Environmental Modelling & Software 69 (2015) 14–22

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

A continuous variable Bayesian networks model for water quality modeling: A case study of setting nitrogen criterion for small rivers and streams in Ohio, USA

Song S. Qian [a, *], Robert J. Miltner [b]

[a] Department of Environmental Sciences, University of Toledo, 2801 W. Bancroft Street, MS# 604, Toledo, OH 43606-3390, USA
[b] Ohio Environmental Protection Agency, 4675 Homer-Ohio Lane, Groveport, OH 43125, USA

ARTICLE INFO

Article history:
Received 2 March 2015
Accepted 2 March 2015
Available online

Keywords:
Bayesian statistics
Nutrient criteria
Biological monitoring
Simulation
Ohio
Clean Water Act
Updating

ABSTRACT

We present a continuous variable Bayesian networks modeling framework that integrates the graphical representation of a Bayesian networks model with empirical model-developing approach. Our model retains the Bayesian networks model's graphical representation of hypothesized causal connections among important variables and employs conventional statistical modeling approaches for establishing functional relationships among these variables. The modeling framework avoids discretizing continuous variables and the resulting models can be updated over time when new data are available or updated using local data to develop a site-specific model. We illustrate the modeling approach using a data for establishing nutrient criteria in streams and rivers in Ohio, U.S.A.

1. Introduction

The Bayesian networks (BN) modeling approach is increasingly used in environmental modeling for supporting management decisions (see Chen and Pollino (2012) and Aguilera et al. (2011) for reviews). One of the most appealing feature of a BN model is likely its graphical approach for illustrating complicated connections among multiple components of a problem thereby facilitating the communication of scientific research to a wide range of stakeholders. Furthermore, the relatively simple graphical structure allows participation of stakeholders in model developments.

An oft-downplayed drawback of the BN approach is the need for discretizing continuous variables. Discretization is a problem not only because of the potential loss of statistical accuracy (Chen and Pollino, 2012), but also because of the difficulty in subsequent interpretation. The difficulty arises because the discretized variable is often categorized as, for example, *low*, *medium*, and *high*. Because there is no consensus on how to properly discretize a continuous variable, the resulting categories are often ambiguous and their meanings depend on the context. For example, Kashuba (2010) developed a BN model for assessing stream ecosystem status in southeast U.S. Using the equal frequency method, the variable defining watershed urban intensity (% developed land) is divided into four bins (*low* 0–11%, *medium low* 12–36%, *medium high* 37–65%, and *high* 66–100%). This division would likely be meaningless if applied to a different region. For example, almost all headwater stream watersheds in Maine, U.S.A. will fall into the category "low" (0–11%) (Susan Davies, 2009, personal communication). In developing a eutrophication BN model for an estuary in North Carolina, Nojavan et al. (2014) divided the variable total nitrogen concentration into three bins (<56, 56–334, and >334 μg/L) based on data from 2007 to 2011. Because the water quality of their study area has been improving since the mid-1990s, the category "low" would include more than half of the observations collected since 2012 (Nojavan, 2014).

Furthermore, how a continuous variable is discretized (discretization methods and number of bins) is directly linked to the subsequent models. Discretization methods and number of bins can change a BN model's structure when using structural learning algorithms (Alameddine et al., 2011), as well as the conditional probability tables when using a fixed model structure (Nojavan, 2014).

* Corresponding author. Tel.: +1 419 530 4230; fax: +1 419 530 4421.
E-mail addresses: song.qian@utoledo.edu (S.S. Qian), bob.miltner@epa.state.oh.us (R.J. Miltner).

http://dx.doi.org/10.1016/j.envsoft.2015.03.001
1364-8152/

Although recognized by many, the problem of discretization is often unmentioned in many applications of BN in the environmental modeling literature (Aguilera et al., 2011). In this paper, we present a framework for developing a BN model without discretizing continuous variables. Our aim is to develop a probabilistic modeling framework that can be used for supporting management and decision-making under uncertainty, specifically, for making operational decisions (Kelly et al., 2013). The approach borrows the basic model-construction strategy used in a BN model (Jensen, 2001), and expands the model specification and fitting using traditional data exploration and regression (Weisberg, 2005; Gelman and Hill, 2007; Qian, 2010) and the model updating using Bayesian computation [Markov chain Monte Carlo simulation, (Qian et al., 2003)]. The model structure is developed through a directed acyclic diagram (DAG) model (Lauritzen, 1996), with connections among nodes represented by empirical models.

We present the general framework of developing a continuous variable Bayesian networks (cBN) model in Section 2.1 and apply it to a data collected for establishing nitrogen criterion for Ohio's small rivers and streams in Section 2.2. We also discuss the process of updating the resulting model using local or regional data for developing local and regional nutrient criteria.

2. Methods

In a traditional environmental modeling domain, we often discuss the trade-off between a mechanistic model and an empirical model. A mechanistic model is a summary of the functional connections among multiple components of a real world problem, reflecting the causal relationship we know. These models are, in theory, suited for supporting management decision-making. But a mechanistic model is often overly complicated and may not be practical for proper model calibration because of the limited availability of appropriate data. Empirical or statistical models establish correlations, which do not necessarily reflect causal relations. Models such as DAG models represent a middle ground between the two approaches. The hypothetical relationships among relevant variables are presented using a graphical model and the functional forms of these relationships (e.g., differential equations) are represented using conditional probability tables. As a result, fast computing algorithms can be applied. However, as noted earlier, discretization can be a source of many problems. We propose a modeling framework to avoid discretization and yet retain the advantages of a BN model.

2.1. A continuous variable Bayesian network model

Graphical models such as the BN models (Pearl, 1986, 1988; Jensen, 2001) take full advantage of the conditional probability structure, not only in model formulation, but also in computation. However, a BN model is limited to using categorical variables. When continuous variables must be used, we either discretize them (BN) or we assume that the continuous variables are normal random variates and connections among nodes are linear [the structural equations model (Bollen, 1989; Grace, 2006)]. With the advent of the Gibbs Sampler (Gelfand and Smith, 1990), the computation requirements are less restrictive. The Bayesian inference software WinBUGS (Gilks et al., 1994; Lunn et al., 2000) (now OpenBUGS (Spiegelhalter et al., 2014)) and JAGS (Plummer, 2003) further popularize the use of Gibbs sampler. We can now conduct the same complex computing with continuous variables without using discretized conditional probability models, without assuming normality, and without being limited to linear models. Consequently, we can take advantage of both the causal relationship (the DAG model) and the available data to build a Bayesian network model using continuous variables.

The basic idea of our modeling approach is to replace the conditional probability tables (CPTs) for factor variables in a BN model with a series of conditional probability distributions for continuous variables. With CPTs, we use the conditional probability formula (the Bayes theorem) to quantify the functional relations among variables and the formula can be readily programmed. When using continuous conditional probability distributions, the computation can be implemented using the Gibbs sampler, or more generally Markov chain Monte Carlo (MCMC) simulation (Gilks et al., 1996; Qian et al., 2003).

We describe the computational strategy using a hypothetical model (Fig. 1(a)). In this simple DAG, two nodes have parent nodes (nodes X_1 and X_2 are parent nodes to Y_1 and nodes Y_1 and X_3 are parents to Y_2). Arrows in the DAG represent the directional dependency among the five variables (e.g., Y_1 is a function of X_1 and X_2). If we consider a DAG such as the one in Fig. 1(a) as a joint probabilistic model of the variables in all nodes, the DAG provides a causal structure which can be translated into conditional probability distributions such that the joint distribution can be simplified. Let V be the collection of all variables represented by the DAG ($V = \{X_1,X_2,X_3,Y_1,Y_2\}$), the joint distribution of V can be represented as $p(V) = \prod_{v\in V} p(v|\text{parents}[v])$, where $p(\cdot)$ represents a probability distribution function. For the model in Fig. 1(a), the joint distribution is $p(V) = p(X_1)p(X_2)p(Y_1|X_1,X_2)p(X_3)p(Y_2|Y_1,X_3)$. When all variables are categorical, these conditional probability distributions are represented by conditional probability tables, and the multiplication operation is done through the Bayes theorem.

When these variables are continuous, we can specify the marginal distributions of X_1,X_2 and X_3 directly (e.g., histograms of data) and the conditional distributions of $Y_1|X_1,X_2$ and $Y_2|Y_1,X_3$ empirically, using perhaps linear or nonlinear regression analysis. For example, if a regression model is used, we may denote the conditional distribution as $Y_1 = f(X_1,X_2,\alpha) + \varepsilon_1$, or $Y_1|X_1,X_2 \sim N(\mu_1,\sigma_1^2)$, where $f(\cdot)$ denotes a linear or nonlinear function of X_1 and X_2, $\mu_1 = f(X_1,X_2,\alpha)$ with coefficients α, and $\varepsilon_1 \sim N(0,\sigma_1^2)$ is the residual random variable. Likewise, exploratory analysis can lead to the conditional distribution of $Y_2|X_3,Y_1 \sim N(\mu_2,\sigma_2^2)$, where $\mu_2 = g(X_3,Y_1,\beta)$. Furthermore, variables in V can be divided into predictors (X's, observed without error) and response variables (Y's, observed with error). When data are available for all variables, the model is reduced to a problem of estimating the probability distribution of model coefficients (α and β) and error variances (σ_1^2,σ_2^2), and the graphical model in Fig. 1(a) is revised to a DAG representing the computational process (Fig. 1(b), where oval nodes are quantities to be estimated, rectangle nodes are variables with observations, and arrows representing the functional dependency). For example, the data node y_1 is a child node of μ_1 and σ_1^2, representing the distributional assumption $y_1 \sim N(\mu_1,\sigma_1^2)$; the parameter node μ_2 has three parent nodes — μ_1,x_3, and model coefficient vector β. If we use a linear regression model, the node μ_2 represents the linear model mean function $\mu_2 = \beta_0 + \beta_1\mu_1 + \beta_2 x_3$.

We note that variable Y_1 in Fig. 1(a) is a parent to Y_2, while in Fig. 1(b) the connection between the two variables are established through their respective means μ_1 and μ_2. In other words, we replace the CPTs in a BN model with functions for calculating μ_1 and μ_2 (f and g in Fig. 1(b)).

Just as eliciting CPTs is an important step of building a BN model, finding the likely functional forms of f and g is an important component for our modeling framework. We can derive the functional forms based on our substantive knowledge (e.g., mechanistic models) or based on empirical modeling through exploratory regression analysis (see Chapter 4 of Gelman and Hill (2007)). When using the empirical modeling approach, we use the DAG as a guide for building component models (finding the likely functional forms of f and g) one at a time. Once these functional forms are established, they can be linked to form the joint distribution of all variables. The resulting model is a continuous variable Bayesian network model (cBN). Model parameters (e.g., α,β in Fig. 1) should be estimated simultaneously using the Gibbs sampler. Once the joint distribution is quantified (all unknown parameters are estimated), statistical inference can be made through Monte Carlo simulations.

Borsuk et al. (2004) proposed a similar modeling approach. However, their model was implemented in Analytica (Lumina, 1997), which requires fixed model coefficients. As a result, conditional models (e.g., $p(Y_1|X_1,X_2)$ are fit independently and cannot be updated upon new observations (nor refit jointly).

2.2. The Ohio example

We illustrate the model-building process using data from wadeable streams and rivers in Ohio collected as part of the effort for setting nutrient criteria. The data is "cross-sectional" in that they represent multiple streams and rivers across the state of Ohio. The objective of the model is to establish a link between nutrient concentration and indicators of stream aquatic ecosystem condition. Through the link, we find the nutrient concentration distribution associated with an aquatic ecosystem that is likely to meet the designated use. Our approach follows the following steps:

1. Developing a conceptual model linking nutrient concentrations and other factors to stream aquatic ecosystems indicators (e.g., Fig. 1(a)),
2. Building empirical (e.g., regression) models among nodes (if such functional relationships are unknown) using available data and expressing these models in terms of conditional probability distributions (probability distribution of a child node conditional on its parent node(s)),
3. Revising the initial graphical model to connect data and unknown parameters (e.g., Fig. 1 (b)), and
4. Estimating all unknown parameters of the joint probabilistic distribution using the Gibbs sampler.

Once the joint distribution of all relevant variables (nodes) are quantified, we can derive the (conditional) distribution of nutrient concentration that is associated with acceptable stream ecosystem indicator values. This conditional distribution can be used to define acceptable nutrient concentrations.

Data collection and the initial nutrient criteria development process are reported by Miltner (2010). The objective of a nutrient criterion is to ensure that streams meet the designated use for aquatic life, which is measured by one or more macroinvertebrate metrics in Ohio. We use the Invertebrate Community Index (ICI) and the EPT taxa richness [EPT, number of Ephemeroptera (mayfly), Plecoptera (stonefly), and Trichopera (caddisfly) taxa in a sample (Ohio EPA, 1978)] in our illustration. Because stream macroinvertebrate community are affected by many other factors (e.g., stream flow, habitat condition, watershed land use, shading, etc.)

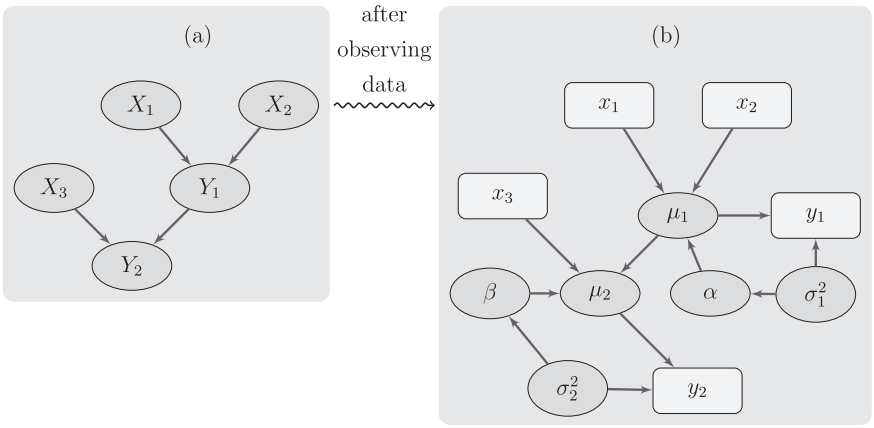

Fig. 1. (a). A hypothetical DAG model shows the causal relationship among five variables. (b). When data from the five variables are available, we may first divide the variables into predictor variables (X's, observed accurately) and response variables (Y's, observed with errors). The revised DAG includes the functional relationships $\mu_1 = f(x_1,x_2,\alpha), y_1 \sim N(\mu_1,\sigma_1^2)$ and $\mu_2 = g(x_3,\mu_1,\beta), y_2 \sim N(\mu_2,\sigma_2^2)$.

in addition to nutrient enrichment, the relationships between nutrient concentration and macroinvertebrate metrics are often noisy (Fig. 2).

This can be a result of (1) a temporal scale mismatch between nutrient concentration data (representing a snapshot in time) and macroinvertebrate metrics (representing an aggregation over the life time of organisms used, weeks or months) and (2) large among site/waterbody variances. As a result, directly building a dose (nitrogen concentration) – response (macroinvertebrate metrics) model is often difficult. Miltner (2010) circumvented the problem by exploring intermediate relationships (e.g., nutrient to benthic chlorophyll *a*, benthic chlorophyll *a* to DO, and DO to ICI and/or EPT) observed from synoptically collected samples of nutrient concentrations (TP and DIN), benthic chlorophyll, hourly dissolved oxygen concentrations, and benthic macroinvertebrates obtained from small rivers and streams (second to fourth order streams with drainage area of ~10–1000 km^2). The modeling approach used by Miltner (2010) can be represented by a graphical model (Fig. 3).

2.3. Probabilistic models

The modeling approach of Miltner (2010) used four independent regression models. The first is a regression model for predicting benthic chlorophyll *a*, where dissolved inorganic nitrogen (DIN), total phosphorus (TP), river canopy angle (Canopy), and watershed agriculture land use (Ag, in percent of total watershed area) were used as predictors. The linear regression model can be expressed as a DAG (Fig. 4, upper-right block). The second model is for predicting a DO variable using benthic chlorophyll *a* concentration. The third and fourth models are for predicting ICI and EPT, respectively, both using DO variables, benthic chlorophyll *a*, and QHEI (an index of habitat quality for aquatic life, (Ohio EPA, 2006)) as predictors.

Instead of following the Occam's razor to make a model as simple as possible (but not simpler), our exploration followed the guidelines in Gelman and Hill (2007) and Qian (2010):

- Including predictors that are substantively relevant,
- Considering compound variables, i.e., creating a variable by combining several variables (e.g., length times width to form area),
- Consider log-transformation and other transformations to normality (based on Box–Cox),
- Including a statistically insignificant predictor if the estimated slope is of the correct sign,

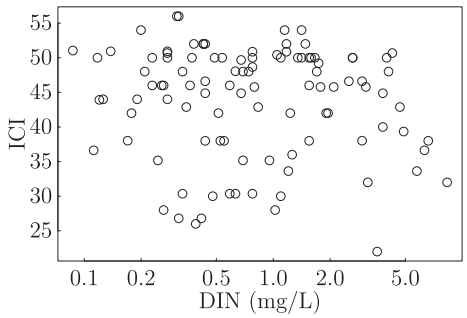

Fig. 2. Scatter plot of ICI against DIN shows no particular pattern.

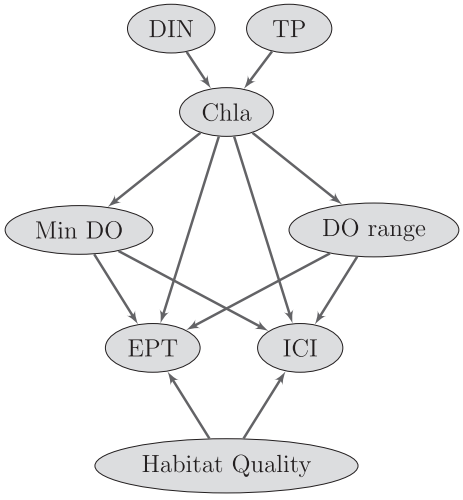

Fig. 3. DAG representation of the Ohio EPA modeling approach based on Miltner (2010).

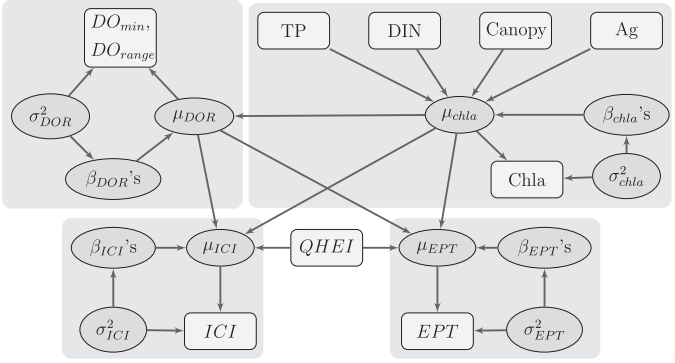

Fig. 4. DAG representing the regression models and there relationship based on the causal model in Fig. 3 is grouped into four blocks (4 shaded rectangles). The chlorophyll *a* regression model (upper-right block) includes watershed agriculture land use (Ag) and river segment canopy cover (Canopy). The DO model (upper left block) takes the estimated mean chlorophyll *a* concentration (μ_{chla}) as input, and the ICI (lower left block) and EPT (lower right block) models are connected to the other two blocks through the estimated μ_{chla} and μ_{DOR}.

- Considering interactions of predictors with strong effects.

Including a statistically insignificant predictor is often considered a bad practice. But in the context of this study, we find this rule to be important. For example, when developing the model for predicting benthic chlorophyll *a* concentration, TP is consistently a weak and insignificant predictor. The estimated TP slope is small but positive (the correct sign). Including TP in the model will not make much difference in predicting chlorophyll *a* concentrations. But conceptually, TP is a nutrient and it should have a positive effect on benthic chlorophyll *a*. A potential reason of an insignificant TP slope is the use of a dataset representing multiple streams and rivers, and the effect of TP varies from stream to stream. When updating our model using stream- or region-specific data, we may find that TP is an important factor in some streams or regions.

Relations among these four regression models are represented in a DAG (Fig. 4). We briefly summarize each of the four regression models.

2.3.1. The benthic chlorophyll a (chla) model

We used all available variables that are important to benthic algal growth in the model. Specifically, dissolved inorganic nitrogen (*DIN*, in mg/L), total phosphorus (*TP*, in mg/L), % agriculture land cover in the watershed (*Ag*), and canopy cover measured as the degree of open arc between the canopy tops of either bank (*Canopy*). Based on our exploratory analysis, we log-transformed all variables except *Ag*. A logit transformation is used for *Ag* [$\text{logit}(Ag) = \log(Ag/100 - Ag)$] because it is a fraction variable (Gelman and Hill, 2007; Qian, 2010).

The resulting model

$$\log(chla) = \beta_0^c + \beta_1^c \log(DIN) + \beta_2^c \log(TP) + \beta_3^c \text{logit}(Ag) + \beta_4^c \log(Canopy) + \varepsilon^c \quad (1)$$

can be presented in terms of the probability distribution of $\log(chla)$

$$\log(chla) \sim N\left(\mu_{chla}, \sigma^2_{chla}\right)$$
$$\mu_{chla} = \beta_0^c + \beta_1^c \log(DIN) + \beta_2^c \log(TP) + \beta_3^c \text{logit}(Ag) + \beta_4^c \log(Canopy)$$

2.3.2. The DO model

In exploring the relation between DO and *chla*, we found that neither DO minimum nor DO range are responsive to changes in *chla* (see On-line Supporting Materials). Furthermore, DO variables also vary as a function of temperature. As a result, we expect that DO changes should not be a function of *chla* when *chla* is low. After exploring multiple combinations of the two DO variables, we found that the ratio of DO minimum over DO range (*DOR*) is a suitable compound variable. Initial exploration suggested that a log–log linear model of *DOR* and *chla* can be an adequate model. However, we find that the possible threshold response as explored in Miltner (2010) can be better interpreted in scientific terms. As a compromise, we used a piece-wise linear model (Qian and Richardson, 1997), also known as the hockey-stick model (Qian, 2010):

$$\log\left(\frac{DO_{min}}{DO_{range}}\right) = \beta_0^d + \beta_1^d \times I(\log(chla) - \phi) \times (\log(chla) - \phi) + \varepsilon^d, \quad (2)$$

where $I(a) = 0$ if $a \le 0$ and $I(a) = 1$ if $a > 0$ and ϕ is the change point of $\log(chla)$ at which the slope changes. In this case, when $\log(chla) \le \phi$, *DOR* is a constant (β_0^d), when $\log(chla) > \phi$, *DOR* is a linear function of $\log(chla)$ with slope β_1^d. Although the diagnostic statistics show that both the log–log linear model and the hockey-stick model are adequate, we used the hockey-stick model because it is more flexible and includes the linear model as a special case (when ϕ approaches one or the other end of the $\log(chla)$ range).

When used in the cBN model, we replace $\log(chla)$ (the observed chlorophyll *a* concentrations) with μ_{chla} (the estimated log mean chlorophyll *a*):

$$\log\left(\frac{DO_{min}}{DO_{range}}\right) \sim N\left(\mu_{DOR}, \sigma^2_{DOR}\right)$$
$$\mu_{DOR} = \beta_0^d + \beta_1^d \times I(\mu_{chla} - \phi) \times (\mu_{chla} - \phi) \quad (3)$$

2.3.3. The EPT and ICI models

The macroinvertebrate indicators are predicted by variables representing water quality (*chla*,*DOR*) and a habitat quality index (QHEI). We log-transformed EPT, a count variable with a skewed distribution. ICI is a sum of multiple indicators and its distribution is roughly symmetric, hence it is not transformed:

$$\log(EPT) \sim N(\mu_{EPT}, \sigma^2_{EPT})$$
$$\mu_{EPT} = \beta_0^e + \beta_1^e \mu_{chla} + \beta_2^e \mu_{DOR} + \beta_3^e QHEI \quad (4)$$

and

$$ICI \sim N(\mu_{ICI}, \sigma^2_{ICI})$$
$$\mu_{ICI} = \beta_0^i + \beta_1^i \mu_{chla} + \beta_2^i \mu_{DOR} + \beta_3^i QHEI \quad (5)$$

2.4. Bayesian updating

Data used to develop nutrient criteria are frequently collected from multiple sites to establish a gradient in nutrient concentrations. Differences among streams further contribute to the uncertainty on the relation between nutrient concentrations and macroinvertebrate metrics. We expect that the effect of nutrient enrichment varies among streams. For example, streams near Lake Erie in northern Ohio are likely slow moving with more sediments in the bottom, while streams in the Appalachian mountains in southern Ohio are likely fast moving with rocky substrates. As a result, empirically estimated model coefficients using a cross-sectional data set represent (weighted) average over rivers and streams included in the data.

This case illustrates the common dilemma of empirical modeling – if we use cross-sectional data to increase sample size and broaden the inference base of the model, the resulting model may be irrelevant to any specific site; if we use site-specific data, we are often limited by small sample sizes and cannot effectively study the effects of variables operating on large spatial and long temporal scales (Qian et al., 2010). Ideally, this dilemma should be addressed by partially pooling data from different sites using a shrinkage estimator (Efron and Morris, 1973; Gelman and Hill, 2007), specifically the Bayesian hierarchical modeling approach (Gelman et al., 2014). When using a Bayesian hierarchical modeling approach for estimating a parameter, we learn about the distribution of parameter values across multiple sites and such information helps improving estimating site-specific parameter.

In most environmental management studies, however, we often have data from the study site only. Efron (1978) and Efron and Morris (1977) suggest that the same improvement can be achieved if we know a priori the parameter distribution across similar sites. That is, we can achieve the same improvement in parameter estimation by using the Bayesian estimator with a known cross-site parameter distribution as the prior. For our example, the model developed using data from the entire state reflects the variation across all small rivers and streams. It is, therefore, such a prior. The resulting model should be used in subsequent studies of individual streams.

Using the Bayesian updating process, the cross-sectional model is updated using data from specific sites or regions (posteriors). Over time, the updated model will become a site- or region-specific model, providing specific information for site-specific management decision-making. In this work, model coefficients distributions estimated using the cross-sectional data are used as the prior distributions for coefficients of a river-specific model. Data from the river of interest are then used to derive posterior distributions for these coefficients. As discussed in Qian and Reckhow (2007), Bayesian updating is an effective means for supporting adaptive management.

3. Results

We present the results in two parts – the fitted joint model and the model updating process using data from a set-aside river.

3.1. The Ohio stream model

The joint model is built upon the four empirical models summarized in a DAG (Fig. 4). Coefficients of these individually fit models (Tables 1–4, under heading "Regression estimates") are compared to the coefficients of the cBN model (Tables 1–4, under heading "cBN estimates"). The differences between the two sets of coefficients are small as both were fit to the same data.

When the response variable is log-transformed, the estimated slope coefficients represent a multiplicative factor change. Specifically, if the predictor is in its original scale (e.g., *QHEI* and β_3^e in equation (4)), the slope represents a factor of $e^{\beta_3^e}$ change in *EPT* per one unit change in *QHEI*. For example, the estimated β_3^e is 0.01, representing a factor of $e^{0.01} \approx 1.01$ or 1% increase in *EPT* for every unit increase in *QHEI*. When both the predictor and the response are log-transformed, the slope represents a factor change in the response per unit factor change in the predictor. For example, the estimated β_1^c is 0.21, representing approximately a 0.21% increase in *chla* for a 1% increase in *DIN* (see page 157 of Qian (2010)).

With the joint cBN model, we are able to connect variables throughout the network. For the purpose of setting a nutrient criterion, we are interested in estimating an acceptable DIN concentration, below which a stream's ICI and EPT are consistent with the designated use for aquatic life. For illustration purposes, we used the 75 percentile of EPT and ICI in our data (21 and 50, respectively) as the lower bound of the desired biological conditions. The

Table 1
The estimated chlorophyll *a* model coefficients (in Equation (1)).

Coef	Regression estimates				cBN estimates				Bayesian updating			
	Mean	sd	2.5%	97.5%	Mean	sd	2.5%	97.5%	Mean	sd	2.5%	97.5%
β_0^c	3.73	0.38	2.99	4.50	3.87	0.26	3.38	4.37	3.82	0.37	3.11	4.55
β_1^c	0.21	0.07	0.08	0.34	0.18	0.04	0.09	0.26	0.19	0.07	0.06	0.32
β_2^c	0.02	0.05	−0.08	0.13	0.03	0.03	−0.04	0.09	0.03	0.05	−0.07	0.13
β_3^c	0.09	0.04	0.02	0.16	0.13	0.03	0.08	0.19	0.10	0.04	0.03	0.17
β_4^c	0.85	0.20	0.46	1.23	0.76	0.13	0.5	1.01	0.80	0.19	0.41	1.17
σ_{chla}	0.53	0.04	0.46	0.61	0.52	0.04	0.45	0.60	0.53	0.04	0.46	0.61

Table 2
The estimated DO model coefficients (in Equation (2)).

Coef	Regression estimates				cBN estimates				Bayesian updating			
	Mean	sd	2.5%	97.5%	Mean	sd	2.5%	97.5%	Mean	sd	2.5%	97.5%
β_0^d	0.88	0.26	0.36	1.40	1.06	0.18	0.71	1.43	0.87	0.26	0.38	1.38
β_1^d	−0.59	0.26	−1.10	−0.07	−0.78	0.17	−1.10	−0.46	−0.58	0.26	−1.10	−0.09
ϕ	4.66	0.61	3.47	5.82	4.48	0.28	3.87	4.98	4.68	0.59	3.51	5.88
σ_{DOR}	0.97	0.08	0.84	1.14	0.91	0.07	0.79	1.06	0.98	0.08	0.84	1.14

boundary for ICI, coincidentally, approximates the bio-criterion for exceptional waters (Ohio Administrative Code 3745-1-07[1]).

In the U.S., two methods are recommended by the U.S. EPA for setting nutrient criteria (U.S. EPA, 2000): the reference condition approach and the dose–response modeling approach. Although we are developing a model to connect nutrient concentration to ICI and EPT, our model is ill-suited for deriving a dose–response model because we used data from streams across the state of Ohio and factors other than nutrient concentration often play a more important role in determining a stream's ecological condition. In Section 3.2 we develop a dose–response model for a specific stream where we can hold other factors constant. The model, however, can be used to derive nutrient concentration distribution associated with desired stream ecological conditions (EPT $\geq$ 21 and ICI $\geq$ 50). This distribution is conceptually similar to EPA's reference distribution (U.S. EPA, 2000).

We use a Monte Carlo simulation algorithm similar to the Hornberger-Spear-Young algorithm (Whitehead and Young, 1979; Hornberger and Spear, 1981; Young, 1983):

1. Random variates of the chlorophyll *a* model input variables (DIN, TP, Ag, and Canopy) are drawn from distributions derived based on the cross-sectional data (log TP and log DIN are assumed to follow a bivariate normal distribution because of the observed Pearson correlation coefficient of 0.68, and Ag and Canopy are uniform with the lower and upper bounds set by the data);
2. A set of chlorophyll *a* model coefficient values are drawn from their joint posterior distribution;
3. Random variates drawn from the previous two steps are used to calculate μ_{clha};
4. A set of DO model coefficients are drawn from the joint posterior distribution, which are then combined with μ_{chla} from step 3 to calculate μ_{DOR};
5. Random variates of the ICI and EPT model coefficients, as well as a random QHEI value from a uniform distribution defined by the data are used to calculate μ_{ICI} and μ_{EPT}.
6. If $\mu_{ICI} \geq 50$ and $\mu_{EPT} \geq 21$, we keep DIN from step 1 as a sample of DIN that resulted in an acceptable biological condition.

This process is repeated until 10,000 DIN values are accepted. These DIN values are used to quantify the conditional distribution of DIN given $\mu_{ICI} \geq 50$ and $\mu_{EPT} \geq 21$. Compared to the marginal distribution of DIN estimated from the data (histogram in Fig. 5), the conditional distribution has a variance similar to the DIN data variance, but a smaller mean (Fig. 5, solid line). If we take this conditional distribution as the "reference" distribution, we can use U.S. EPA's recommendation and set our DIN criterion at the 75 percentile (U.S. EPA, 2000), which is 0.625 mg/L. Ohio uses an ICI standard of 46 and EPT of 10 to define a water as in compliance, which resulted in a conditional distribution of DIN very similar to its marginal distribution (dashed line in Fig. 5). (In our data, only 22 of the 109 observations have EPT less than 10, the median of ICI in our data is 46.) The second conditional distribution has a 75 percentile of 1.251 mg/L. U.S. EPA recommends that when reference sites are unavailable, the 25 percentile of all data should be used as the nutrient criterion (U.S. EPA, 2000). DIN in our data has a 25 percentile of 0.36 mg/L, much smaller than our estimates. The comparable variances in the two conditional and the marginal DIN distributions is a result of using the cross-sectional data. In our data set, each river is represented by 1–3 observations (the set-aside river has 4 observations), as a result, the estimated variances represent mostly between river/site variances.

3.2. Model updating

When fitting the model, we set aside data from one river (River code 02–500) for illustrating Bayesian model updating. Four observations were taken from two segments of the river in two years. But only two observations had DO values. Using the cBN model coefficient distributions as priors, the updating process is essentially the same as fitting the original model, but with more confined prior distributions and only 4 observations. The resulting posterior distributions of model coefficients are not very different from their priors (the cBN estimated coefficients) (Tables 1–4, under heading "Bayesian updating"), because very small sample size for the river resulted in a posterior dominated by the prior.

The updated model for the set-aside river can be used to develop river-specific nutrient criteria using the conditional distribution approach as in Fig. 5 or more directly using the probability of achieving the desired ecological conditions (e.g., ICI $\geq$50 and EPT $\geq$21) at a series of nutrient concentration values (i.e.,

[1] http://www.epa.ohio.gov/portals/35/rules/01-07.pdf.

Table 3
The estimated ICI model coefficients (in Equation (5)).

Coef	Regression estimates				cBN estimates				Bayesian updating			
	Mean	sd	2.5%	97.5%	Mean	sd	2.5%	97.5%	Mean	sd	2.5%	97.5%
β_0^i	36.85	7.56	22.05	51.55	32.64	6.49	19.97	44.82	35.99	7.35	21.77	51.09
β_1^i	−1.63	1.19	−3.92	0.75	−1.29	1.09	−3.38	0.82	−1.54	1.15	−3.89	0.75
β_2^i	2.84	0.73	1.38	4.24	2.67	0.82	1.12	4.26	2.81	0.72	1.38	4.20
β_3^i	0.20	0.05	0.11	0.30	0.23	0.04	0.16	0.31	0.21	0.05	0.12	0.31
σ_{ICI}	6.09	0.50	5.20	7.15	6.79	0.47	5.94	7.79	6.00	0.47	5.13	7.04

Table 4
The estimated EPT model coefficients (in Equation (4)).

Coef	Regression estimates				cBN estimates				Bayesian updating			
	Mean	sd	2.5%	97.5%	Mean	sd	2.5%	97.5%	Mean	sd	2.5%	97.5%
β_0^e	2.31	0.53	1.26	3.36	1.86	0.47	0.93	2.83	2.16	0.52	1.14	3.17
β_1^e	−0.12	0.08	−0.29	0.04	−0.08	0.08	−0.24	0.07	−0.11	0.08	−0.28	0.06
β_2^e	0.13	0.05	0.04	0.23	0.12	0.05	0.01	0.22	0.13	0.05	0.03	0.22
β_3^e	0.01	0.003	0.01	0.02	0.02	0.002	0.01	0.02	0.01	0.003	0.01	0.02
σ_{EPT}	0.41	0.03	0.35	0.48	0.46	0.03	0.40	0.53	0.41	0.03	0.35	0.48

$\Pr(EPT \geq 21 \& ICI \geq 50 | DIN = x)$). However, because nutrient is not necessarily the most important factor for macroinvertebrates, this "reverse" conditional probability can be highly variable for the same DIN value because of changes of other factors such as habitat condition. We used the updated model coefficients to estimate the probability of ICI ≥50 and EPT ≥21 as a function of both DIN and QHEI with all other input variables taken their respective observed means (Fig. 6). Although we consider Fig. 6 as tentative because of a very small site-specific sample size, these probabilities show that habitat quality may be more important than nitrogen concentration in achieving the desired ecological condition for this particular river.

3.3. Model assessment

When developing individual component models, we used typical statistical model diagnostic methods to ensure that regression model assumptions are met (see Section 5.2.6 of Qian (2010)). As a result, properly fit statistical models can almost always show favorable model evaluations using commonly used statistical model assessment methods. In our example, the model is fit to a dataset of many rivers and streams. Therefore, model predictions represent a state average, not for a specific river or stream. As discussed in Section 2.4, we see the role of a model based on a cross-sectional dataset as the basis for developing informative prior distributions for site-specific models. Consequently, we are more interested in whether the model can be properly updated when site-specific data are available. In this regard, we repeat the model updating process two more times, each time setting aside data from one variable (DIN or QHEI). A uniform prior is used for the "set-aside" variable. Treating the set-aside variable as unknown, we estimate its posterior distribution. This process is to evaluate whether functional connections established using data from multiple rivers and streams are appropriate for a single river.

The posterior distribution of DIN (Fig. 7) is centered around the observed values. The posterior distribution of QHEI (Fig. 8) is skewed to the left, properly represents the observed distribution –

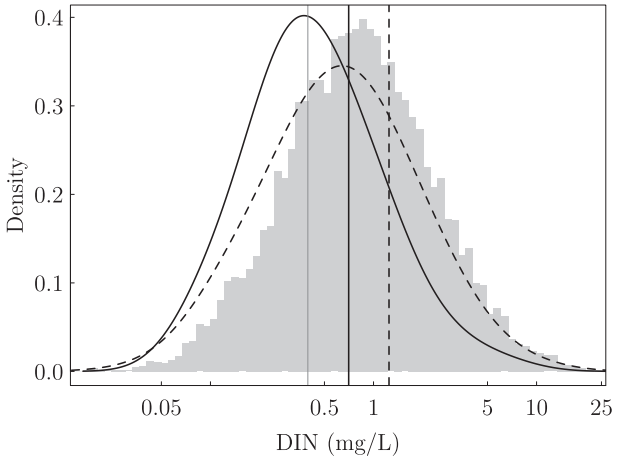

Fig. 5. Marginal distribution of DIN (histogram) is compared to two conditional DIN distributions – the DIN distribution corresponding to ICI ≥50 and EPT ≥21 (representing the respective 75 percentiles in the data) (black solid line) and the DIN distribution corresponding to ICI ≥46 and EPT ≥10 (Ohio's bio-criteria for small rivers and streams) (black dashed line). The 75 percentiles of the two DIN conditional distributions are shown as vertical lines with the same line types as the density curve, and the 25 percentile of the DIN marginal distribution is shown by the thin gray line.

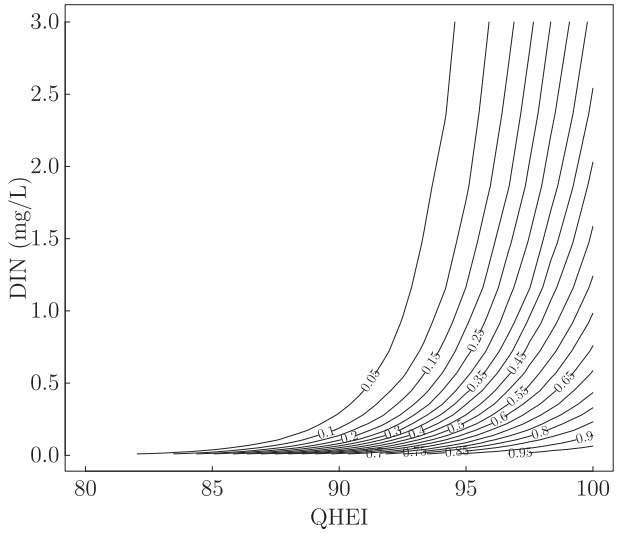

Fig. 6. Predicted probabilities of ICI ≥50 and EPT ≥21 are shown as a function of QHEI and DIN.

one of the four observed values is much lower than the other three. In both cases, centers of the posterior marginal distributions are consistent with the observed data, while the spreads of the two posteriors are still large, reflecting the large among sites variance in the prior model.

In both cases, the posterior distributions of these two set-aside variables are much closer to the data than their uniform priors are. Although the correlation between DIN and ICI is weak in the cross-sectional data (Fig. 2), Figs. 7 and 8 indicate that connections among nodes shown in Fig. 4 are well established and can be updated when new data are available. We note that intentionally setting aside the values of one or two variables is the same as a missing data problem.

4. Discussion

We cited Efron and Morris (1977) in Section 2.4 to justify the use of Bayesian updating. Without the updating process, our modeling framework is simply a Bayesian parameter estimation and inference on a series of linked empirical models. In other words, we are not presenting the modeling approach as a technological advance, rather as an example of an environmental modeling philosophy on how to appropriately integrate information from different sources for improving a model's accuracy. Specifically, we want to properly integrate information represented in data from multiple sites and information represented in data from a single site. Using the empirical Bayes interpretation of a shrinkage estimator (Efron, 1978), we suggest that site-specific estimation can be improved without data from multiple sites as long as information represented in the cross-sectional data is properly used through the application of the Bayes theorem.

Properly integrating information from different sources has practical implications. Many BN models reviewed by Aguilera et al. (2011) were developed either entirely based on expert opinion elicitation or a combination of limited data and expert opinions. Even when developing a BN model for a specific system (e.g., the Neuse River Estuary in Borsuk et al. (2004)), information representing multiple systems are often used (e.g., the benthic oxygen demand modal coefficients in Borsuk et al. (2004) are weighted averages of coefficients from a number of estuaries and coast zones based on a Bayesian hierarchical model (Borsuk et al., 2001)). As a result, these models are often not specific for an individual waterbody. Most BN models can be updated upon observing new data,

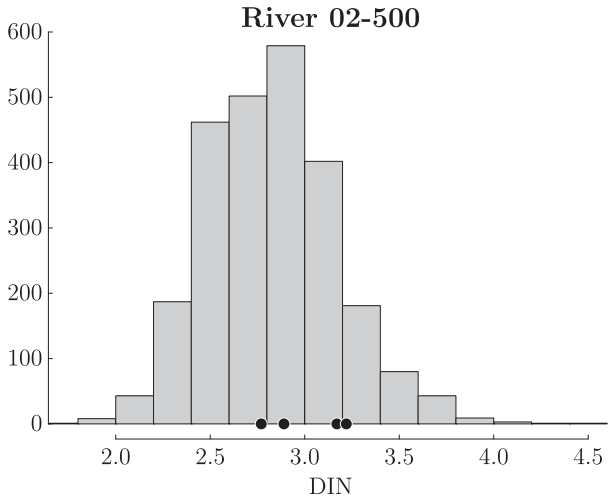

Fig. 7. Marginal distribution of DIN predicted for the set-aside river without using the river-specific DIN data (histogram) is compared to the observed DIN data (solid dots).

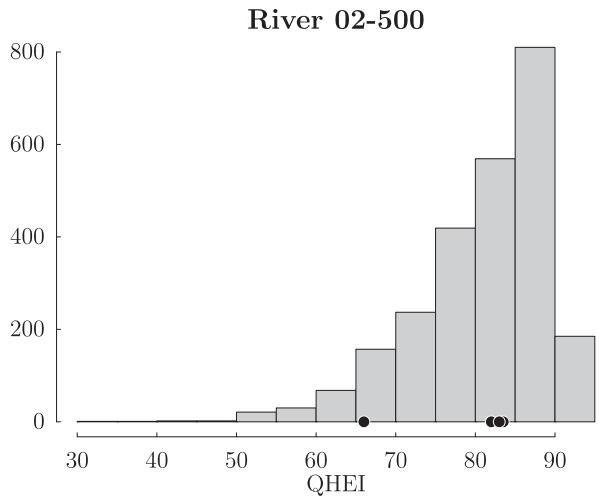

Fig. 8. Marginal distribution of QHEI predicted for the set-aside river without using the river-specific QHEI data (histogram) is compared to QHEI data (solid dots).

thereby, increasingly refined to represent the target system. However, such refinement is greatly limited because discretization boundaries are fixed.

Using cBN, the updating process is reflected in changes of model coefficients. As long as the functional forms of the component models stay the same, updating is informative and meaningful. In our case, we chose models with more terms (e.g., including non-significant TP in Equation (1)) or a more flexible functional form (e.g., using the hockey-stick model of Equation (2) instead of a simpler linear model) to increase the flexibility of subsequent updating.

For the Ohio streams example, we intended to develop a model to improve the process of establishing nutrient criteria for Ohio's small rivers and streams. As we learned more about the links among various variables through our model, we realized that the goal of setting a single nutrient criterion for the entire state of Ohio is likely impractical. A single criterion derived using a model based on cross-sectional data (e.g., Fig. 5) inevitably represent the average condition of the state, both in terms of nutrient levels and other factors, such as habitat quality. An average is almost always different from individual rivers. As a result, the established nutrient criterion will be too low or too high. In some cases, a nutrient criterion can be irrelevant when factors other than nutrient concentration prevent a stream from achieving the desired ecological condition (Fig. 6).

As a result, a flexible and updatable model is particularly important in cases where initial model development is based on data from multiple sites. In our case study, we treat the estimated nutrient criteria as tentative because the model represents a state average, not a particular water. Region- or river-specific criteria should be developed, because the effect of nutrient enrichment may depend on other factors. For example, habitat quality (represented by QHEI) is an important predictor of EPT and ICI. The conditional DIN distributions shown in Fig. 5 are derived without considering QHEI (its effect is integrated out). The conditional DIN distribution varies when QHEI varies (Fig. 9). We note that we are not suggesting to use local nutrient criteria to compensate for all other insults to the system Fig. 9 is an illustrative example of how other factors may affect the influence of nutrient on stream ecosystem. In some cases, stream habitat destruction alone can make achieving the desired ecological goal impossible (Fig. 6). If nutrient criteria are set based on cross-sectional data, we are setting nutrient criteria by averaging out all other factors. Our

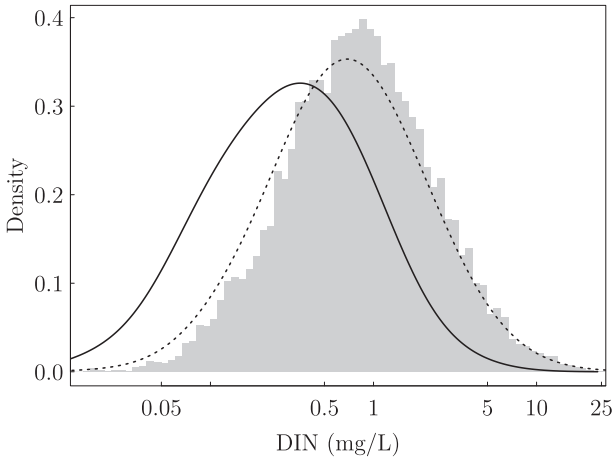

Fig. 9. Marginal distribution of DIN (histogram) is compared to two conditional DIN distributions – corresponding to ICI ≥46 and EPT ≥10 – estimated by setting QHEI = 65 (25 percentile of the data, solid line) and QHEI = 85.5 (95 percentile, dashed line). When setting QHEI at its 5 percentile, we failed to obtain simulation results that meet the ICI and EPT conditions.

model includes only variables with available data, and we expect other factors not included may also confound the nutrient-biological metric relationship. As a result, nutrient criteria should be established locally.

Model performance assessment is an important aspect of environmental modeling. Bennett et al. (2013) discussed a series of performance metrics for conventional environmental modeling. In Bayesian statistics, performance assessment is often done by using posterior simulation (Gelman, 2003; Gelman et al., 1996, 2014) because a Bayesian model results in probability distributions of model parameters, rather than point estimates. Consequently, the comparison of a predicted value and an observed value becomes a comparison of a probability distribution and an observation. We use the Bayesian posterior model to produce probability distributions of important quantities derived from the model (posterior predictive distributions) and compare these distributions to the respective observed values. In our example, we evaluated our model by using the posterior model, updated by a subset of the data from the set-aside river, to predict DIN and QHEI (not included in the updating process). The comparison of the posterior predictive distributions and the observed values (Figs. 7 and 8) can often reveal weakness of the model and lead to strategies for model improvement. In our example, the comparisons suggested that the model captures the center of the distributions of the two variables well. We selected DIN and QHEI for model assessment because reducing nitrogen load and improving stream habitat are often the goals of management practices.

Model structure uncertainty can be (and often is) a dominant source of uncertainty. In theory any specific model form is a simplification of the real world. Furthermore, separating model structure uncertainty in a statistical model from random noise is often difficult as both sources of uncertainty are reflected in the discrepancy between the model predicted and observed (the residuals). With additional information, we can formulate a number of statistical models to separate these two uncertainties. For mechanistic models, Reichert and Schuwirth (2012) discussed a method for quantifying model bias using a multiobjective model calibration in conjunction with a Bayesian description of model bias. In statistical modeling, the state-space modeling approach (Petris et al., 2009) (including the Kalman filter model used by Beck (1987)) separates the total error into “observational error” and “model error” in time-series data. The state-space modeling approach is used in Qian et al. (2005) to model nutrient loadings from a series of connected watershed. In our case study, we cannot explicitly separate the two sources of uncertainty based on the data we have. However, by modeling regression model coefficients of each component model as random variates from a multivariate normal distribution, we implicitly included model structure uncertainty (in a limited way). Specifically, in the Monte Carlo simulation, we draw random values of model coefficients, each set of coefficients represents a different model. As a result, the Bayesian estimated posterior distribution of a mean variable (e.g., μ_{chla}) is represented by a distribution, not a point estimate.

Acknowledgements

We thank Kenneth H. Reckhow, Thomas Cuffney, Jerry McMahon, and students in a Bayesian modeling class at Duke University for their initial feedback on this work. Comments and suggestions from three anonymous reviewers and the editor are greatly appreciated.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.envsoft.2015.03.001.

References

Aguilera, P., Fernández, A., Fernández, R., Rumí, R., Salmerón, A., 2011. Bayesian networks in environmental modelling. Environ. Model. Softw. 26, 1376–1388.
Alameddine, Ibrahim, Cha, YoonKyung, Reckhow, Kenneth H., 2011. An evaluation of automated structure learning with Bayesian networks: an application to estuarine chlorophyll dynamics. Environ. Model. Softw. 26 (2), 163–172.
Beck, M.B., 1987. Water quality modeling: a review of the analysis of uncertainty. Water Resour. Res. 23 (8), 1393–1442.
Bennett, Neil D., Croke, Barry F.W., Guariso, Giorgio, Guillaume, Joseph H.A., Hamilton, Serena H., Jakeman, Anthony J., Marsili-Libelli, Stefano, Newham, Lachlan T.H., Norton, John P., Perrin, Charles, Pierce, Suzanne A., Robson, Barbara, Seppelt, Ralf, Voinov, Alexey A., Fath, Brian D., Andreassian, Vazken, 2013. Characterising performance of environmental models. Environ. Model. Softw. 40, 1–20.
Bollen, K.A., 1989. Structural Equation with Latent Variables. Wiley, New York.
Borsuk, M.E., Higdon, D., Stow, C.A., Reckhow, K.H., 2001. A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecol. Model. 143 (3), 165–181.
Borsuk, M.E., Stow, C.A., Reckhow, K.H., 2004. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. Ecol. Model. 173 (2–3), 219–239.
Chen, S.H., Pollino, C.A., 2012. Good practice in Bayesian network modelling. Environ. Model. Softw. 37, 134–145.
Efron, B., 1978. Controversies in the foundations of statistics. Am. Math. Mon. 85 (4), 231–246.
Efron, B., Morris, C., 1973. Combining possibly related estimation problems. J. R. Stat. Soc. Ser. B Methodol. 35 (3), 379–421.
Efron, B., Morris, C., 1977. Stein's paradox in statistics. Sci. Am. 236, 119–127.
Gelfand, A.E., Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities. J. Am. Stat. Assoc. 85, 398–409.
Gelman, A., 2003. A Bayesian formulation of exploratory data analysis and goodness-of-fit testing. Int. Stat. Rev. 71 (2), 369–382.
Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York.
Gelman, A., Meng, X.L., Stern, H.S., 1996. Posterior predictive assessment of model fitness via realized discrepancies (with discussions). Stat. Sin. 6, 733–807.
Gelman, A., Carlin, J.B., Stern, H.S., Dunson, David B., Vehtari, Aki, Rubin, D.B., 2014. Bayesian Data Analysis, third ed. CRC Press, Boca Raton, Florida.
Gilks, W.R., Thomas, A., Spiegelhalter, D.J., 1994. A language and program for complex bayesian modelling. Statistician 43 (1), 169–177.
Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (Eds.), 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London. UK.
Grace, J.B., 2006. Structural Equation Modeling and Natural Systems. Cambridge University Press, New York.
Hornberger, G.M., Spear, P.A., 1981. An approach to the preliminary analysis of environmental systems. J. Environ. Manag. 12, 7–18.
Jensen, F.V., 2001. Bayesian Networks and Decision Graphs. Springer, New York.
Kashuba, R.O., 2010. Bayesian Methods to Characterize Uncertainty in Predictive Modeling of the Effect of Urbanization on Aquatic Ecosystems. Duke University, Durham, NC (PhD thesis). URL. http://hdl.handle.net/10161/2366.

Kelly, Rebecca A., Jakeman, Anthony J., Barreteau, Olivier, Borsuk, Mark E., ElSawah, Sondoss, Hamilton, Serena H., Henriksen, Hans Jørgen, Kuikka, Sakari, Maier, Holger R., Emilio Rizzoli, Andrea, Delden, Hedwig van, Voinov, Alexey A., 2013. Selecting among five common modelling approaches for integrated environmental assessment and management. Environ. Model. Softw. 47, 159–181.

Lauritzen, S.L., 1996. Graphical Models. Oxford University Press, Oxford.

Lumina, 1997. Analytica for Windows, User's Guide, Version 1.1. Lumina Decision Systems, Denvor, CO, USA.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Stat. Comput. 10 (4), 325–337.

Miltner, R., 2010. A method and rationale for deriving nutrient criteria for small rivers and streams in Ohio. Environ. Manag. 45, 842–855. ISSN: 0364-152X.

Nojavan, A.F., August 2014. Bayesian Statistical Analysis in Coastal Eutrophication Models: Challenges and Solutions. Duke University, Durham, NC, U.S.A. (PhD thesis).

Nojavan, A.F., Qian, S.S., Paerl, H.W., Reckhow, K.H., Albright, E.A., 2014. A study of anthropogenic and climatic disturbance of the new river estuary using a Bayesian belief network. Mar. Pollut. Bull. 83, 107–115.

Ohio EPA, 1978. Users Manual for Biological Field Assessment of Ohio Surface Waters, vol. 2. URL. http://www.epa.ohio.gov/Portals/35/documents/Vol2.pdf.

Ohio, E.P.A., 2006. Methods for Assessing Habitat in Flowing Waters: Using the Qualitative Habitat Evaluation Index (QHEI). OHIO EPA Technical Bulletin EAS/2006-06-1. URL. http://www.epa.state.oh.us/portals/35/documents/qheimanualjune2006.pdf.

Pearl, J., 1986. Fusion, propagation, and structuring in belief networks. Artif. Intell. 29 (3), 241–288.

Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. In: Series in Representation and Reasoning. Kaufmann Publishers, San Francisco, CA.

Petris, G., Petrone, S., Campagnoli, P., 2009. Dynamic Linear Models with R. Springer.

Plummer, Martyn, March 20-22 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), Vienna, Austria. ISSN: 1609-395X.

Qian, S.S., 2010. Environmental and Ecological Statistics with R. Chapman and Hall/CRC Press.

Qian, S.S., Reckhow, K.H., 2007. Combining model results and monitoring data for water quality assessment. Environ. Sci. Technol. 41 (14), 5008–5013.

Qian, S.S., Richardson, C.J., 1997. Estimating the long-term phosphorus accretion rate in the Everglades: a Bayesian approach with risk assessment. Water Resour. Res. 33 (7), 1681–1688.

Qian, S.S., Stow, C.A., Borsuk, M.E., 2003. On Monte Carlo methods for Bayesian inference. Ecol. Model. 159, 269–277.

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.

Qian, S.S., Cuffney, T.F., Alameddine, I., McMahon, G., Reckhow, K.H., 2010. On the application of multilevel modeling in environmental and ecological studies. Ecology 91, 355–361.

Reichert, P., Schuwirth, N., 2012. Linking statistical bias description to multi-objective model calibration. Water Resour. Res. 48, W09543.

Spiegelhalter, David, Thomas, Andrew, Best, Nicky, Lunn, Dave, 2014. Open BUGS User Manual, Version 3.2.3. URL. http://www.openbugs.net/Manuals/Manual.html.

U.S. EPA, 2000. Nutrient Criteria Technical Guidance Manual. Technical Report EPA-822-B00-022. United States Environmental Protection Agency, Office of Water, Washington, D.C.

Weisberg, S., 2005. Applied Linear Regression. Wiley.

Whitehead, P.G., Young, P.C., 1979. Water quality in river systems: Monte Carlo analysis. Water Resour. Res. 15, 451–459.

Young, P.C., 1983. The validity and credibility of models for badly-defined systems. In: Beck, M.B., van Straten, G. (Eds.), Uncertainty and Forecasting of Water Quality. Springer-Verlag, Berlin, pp. 69–98.