Combining Model Results and Monitoring Data for Water Quality Assessment

ARTICLE *in* ENVIRONMENTAL SCIENCE AND TECHNOLOGY · AUGUST 2007

Impact Factor: 5.33 · DOI: 10.1021/es062420f · Source: PubMed

2 AUTHORS:

Song S. Qian
University of Toledo
99 PUBLICATIONS **1,817** CITATIONS
SEE PROFILE

Kenneth H Reckhow
Duke University
140 PUBLICATIONS **3,219** CITATIONS
SEE PROFILE

Combining Model Results and Monitoring Data for Water Quality Assessment

SONG S. QIAN* AND
KENNETH H. RECKHOW
Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina 27708-0328

A Bayesian approach is used to update and improve water quality model predictions with monitoring data. The objective of this work is to facilitate adaptive management by providing a framework for sequentially updating the assessment of water quality status, to evaluate compliance with water quality standards, and to indicate if modification of management strategies is needed. Currently, most water quality or watershed models are calibrated using historical data that typically reflect conditions different from those being forecast. In part because of this, predictions are often subject to large errors. Fortunately, in many instances, postmanagement implementation monitoring data are available, although often with limited spatiotemporal coverage. These monitoring data support an alternative to the one-time prediction: pool the information from both the initial model prediction and postimplementation monitoring data. To illustrate this approach, a watershed nutrient loading model and a nitrogen−chlorophyll *a* model for the Neuse River Estuary were applied to develop a nitrogen total maximum daily load program for compliance with the chlorophyll *a* standard. Once management practices were implemented, monitoring data were collected and combined with the model forecast on an annual basis using Bayes Theorem. Ultimately, the updated posterior distribution of chlorophyll *a* concentration indicated that the Neuse River Estuary achieved compliance with North Carolina's standard.

1 Introduction

If a total maximum daily load program (TMDL) or other water quality management plan is developed and implemented to meet a water quality criterion, monitoring is usually the basis for assessing compliance and determining if any management modifications are needed. Yet, we know that lags in implementation of plans, lags in pollutant concentration change and/or in biotic response, measurement uncertainty, and natural variability all may lead to errors in inferences based on measurements. This has led some in the water quality modeling community to recommend the use of models to assess progress.

An example of this conflict between model and data is the TMDL program implemented to reduce nitrogen loading to the Neuse River Estuary in North Carolina. The Neuse River Estuary was placed in State of North Carolina's Water Quality Assessment and Impaired Waters List (305(b) and 303(d) Report) in 2000 for chlorophyll *a* standard violation (*1*). However, the available chlorophyll *a* data from the Estuary do not support the conclusion that the Neuse River Estuary is in danger of water quality standard violation (see the Supporting Information). Furthermore, an earlier study of the estuary and the river basin (*2*) indicated that nutrient inputs to the estuary have been either steady or decreasing from early 1970 to 2001. The decision to prepare a TMDL for the Neuse River basin apparently was made based on a combination of fish-kill incidents in mid-1990s and modeling studies.

This conflict between the TMDL decision and the compliance of water quality reflects the different interpretations of what is the "true" water quality status. All water quality models have prediction uncertainties, some of which can be quite large. So, which assessment is more reliable: the model forecast or the monitoring data? On the one hand, routine monitoring data are often regarded as "happenstance data" (*3*) not designed for inference about the water quality status of an entire basin. In other words, they may not have the proper spatiotemporal coverage to reflect the true concentration conditions. In addition, water quality response to management actions tends to have a time lag. As a result, using monitoring data alone will not give us confidence that the estimated water quality status is close to the truth. On the other hand, models are simplifications of the real world and a model's predictions are limited by the model's own structural limitations, as well as by the data used to calibrate the model. For the Neuse River Estuary, we observed large scale fish-kill incidents in the mid-1990s' and the rapid increase in concentrated animal farming in the river basin. These observations along with model predictions indicated a need for reducing nutrient input to the estuary to keep water quality in compliance. Both approaches (using data or using a model) can be justified, and both approaches can be justifiably criticized. It is often impossible to determine which source of information is more reliable. Consequently, the question should be whether we can combine these two sources of information to better support the water quality management decision making process. We believe that both assessments can, and should, be used to evaluate compliance and the adequacy of management actions. That is, even though the model is just that: a model, and even though it will always yield uncertain predictions, it has value in forecasting impacts (otherwise we would not be using it to develop the management plan). Likewise, lags, natural variability, and measurement uncertainty do not prevent useful inferences to be derived from measurements, even though adjustments to account for these assessment shortcomings may be required.

The objective of this study is not to compare the two approaches, but rather to introduce a Bayesian approach for combining preimplementation model forecasts with postimplementation measurements for assessing water quality standard compliance. In its simplest form, this Bayesian approach involves a variance-weighted combination of the model forecast and the postimplementation monitoring data. There is a well-established analytic basis for combining (or pooling) information (*4*); however, it requires measures of uncertainty (or variability) to weight each contributing piece of information by its precision. The logic of this approach is unassailable; the more precise (less uncertain) the information, the higher that information is weighted. So, measures of uncertainty are necessary to yield the most defensible basis for pooling information.

We illustrate the process of combining modeling results and monitoring data to obtain a better understanding of the

* Corresponding author phone: 919-613-8105; fax: 919-684-8741; e-mail: song@duke.edu.

Published on Web 06/13/2007

water quality under study. For this paper, we collected available monitoring data for total nitrogen and chlorophyll *a* concentrations in the entire estuary from 1992 to 2000. Our example is designed to illustrate (1) the process of estimating prior distributions based on model output, (2) the sequential updating process to combine model output and monitoring data, and (3) the use of a predictive distribution for water quality standard compliance assessment.

Prior information on nitrogen concentration was based on the Bayesian SPARROW model, which produces estimates of TN loading (with uncertainty) from the watershed. The estimated TN loadings were converted to the estimated TN concentration using (1) historical flow distribution information (as suggested by Qian et al. (*5*)) and (2) the flow-TN concentration relationship developed in ref *6*. These two prior distributions of estuary TN concentrations are evaluated using the procedure described in Section 2.5. The selected prior distribution for TN inflow concentration was then used in a chlorophyll *a* concentration model developed in (*7*) to produce a model estimated estuary chlorophyll *a* concentration distribution (the prior distribution). Monitoring data were used to estimate the posterior predictive distribution of chlorophyll *a* concentrations in the estuary. The posterior distribution is used to assess water quality standard compliance.

Our study has three steps: (1) deriving prior distributions of TN concentration distribution parameters (Sections 3.1.1 and 3.1.2); (2) deriving prior distributions of chlorophyll *a* concentration distribution parameters based on the priors derived in (1) (Section 3.2.1); and (3) updating the priors from (1) and (2) using observations (Sections 3.1.3 and 3.2.2). The sequentially updated posterior distributions of chlorophyll *a* concentration are used as basis for water quality assessment. In the rest of this paper, we describe the study area, the models, and data set in section 2.1, including the set up of our example. The Bayesian updating procedure is briefly discussed in section 2.3. Section 3 presents the results of the example, followed by a discussion. Because prior distributions can come from competing water quality models, we discuss a process of evaluating these models based on the posterior distributions.

2. Materials and Methods

2.1. Study Area. Data for our example are from the Neuse River Estuary, in North Carolina. The Neuse River Basin has been the subject of intense study over the last 10 years due to large scale fish-kill events in the mid 1990s. The Neuse River estuary is seen as threatened by cultural eutrophication from the rapid urban development in the upper Neuse River Basin and expanded agriculture activity (especially hog farming) in the middle and lower Neuse River Basin. The state of North Carolina has a chlorophyll *a* concentration criterion of 40 μg/L to protect designated uses associated with eutrophication. Since in the mid-1990s, chlorophyll *a* had, on occasion, exceeded this criterion, a TMDL was required for the Neuse with an objective of reducing total nitrogen (TN) loading to the estuary so that compliance with the criterion would be achieved. As part of the TMDL work, a spatially referenced regressions on watershed attributes (SPARROW) model was developed for the Basin (*8*); subsequently, a Bayesian SPARROW model was developed for the same study area (*5*). The impact of watershed nitrogen loading to the Neuse River Estuary was assessed largely based on several modeling studies.

2.2. Notation. We will use the following convention for our mathematical notation. The response variable is represented by *Y*. We will use three water quality variables as response (TN loading, TN concentration and chlorophyll *a* concentration, all in natural logarithm). In all cases, the response variable *Y* is assumed to have a normal distribution, i.e., $Y \sim N(y|\mu, \lambda)$, where μ is the mean and λ is the precision (reciprocal of variance). We use $\{\mu_0, n_0, \alpha, \beta\}$ and $\{\mu_n, n_n, \alpha_n, \beta_n\}$ to denote parameters from the prior and posterior distributions of (μ, λ), respectively. When necessary, subscripts (*L*, *N*, *a* for TN loading, TN concentration, and chlorophyll *a* concentration, respectively) will be used to qualify μ and λ. We use $E(\cdot)$ and $V(\cdot)$ to represent the expectation (or mean) and variance of a random variable, and $\hat{E}(\cdot)$ and $\hat{V}(\cdot)$ are the estimated mean and variance from previous modeling results.

2.3. Bayesian Updating. The Bayesian approach provides a rigorous way for information updating, through repeated application of Bayes theorem:

$$\pi(\theta|Y) = \frac{\pi(\theta)L(Y|\theta)}{\int_\Theta \pi(\theta)L(Y|\theta)d\theta} \quad (1)$$

where θ is the unknown parameter of interest, $\pi(\theta)$ is the prior probability density function, summarizing our knowledge of the parameter before observing the data *Y*, $L(Y|\theta)$ is the likelihood function, a function of the unknown parameter θ. The likelihood function summarizes the levels of support from the data to various θ values. The posterior distribution combines the prior knowledge and the likelihood, representing our knowledge about θ after observing the data. When important prior information is available and is desired to be included in the decision process, the Bayesian approach is much more effective than the classical statistics approach (*9*). The Bayesian approach also provides a framework for effective information accumulation by repeated application of Bayes theorem each time using the posterior from the previous application as the prior. This sequential updating process is the main focus of this study.

There are significant challenges in using Bayes theorem. First, many scientists are unaccustomed to using probability distributions to summarize knowledge. Although it is argued that uncertainty should be quantified probabilistically for both philosophical and practical reasons (*9*), it is nevertheless an unfamiliar tool for most of us. Second, Bayesian computation is often complicated.

Fortunately, water quality concentration variables can often be approximated by the log-normal distribution (see justifications by Ott (*10*)). As a result, the natural logarithm of concentration data can be approximated by the normal distribution and it is possible to use the natural conjugate family of prior distributions for a probabilistic water quality standard compliance study. Because the log-normal distribution is defined in terms of the natural logarithm, we use the natural logarithm exclusively and the notation log.

The basic setup for Bayesian updating using normal conjugate priors is as follows. Suppose the log concentration distribution is normal with unknown mean (μ) and standard deviation (σ). To simplify the notation, we will use precision $\lambda = 1/\sigma^2$ instead of σ. The conjugate prior distribution of μ, λ is the normal-gamma distribution (*11*), defined as follows:

$$\pi(\mu, \lambda) = N(\mu|\mu_0, n_0\lambda)Ga(\lambda|\alpha, \beta) \quad (2)$$

This distribution can be explained in the following way. First, the unknown precision is assumed to have a Gamma distribution with parameters α, β, i.e., $Ga(\lambda|\alpha, \beta)$. Conditional on λ, the unknown mean μ is assumed to have a normal distribution with mean μ_0 and precision $n_0\lambda$: $N(\mu|\mu_0, n_0\lambda)$. The prior information is summarized in four parameters: μ_0 (the prior guess of the mean), n_0 ("prior sample size," reflecting the confidence we have on μ_0), and α, β (describing our knowledge of λ). With uncertainty about μ and λ defined by the prior distribution in eq 2, the predictive distribution of *Y* is now a Student *t* distribution with location parameter

$\mu = \mu_0$, a scale parameter $\tau = n_0/(n_0 + 1)\ \alpha/\beta$, and degrees of freedom df $= 2\alpha$, or $\tilde{y} \sim St(\bar{y}|\mu,\ \tau,\ \mathrm{df})$. This Student t distribution has a mean μ_0 and variance $\mathrm{df}/(\mathrm{df}-2)\ 1/\tau$, which summarizes our prior uncertainty about the concentration distribution (see ref *11* for details).

When observing data Y (summarized in terms of sufficient statistics: sample size n, sample mean $\bar{x}$, and sample standard deviation s), the posterior distribution of μ, λ is also a normal-gamma distribution:

$$\pi(\mu,\ \lambda|Y) = N(\mu|\mu_n,\ n_n\lambda)Ga(\lambda|\alpha_n,\ \beta_n) \tag{3}$$

where $\mu_n = (n\bar{x} + n_0\mu_0/n + n_0)$ (weighted average of the prior mean and the sample mean), $n_n = n + n_0$, $\alpha_n = \alpha + n/2$, and $\beta_n = \beta + (1/2)ns^2 + (1/2)(n_0 n(\mu_0 - \bar{x})^2/n_0 + n)$. Equation 3 defines the distribution of the mean and precision of a log concentration variable. The predictive distribution of future log concentration is again a Student t distribution with location parameter $\mu = \mu_n$, a scale parameter $\tau = (n_n/n_n + 1)\ (\alpha_n/\beta_n)$, and degrees of freedom df $= 2\alpha_n$. This posterior predictive distribution is the updated distribution of the log concentration combining both the prior knowledge and information represented in the data.

When new data are available, the current posterior in eq 3 becomes the new prior, and a new posterior distribution of μ and λ and the predictive distribution of the log concentration are derived combining the new prior and the new data.

2.4. Selection of the Prior. Selecting the prior involves determining the values of those four parameters ($\theta = \{\mu_0, n_0, \alpha, \beta\}$) defining the prior distribution in eq 2. Our knowledge of the concentration distribution can help us choose the appropriate values of θ. For example, μ_0 represents the prior estimate of the mean. We can assess the quality of the initial estimate by using n_0. If we are uncertain about the estimate, a small n_0 should be used, and vice versa. Often we can assess the uncertainty on λ (or specifically on σ) using judgment on the range of the log concentration values and assuming the range is close to the 95% credible interval. The distance of the range is close to $4 \times \sigma$, which provides a rough estimate of the mean of $\lambda = 1/\sigma^2$. Under the conjugate prior, $\lambda \sim Ga(\lambda|\alpha, \beta)$, the expected value of λ is α/β. Knowing the mean of λ leads to an estimate of the ratio α/β. Larger values of α, β indicate higher confidence on the estimate and vice versa. In general, we suggest that values of α, β be chosen to be in the same order of magnitude as the "prior sample size" used to evaluate the precision. This is because the posterior mean precision α_n/β_n is a balance between the prior mean precision and the sample precision. The sample precision is weighted by the data sample size. Therefore, values of α, β comparable to the prior sample size are conceptually in agreement with the mathematics. However, since the prior sample size is often a concept rather than a real number, we concede that some judgment is inevitable. We suggest, again based on our experience, the values of α, β be similar to the data sample size if we want to give the prior information a weight similar to that given to the data.

When a water quality model is used to predict the water quality constituent concentration distribution, more information can be extracted as the basis for the prior parameters. For example, Qian et al. (*12*) discussed the use of Markov chain Monte Carlo simulation as a parameter estimation method which leads to the marginal posterior distributions of the mean water quality constituent concentration and its variance. These two marginal distributions are presented in terms of random samples of μ and λ, which can be used to calculate the respective means and variances $\hat{E}(\lambda)$, $\hat{V}(\lambda)$, $\hat{E}(\mu)$, and $\hat{V}(\mu)$. Using the method of moments, we can estimate the unknown prior distribution parameters. That is, using the four formulas of the expected values and the variances of μ and λ to solve for the four unknown parameters μ_0, n_0, α, β in

$$E(\mu) = \mu_0, \quad V(\mu) = \beta/(n_0(\alpha - 1))$$
$$E(\lambda) = \alpha/\beta, \quad V(\lambda) = \alpha/\beta^2$$

(see ref *11*, pages 136 and 434) will yield estimates:

$$\hat{\beta} = \hat{E}(\lambda)/\hat{V}(\lambda), \quad \hat{\alpha} = \hat{E}(\lambda)\hat{\beta}$$
$$\hat{\mu}_0 = \hat{E}(\mu), \qquad \hat{n_0} = \hat{\beta}/[(\hat{\alpha} - 1)\,\hat{V}(\mu)] \tag{4}$$

2.5. Evaluating the Prior. Competing prior distributions can come from competing water quality models. When two or more models are used to assess water quality it is important to know which model is better supported by the data. The traditional approach of model evaluation and assessment can be difficult with limited monitoring data. From a practical perspective, a good model should provide adequate prediction of the water quality parameter distribution. The most direct way of comparing two models is through the use of the likelihood ratio. In this particular case, our prior models are summarized by the prior predictive distribution (the student t distribution). The likelihood ratio is simply the ratio of the density functions evaluated at the values of observed concentration data. A likelihood ratio is the basis for many Bayesian model comparison methods. Because of the simplicity of our prior model, a likelihood ratio would be enough.

3. Results

The results are presented in two parts. In the first part, we present the process of prior selection under two scenarios, followed by the updating results for the chlorophyll *a* concentration distribution.

3.1. Priors for log TN Concentration Distribution Parameters. The basis for the prior probability determination is a nonlinear regression model (the Bayesian SPARROW model in ref *5*), which predicts TN loading to the Neuse River Estuary. The model-predicted TN loading is represented as a conditional log-normal distribution, with log mean and log precision estimated using a Markov chain Monte Carlo (MCMC) simulation method. In other words, the log TN loading is predicted as a normal distribution with mean (μ_L) and precision (λ_L) estimated using MCMC. These MCMC samples of log TN load mean and TN load precision are the basis for developing priors of TN concentration in the Neuse River Estuary. We present two competing priors, one directly from the Bayesian SPARROW model output (assuming independence between TN concentration and flow) in Section 3.1.1 and the other derived by combing the Bayesian SPARROW model output and a regression model of the TN concentration−flow relationship in Section 3.1.2.

3.1.1. Priors from a TN loading model. Assuming the natural logarithm (log) of TN concentration [log(TN)] follows a normal distribution with unknown mean μ_N and precision λ_N: $\log(\mathrm{TN}) \sim N(\mu_N, \lambda_N)$. From the Bayesian SPARROW model, we have a series of MCMC samples of the mean log nitrogen loading μ_L and its precision λ_L (see ref *5*). Considering a typical flow condition designated by log mean of q and precision of λ_q, we can convert the MCMC samples of log loading to samples of mean nitrogen concentrations ($\mu_N = \mu_L - q$ with variance $1/\lambda_N = 1/\lambda_L + 1/\lambda_q$). Using these samples of μ_N and λ_N we calculate the sample means and variances, or, $\hat{E}(\mu_N)$, $\hat{V}(\mu_N)$, $\hat{E}(\lambda_N)$, $\hat{V}(\lambda_N)$ and the prior parameters are calculated using eq 4. This approach ignores the (positive) correlation between flow and concentration and thereby overestimates the variance. The estimated prior distribution parameters are $\mu_0 = 7.63$, $n_0 = 45.10$, $\alpha = 317$, $\beta = 275$. The prior sample size (n_0) is close to the number of TN load observations ($n = 44$) used to develop the Bayesian SPARROW model. The

large values of α and β resulted in a highly concentrated prior distribution on the precision (the 95% credible interval of λ is between 1.03 and 1.28; or, the same interval for the standard deviation is between 0.88 and 0.98). We used α = 31.7 and β = 27.5 to reflect the fact that we ignored the correlation between flow and concentration. The resulting α, β values are close to the sample size used to fit the SPARROW model (44), but much smaller than the TN observation sample size. This prior has a 95% credible interval for the standard deviation of (0.79, 1.13). Dividing both α, β by 10 is equivalent to multiplying the prior variance of λ_c by 10, resulting in a discount in the estimated average precision of 1.15 (=α/β). Using this discount allows us to recognize the possible overestimation of the variance and reduce the influence of the prior on the posterior. The prior distribution derived from SPARROW model output will be referred to as the *SPARROW prior* in the rest of the paper. See the Supporting Information for additional discussion on some issues of using model output for generating prior distribution.

3.1.2. Priors from a TN Loading Model and a Flow-TN Model. Stow and Borsuk (*6*) analyzed the flow and TN concentration relationship for the Neuse River and determined that log(TN) can be predicted by a linear function of log flow

$$\log(\mathrm{TN}) = \beta_0 + \beta_1 \log(\mathit{flow}) + \epsilon \tag{5}$$

Adding β_1 log(TN) to both sides of eq 5 we have

$$(\beta_1 + 1)\log(\mathrm{TN}) = \beta_0 + \beta_1 \log(L) + \epsilon$$

or

$$\log(\mathrm{TN}) = \beta_0^* + \beta_1^* \log(L) + \epsilon^*$$

where L = flow × TN is the TN loading, $\beta_0^* = \beta_0/(\beta_1 + 1)$, $\beta_1^* = \beta_1/(\beta_1 + 1)$, and $\epsilon^* = \epsilon/(\beta_1 + 1)$. Using this relationship, we show that the conditional mean of the log TN concentration is

$$E(\log(\mathrm{TN})) = \beta_0^* + \beta_1^* \log(L) \tag{6}$$

and the conditional variance of the log TN concentration is

$$V(\log(\mathrm{TN})) = V(\epsilon^*) \tag{7}$$

Both E(log(TN)) and V(log(TN)) are random variables (because β_0^*, β_1^*, L, and $V(\epsilon^*)$ are random) and $\mu_N = E(\log(\mathrm{TN}))$, $\lambda_N = 1/V(\log(\mathrm{TN}))$. To combine the concentration-flow model of eq 5 and the Bayesian SPARROW model, we pair each MCMC sample of μ_L [as log(L)] and λ_L [as $1/V(\log(L))$] from SPARROW with a set of randomly generated coefficients of β_0^*, β_1^*, and $V(\epsilon^*)$ to generate a set of random samples of E(log(TN)) and V(log(TN)) to obtain a set of samples of μ_N and λ_N. The sample means and variances of the generated μ_N and λ_N are used to estimate the prior parameters using eq 4. The estimated prior distribution parameters are $\mu_0 = 4.98$, $n_0 = 2.42$, α = 27380, β = 6744. The prior reflects the strong belief that the log(TN) distribution has a standard deviation of 0.495 (with a 95% credible interval of (0.493, 0.499)). Because the regression model was developed using the data from a station upstream of the estuary, we used α = 2738, β = 674, which still has a 95% interval for σ to be (0.487, 0.506). The α, β values are the same order of magnitude as the sample size used by Stow and Borsuk (*6*). They are larger than the TN observation sample size in our study. This prior will be referred to as the *flow-adjusted prior* in the rest of the paper.

The predictive distributions of log(TN) represented by the two prior distributions (Figure 1, left panel) show notable differences. The use of the flow-concentration relationship reduced the variance of the predictive distribution. However, because of the high level of uncertainty in the regression coefficients, n_0 for the second prior is small. In comparison, the SPARROW prior is highly certain about the estimated location and the flow-adjusted prior has a higher level of certainty on the spread.

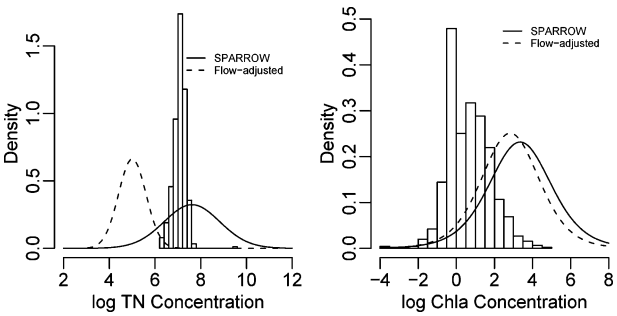

FIGURE 1. Prior predictive distributions: the left panel shows the two priors of log TN concentration (μg/L); the right panel shows the two priors of log chlorophyll *a* concentration (in μg/L) derived using respective log TN priors. The priors are compared to data collected from 1992 to 2000, shown by respective histograms.

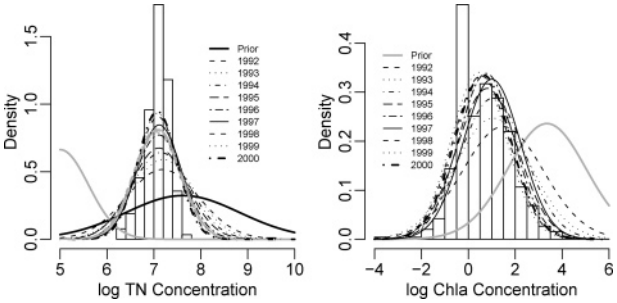

FIGURE 2. Sequentially updated posterior predictive distributions: the left panel shows the sequentially updated posterior distributions of log TN concentration (the black lines are the results using the SPARROW prior, and the gray lines are the results using the flow-adjusted prior); the right panel shows the sequentially updated posterior distributions of log chlorophyll *a* concentration (using the chlorophyll *a* prior derived from the SPARROW prior). Both are compared to their respective data shown in histograms.

3.1.3. Posterior Distributions of TN. The TN concentration data were collected in the upper section of the Neuse River Estuary near New Bern, North Carolina from 1992 to 2000. Using the two sets of prior distribution parameters, we sequentially updated the log TN concentration distribution using data from each calender year (Figures 2, left panel; and Figure S1 of Supporting Information). The prior distributions either overestimate or underestimate the TN concentrations. However, the posterior mean log TN concentration shifted toward the center of the data quickly for both prior distributions. Sample standard deviations of the log TN concentrations vary from year to year between 0.13 and 0.39 (or 1.14 to 1.48 μg/L), closer to the flow-adjusted prior standard deviation (0.495 or 1.64 μg/L) than to the SPARROW prior (0.92 or 2.51 μg/L). The posterior predictive distributions from using the SPARROW prior distribution (Figures 2 left panel, and Figure S1 of Supporting Information) converges slowly in terms of the spread. Although the prior mean μ_0 from the flow-adjusted prior is farther away from the data samples means (between 6.9 and 7.2 in logarithm, or 992 and 1139 μg/L), the posterior predictive mean quickly converged to the data mean because of a small n_0. The prior parameters α and β have a subtle impact on the posterior predictive distributions. We note that the spread of the predictive distributions based on the SPARROW prior distribution gradually decreases as more data were included,

while the spread of the predictive distributions using the flow-adjusted prior distribution barely changes.

The likelihood ratio of the two priors are calculated for each of the 9 years. The log likelihood ratio varies from 68 to 433, indicating that the SPARROW prior has stronger support from the data.

3.2. Updating Chlorophyll *a*. Because the Neuse River Estuary is subject to a nitrogen TMDL to achieve compliance with a chlorophyll *a* criterion, it is important that information with regard to total nitrogen be translated to chlorophyll *a* concentration. To accomplish this, we use the simple linear regression chlorophyll *a* model developed in (*7*) to convert the TN prior information into that of chlorophyll *a*.

3.2.1. Converting Priors on TN to Priors on Chlorophyll a. The chlorophyll *a* regression model developed by Borsuk et al. (*7*) is of the following form:

$$\log(\text{chla}) = \begin{cases} \beta_1 + \log(\vartheta)(T-20) + \beta_2(15.7 - \log(\text{flow})) + \beta_4 e^{\text{TN}} & \text{if } \log(\text{flow}) \le 15.7 \\ \beta_1 + \log(\vartheta)(T-20) + \beta_3(15.7 - \log(\text{flow})) + \beta_4 e^{\text{TN}} & \text{if } \log(\text{flow}) > 15.7 \end{cases}$$

Although a log-normally distributed TN will strictly not result in a normally distributed log(chla), the log(chla) distribution is dominated by the distribution of log(flow) and is approximately normal. The prior distribution on log chlorophyll *a* concentration (log(chla)) was developed using Monte Carlo simulation. TN concentration samples are from Section 3.1.1, samples of log flow (log(flow)) and water temperature (T) were generated based on historical data, samples of the coefficient ϑ were generated based on the prior distribution suggested by Borsuk et al. (*7*). Random samples of the regression model coefficients (β_1–β_4) were generated based on the original model fit information (a multivariate normal distribution). These random samples were then entered into the chlorophyll *a* model to produce 10 000 Monte Carlo samples of log(chla) values, from which we estimated the log mean and log variance. Assuming chlorophyll *a* concentration is a log-normal random variate, the log mean and log variance provided an estimate of μ_0 (=3.37) and the ratio of β/α (=1/$\hat{E}(\lambda) = 0.4$, hence $\hat{E}(\lambda) = 2.5$). We are fairly confident about the estimated log variance because of the large number of Monte Carlo samples. Consequently, we used $\beta = 40$ and $\alpha = 100$ to reflect our confidence on the estimated variance of 0.4 (estimated mean precision 2.5). The prior distribution on the precision has a standard deviation of 0.06. However, we have less confidence in the estimated log mean, a combination of results from two empirically based models. We chose to use $n_0 = 40$, a value slightly smaller than the number of observations used to fit the Bayesian SPARROW model. The resulting prior predictive distribution of log(chla) is compared to the observed chlorophyll *a* data (Figure 1, right panel). An alternative prior distribution can be generated using the TN concentration samples from Section 3.1.2 (Figure 1, right panel, dashed line). Both prior distributions overestimate the chlorophyll *a* concentrations. The alternative prior was not used in the subsequent updating because the difference between the two priors is small.

3.2.2 Posterior Distributions of Chlorophyll a Concentrations. The sequentially updated posteriors converge to the center of the data mass (Figures 2, right panel). When plotting these posterior distributions against data from individual years (Figure S2, Supporting Information), we note that chlorophyll concentrations have a significant year to year fluctuation. In the first 2 years (1992 and 1993), all observed log chlorophyll concentrations were above 0 (or 1 μg/L). In the next 3 years, the concentration shifted lower. In 1997 the concentration rebounded slightly and in 1998 we observed the lowest concentration levels. In the last 2 years, the chlorophyll *a* concentrations increased rapidly. These annual changes resulted in obvious differences between the priors and the posteriors.

4. Discussion

We presented a Bayesian process of updating model output using subsequent monitoring data. The updated results are a combination of our understanding of the system (reflected in model results) and the information in the data. Our method requires a water quality model and historical data necessary to convert output from the model to the metric of interest. In our example, we used the SPARROW model output to produce a prior distribution for TN loading, which is then modified using historical data (or another model) to obtain a prior for chlorophyll *a* concentration, the water quality constituent regulated under the TMDL program for the estuary. These two information requirements will be present in almost every application using our method.

Although the SPARROW prior only slightly overestimated the TN concentration, the derived chlorophyll *a* prior overestimates the chlorophyll *a* concentration substantially, even when the TN is underestimated as in the flow-adjusted prior (Figure 1). This overestimation is not unique to the regression model we used. In fact, the Neuse Estuary TMDL program was established in part because of the consistent overestimation of chlorophyll *a* concentration by many modeling studies. Retrospectively, this overestimation is obvious in our study because of the availability of a large amount of data, an anomaly of water quality assessment studies. Although, as one reviewer suggested, we would conclude that the Neuse River Estuary is in compliance faster without using the model derived prior information, it is impossible to ignore this source of information. More importantly, as we discussed in introduction, although monitoring data from the Neuse River Estuary did not suggest a consistent violation of the North Carolina's chlorophyll *a* criterion, circumstantial evidence (e.g., large scale fish-kills and rapid growth of concentrated animal farming in the watershed) suggested that nitrogen enrichment is a potential threat. This evidence is further reinforced by modeling results. As a result, a nitrogen TMDL was developed to meet the chlorophyll *a* criterion.

The study presented in this paper is a first step in our adaptive management research. The results presented in this paper will be adapted into our subsequent studies of combining information to updating watershed model parameters as a means to quantify the effects of various water quality management plans implemented in the Neuse River watershed through the TMDL program. Obviously, neither the SPARROW model nor the Estuary chlorophyll *a* regression model adequately predicted nitrogen and chlorophyll *a* concentrations in the Neuse River Estuary. The logical question is then why. By coupling the watershed nitrogen model and the estuary chlorophyll *a* model, we hope to quantitatively identify weaknesses in our models. The improved models should be able to better assess the effectiveness of watershed water quality management strategies on estuary water quality.

Acknowledgments

We thank the associate editors, George Arhonditsis and Craig A. Stow, and four reviewers for their constructive comments and suggestions. This research was funded under EPA STAR grant no. R830883.

Supporting Information Available

(1) Considerations of generating prior from model output. (2) The use of our Bayesian updating process for water quality

standard compliance assessment. (3) Additional figures illustrating the Bayesian updating process. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) North Carolina Division of Water Quality; The N.C. Water Quality Assessment and Impaired Waters List (305(b) and 303(d) Report): Final 303(d) List, 2000. http://h2o.enr.state.nc.us/tmdl/Docs_303/2000/nc2k_list.pdf

(2) Qian, S. S.; Borsuk, M. E.; Stow, C. A. Seasonal and long-term nutrient trend decomposition along a spatial gradient in the Neuse River watershed. *Environ. Sci. Technol.* **2000**, *34*, 4474−4482.

(3) Berthouex, P. M.; Brown, L. C. *Statistics for Environmental Engineers*; Lewis Publishers: Boca Raton, FL, 1994.

(4) Draper, D.; Gaver, D. P.; Goel, P. K.; Greenhouse, J. B.; Hedges, L. V.; Morris, C. N.; Tucker, J. R.; Waternaux C. M. *Combining Information − Statistical Issues and Opportunities for Research*; National Academy Press: Washington, DC, 1992.

(5) Qian, S. S.; Reckhow, K. H.; Zhai, J.; McMahon, G. Nonlinear regression modelling of nutrient loads in streams—A Bayesian approach. *Water Resour. Res.* **2005**, *41* (7), W07012.

(6) Stow, C. A.; Borsuk, M. E. Assessing TMDL effectiveness using flow-adjusted concentrations: a case study of the Neuse River, North Carolina. *Environ. Sci. Technol.* **2003**, *37*, 2043−2050.

(7) Borsuk, M. E.; Stow, C. A.; Reckhow, K. H. Confunding effect of flow on estuarine response to nitrogen loading. *J. Environ. Eng.* **2004**, *130* (6), 605−614.

(8) McMahon, G.; Alexander, R. B.; Qian, S. S. Support of TMDL programs using spatially referenced regression models. *J. Water Resour. Plan. Manage.* **2003**, *129* (4), 315−329.

(9) Berger, J. O. *Statistical Decision Theory and Bayesian Analysis*; Springer-Verlag: New York, 1985.

(10) Ott, W. R. *Environmental Statistics and Data Analysis*; Lewis Publishers: Boca Raton, FL, 1995

(11) Berdardo, J. M.; Smith, A. F. M. *Bayesian Theory*; Wiley: New York, 1994.

(12) Qian, S. S.; Stow, C. A.; Borsuk, M. E. On Monte Carlo methods for Bayesian inference. *Ecol. Modell.* **2003**, *159*, 269−277.

(13) Smith, E. P.; Ye, K.; Hughes, C.; Shabman, L.; Statistical assessment of violations of water quality standards under section 303(d) of the Clean Water Act. *Environ. Sci. Technol.* **2001**, *35* (3), 606−612.

Received for review October 9, 2006. Revised manuscript received May 5, 2007. Accepted May 15, 2007.

ES062420F