nitrate concentrations. The DPM estimates of developed and grass land exports are low but slightly higher than the SERC model. This difference may arise because the SERC model did not accommodate differences in coefficients among watersheds within a physiographic province and used one temporally averaged nitrate concentration value for each watershed rather than the weekly nitrate time series (Weller et al., 2011). Our study extended the understanding from the SERC model by quantifying spatial and temporal variations in nitrate sources and sinks (Fig. 8).

We found that the nitrate export coefficients for all cropland and unbuffered cropland rise steeply with increasing stream flow at low flow rates, but level off at higher flow rates (Fig. 6). Pionke et al.

(1996) report a similar pattern for other agricultural watersheds in the Chesapeake drainage. Our results indicate that buffer removal increased at low stream flow rates and decreased during high flows, probably because of buffer retention is less effective during the high flows of storm events. We also found statistically significant relationships between watershed average slope and the coefficients for nitrate export from all cropland and from unbuffered cropland (Fig. 6). Higher slopes may result in higher subsurface flow rates and in greater amounts of overland flow, which could both reduce opportunities for uptake and denitrification (Preston and Brakebill, 1999). In-stream attenuation coefficients were quite variable in space and time (Fig. 5) Stream denitrification, which is the main mechanism of in-stream attenuation, is known to vary with temperature and with the exchange between the water column and sediment (Zhao et al., 2015; Burgin and Hamilton, 2007). We found a statistically significant positive correlation between in-stream attenuation and nitrate concentration (Fig. 7), which is consistent with the kinetics of denitrification (Seitzinger, 1988) and with tracer experiments reporting increasing denitrification with higher nitrate concentration (Mulholland et al., 2008).

4.3. Directions for future model development and application

We focused initially on cropland because it is the dominant source of nitrate to streams (Weller et al., 2003), but future applications could explicitly model spatial and temporal variations in the developed land coefficient. Such an analysis should include more watersheds with higher proportions of developed land than in our proof of concept study here.

We found statistically significant relationships in *post hoc* analyses relating model parameters to environmental factors like watershed slope, stream flow, and nitrate concentration. Future models could seek to directly incorporate those environmental variables into the Bayesian hierarchical framework as predictors of model parameters.

Our analysis quantified prediction uncertainty within the calibration dataset (Fig. 4), but it is also important to understand how models perform in predicting novel cases outside the training data. A recent study reports that current approaches to quantifying predictive model uncertainty remain unsatisfying, but cross validation or adjustments based on the Watanabe-Akaike information criterion (WAIC) have some value (Gelman et al., 2014). Future analysis applying our models (or other Bayesian loading models) in predictive mode should consider applying these methods.

Finally, we must acknowledge that Bayesian methods are difficult to learn and are less accessible to many researchers and environmental managers. Widespread adoption of Bayesian modeling in water quality management may require training opportunities and the development of more user-friendly tools for developing and applying Bayesian models.

4.4. Management implications

We have demonstrated that including parameter uncertainty greatly improves model performance; so we suggest that temporally and spatially variable parameters should be included in export coefficient models used for management decisions. Our proof of concept study demonstrated how our best model (DPM) can provide knowledge to help guide nitrogen management for reducing stream nitrate levels and nitrate delivery to receiving waters like Chesapeake Bay. Five of the study watersheds (147, 156, 162.5, 163, and 167) have relatively high cropland export coefficients (Fig. 5), so reducing N input in these watersheds will have greater leverage to reduce overall nitrate delivery.

Our results can also help guide buffer management. Previously, (Weller et al., 2011; Weller and Baker, 2014) compared buffer effects on annual average N discharge from watersheds in different physiographic provinces and suggested that widespread restoration of Coastal Plain buffers could result in significant reductions in N discharge. Our new results reveal important differences in buffer effects among watersheds within the Coastal Plain. Five watersheds (147, 156, 162.5, 163 and 167) were very efficient at absorbing nitrate (Fig. 6a) due to the effects of buffers. Watershed 147 and 167 also have the greatest potential for further improvement with buffer restoration, suggesting that buffer restoration there should have high priority. However, buffer restoration in four watersheds (153, 154, 155, and 157.5) should have lower priority because of low nitrate delivery, buffer capacity, and lack of cropland (Fig. 8).

5. Conclusion

We developed and tested a Bayesian hierarchical framework that accounts for uncertainties in model structure and parameters and quantifies how they affect the prediction uncertainty of export coefficient models. The Dynamic Parameter Model (DPM), which incorporates a first-order random walk to represent spatial correlation among parameters and a dynamic linear model to accommodate temporal correlation, was the best of the alternative models we tested The DPM provided a far better fit, much narrower prediction intervals, and a much more effective tradeoff between fit and complexity than the alternative models we tested. The DPM resolves the problems with the basic export coefficient that we sought to address. We also demonstrated that incorporating temporally and spatially variability in model parameters can improve understanding of the nitrate sources and sinks. In particular, the variability in cropland export coefficients was logically related to stream flow and watershed slope, while instream retention coefficients were related to nitrate concentrations. In management applications, the models can help identify which watersheds should have priority for buffer restoration efforts and which watersheds and times are the strongest sources of nitrate exports. The methods developed here can be broadly applied to other materials and other watersheds where pollutant sources and sinks vary spatially and temporally.

Acknowledgements

The measurements of watershed discharges were supported by a contract from Charles County, Maryland, USA and by a grant from the U. S. National Science Foundation (DEB-93-17968). The model comparison was financially supported by the National Natural Science Foundation of China through Grant (41471238), and Young Talents in State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science (No. Y212000013). We thank the China Scholarship Council (No. 201404910320) for supporting study abroad.

References

Ajami, N.K., Duan, Q., Sorooshian, S., 2007. An integrated hydrologic Bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resour. Res. 43 (1).

Alexander, R.B., Smith, R.A., Schwarz, G.E., 2006. Comment on “in-stream nitrogen attenuation: model-aggregation effects and implications for coastal nitrogen impacts”. Environ. Sci. Technol. 40 (7), 2485–2486.

Arhonditsis, G.B., Perhar, G., Zhang, W., Massos, E., Shi, M., Das, A., 2008. Addressing equifinality and uncertainty in eutrophication models. Water Resour. Res. 44 (1).

Baker, M.E., Weller, D.E., Jordan, T.E., 2006a. Improved methods for quantifying potential nutrient interception by riparian buffers. Landsc. Ecol. 21 (8), 1327–1345.

Baker, M.E., Weller, D.E., Jordan, T.E., 2006b. Comparison of automated watershed delineations. Photogrammetric Eng. Remote Sens. 72 (2), 159–168.

Boomer, K., Weller, D.E., Jordan, T.E., Linker, L., Liu, Z.J., Reilly, J., Shenk, G., Voinov, A.A., 2013. Using Multiple Watershed Models to Predict Water, Nitrogen, and Phosphorus Discharges to the Patuxent Estuary1. Wiley Online Library.
Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Statistics 7 (4), 434–455.
Burgin, A.J., Hamilton, S.K., 2007. Have we overemphasized the role of denitrification in aquatic ecosystems? A review of nitrate removal pathways. Front. Ecol. Environ. 5 (2), 89–96.
Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: a Practical Information-theoretic Approach. Springer Science & Business Media.
Congdon, P., 2007. Bayesian Statistical Modelling. John Wiley & Sons.
Endreny, T.A., Wood, E.F., 2003. Watershed weighting of export coefficients to map critical phosphorous loading areas. J. Am. Water Resour. Assoc. 39 (1), 165–181.
ESRI, 2011. ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.
Galloway, J.N., Townsend, A.R., Erisman, J.W., Bekunda, M., Cai, Z., Freney, J.R., Martinelli, L.A., Seitzinger, S.P., Sutton, M.A., 2008. Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. Science 320 (5878), 889–892.
Gao, C., Zhu, J., Zhu, J., Gao, X., Dou, Y., Hosen, Y., 2004. Nitrogen export from an agriculture watershed in the Taihu Lake area, China. Environ. Geochem. Health 26 (2), 199–207.
Gardner, K.K., McGlynn, B.L., Marshall, L.A., 2011. Quantifying watershed sensitivity to spatially variable N loading and the relative importance of watershed N retention mechanisms. Water Resour. Res. 47 (8).
Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. Statistics Comput. 24, 997–1016.
Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., Tyler, D., 2002. The national elevation dataset. Photogrammetric Eng. Remote Sens. 68 (1), 5–32.
Gilks, W.R., 2005. Markov Chain Monte Carlo. Wiley Online Library.
Gronewold, A.D., Qian, S.S., Wolpert, R.L., Reckhow, K.H., 2009. Calibrating and validating bacterial water quality models: a Bayesian approach. Water Res. 43 (10), 2688–2698.
Huard, D., Mailhot, A., 2006. A Bayesian perspective on input uncertainty in model calibration: application to hydrological model “abc”. Water Resour. Res. 42 (7).
Jobson, H.E., 1996. Prediction of Traveltime and Longitudinal Dispersion in Rivers and Streams. Citeseer.
Johnes, P.J., 1996. Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach. J. Hydrology 183 (3), 323–349.
Jordan, T.E., Correll, D.L., Weller, D.E., 2000. Mattawoman Creek Watershed Nutrient and Sediment Dynamics: Final Contract Report to Charles County, Maryland. Smithsonian Environmental Research Center, Edgewater, MD, pp. 1–73.
Jordan, T.E., Correll, D.L., Weller, D.E., 1997a. Relating nutrient discharges from watersheds to land use and streamflow variability. Water Resour. Res. 33 (11), 2579–2590.
Jordan, T.E., Correll, D.L., Weller, D.E., 1997b. Effects of agriculture on discharges of nutrients from coastal plain watersheds of Chesapeake Bay. J. Environ. Qual. 26 (3), 836–848.
Jordan, T.E., Correll, D.L., Weller, D.E., 1997c. Noinpoint source discharges of nutrients from pidmnont watersheds of Chesapeake Bay. JAWRA J. Am. Water Resour. Assoc. 33 (3), 631–645.
Khadam, I.M., Kaluarachchi, J.J., 2006. Water quality modeling under hydrologic variability and parameter uncertainty using erosion-scaled export coefficients. J. Hydrology 330 (1), 354–367.
Lagazio, C., Dreassi, E., Biggeri, A., 2001. A hierarchical Bayesian model for space-time variation of disease risk. Stat. Model. 1 (1), 17–29.
Lamon III, E.C., Carpenter, S., Stow, C., 1998a. Forecasting pcb concentrations in lake Michigan salmonids: a dynamic linear model approach. Ecol. Appl. 8 (3), 659–668.
Lamon, E.C., Carpenter, S.R., Stow, C.A., 1998b. Forecasting pcb concentrations in lake Michigan salmonids: a dynamic linear model approach. Ecol. Appl. 8 (3), 659–668.
Li, X., Xia, Y., Li, Y., Kana, T.M., Kimura, S.D., Saito, M., Yan, X., 2013. Sediment denitrification in waterways in a rice-paddy-dominated watershed in eastern China. J. Soils Sediments 13, 783–792.
Liu, Z.J., Weller, D.E., Correll, D.L., Jordan, T.E., 2000. Effects of land cover and geology on stream chemistry in watersheds of Chesapeake Bay. J. Am. Water Resour. Assoc. 36 (6), 1349–1366.
Lu, J., Gong, D., Shen, Y., Liu, M., Chen, D., 2013. An inversed Bayesian modeling approach for estimating nitrogen export coefficients and uncertainty assessment in an agricultural watershed in eastern China. Agric. Water Manag. 116, 79–88.
Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. Statistics Comput. 10 (4), 325–337.
Maillard, P., Santos, N.A.P., 2008. A spatial-statistical approach for modeling the effect of non-point source pollution on different water quality parameters in the Velhas river watershed–Brazil. J. Environ. Manag. 86 (1), 158–170.
McClain, M.E., Boyer, E.W., Dent, C.L., Gergel, S.E., Grimm, N.B., Groffman, P.M., Hart, S.C., Harvey, J.W., Johnston, C.A., Mayorga, E., McDowell, W.H., 2003. Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. Ecosystems 6 (4), 301–312.
Meinhold, R.J., Singpurwalla, N.D., 1983. Understanding the kalman filter. Am. Statistician 37 (2), 123–127.
Mulholland, P.J., Helton, A.M., Poole, G.C., Hall, R.O., Hamilton, S.K., Peterson, B.J., Dodds, W.K., 2008. Stream denitrification across biomes and its response to anthropogenic nitrate loading. Nature 452 (7184), 202–205.
Norvell, W., Frink, C., Hill, D., 1979. Phosphorus in Connecticut lakes predicted by land use. Proc. Natl. Acad. Sci. 76 (11), 5426–5429.
Obenour, D.R., Gronewold, A.D., Stow, C.A., Scavia, D., 2014. Using a Bayesian hierarchical model to improve Lake Erie cyanobacteria bloom forecasts. Water Resour. Res. 50 (10), 7847–7860.
O'Callaghan, J.F., Mark, D.M., 1984. The extraction of drainage networks from digital elevation data. Comput. Vis. Graph. image Process. 28 (3), 323–344.
Office of Water, 1997. Guidelines for Preparation of the Comprehensive State Water Quality Assessments. Environmental Protection Agency, U.S. Washington, DC.
Pionke, H., Gburek, W., Sharpley, A., Schnabel, R., 1996. Flow and nutrient export patterns for an agricultural hill-land watershed. Water Resour. Res. 32 (6), 1795–1804.
Preston, S.D., Brakebill, J.W., 1999. Application of Spatially Referenced Regression Modeling for the Evaluation of Total Nitrogen Loading in the Chesapeake Bay Watershed. USGS.
Sadraddini, S., Azim, M.E., Shimoda, Y., Mahmood, M., Bhavsar, S.P., Backus, S.M., Arhonditsis, G.B., 2011. Temporal PCB and mercury trends in Lake Erie fish communities: a dynamic linear modeling analysis. Ecotoxicol. Environ. Saf. 74 (8), 2203–2214.
Seitzinger, S.P., 1988. Denitrification in freshwater and coastal marine ecosystems: ecological and geochemical significance. Limnol. Oceanogr. 33 (2), 702–724.
Seitzinger, S.P., 1994. Linkages between organic matter mineralization and denitrification in eight riparian wetlands. Biogeochemistry 25 (1), 19–39.
Shaddick, G., Wakefield, J., 2002. Modelling daily multivariate pollutant data at multiple sites. J. R. Stat. Soc. Ser. C Appl. Statistics) 51 (3), 351–372.
Smith, R.A., Schwarz, G.E., Alexander, R.B., 1997. Regional interpretation of water-quality monitoring data. Water Resour. Res. 33 (12), 2781–2798.
Simley, J.D., Carswell, W.J., 2009. The National Map-hydrography:fact Sheet 2009-3054. Geological Survey, Washington, D.C.
Spiegelhalter, D.J., Best, N.G., Carlin, B.R., van der Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. Ser. B-Statistical Methodol. 64, 583–616.
Stow, C.A., Reckhow, K.H., Qian, S.S., Lamon, E.C., Arhonditsis, G.B., Borsuk, M.E., Seo, D., 2007. Approaches to evaluate water quality model parameter uncertainty for adaptive tmdl implementation. J. Am. Water Resour. As 43 (6), 1499–1507.
USEPA, 2007. Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, PM2:5, and Regional Haze. Final Report, EPA-454/B-07–002. US Environmental Protection Agency.
Valett, H.M., Morrice, J.A., Dahm, C.N., Campana, M.E., 1996. Parent lithology, surface–groundwater exchange, and nitrate retention in headwater streams. Limnol. Oceanogr. 41 (2), 333–345.
Vassiljev, A., Blinova, I., Ennet, P., 2008. Source apportionment of nutrients in Estonian rivers. Desalination 226 (1), 222–230.
Vitousek, P.M., Aber, J.D., Howarth, R.W., Likens, G.E., Matson, P.A., Schindler, D.W., Schlesinger, W.H., Tilman, D.G., 1997. Human alteration of the global nitrogen cycle: sources and consequences. Ecol. Appl. 7 (3), 737–750.
Vogelmann, J., Sohl, T., Howard, S., 1998b. Regional characterization of land cover using multiple sources of data. Photogrammetric Eng. Remote Sens. 64 (1), 45–57.
Vogelmann, J., Sohl, T., Campbell, P., Shaw, D., 1998a. Regional land cover characterization using Landsat Thematic Mapper data and ancillary data sources. Environ. Monit. Assess. 51 (1–2), 415–428.
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. 10 (48).
Weller, D.E., Baker, M.E., 2014. Cropland riparian buffers throughout Chesapeake Bay watershed: spatial patterns and effects on nitrate loads delivered to streams. JAWRA J. Am. Water Resour. Assoc. 50 (3), 696–712.
Weller, D.E., Baker, M.E., Jordan, T.E., 2011. Effects of riparian buffers on nitrate concentrations in watershed discharges: new models and management implications. Ecol. Appl. 21 (5), 1679–1695.
Weller, D.E., Jordan, T.E., Correll, D.L., Liu, Z.J., 2003. Effects of land-use change on nutrient discharges from the Patuxent River watershed. Estuaries 26 (2), 244–266.
Zhao, Y., Xia, Y., Ti, C., Shan, J., Li, B., Xia, L., Yan, X., 2015. Nitrogen removal capacity of the river network in a high nitrogen loading region. Environ. Sci. Technol. 49 (3), 1427–1435.
Zobrist, J., Reichert, P., 2006. Bayesian estimation of export coefficients from diffuse and point sources in Swiss watersheds. J. Hydrology 329 (1), 207–223.