Bayesian Framework for Water Quality Model Uncertainty Estimation and Risk Management

Mohamed M. Hantush, Ph.D., A.M.ASCE[1]; and Abhishek Chaudhary[2]

Abstract: A formal Bayesian methodology is presented for integrated model calibration and risk-based water quality management using Bayesian Monte Carlo simulation and maximum likelihood estimation (BMCML). The primary focus is on lucid integration of model calibration with risk-based water quality management and total maximum daily load (TMDL) estimation under conditions of uncertainty. The sources of uncertainty considered in the analysis are modeling errors, observational data errors and fuzziness of the water quality standard. The difference between observed data or transformation thereof and corresponding model response is assumed to follow first-order Markov process, a specific case of which is statistically independent Gaussian errors. The BMCML method starts with sampling parameter sets from prior probability distributions of the model parameters and uses Bayes theorem and the maximum likelihood technique to estimate the triplicate (variance of residual errors, bias and autocorrelation coefficient of total errors) for each parameter set and the corresponding likelihood value. By approximating integration over the entire parameter space discretely, analytical expressions are derived for the cumulative probability distributions of model outputs and probability of violating water quality standards. The solution of the TMDL problem and related margin of safety (MOS) is then framed in the context of the developed Bayesian framework. Three example applications of varying complexities are utilized to demonstrate the versatility of the Bayesian methodology for water quality management. The BMCML methodology is validated using a hypothetical lake-phosphorus model and familiar statistical benchmarks. It is shown that the risk-based framework can estimate the reliability of an arbitrarily selected MOS as demonstrated in the Fork Creek bacteria and Shunganunga Creek dissolved oxygen TMDL case-studies. It is also shown that neglecting covariation among model parameters (i.e., by sampling parameter values from their posterior marginal distributions) influences the estimation of probability of exceedance and could potentially lead to the overestimation of the MOS at low risk levels. **DOI: 10.1061/(ASCE)HE.1943-5584.0000900.** *© 2014 American Society of Civil Engineers.*

Author keywords: Water quality; Model; Uncertainty; Bayes; Monte Carlo; Calibration; Maximum likelihood; Risk management; Total maximum daily load (TMDL); Margin of safety.

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

Introduction

Despite the proliferation of methods for uncertainty estimation of complex hydrologic and water quality models over the last two decades, these efforts have disproportionally focused on model calibration and forecast, yet much less effort has been devoted to the implications of uncertainty on risk-based water quality management in general and the Clean Water Act Total Maximum Daily Loading (TMDL) program in specific. Computer modeling is an essential component of TMDL development typically used for predicting the required reduction in pollutant loadings such that water quality standards are met with an adequate margin of safety (MOS). The MOS is introduced to account for uncertainty of observational data and uncertainty in the relationship between pollutant load and quality of the receiving water body (U.S. EPA 1999). Despite the fact that most TMDLs lacked rigorous uncertainty analysis and that the MOS is often arbitrarily selected in these TMDLs (Dilks and Freedman 2004), a few investigators nevertheless have attempted to link model forecast uncertainty to risk-based water quality management and TMDL development (Borsuk et al. 2002; Zhang and Yu 2004; Zheng and Keller 2007; Liu et al. 2008; Faulkner 2008; Zheng and Keller 2008; Chin 2009; Shen and Zhao 2010; Patil and Deng 2011). These efforts were research oriented and of limited scope. Borsuk et al. (2002) were among the first to apply the Monte Carlo (MC) procedure for TMDL development and the estimation of the MOS on the basis of observational data and modeling errors. Preserving generalized likelihood uncertainty estimation (GLUE) equifinality concept, Zheng and Keller (2007) presented what is referred to as the management objectives constrained analysis of uncertainty (MOCAU), which emphasizes the likelihoods of multiple management variables, and applied the methodology to develop a diazinon TMDL in a southern California watershed (Zheng and Keller 2008).

In typical TMDLs, the MOS is either implicitly accounted for by means of conservative assumptions or explicitly computed as some fraction of water body loading capacity (Dilks and Freedman 2004; Shirmohammadi et al. 2006; Novotny 2003). Zhang and Yu (2004) and Patil and Deng (2011) considered model uncertainty and related the MOS to the standard deviations of simulated water quality parameters. The shortcomings in these approaches are an opaque degree of protection and an ambiguous relationship between analysis uncertainty and the MOS (Dilks and Freeman 2004). The National Research Council (NRC 2001) cautioned against the practice of arbitrary selection of the MOS and recommended uncertainty analysis as the base for MOS determination (see also

[1]Research Hydrologist, Land Remediation and Pollution Control Division, National Risk Management Research Laboratory, ORD, USEPA, 26 West Martin Luther King Dr., Cincinnati, OH 45268 (corresponding author). E-mail: hantush.mohamed@epa.gov

[2]Ph.D. Candidate, Institute for Environmental Engineering, ETH Zürich, HPZ 32.2, John-von-Neumann-Weg 9, 8093 Zürich, Switzerland; formerly, Environmental Engineer, Pegasus Technical Services Inc., 46 E. Hollister St., Cincinnati, OH 45219.

Note. This manuscript was submitted on August 23, 2012; approved on August 30, 2013; published online on September 2, 2013. Discussion period open until November 19, 2014; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, © ASCE, ISSN 1084-0699/04014015(14)/$25.00.

Reckhow 2003; Novotny 2004). A probabilistic framework therefore becomes more than just a convenient alternative approach. Rather than making conservative assumptions that often result in an overdesign or arbitrarily selecting a safety factor, a more transparent approach is to fully utilize uncertainty in measured or model predicted values of water quality parameters and insure compliance with adequate reliability or, equivalently, acceptable risk of violating the standard. A byproduct of such a probabilistic approach is a formal means for factoring uncertainty into the calculation of the MOS and an explicit statement of the magnitude of protection provided by the MOS. Conversely, the TMDL and associated MOS can be estimated given a desirable degree of protection (e.g., based on some policy decision).

Although selecting a fixed numerical water quality criterion is a standard practice in TMDLs development, in the classification of the trophic status of water bodies, the trophic conditions generally tend to be fuzzy around their boundaries (e.g., Wang et al. 2009; Vollenweider et al. 1998). The border values separating the trophic status typically vary within ranges of values (e.g., Chapra and Tarapchak 1976); consequently, it makes more sense to treat them as random variables constrained within realistic range values.

In light of this rationalization, a formal methodology based on mathematical rigor and consistent probabilistic framework is obviously needed to unambiguously relate a TMDL to analysis uncertainty while explicitly relating the MOS to the underlying risk. Uncertainty in numerical criteria or classification standards is also considered. This paper presents a risk-based approach wherein the reliability of the TMDL, and associated MOS is determined on the basis of observational data and modeling uncertainty.

The presented approach combines Bayesian Monte Carlo (BMC) simulation (e.g., Dilks et al. 1992) with maximum likelihood estimation and borrows from the generalized likelihood uncertainty estimation (GLUE) methodology the concept of "equifinality" and importance of parameter sets (e.g., Beven and Binley 1992; Beven 2006) as opposed to utilizing the posterior marginal distributions of individual parameters as the base for constructing model forecasts and risk assessment. Previous efforts linking model calibration and uncertainty estimation to water quality risk assessment have either used first-order approximation (FOA) of the first-two moments of model outputs by focusing exclusively on parametric uncertainty (e.g., Melching and Bauwens 2001; Zhang and Yu 2004), or have utilized more rigorous probabilistic approaches (e.g., Markov Chain Monte Carlo (MCMC) and other sampling techniques) to estimate the risk on the basis of posterior marginal distributions of the parameters (Borsuk et al. 2002; Zheng and Keller 2007; Liu et al. 2008; Shen and Zhao 2010; Patil and Deng 2011). Here, the traditional approach to likelihood estimation is adopted whereby observation-model errors are assumed to be autocorrelated and have a bias. We demonstrate that Bayesian formulation facilitates not only a complete forecast (i.e., cumulative distribution of model outputs), but also integrates total uncertainty to risk management on firm application of the rules of mathematical probability. As with the GLUE methodology, an important trait of the presented Bayesian methodology is that interactions among parameters in a Monte Carlo (MC) simulation are implicitly reflected by the likelihood weight associated with each sampled parameter set. The objectives of this paper are to (1) present a new model calibration and uncertainty estimation methodology using BMC and maximum likelihood estimation; (2) succinctly integrate model calibration and uncertainty estimation to risk management; (3) mathematically frame the concept of MOS; and (4) verify the methodology and demonstrate its applicability to a synthetic water quality problem and case-study TMDLs.

The manuscript is organized as follows. First, the risk-management problem is posed mathematically. The subsequent section is divided into several subsections that present the BMCML methodology and a comprehensive theoretical framework integrating modeling uncertainty to risk-based water quality management. A full subsection is devoted to the estimation of TMDL and MOS on the basis of analysis uncertainty. Three subsections validate and demonstrate applicability of the probabilistic framework to hypothetical and real quality management problems including two case-study TMDLs. The applications are followed by a discussion of the results. The manuscript ends with a summary and conclusions.

Theoretical Framework

Risk-Based Compliance

Management of the risk of violating water quality standards can be posed mathematically as follows:

$$P\{Y > Y^*\} \leq \beta \tag{1}$$

in which P = probability of the argument between parentheses; Y = water quality parameter or transformation thereof (e.g., log-transformed concentration) describing the state of the water body measured or estimated by an empirical or a mathematical model; Y^* = water quality numeric target or transformation thereof chosen to maintain designated use of the water body; and β = acceptable risk or probability of failure. Eq. (1) is a mathematical statement of the $(1-\beta)$ confidence of compliance first presented by Borsuk et al. (2002).

Y could also be a statistical metric of the water quality parameter (e.g., maximum or minimum daily/annual concentration, monthly average concentration, monthly median concentration, annual average concentration, 90th percentile of sampled concentrations, etc.) monitored over a preselected period in time (e.g., a month, three years, low flow periods, etc.), or a management variable such as nonattainment frequency or severity of exceedance (Zheng and Keller 2007). For instance, according to the U.S. EPA (1997), for the water body to be listed as unimpaired, no more than 10% of the samples ($\beta = 0.1$) collected from the water body over a specified time period should violate water quality standards. Y is random due to combinations of model inadequacies, natural variability, parametric uncertainty, and measurement errors. If Y is purely sampled values of the water quality parameter, then uncertainty in this case stems from natural (spatiotemporal) variability and measurement errors. In current TMDL practices, $Y^* = y^*$ is a state's designated numerical target, which is a fixed deterministic value. However, as argued in preceding paragraphs, in water quality assessment problems (e.g., classification of eutrophic status of water bodies), the use of crisp values of a water quality parameter (e.g., *chlorophyll-a* concentration) as border values separating different "trophic statuses" might not be prudent due to inherent uncertainties. In such cases, it makes more sense to treat the border value separating two trophic states, Y^*, as a random variable (Wang et al. 2009). A uniform distribution might be used when the best available information permits only a minimum and maximum value of Y^*. If the available information reveals a most likely value in addition to a range for the parameter, then a triangular distribution could be used.

Selection of an appropriate value of β is not a straightforward process; it involves a tradeoff between the cost of pollutant load reduction and the socioeconomically acceptable risk of violating water quality standards. A prescribed low value of β could lead

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

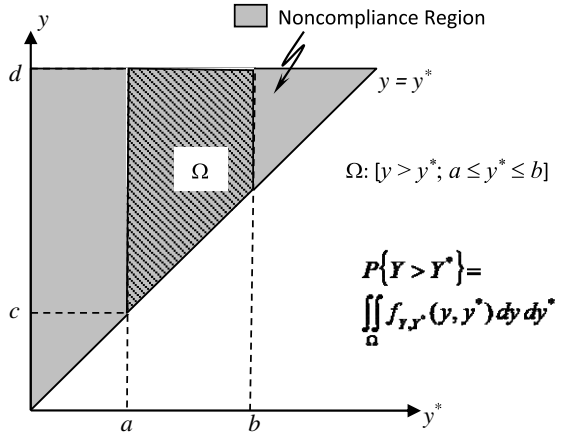

Fig. 1. Noncompliance domain for integration of joint probability density function, $f_{Y,Z}$

to conservative management actions that are too costly to sustain both economically and socially, whereas a high β runs the risk of compromising the integrity of the water body, which could lead to adverse human health conditions and undesirable environmental consequences.

We consider the general case of a random Y^* whose values are nonetheless constrained within the interval $[a, b]$. Let $Y \in [c, d]$. The probability of compliance $P\{Y \le Y^*\}$ can be obtained by integrating the joint probability density function $p^*_{Y,Y}(y, y^*)$ over the domain where $y \le y^*$ (shaded area in Fig. 1)

$$P[Y > Y^*] = \int_a^b \int_{y^*}^d p_{Y,Y^*}(y, y^*) dy dy^* \tag{2}$$

Noting the statistical independence of Y and Y^*, we can express the joint distribution, $p^*_{Y,Y}(y, y^*)$, as the product of the corresponding marginal probability density functions, $p_{Y,Y^*}(Y, Y^*) = p_Y(y) p_{Y^*}(y^*)$, and integrate over the shaded area in Fig. 1 to obtain a simpler form

$$P[Y > Y^*] = 1 - \int_a^b F_Y(y) p_{Y^*}(y) dy \tag{3}$$

where $F_Y(y)$ = cumulative density function of Y (CDF); and $p_{Y^*}(y)$ = probability density function (PDF) of Y^*.

Certain water quality problems require a lower bound (e.g., oxygen levels), in which case the probability of violation is

$$P[Y \le Y^*] = \int_a^b F_Y(y) p_{Y^*}(y) dy \tag{4}$$

For a fixed y^* value, $Y = y^*$, we have $p_{Y^*}(y^*) = \delta(y - y^*)$, which when substituted into Eq. (4) yields the familiar result $P\{Y \le y^*\} = F_Y(y^*)$. Eqs. (3) and (4) can be evaluated analytically for specific probability distributions. In general, $P[Y > Y^*]$ and $P[Y \le Y^*]$ can be evaluated using the MC method.

Modeling Uncertainty

Let $f(\mathbf{\Theta}, \mathbf{x})$ be the model that simulates the water quality parameter (also referred to as the state variable), Y, where $\boldsymbol{x}$ is a vector of input forcing (e.g., inflow, pollutant loading, precipitation, ET, etc.) and $\mathbf{\Theta}$ is a vector of the model parameters. In spatially explicit models, $\boldsymbol{x}$ also includes initial and boundary conditions. The state-model process, possibly after some transformation, can be mathematically posed as

$$g[Y(t)] = g[f(t; \mathbf{\Theta}, \boldsymbol{x})] + \epsilon(t) \tag{5}$$

where $Y(t)$ = true value of the state variable, Y, at time t; $g[\cdot]$ = transformation such that model residual errors are normally distributed; and ϵ = random variable representing model structural errors.

The corresponding measurement process is given by

$$g[O(t)] = g[Y(t)] + v(t) \tag{6}$$

in which $O(t)$ = observed value of the state variable; and $v(t)$ = measurement-induced random error assumed to be statistically independent of $\epsilon(t)$. The state and measurement processes [Eqs. (5) and (6)] can be combined to yield the overall observation-model relationship

$$g[O(t)] = g[f(t; \mathbf{\Theta}, \boldsymbol{x})] + \varepsilon(t) = v(t) + \epsilon(t) \tag{7}$$

where $\varepsilon(t)$ = overall error in the relationship between observed and model computed values of the state variable Y, often considered a Gaussian process. A specific case is when transformation is not needed, i.e., $g(Y) = Y$, whereby $\varepsilon(t)$ is often assumed to be zero-mean, independent, and normally distributed. However, the time series of residual errors generally exhibit non-stationarity and autocorrelation (e.g., Sorooshian and Dracup 1980; Vrugt et al. 2009) caused primarily by model structural errors. Further, residual errors might be biased (i.e., have a non-zero mean) indicating a model consistently overpredicting or underpredicting the observed data. For biased and auto-correlated $\varepsilon(t)$, the simplest model would be a first-order Markov (or autoregressive) process

$$\varepsilon(t) - \mu = \rho[\varepsilon(t-1) - \mu] + \omega(t), \omega(t) \sim N(0, \sigma_\omega^2) \tag{8}$$

where $\mu = E[\varepsilon(t)]$ is the bias of the overall error; $\rho = E[(\varepsilon(t) - \mu)(\varepsilon(t-1) - \mu)]$ is lag-one autocorrelation of the overall error; $\omega(t)$ is zero-mean, independent and normally-distributed residual error; and σ_ω^2 is variance of residual errors.

The likelihoods and Bayesian estimates of $\mathbf{\Theta}$ and $\mathbf{\vartheta} = (\mu, \rho, \sigma_\omega^2)$ are obtained below.

Bayesian Monte Carlo and Maximum Likelihood Method (BMCML)

In this section, we present a Bayesian methodology for updating prior parameter distributions and determining posterior estimates of probability of violation conditional on the observations. The methodology serves three purposes simultaneously: (1) model calibration; (2) uncertainty estimation; and (3) improved risk evaluation.

The Bayesian theorem furnishes a mathematical relationship for improving prior estimates of parameter distributions by combining preliminary information about those parameters with site-specific observational data that describe system behavior. Conditional on a set of m observed values of the state variable, $\boldsymbol{o} = (\boldsymbol{o}_1, \boldsymbol{o}_2, \ldots, \boldsymbol{o}_m)$, the posterior joint probability density function of model parameters, $\mathbf{\Theta}$, is given by the Bayes theorem (e.g., Gelman et al. 2003)

$$p_{\mathbf{\Theta}|\mathbf{O},\mathbf{X}}(\boldsymbol{\theta}|\boldsymbol{o}, \mathbf{x}) = \frac{p_{\mathbf{O},\mathbf{X}|\mathbf{\Theta}}(\boldsymbol{o}, \mathbf{x}|\boldsymbol{\theta}) p_{\mathbf{\Theta}}(\boldsymbol{\theta})}{\int_{\Omega_{\mathbf{\Theta}}} p_{\mathbf{O},\mathbf{X}|\mathbf{\Theta}}(\boldsymbol{o}, \mathbf{x}|\boldsymbol{\theta}) p_{\mathbf{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}} = \frac{p_{\mathbf{O}|\mathbf{X},\mathbf{\Theta}}(\boldsymbol{o}|\mathbf{x}, \boldsymbol{\theta}) p_{\mathbf{X}}(\mathbf{x}) p_{\mathbf{\Theta}}(\boldsymbol{\theta})}{\int_{\Omega_{\mathbf{\Theta}}} p_{o,x|\mathbf{\Theta}}(\boldsymbol{o}, \boldsymbol{x}|\boldsymbol{\theta}) p_{\mathbf{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{9}$$

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

where $p_{\Theta|O,X}(\boldsymbol{\theta}|\boldsymbol{o},\boldsymbol{x})$ = conditional probability density function of $\boldsymbol{\Theta}$ given the observation (i.e., the posterior joint density function of model parameters) and $\boldsymbol{x}$; $p_{O,X|\Theta}(\mathbf{o},\mathbf{x}|\boldsymbol{\theta})$ = conditional joint PDF of $\boldsymbol{O}$ and $\boldsymbol{x}$ given a parameter set $\boldsymbol{\Theta}=\boldsymbol{\theta}$; $p_{\Theta}(\boldsymbol{\theta})$ = joint PDF of model parameters; $p_X(\boldsymbol{x})$ = the joint PDF of model input forcing; and Ω_{Θ} = multidimensional parameter space. The second line in Eq. (9) follows by noting that x is statistically independent of model θ. The integration in the denominator is multidimensional, carried over the entire parameter space, Ω_{Θ}. We can also refer to $p_{O|X,\Theta}(\boldsymbol{o}|\boldsymbol{x},\boldsymbol{\theta})$ as the likelihood function—that is, the likelihood of having observed the data given $\boldsymbol{\theta}$. For practical application involving highly complex models with sizable number of parameters, analytical joint probability densities $p_{\Theta}(\boldsymbol{\theta})$ and $p_{\Theta|O,X}(\boldsymbol{\theta}|\boldsymbol{o},\boldsymbol{x})$ are near impossible to obtain, which makes Eq. (9) too difficult to manipulate in the multidimensional and continuous parameter space, Ω_{Θ}. As in the GLUE methodology, or any other sampling technique, we evaluate Eq. (9) in terms of discrete parameter sets $\boldsymbol{\theta}_i, i=1,2,\ldots,n$, with prior and posterior probability mass functions, $P(\boldsymbol{\theta}_i)$ and $P(\boldsymbol{\theta}_i|\boldsymbol{o},\boldsymbol{x})$, instead of the corresponding continuous PDFs, $p_{\Theta}(\boldsymbol{\theta})$ and $p_{\Theta|O,X}(\boldsymbol{\theta}|\boldsymbol{o},\boldsymbol{x})$,

$$P(\boldsymbol{\theta}_i|\boldsymbol{o},\boldsymbol{x})=\frac{p_{O|X,\Theta}(\boldsymbol{o}|\boldsymbol{x},\boldsymbol{\theta}_i)p_X(\boldsymbol{x})P(\boldsymbol{\theta}_i)}{\sum_{i=1}^{n}p_{O,X|\Theta}(\boldsymbol{o},\boldsymbol{x}|\boldsymbol{\theta}_i)P(\boldsymbol{\theta}_i)}=kl(\boldsymbol{\theta}_i)P(\boldsymbol{\theta}_i),\quad i=1,2,\ldots,n \tag{10}$$

where $P(\boldsymbol{\theta}_i|\boldsymbol{o},\boldsymbol{x})$ = posterior probability mass of $\boldsymbol{\theta}_i=(\boldsymbol{\theta}_1^i,\boldsymbol{\theta}_2^i,\ldots,\boldsymbol{\theta}_r^i)^T$; r = number of model parameters; $l(\boldsymbol{\theta}_i)=p_{O|X,\Theta}(\boldsymbol{o}|\boldsymbol{x},\boldsymbol{\theta}_i)$ = likelihood function evaluated at $\boldsymbol{\Theta}=\boldsymbol{\theta}_i$ and $\boldsymbol{X}=\boldsymbol{x}$; $P(\boldsymbol{\theta}_i)$ = prior probability mass of $\boldsymbol{\theta}_i$; k = normalizing factor such that $\sum_{i=1}^{n}P(\boldsymbol{\theta}_i|\boldsymbol{o},\boldsymbol{x})=1$, that is, $k=[\sum_{i=1}^{n}l(\boldsymbol{\theta}_i)P(\boldsymbol{\theta}_i)]^{-1}$; and n = number of randomly drawn parameter sets. As in the GLUE methodology, some prior information about feasible ranges of parameter values is used to generate independent parameter sets $\boldsymbol{\theta}$ for use in the given model. The prior parameter distributions should be broad enough to insure model outputs will capture the observed data. We will also assume equally likely parameter sets prior to the introduction of measurements, i.e., $P(\boldsymbol{\theta}_i)=1$, $i=1,2,\ldots,n$. Henceforth, for notational convenience, we drop the dependence on x and refer to $P(\boldsymbol{\theta}_i|\boldsymbol{o})$ instead of $P(\boldsymbol{\theta}_i|\boldsymbol{o},\boldsymbol{x})$, with implicit dependence on x. Eq. (10) can be applied sequentially when multiple output variables are observable; this is in lieu of the otherwise batch approach that involves simultaneous conditioning on all observable data and the use of the joint likelihood function (e.g., Dilks et al. 1992).

Let $Z=g(O)$, and $M(t;\boldsymbol{\Theta},\boldsymbol{x})=g[f(t;\boldsymbol{\Theta},\boldsymbol{x})]$. Without loss of generality, conditioning in Eq. (10) is equally valid to any transformation of O; hence, O can be replaced with Z throughout, of course, with a proper definition of the densities. With the newly defined variables substituted into Eqs. (7) and (8), and assuming random sampling, the likelihood of obtaining a set of m independent observations, $o_1,o_2,\ldots,o_m$, given $\boldsymbol{\Theta}=\boldsymbol{\theta}_i$ is

$$l(\boldsymbol{\theta}_i)=(2\pi)^{-m/2}(\sigma_\omega)^{-m}\times\exp\left[-\frac{1}{2}\sum_{k=1}^{m}\left(\frac{z_k-M_k(\boldsymbol{\theta}_i,\boldsymbol{x})-(1-\rho)\mu-\rho\varepsilon_{k-1}}{\sigma_\omega}\right)^2\right] \tag{11}$$

where $z_k\equiv z(t_k)$; $M_k(\boldsymbol{\theta}_i,\boldsymbol{x})\equiv M(t_k;\boldsymbol{\theta}_i,\boldsymbol{x})$; and $\varepsilon_k\equiv\varepsilon(t_k)=z_k-M_k(\boldsymbol{\theta}_i,\boldsymbol{x})$. For each randomly sampled $\boldsymbol{\theta}_i$, $l(\boldsymbol{\theta}_i)$ is optimized by conveniently maximizing the logarithm of the likelihood function; i.e.,

$$\ln l(\boldsymbol{\theta}_i)=-\frac{m}{2}\ln(2\pi)-m\ln\sigma_\omega-\frac{1}{2}\sum_{k=1}^{m}\left[\frac{z_k-M_k(\boldsymbol{\theta}_i,\boldsymbol{x})-(1-\rho)\mu-\rho\varepsilon_{k-1}(\boldsymbol{\theta}_i)}{\sigma_\omega}\right]^2 \tag{12}$$

Sorooshian and Dracup (1980) were among the first investigators to consider first-order autoregressive residual errors and incorporate the model into the formulation of the log-likelihood function. However, here we extend the model to include potential bias in model errors and derive explicit algebraic relationships for the maximum likelihood estimates (MLE) of $\boldsymbol{\vartheta}=(\mu,\rho,\sigma_\omega^2)$ expressed in terms of observed data. The maximum likelihood estimator $\hat{\boldsymbol{\vartheta}}=(\hat{\mu},\hat{\rho},\hat{\sigma}_\omega^2)$ is the value of ϑ that maximizes the likelihood function [Eq. (11)] or its logarithm, and can be obtained from the solution to the following set of simultaneous equations

$$\frac{\partial}{\partial\mu}\ln l(\boldsymbol{\theta}_i)=0 \tag{13a}$$

$$\frac{\partial}{\partial\rho}\ln l(\boldsymbol{\theta}_i)=0 \tag{13b}$$

$$\frac{\partial}{\partial\sigma_\omega^2}\ln l(\boldsymbol{\theta}_i)=0 \tag{13c}$$

The solutions, albeit after some algebraic effort, are

$$\hat{\mu}_i=\frac{1}{m[1-\hat{\rho}_i]}\sum_{k=1}^{m}(\varepsilon_k-\hat{\rho}_i\varepsilon_{k-1}) \tag{14a}$$

$$\hat{\sigma}_{\omega i}^2=\frac{1}{m}\sum_{k=1}^{m}[(\varepsilon_k-\hat{\mu}_i)-\hat{\rho}_i(\varepsilon_{k-1}-\hat{\mu}_i)]^2 \tag{14b}$$

and $\hat{\rho}_i$ is obtained by solving a quadratic equation $A\rho^2+B\rho+C=0$, $|\rho|<1$, to yield

$$\hat{\rho}_i=-\frac{B}{2A}+\frac{\sqrt{B^2-4AC}}{2A} \tag{14c}$$

in which

$$A=\left(\sum_{\mathbf{k}=1}^{\mathbf{m}}\varepsilon_{\mathbf{k}-1}\right)^2-m\sum_{k=1}^{m}(\varepsilon_{k-1})^2$$

$$B=m\sum_{k=1}^{m}(\varepsilon_{\mathbf{k}-1})^2-\left(\sum_{k=1}^{m}\varepsilon_{\mathbf{k}-1}\right)\left(\sum_{k=1}^{m}\varepsilon_{\mathbf{k}}+\sum_{k=1}^{m}\varepsilon_{\mathbf{k}-1}\right)+m\sum_{k=1}^{m}\varepsilon_k\varepsilon_{k-1}$$

$$C=\left(\sum_{k=1}^{m}\varepsilon_k\right)\left(\sum_{k=1}^{m}\varepsilon_{k-1}\right)-m\sum_{k=1}^{m}\varepsilon_k\varepsilon_{k-1} \tag{15}$$

where subscript i = dependence of the maximum likelihood estimators on the parameter set $\boldsymbol{\theta}_i$.

Each $\boldsymbol{\theta}_i$ has a corresponding $\hat{\boldsymbol{\vartheta}}_{\mathbf{i}}=(\hat{\mu}_i,\hat{\rho}_i,\hat{\sigma}_{\omega i}^2)$ whose likelihood $\hat{l}(\boldsymbol{\theta}_i)$ is obtained by substituting Eq. (14b) into Eq. (11)

$$\hat{l}(\boldsymbol{\theta}_i)=(2\pi e\hat{\sigma}_{\omega i}^2)^{-\frac{m}{2}} \tag{16}$$

Hence, weighted estimates of the first-order Markov process parameters are

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

$$\hat{\mu} = \sum_{i=1}^{n} \hat{\mu}_i P(\boldsymbol{\theta_i}|\boldsymbol{o}) \qquad \hat{\rho} = \sum_{i=1}^{n} \hat{\rho}_i P(\boldsymbol{\theta_i}|\boldsymbol{o})$$

$$\hat{\sigma}_\omega^2 = \sum_{i=1}^{n} \hat{\sigma}_{\omega i}^2 P(\boldsymbol{\theta_i}|\boldsymbol{o}) \tag{17}$$

where $P(\boldsymbol{\theta}_i|\boldsymbol{o})$ is computed by substituting $\hat{l}(\boldsymbol{\theta}_i)$ for $l(\boldsymbol{\theta}_i)$ in Eq. (10) and noting that $k = [\sum_{i=1}^{n} \hat{l}(\boldsymbol{\theta}_i) P(\boldsymbol{\theta}_i)]^{-1}$.

The variance of the overall error, ε_k, can be obtained from the following recursive form of Eq. (8):

$$\varepsilon_n - \mu = \rho^n(\varepsilon_0 - \mu) + \sum_{k=1}^{n} \rho^{n-k}\omega_k \tag{18}$$

Switching to summation index $m = n - k$ and taking the variance of ε_n, one can show that

$$\sigma_{\varepsilon_n}^2 = \rho^{2n}\sigma_{\varepsilon_0}^2 + \sigma_\omega^2 \frac{\rho^{2n} - 1}{\rho^2 - 1}, \qquad |\rho| > 1 \tag{19}$$

For sufficiently large n and finite $\sigma_{\varepsilon_0}^2$, the variance of a stationary sequence is the limit value of $\sigma_{\varepsilon_n}^2$ as $n \to \infty$

$$\sigma_\varepsilon^2 = \frac{\sigma_\omega^2}{1 - \rho^2}, \qquad |\rho| < 1 \tag{20}$$

Eq. (19) shows that residual errors are generally non-stationary and only after a sufficiently long time the process becomes second-order stationary. From Eqs. (18) and (20) and noting that ω_n is Gaussian, one concludes that for stationary errors [i.e., sufficiently large and stable sequence, $\varepsilon_n \sim N(\mu, \sigma_\omega^2/(1 - \rho^2))$]; this result and Eq. (20) will be used in the numerical experiment under the section "Applications". Eq (20) relates an estimate of the variance of the overall error, $\hat{\sigma}_{\varepsilon i}^2$, given $\boldsymbol{\theta}_i$, to the MLEs $\hat{\sigma}_{\omega i}^2$ and $\hat{\rho}_i$; it can also be used to relate $\hat{\sigma}_\varepsilon^2$ to $\hat{\sigma}_\omega^2$ and $\hat{\rho}$.

If the variance of measurement errors, σ_v^2, is known, the variance of model structural errors, σ_ϵ^2, for each $\boldsymbol{\theta}_i$ can be determined by noting that $\hat{\sigma}_{\epsilon i}^2 = \hat{\sigma}_{\varepsilon i}^2 - \sigma_v^2$ [Eq. (7)]. Information on the standard deviation of instrumentation errors might be available and laboratory analysis uncertainty is quantifiable by conducting replicate tests. However, errors introduced by scale-dependent averaging (spatially and temporally) and interpolation are too difficult to determine and make obtaining an accurate prior estimate of σ_v^2 unfeasible. This problem is further confounded by the lack of correlation between observation and model structural errors and the additive nature of the noise terms, thereby making it difficult to resolve the total variance, σ_ω^2, to σ_v^2 and σ_ϵ^2 components. Consequently, the analysis becomes that of using the model to predict the observed value, O, rather than the true value of the water quality parameter being observed, Y. In other words, when addressing a pollution management problem, O (instead of Y) is compared against the regulatory water quality criterion, Y^*. Unfortunately, this is a common practice in the application of water quality models to risk management, whereby little or no information is given on whether or not the derived or consensus Y^* accounts for observation errors. Nevertheless, Zheng and Keller (2007, 2008) presented an approach using water quality data monitored during several storm events at different sites to obtain a relationship between true watershed response (average daily concentrations) and observed (instantaneous) value. The MOCAU procedure they have developed relied on true watershed response instead of error-prone observation to estimate management variables.

Henceforward, we focus on Y with the understanding that in real world applications, the analysis would proceed with Ó (future observations) replacing Y (true value or unknown observable), and total (model-observation) error variance (σ_ε^2) replacing model structural error variance (σ_ϵ^2) in the equations presented henceforth, as applicable.

Prediction

We derive explicit expressions for the posterior CDF of the water quality parameter (Y) and the corresponding Bayesian estimates. Henceforth, unless otherwise stated, conditioning on $\boldsymbol{x}$ is implicitly understood. While the simple case of univariate, statistically independent random forcing inputs could be treated by augmenting $\boldsymbol{\Theta}$ with $\boldsymbol{x}$ and treating the elements of $\boldsymbol{x}$ as unknown model parameters, the case of stochastic $\boldsymbol{x}$ requires extended analysis, which is beyond the scope of this effort.

Posterior Cumulative Distribution

If we let $U = g(Y)$, i.e., $Y = g^{-1}(U)$, the conditional cumulative distribution (CDF) of U given the observation, $F_{(U|\boldsymbol{O})}(u|\boldsymbol{o})$, can be obtained as follows:

$$F_{(U|\boldsymbol{O})}(u|\boldsymbol{o}) = \int_{\Omega_U} p_{(U|\boldsymbol{O})}(u|\boldsymbol{o})du = \int_{\Omega_U}\int_{\Omega_{\boldsymbol{\Theta}}} p(u|\boldsymbol{\theta},\boldsymbol{o})p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta}du$$
$$= \int_{\Omega_{\boldsymbol{\Theta}}} F(u|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta} \tag{21}$$

where $p_{(U|\boldsymbol{O})}(u|\boldsymbol{o})$ and $F_{(U|\boldsymbol{O})}(u|\boldsymbol{o})$, respectively, = PDF and CDF of u given the observation; $F(u|\boldsymbol{\theta})$ = CDF of u given $\boldsymbol{\theta}$; Ω_U = domain of $U = g(Y)$; $\Omega_{\boldsymbol{\Theta}}$ = p-dimensional parameter space; and $\int_{\Omega_{\boldsymbol{\Theta}}} d\boldsymbol{\theta} = \int_{\theta_1}\int_{\theta_2}, \ldots, \int_{\theta_p} d\boldsymbol{\theta}_p, \ldots, d\boldsymbol{\theta}_2 d\boldsymbol{\theta}_1$ = p-dimensional integration. The third line follows because U and $\boldsymbol{O}$ are independent given $\boldsymbol{\theta}$ [Eq. (5)]; it displays the posterior predictive CDF as an average of conditional predictions over the posterior distributions of $\boldsymbol{\Theta}$. We have dropped subscripts from p and F for notational convenience. Eq. (21) can be evaluated further by noting that U is normally distributed given $\boldsymbol{\theta}$

$$F(u|\boldsymbol{o}) = \frac{1}{2} + \frac{1}{2}\int_{\Omega_{\boldsymbol{\Theta}}} \text{erf}\left\{\frac{\mu - g[f(t;\boldsymbol{\theta},\boldsymbol{x})]}{\sqrt{2}\sigma_\epsilon}\right\} p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta} \tag{22}$$

$F_Y(y|\boldsymbol{o})$, therefore, is obtained by noting that for $U = g(Y)$, $F_U[u] = F_Y(y)\{\text{or}, F_Y(y) = F_U[g(y)]\}$. Hence

$$F(y|\boldsymbol{o}) = \frac{1}{2} + \frac{1}{2}\int_{\Omega_{\boldsymbol{\Theta}}} \text{erf}\left\{\frac{g(y) - g[f(t;\boldsymbol{\theta},\boldsymbol{x})]}{\sqrt{2}\sigma_\epsilon}\right\} p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta} \tag{23}$$

Again, we approximate the integral in Eq. (23) in terms of posterior probability mass functions, $P(\boldsymbol{\theta}_i|\boldsymbol{o})$, of $\boldsymbol{\theta}_i$, $i = 1, 2, \ldots, n$

$$F(y|\boldsymbol{o}) \approx \frac{1}{2} + \frac{1}{2}\sum_{i=1}^{n} \text{erf}\left\{\frac{g(y) - g[f(t;\boldsymbol{\theta_i},\boldsymbol{x})]}{\sqrt{2}\hat{\sigma}_\epsilon}\right\} p(\boldsymbol{\theta_i}|\boldsymbol{o}) \tag{24}$$

where $F(y|\boldsymbol{o})$ = posterior CDF of Y. Eq. (24) serves the base for estimating confidence limits, median, mean, variance, and higher order moments of Y.

By substituting $\boldsymbol{o}$ for y and $\hat{\sigma}_\omega$ for $\hat{\sigma}_\epsilon$, Eq. (24) could also be used to construct predictions (i.e., median, confidence limits) for future observed values of the variable Y given the observed record, $\boldsymbol{o}$.

Bayesian Estimate

The Bayesian estimate of Y at any point in time is the conditional mean of Y given $\boldsymbol{o}$, $E(Y|\boldsymbol{O})$

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

$$E(Y|\boldsymbol{O}) = \int_{\Omega_Y} yp(y|\boldsymbol{o})dy = \int_{\Omega_Y}\int_{\Omega_\Theta} yp(y|\boldsymbol{\theta},\boldsymbol{o})p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta}dy$$
$$= \int_{\Omega_\Theta} E(Y|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta} \quad (25)$$

where $E(Y|\boldsymbol{\theta})$ = conditional mean of Y given $\boldsymbol{\theta}$; and Ω_Y = domain of Y (e.g., $[0,\infty]$ for concentrations). The third line shows the posterior mean of Y as an average of conditional mean over the posterior distributions of $\boldsymbol{\Theta}$.

For lognormal Y where $g(Y) = \ln(Y) = U$ is normally distributed, we have $E(Y|\boldsymbol{\theta}) = \exp\{(\sigma_\epsilon^2/2) + g[f(t;\boldsymbol{\theta},\boldsymbol{x})]\} = f(t;\boldsymbol{\theta},\boldsymbol{x})e^{\sigma_\epsilon^2/2}$; the R.H.S of Eq. (25) thus becomes

$$E(Y|\boldsymbol{O}) = e^{\sigma_\epsilon^2/2}\int_{\Omega_\Theta} f(t;\boldsymbol{\theta},\boldsymbol{x})p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta} \quad (26)$$

Similarly, the discrete form approximation of Eq. (25) (assuming uniformly sampled parameter space) is

$$E(Y|\boldsymbol{O}) \approx \sum_{i=1}^{n} E(Y|\boldsymbol{\theta}_i)P(\boldsymbol{\theta}_i|\boldsymbol{o}) \quad (27)$$

and for lognormally distributed Y,

$$E(Y|\boldsymbol{O}) \approx e^{\sigma_\epsilon^2/2}\sum_{i=1}^{n} f(t;\boldsymbol{\theta}_i,\boldsymbol{x})P(\boldsymbol{\theta}_i|\boldsymbol{o}) \quad (28)$$

In general, $E(Y|\boldsymbol{\theta}_i)$ can be computed numerically by noting that given $\boldsymbol{\theta}_i$, $U = g[f(t;\boldsymbol{\theta}_i,\boldsymbol{x})] + \epsilon$ is normally distributed with a mean of $g[f(t;\boldsymbol{\theta}_i,\boldsymbol{x})]$ and variance of $\sigma_{\epsilon i}^2$, i.e., $U - N\{g[f(t;\boldsymbol{\theta}_i,\boldsymbol{x})],\hat{\sigma}_{\epsilon i}^2\}$, For each parameter set, $\boldsymbol{\theta}_i$, normally distributed u values, u_n, can be generated; the corresponding y_n values (recall, conditional on $\boldsymbol{\theta}_i$) are subsequently obtained from the inverse relationship $y_n = g^{-1}(u_n)$. The sample mean of the randomly generated y_n values could then be used to infer $E(Y|\boldsymbol{\theta}_i)$.

Water Quality Management

Posterior Risk Estimation

We consider the general case of a random Y^* where the conditional probability of Y^* exceeding Y given $O = o$ is given by Eq. (4)

$$P[Y \le Y^*|\boldsymbol{O}] = \int_{\Omega_{Y^*}} F_{(Y|\boldsymbol{O})}(y|\boldsymbol{o})p_{Y^*}(y)dy \quad (29)$$

Above, we have established that

$$F_{(Y|\boldsymbol{O})}(y|\boldsymbol{o}) = \int_{\Omega_{\boldsymbol{\Theta}}} F(y|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta} \quad (30)$$

Inserting Eq. (30) in Eq. (29) leads to

$$P[Y \le Y^*|\boldsymbol{O}] = \int_{\Omega_{Y^*}}\int_{\Omega_\Theta} F(y|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta}p_{Y^*}(y)dy \quad (31)$$

From the fact that $F_Y(y|\boldsymbol{\theta}) = F_U(u = g(y)|\boldsymbol{\theta})$ and that U is normally distributed given $\boldsymbol{\Theta}$, it follows from Eq. (31) that

$$P[Y \le Y^*|\boldsymbol{O}] = \frac{1}{2} + \frac{1}{2}\int_{\Omega_{Y^*}}\int_{\Omega_{\boldsymbol{\theta}}} \text{erf}\left\{\frac{g(y) - g[f(t;\boldsymbol{\theta},\boldsymbol{x})]}{\sqrt{2}\sigma_\epsilon}\right\} \times p(\boldsymbol{\theta}|\boldsymbol{o})d\boldsymbol{\theta}p_{Y^*}(y)dy \quad (32)$$

which in discrete form can be approximated as

$$P[Y \le Y^*|\boldsymbol{O}] \approx \frac{1}{2} + \frac{1}{2}\sum_{i=1}^{n}\int_{\Omega_{Y^*}} \text{erf}\left\{\frac{g(y) - g[f(t;\boldsymbol{\theta}_i,\boldsymbol{x})]}{\sqrt{2}\hat{\sigma}_\epsilon}\right\} \times p_{Y^*}(y)dyP(\boldsymbol{\theta}_i|\boldsymbol{o}) \quad (33)$$

Finally, probability of exceeding Y^* is

$$P[Y > Y^*|\boldsymbol{O}] \approx \frac{1}{2} - \frac{1}{2}\sum_{i=1}^{n}\int_{\Omega_{Y^*}} \text{erf}\left\{\frac{g(y) - g[f(t;\boldsymbol{\theta}_i,\boldsymbol{x})]}{\sqrt{2}\hat{\sigma}_\epsilon}\right\} \times p_{Y^*}(y)dyP(\boldsymbol{\theta}_i|\boldsymbol{o}) \quad (34)$$

For the specific case of $p_{Y^*}(y) = \delta(y - y^*)$, Eq. (34) reduces to

$$P[Y > y^*|\boldsymbol{O}] \approx \frac{1}{2} - \frac{1}{2}\sum_{i=1}^{n} \text{erf}\left\{\frac{g(y^*) - g[f(t;\boldsymbol{\theta}_i,\boldsymbol{x})]}{\sqrt{2}\hat{\sigma}_\epsilon}\right\}P(\boldsymbol{\theta}_i|\boldsymbol{o}) \quad (35)$$

Closed form expressions to the integrals in Eq. (34) are obtained in the Appendix for uniform and triangular distributions of Y^*.

A more general approach is to compute the probability of exceedance using MC simulation assuming that $g(y)$ is monotonically increasing with y:

$$P[Y > Y^*|\boldsymbol{O}] \approx \sum_{i=1}^{n} P\{g[Y(\boldsymbol{\theta}_i,\boldsymbol{x})] > g[Y^*]\}P(\theta_i|\boldsymbol{o}) \quad (36)$$

The first probability inside the summation can be computed numerically by sampling $g[Y(\boldsymbol{\theta}_i,\boldsymbol{x})]$ from $N(g[f(t;\boldsymbol{\theta}_i,\boldsymbol{x})],\sigma_\epsilon^2)$, y^* from its respective distribution, and computing the fraction of times the inequality holds.

The primary advantage of using the MC method in Eq. (36) over analytical expressions of Eq. (34) is that it fully accounts for the uncertainty of $\mathbf{X}$; i.e., Eq. (36) samples the full range of $\boldsymbol{X}$ [by random sampling from $p_X(\boldsymbol{x})$] when $P[g(Y) > g(Y^*)]$ computing for each parameter set. However, this comes at the expense of prohibitively large MC samples and hence its evaluation is computationally demanding, especially for complex models. One other advantage of Eq. (36) is that model residual errors (or transformation thereof) need not be Gaussian. Eq. (34), on the other hand, is exact if the input forcing is known a priori and residual errors are Gaussian. For the simple case in which the PDF of the input forcing is known, uncertainty in $\boldsymbol{X}$ can be accounted for in Eq. (34) by extending the parameter vector $\boldsymbol{\Theta}$ to include $\boldsymbol{X}$ as an additional parameter, i.e., $\boldsymbol{\Theta}' = (\boldsymbol{\Theta}\boldsymbol{X})$.

Recognizing Eqs. (34)–(36) are essentially discrete approximations of Eq. (29), henceforth we refer to the evaluation of probability of exceedance by either of these equations as the formal approach since they follow from firm application of laws of probability (total probability and Bayes theorems). The informal approach, on the other hand, is more commonly used whereby parameter values are sampled independently from their posterior marginal distributions then used to compute probability of exceedance, but without due consideration of potential parameter interactions and the importance of parameter sets as reflected by the corresponding likelihood function estimates.

TMDL and MOS

TMDL is the maximum pollutant load a water body can receive that will not violate the state water quality standards with and adequate margin of safety (MOS). This maximum load is the loading capacity of the water body (LC) minus the MOS:TMDL = LC − MOS. The LC includes all point sources (waste load discharge) and nonpoint sources (stormwater runoff, groundwater discharge); it is typically obtained by insuring model-simulated water quality metrics honor the numerical targets for the water body. The MOS is

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

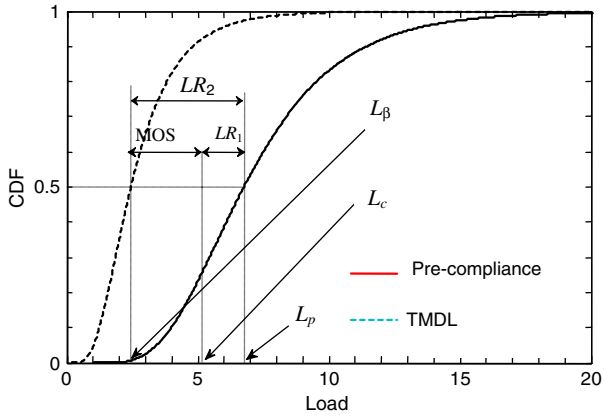

Fig. 2. Estimation of margin of safety (MOS); solid line corresponds to pre-compliance CDF of loading and dashed line corresponds to compliance loading (TMDL); LR_1 is required load reduction for the reference simulation; LR_2 is β-compliance mean-load reduction

introduced implicitly by means of conservative assumptions (e.g., smaller than expected assimilative capacity, more conservative standard) or explicitly by assigning a fraction of the computed LC as a protective cushion in account of the uncertainty in the analysis. As previously stated, both the relationship between uncertainty and the MOS and the level of protection thus provided would be opaque if uncertainty is not rigorously accounted for in the TMDL analysis. Eq. (1) and its working forms (i.e., BMCML approximation) of Eqs. (34) and (35), however, provide the basis for estimating the required PDF of the TMDL, L, and related MOS given the observation and acceptable risk of failure, β.

TMDL can be determined by identifying the critical distribution of L such that $P[Y > Y^*|\boldsymbol{O}] \equiv G(L) = \beta$, whereby $G(L)$ can be evaluated by means of Eq. (34) or (35). For instance, assuming a normally or lognormally distributed L, the compliance mean of L, μ_L, can be obtained by solving the equation $G(\mu_L) - \beta = 0$ for μ_L, while assuming a constant variance (σ_L^2) fixed at the value obtained from observed or model simulated loading values. A computationally simpler approach is to evaluate $G(\mu_L)$ over the range of μ_L and identify the value of μ_L corresponding to $G(\mu_L) = \beta$. The scope of this presentation is that of identifying permissible total loading to an impaired stream segment or a water body, rather than resolving L into point and nonpoint source contributions. The latter requires complex watershed models with point and diffuse sources accounted for explicitly (e.g., Zheng and Keller 2008; Patil and Deng 2011); in which case the problem of apportioning required loading capacity to various sources as well as identifying proper BMPs for achieving desired load reductions probabilistically is beyond the scope of this treatment.

The MOS can be computed as the difference between required load reduction corresponding to a reference deterministic simulation and required mean load reduction taking into consideration all sources of uncertainty. In a normalized form, MOS may be estimated as follows (Fig. 2)

$$\text{MOS} = \frac{LR_2 - LR_1}{L_p} \times 100, \qquad LR_1 = L_p - L_c,$$
$$LR_2 = L_p^* - L_\beta \tag{37}$$

where L_p = pre-compliance load corresponding to a reference simulation in which some average values of the parameters are used; L_c = compliance load corresponding to the reference simulation; L_p^* = mean precompliance load considering random parameters and analysis uncertainty; and $L_\beta = \beta$-compliance mean load considering random parameters and analysis uncertainty. Noting that L_p is a first-order approximation of $L_p^*, L_p \approx L_p^*$, we have

$$\text{MOS}(\beta) \approx \frac{L_c - L_\beta}{L_p} \times 100 \tag{38}$$

If LR_1 corresponds to the load reduction necessary to achieve compliance with 50% confidence, then Eq. (37) and its approximation Eq. (38) essentially yield the MOS proposed by Borsuk et al. (2002). This equation reveals three points worth emphasizing. Firstly, the MOS is computed relative to a reference simulation and thus differs from one application to another. For example, in the absence of observational data, a reference simulation is that which is based on some average of default model parameter values or the means of a priori parameter distributions constructed from literature tabulations and expert judgment. In the presence of observational data, a reference simulation is based on calibrated values of model parameters (as in traditional model calibration) or the means of the posterior parameter distributions obtained after Bayesian conditioning. Secondly, the dependence of the MOS on β follows from the functional relationship between L_β and β. Thirdly, for TMDLs with MOS implicitly or explicitly accounted for, Eq. (38) can be used to calculate the risk (β), alternatively, the reliability $(1 - \beta)$ corresponding to the load reduction recommended for the TMDL. This will be accentuated subsequently in the Fork Creek Bacterial and Shunganunga Creek TMDL applications.

Applications

Application I: Hypothetical Lake-Phosphorus TMDL

We consider the lake-sediment-phosphorus model presented by Hantush (2009). The steady-state solution for lake-averaged sediment and total phosphorus concentrations are

$$m_w = m_{w,\text{in}} - \frac{v_b A_s}{Q}(1 - \phi)\rho_s \tag{39a}$$

$$C = \frac{\alpha f_d + v_r f_s + v_b}{v_b A_s (v_s F_s + \alpha F_d) + Q(\alpha f_d + v_r f_s + v_b)} L,$$
$$L = Q c_{d,\text{in}} + Q m_{w,\text{in}} c_{s,\text{in}} \tag{39b}$$

where m_w = sediment concentration in lake water (ML^{-3}); $m_{w,\text{in}}$ = inflow sediment concentration (ML^{-3}); C = total phosphorus concentration in lake water (ML^{-3}); Q = lake inflow rate (L^3T); L = total phosphorus loading rate (MT^{-1}); $c_{d,\text{in}}$ = inflow dissolved phosphorus concentration (ML^{-3}); and $c_{s,\text{in}}$ = inflow sediment-bound phosphorus concentration (MM^{-1}). Definitions of all other parameters can be found in Hantush (2009).

Table 1 lists all model parameters and inputs, their units, and the assumed PDFs distributions selected from typical values and ranges reported in the literature. True parameter values and input time series are sampled randomly from these PDFs. We refer to the true parameter values as reference values. A hypothetical bottom lake surface area, A_s, of 5×10^6 m^2 is assumed. Dissolved phosphorus is assumed to be 5% of the sorbed phosphorus in (g/m^3): $c_{d,\text{in}} = 0.05 m_{w,\text{in}}\ c_{s,\text{in}}$. C^* is the total phosphorus target concentration assumed to be a random variable and follows a triangular distribution with minimum, most probable (mode), and maximum values of 16, 20, and 96 mg/m^3, respectively. As argued earlier, shifting border values separating different trophic conditions is accounted for more realistically by treating C^* as a random variable.

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

Table 1. Lake-Sediment Phosphorus Model Parameters and Inputs, and Corresponding PDFs for Generating the True Values of the Parameters and Inputs

Parameter	Unit	Distribution	Parameter	Unit	Distribution
Q	m^3/year	$\sim \log -N(2 \times 10^9, 8 \times 10^8)$[a]	D^*	m^2/year	$\sim U(1.8, 2)$[b]
$m_{w,\text{in}}$	g/m^3	$\sim \log -N(30, 10)$	δ	m	$\sim \log -N(0.005, 0.0002)$
$c_{s,\text{in}}$	g/g	$\sim \log -N(1.5 \times 10^{-3}, 10^{-3})$	ϕ	$m^3\ m^{-3}$	$\sim T(0.68, 0.7, 0.72)$
v_s	m/year	$\sim \log -N(32, 1)$	ρ_s	g/cm^3	$\sim T(2.4, 2.5, 2.6)$
K_d	m^3/g	$\sim \log -N(3 \times 10^{-4}, 10^{-5})$	Y^*	mg/m^3	$\sim T(16, 20, 96)$
v_b	m/year	$\sim T(0.0047, 0.0049, 0.0052)$[c]	—	—	—

[a]$\log -N(\lambda, \sigma)$ denotes log-normally distributed random variable with mean λ and standard deviation σ.
[b]$U(a, b)$ denotes a uniformly distributed random variable in the interval $[a, b]$.
[c]$T(a, b, c)$ denotes triangular distribution of random variable with minimum value a, mode c, and maximum value c.

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

In this example illustration, we considered the transformation $g(y) = \ln y$ in Eq. (7) and made use of Eqs. (8), (39*a*), and (39*b*) to synthesize observations for m_w and C. This was achieved by generating a random sequence of total errors, ε, for each constituent concentration and superimposing the errors so generated over log-transformed model simulated values of the constituent concentration using actual (reference) parameter values sampled from the prior distributions in Table 1. Table 2 displays values of μ, ρ, and σ_ω^2 that were used to generate ε for each log-transformed constituent concentration. The variances of measurement errors (σ_v^2) of log-m_w and log-C were assumed to be 6.5×10^{-3} and 6.84×10^{-3}, respectively and 2.6×10^{-2} and 2.74×10^{-2}, respectively, for the variances of model structural errors (σ_ϵ^2). The values of σ_ω^2 shown in Table 2 were then obtained from Eq. (20) in which $\sigma_\varepsilon^2 = \sigma_v^2 + \sigma_\epsilon^2$. The procedure for generating correlated and biased ε sequence is as follows. For each constituent, a relatively large number of independent values of ω were sampled from $N(0, \sigma_\omega^2)$. A corresponding sequence of ε was then generated from the First-order Markov Process in Eq. (8) starting with the initial condition $\varepsilon(0) = 0$, and using the values of μ and ρ in Table 2 and the randomly sampled ω values. The generated sequence was truncated to eliminate errors associated with the arbitrarily assumed zero initial condition and to insure a stationary and stable error sequence [recall, Eq. (20)]. The last 80 values of ε were retained as the correlated and biased observation-model errors. The adequacy of the procedure for generating the ε sequence was tested by comparing reference values of μ, ρ, and σ_ω^2 to those obtained by regressing Eq. (8) on the generated ε values (Table 2).

Another source of uncertainty is the natural variability of x, which in this particular example is attributed to inflow and input concentrations: $\boldsymbol{x} = (Q, m_{w,\text{in}}, c_{s,\text{in}}, c_{d,\text{in}})^T$. Table 1 lists hypothetical lognormal distributions of x. For example, the PDFs describe interannual or year-to-year seasonal variations of x. For each of $m_{w,\text{in}}$, $c_{s,\text{in}}$, and Q, a sequences of 80 statistically independent values were sampled from the corresponding distributions in Table 1 to yield actual (reference) input time series. The log-transformed sampled values of Q, $m_{w,\text{in}}$, and $c_{s,\text{in}}$ were slightly corrupted by superimposing normally distributed, zero-mean measurement noise of variances equal to 1.15×10^{-2}, 2.89×10^{-4}, and 1.06×10^{-3}, respectively. The resulting values were transformed back by taking the exponent to yield the observed input sequence.

Eighty zero-mean, Gaussian ϵ values were generated for each of m_w and C using the assumed σ_ϵ^2 values. The reference time series of $\boldsymbol{x}$ and the reference $\boldsymbol{\Theta}$ values sampled from the distributions in Table 1 were fed into Eqs. (39) to yield reference model simulated time series [$f(t; \boldsymbol{\Theta}, \boldsymbol{x})$ in Eq. (5)]. The resulting values were log-transformed (natural logarithm), then corrupted by adding the generated ϵ values to obtain the corresponding reference sequences of log-transformed m_w and C ($g[Y(t)]$ in Eq. (5), $g(Y) = \ln Y$). The exponent of these values is the reference time series (true values) of the variables (m_w and C). For each variable, the time series of observed values, $\boldsymbol{O}$, was obtained by superimposing the generated ε sequence on the log-transformed reference model simulated time series [Eq. (7)] and taking the exponent of the resulting values. Only the first 40 values were retained as the observations. Each of the generated 80-long input forcing (x) and reference output (Y) time series was split into two halves; the first is used for likelihood estimation and parameter estimation, and the second is retained for methodology validation.

The implementation of the Bayesian methodology starts with sampling 100,000 independent parameter sets $\boldsymbol{\theta} = (\phi, \rho_s, v_b, v_s, K_d, \alpha)'$ from the prior PDFs in Table 3. In this example, $n = 100{,}000$, $m = 40$, and $r = 6$. Note that m_w depends on the first three parameters and C on all parameters. A priori, each of the randomly drawn $\boldsymbol{\theta}_i$ is considered equally-likely a good simulator; i.e., in Eq. (10), $P(\boldsymbol{\theta}_i) = 1$. We implement Bayes relationship given in Eq. (10) sequentially; first, by conditioning $\boldsymbol{\theta}$ on the first 40 observed sediment concentrations, then on corresponding observed total phosphorus concentrations (first 40). For instance, when conditioning on observed total phosphorus concentrations, $P(\boldsymbol{\theta}_i)$ in Eq. (10) is set equal to the posterior probability mass computed by conditioning on observed sediment concentrations. At each of the two stages, $P(\boldsymbol{\theta}_i|\boldsymbol{o})$ was computed from Eq. (10) using $l(\boldsymbol{\theta}_i)$ determined from Eqs. (16), (14), and (15). The corresponding prediction of m_w or C (Bayesian estimate and 95% confidence limits) for the validation period was then obtained from Eqs. (28) and (24). The generated 100,000 values for each parameter in $\boldsymbol{\theta}$ and the corresponding $P(\boldsymbol{\theta}|\boldsymbol{o})$ values together were used to construct the 95% confidence interval and compute the median for that parameter. Table 2 compares BMCML estimates of overall

Table 2. Total Error Statistical Parameters

Variable	Model error parameter	Reference (true)	Observations[a]	Bayesian estimation
m_w	μ	0.32	0.32	0.37
	ρ	−0.70	−0.70	−0.37
	σ_ω^2	0.017	0.017	0.018
C	μ	−0.17	−0.16	−0.21
	ρ	0.50	0.60	0.31
	σ_ω^2	0.027	0.028	0.025

[a]Obtained by linear regression using Eq. (8) and the generated ε values.

Table 3. Prior PDFs of the Lake-Sediment Phosphorus Model Parameters

Parameter	Unit	Distribution
v_s	m/year	$\sim \log -N(37, 40)$
K_d	m^3/g	$\sim \log -N(4.87 \times 10^{-4}, 6 \times 10^{-4})$
v_b	m/year	$\sim T(6 \times 10^{-4}, 0.005, 0.015)$
α	m/year	$\sim U(100{,}800)$
ϕ	$m^3\ m^{-3}$	$\sim T(0.2, 0.7, 0.99)$
ρ_s	g/cm^3	$\sim T(1.5, 2.6, 3)$

Table 4. 95% Confidence Limits, Medians, BMCML Estimates, and Reference (True) Values of Lake-Sediment-Phosphorus Model

Parameter	2.50%	50%	97.50%	BMCML	Reference
v_b	0.0016	0.0041	0.0104	0.0046	0.0021
ϕ	0.446	0.778	0.918	0.754	0.664
ρ_s	1668387	2071474	2703376	2302402	2398093
K_d	4.58×10^{-5}	0.0003	0.0019	0.00047	0.0006
α	119.06	450.78	781.598	451.358	312.29
v_s	4.37	24.88	145.79	36.95	38.75

error (ε_t) statistics to the corresponding reference values. With the exception of ρ, the reference values in general are adequately estimated by Eq. (17). BMCML estimates of the parameters; their median values and 95% confidence limits; and corresponding reference values are shown in Table 4. Given the uncertainty in the input data as well as model structural and observational data errors, the Bayesian estimates of the parameters compared generally well with the reference (true) values, which are bracketed by the corresponding 95% confidence intervals.

Figs. 3 and 4 compare predictions (95% confidence limits and Bayesian estimates) to reference time series of m_w and C, respectively, using the set of reference values reserved for validation. Over this evaluation period, the efficiency Nash-Sutcliffe (E_{NS}) coefficient computed from Bayesian estimates and reference values are 0.81 and 0.93 for m_w and C, respectively. The corresponding values of the coefficient of determination (R^2) are 0.95 and 0.94. As shown in the figures, the variability and magnitude of the reference values of m_w and C are adequately reflected by the BMCML estimates. With few exceptions, the reference time series of both variables fell within their estimated 95% confidence limits.

To illustrate the integrated model calibration and risk-management approach, we consider a total phosphorus load reduction scenario such that the lake would hypothetically no longer be eutrophic (i.e., mesotrophic) with $(1-\beta)$ level of confidence: $P[Y > Y^*|\mathbf{O}] \le \beta$, where here Y denotes total phosphorus concentration. Without loss of generality and for the sake of illustration, we assume a deterministic $m_{w,\text{in}}$ but keep $c_{s,\text{in}}$ and Q_{in} random. The method described in the Section "TMDL and Margin of Safety" was implemented along with Eqs. (34) and (38) to identify needed reduction in L and corresponding MOS for various β values. The reduction in L is achieved by reducing $m_{w,\text{in}}$ from its current mean value. Eq. (34) was evaluated analytically noting that $p_{Y^*}(y)$ is given by a triangular PDF and using the closed-form expressions

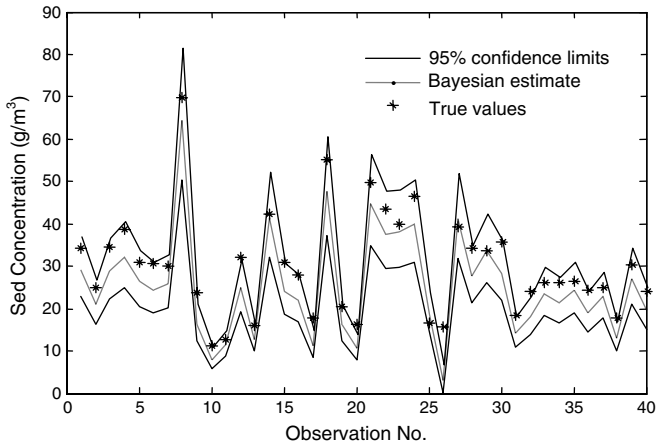

Fig. 3. BMCML estimates and 95% confidence limits of true sediment concentrations

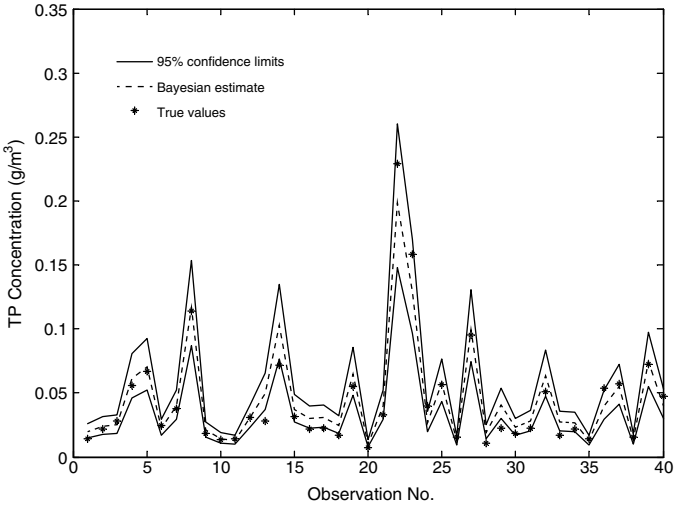

Fig. 4. BMCML estimates and 95% confidence limits of true total phosphorus concentrations

(A-7 and A-8) in the Appendix. The informal approach involves sampling parameter values from their posterior marginal distributions and noting that $P[Y > Y^*|\mathbf{O}] = P[\ln Y > \ln Y^*|\mathbf{O}] = P[\ln C + \epsilon > \ln Y^*|\mathbf{O}]$. Independent sampling of parameter values from posterior marginal distributions overlooks potential interactions (covariation) among the parameters, whereas Eq. (34), as often ascribed to the GLUE methodology, accounts implicitly for parameter interactions and focuses on the performance of the parameter set as opposed to individual parameter values (Beven and Binley 1992; Beven and Freer 2001).

Fig. 5 compares the relationship between MOS and β computed formally in Eq. (34) to the results obtained by the informal approach. An almost identical MOS-β relationship is obtained by the two approaches. As expected, the MOS decreases with increasing β (risk of the water quality standard violation). To examine the adequacy of Eq. (36) in computing the MOS, Fig. 6 compares the MOS-β relationship obtained by Eq. (34) and the analytical expression in the Appendix to that computed using Eq. (36). The results are nearly identical with $n = 100{,}000$. Although not shown in a figure, $n = 10{,}000$ still yielded close comparison, however, with the MC method slightly but consistently overestimating the analytical results.

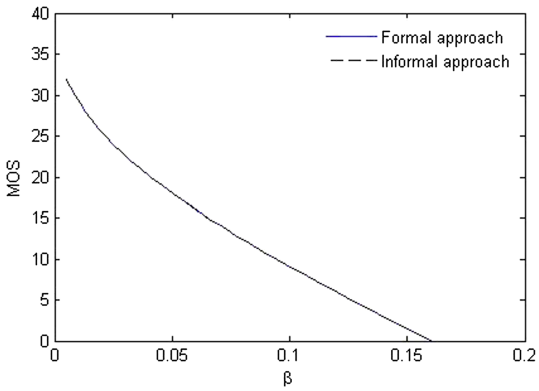

Fig. 5. MOS versus risk (β); the solid curve corresponds to probability of exceedance computed analytically by Eq. (34), and the dashed corresponds to probability of exceedance computed by sampling parameter values independently from posterior marginal PDFs

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

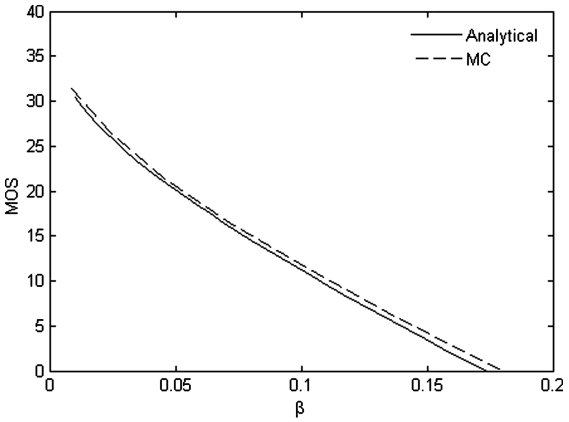

Fig. 6. MOS versus risk (β); the solid curve corresponds to probability of exceedance computed analytically by Eq. (34), and the dashed corresponds to probability of exceedance computed by means of MC using Eq. (36)

Application II: Fork Creek Bacteria TMDL

Fork Creek is a fecal coliform (FC) impaired stream in the Pee Dee river basin of South Carolina, where the applicable water quality standard is that FC should not exceed a geometric mean of 200/100 mL based on at least five samples in a 30-day period, and not more than 10% of the total samples during any 30-day period should exceed 400/100 mL. We use the 63 pathogen-indicator measurements of FC (in #/100 mL) reported by Chin (2009). Bacterial counts ranged from as low as 40/100 mL to as high as 6,600/100 mL. Among the guidelines commonly used to develop TMDL and determine load reductions is to compare 90th percentile of the water quality data, $\hat{y}_{90}$, to the 90-percentile water-quality standard, y^*_{90}, (e.g., Novotny 2003). We found that $\hat{y}_{90} = 744/100$ mL for the data, which exceeds the required 400/100 mL. Thus, a reduction in FC of magnitude 344/100 mL is needed. The shortcoming of this approach is that uncertainty in the sample estimate of percentile was not accounted for and hence the adequacy of reduced instream concentrations based solely on specified margins of safety cannot be assured. We consider the bacterial counts data as one possible sample from a lognormal population with mean μ and standard deviation σ and hence treat y_{90} as a random variable. Indeed, K-S test revealed a lognormally distributed sample with 95% confidence. We use the Bayesian approach to make statistical inferences on μ and σ, then develop a CDF for y_{90} to compute its percentiles. The model is $Y = \ln O \sim N(\mu, \sigma)$, where $\boldsymbol{\theta} = (\mu, \sigma)$ is the unknown parameter vector. It is well known that the maximum likelihood estimator (MLE) of μ is the sample mean of log-transformed data, $\bar{Y}$, and for σ, the MLE is the biased form of the sample variance, i.e., $[(n-1)/n]S$, where S is the sample variance of Y; here n is the sample size. Moreover, when the random sample is from a normal population, $[(\bar{Y} - \mu)/S]\sqrt{n}$ has the Student's t-distribution with $(n-1)$ degrees-of-freedom, and the distribution of nS^2/σ^2 is a chi-square distribution with $(n-1)$ degrees-of-freedom; i.e., χ^2_{n-1} (Neter et al. 1985). We use these benchmarks to validate the Bayesian approach.

One hundred thousand parameter sets $\boldsymbol{\theta} = (\mu, \sigma)$ were sampled from the priori distributions $\log_{10}, \log_{10} \sim U(1, 4)$. The likelihoods of observed bacterial counts, $\boldsymbol{O}$, conditional on the generated parameter sets, $\boldsymbol{\theta}$, were computed to estimate $P(\boldsymbol{\theta}|\boldsymbol{O})$ assuming an equally likely prior $\boldsymbol{\theta}$. The likelihood function of m observed bacterial counts is $l(\boldsymbol{\theta}_i) = (2\pi)^{-m/2}(\sigma_i)^{-m} \exp\{-\frac{1}{2}\sum_{k=1}^{m}[(\ln o_k - \mu_i)/\sigma_i]^2\}$. Table 5 compares Bayesian estimates of μ and σ to their MLE estimators and lists the median and 95% confidence limits for the two parameters. The theoretical confidence limits for μ computed from the t_{n-1} distribution and σ from χ^2_{n-1} distribution are (5.23,5.64) and (0.78,1.11), respectively.

Table 5. 95% Confidence Limits, Medians, Bayesian, and Maximum Likelihood Estimates of Mean and Standard Deviation of Log Bacteria Concentration in Fork Creek

Parameter	2.50%	50%	97.50	Bayesian	MLE
μ	5.22 (5.23)	5.459	5.69 (5.64)	5.45	5.45
σ	0.78 (0.78)	0.92	1.12 (1.11)	0.93	0.91

Note: Values between parentheses correspond to theoretical distributions.

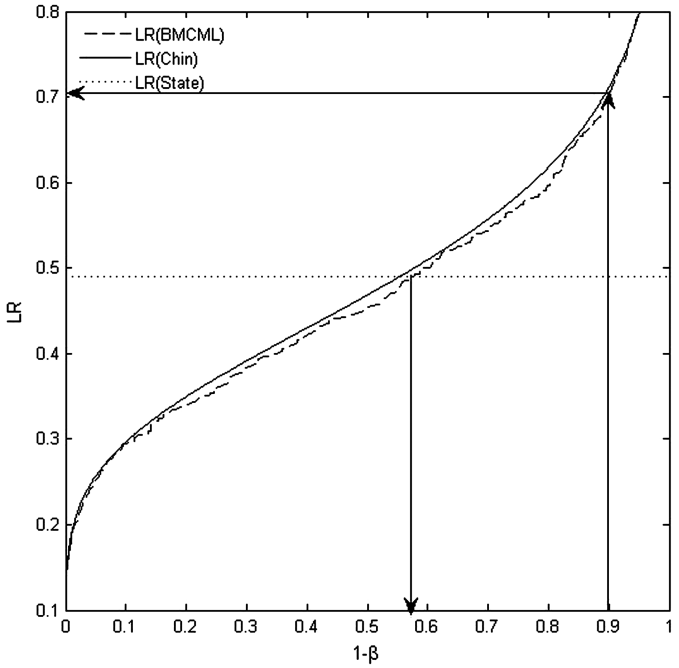

Fig. 7. Required load reduction as a function of confidence of compliance $(1-\beta)$; the 49% LR (as adopted by county) is 68% reliable; a load reduction of 80% would be required to achieve 90% confidence of compliance with the state adopted standard

Each $\boldsymbol{\theta} = (\mu, \sigma)$ defines a lognormal distribution and a unique α-percentile. The α-percentile value $(y_{\alpha i})$ computed for each generated $\boldsymbol{\theta}_i$ has the corresponding posterior probability or likelihood $l(\boldsymbol{\theta}_i)$, $i = 1, \ldots n$. The array y_α and corresponding likelihoods ultimately were used to construct a CDF for the α-percentile, from which all moments and β-percentile of the α-percentile $(y_{\alpha,\beta})$ were inferred.

The relationship between required normalized bacterial concentration reduction $LR = (y_{90,(1-\beta)} - y^*_{90})/\hat{y}_{90}$ and confidence of compliance $(1-\beta)$ is depicted in Fig. 7 using the Bayesian method and the analytical expression used by Chin (2009) for a lognormal data (Hahn and Meeker 1991).

The plotted concentration reductions obtained by the Bayesian methodology fall very close to the plotted reductions computed by the percentile relationship reported by Chin (2009). The adopted TMDL in Fork Creek was that of reducing current loading (i.e., corresponding to $\hat{y}_{90} = 744/100$ mL) to a load allocation that is consistent with meeting the 90-percentile standard ($y^*_{90} = 400/100$ mL) minus a 5% assigned as a MOS. The recommended concentration reduction therefore was $744 - (400 \times 0.95) = 364(/100$ mL). From Fig. 7, $LR = 364/744 = 0.49$ will provide about 57% confidence of compliance as calculated by the Bayesian methodology. Concentration reduction

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

of 70% would be required to achieve 90% confidence of compliance. Defining the normalized MOS $= (y_{90,(1-\beta)} - \hat{y}_{90})/\hat{y}_{90}$, which is consistent with Eq. (38), we find the 43% risk of violating the standard ($\beta = 1 - 0.57 = 0.43$) corresponds to MOS $\approx 3\%$. Note that MOS here is directly tied to the risk of the water-quality standard violation.

Application III: Shunganunga Creek BOD TMDL

Shunganunga Creek, which is located in middle Kansas, has dissolved oxygen as water quality impairment; consequently, Expected Aquatic Life Support is declared as an impaired designated use (U.S. EPA 2007). The creek drains a 73 square miles area to the Kansas River. Most of the watershed is urban (48%), grassland (29%), and cropland (18%). A TMDL model was developed linking dissolved oxygen (DO) to pH, total suspended solids (TSS), biochemical oxygen demand (BOD), and temperature. Multivariate linear regression was implemented to construct the model using 61 measurements of DO and the explanatory water quality parameters (pH, TSS, BOD, T) after removing a few outliers. The measurements were collected from a monitoring site (Station 238) near Topeka, Kansas, not far from the confluence with the Kansas River, over a time period extending from May 1990 to March 2004. We revisit the same model and use the multivariate linear regression results to validate the BMCML method, then use the Bayesian method to estimate the risk associated with the required BOD estimated in the TMDL document (TMDL Document for Shunganunga Creek 2007). The model is

$$DO = \beta_0 + \beta_1 pH + \beta_2 \ln \mathrm{TSS} + \beta_3 \mathrm{BOD}T + \varepsilon \qquad (40)$$

where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. Here, $g(Y) = Y$; $f(t; \mathbf{\Theta}, L) = \beta_0 + \beta_1 pH + \beta_2 \ln \mathrm{TSS} + \beta_3 \mathrm{BOD}T$, and $\mathbf{\Theta} = (\beta_0, \beta_1, \beta_2, \beta_3)$. The loading, which could be controlled to achieve ecologically acceptable oxygen level, is the in-stream BOD, $L = \mathrm{BOD}$. In addition to the 61 sets of observed water quality parameters at Station 238, the priori distributions and size of MC simulation used to conduct the BMCML analysis are $\beta_0 \sim U(-20, 10), \beta_1 \sim U(0.5, 10), \beta_2 \sim U(-5, -0.1), \beta_3 \sim U(-0.36, -0.0036)$, and $n = 100{,}000$. The results are shown in Table 6 and Fig. 8.

The Bayesian estimates of the parameters are remarkably close to their values obtained by regression, with the exception of β_0 (the intercept), which showed a slight difference between the two methods (Table 6). The medians and Bayesian estimates of the parameters are almost similar. The theoretical 95% confidence limits for $\beta_0, \beta_1, \beta_2$, and β_3 are, respectively, $(-19.6, -2.01)$, $(1.91, 4.45)$, $(-2.16, -0.87)$, and $(-0.05, -0.022)$, which also are close to the values obtained by the BMCML technique. The estimate of σ_ε^2 obtained by regression analysis is 2.34 compared to the Bayesian estimate of 2.29. Cross correlation among all parameters was notable. The likelihood-weighted correlation between β_0 and β_1 is -0.99 (almost perfectly, but inversely correlated), -0.66 between β_0 and β_2, and -0.77 between β_2 and β_3.

Fig. 8 shows the BMCML prediction (Bayesian estimate) of DO including observed values at station 238 and 95% prediction confidence limits. Each observation point at the abscissa corresponds to a quadruple of observed values of the predictor variables (pH, TSS, BOD, T). The corresponding results obtained by multivariate linear regression are not shown to prevent clustering. However, the BMCML estimates and 95% confidence limits replicated almost exactly the regressed DO values and the theoretical 95% confidence limits. The E_{NS}, R^2, and SSE for both the BMCML and regression methods were 0.75, 0.75, and 1.48.

Table 6. 95% Confidence Intervals, Medians, BMCML, and Regression Estimates of Model Parameters for Shunganunga Creek BOD TMDL

Parameters	2.5%	50%	97.5%	BMCML	Regression
β_0	−17.35	−7.33	−2.1	−7.63	−8.8
β_1	1.88	2.97	4.24	3.05	3.18
β_2	−2.22	−1.59	−0.93	−1.56	−1.52
β_3	−0.05	−0.036	−0.023	−0.036	−0.036

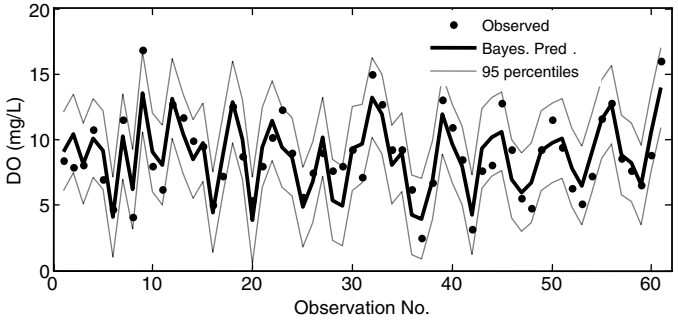

Fig. 8. BMCML estimates compared to observed DO values and their 95% confidence limits (Shunganunga Creek)

Fig. 9(a) displays the BOD loading as a function of the risk (β) of not meeting the DO threshold value of 5 mg/L, with the $P(\mathrm{DO} < 5)$ computed using the formal approach in Eq. (35) and the informal approach. It can be observed that the informal approach tolerates substantially greater BOD loading at higher risk values ($\beta > 0.6$), but tolerates no BOD at all at lower risk values ($\beta < 0.4$). These differences are likely to be caused by significant cross-correlations among some of the parameters as described

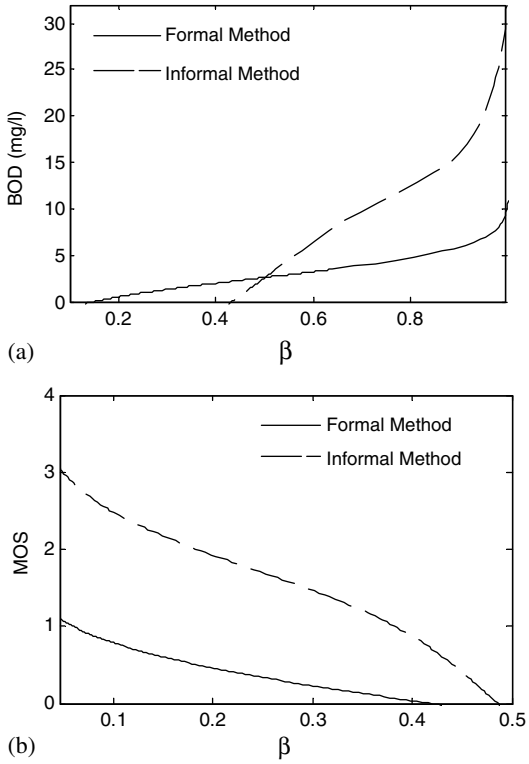

Fig. 9. (a) BOD loading; (b) MOS as a function of risk (β); the solid curve is obtained using the formal method [Eq. (35)], and the dashed corresponds to the informal method (i.e., sampling parameter values from posterior marginal PDFs)

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

above. The relationship between MOS and β computed using Eq. (38) for both the formal and informal approaches is depicted in Fig. 9(b). It is clear that neglecting parameter covariation (informal approach) yields significantly higher MOS than the formal approach.

Using Eq. (40), a BOD TMDL was calculated for the Shunganunga Creek watershed and a target DO value of 5.5 mg/L (instead of the 5 mg/L) was used as a MOS. The following data were also used to obtain a conservative estimate: a median pH = 7.25, maximum $T = 27$°C, and average ln TSS = 4.45. These data and the regression model were used to obtain a target BOD of 2.1 mg/L for the Spring and Summer-Fall seasons. The MOS implemented in this TMDL is an implicit one because it is based on a conservative DO target value, and the safety cushion provided by the inflated target BOD remains unknown. However, from Fig. 9(a), one can quantify the risk associated with a BOD = 2.1 mg/L, which is $\beta = 0.42$. In other words, the adopted implicit MOS yielded a TMDL with 58% chance of yielding DO concentration equal to or above 5 mg/L. Even though conservative values of the input water quality parameters (pH, TSS, T) were used, we were able to account for the uncertainty in the model parameters, model, and measurement errors rather systematically using the BMCML technique and quantify the obscured degree of protection provided by the target BOD recommended in the Shunganunga Creek TMDL document.

Discussion of Results

The lake-sediment phosphorus model application indicates that the BMCML methodology adequately identified the stochastic-error model parameters and physical model parameters given the uncertainty in the input, model structure, and observations. The magnitude of error in the generated observational data was 35% for m_w and 22% for C. The relatively high E_{N-S} and R^2 values showed that the Bayesian methodology performed relatively well in predicting m_w and C during the validation period. True (reference) values of the outputs fell well within the 95% confidence intervals computed by the Bayesian (BMCML) methodology. However, the randomly generated inputs $m_{w,\text{in}}$ and $c_{s,\text{in}}$ were only slightly perturbed and had errors of magnitude 2 and 3%, respectively, and the performance measures deteriorated as the standard deviations of the input variables increased by one to two orders of magnitude, thus reflecting the sensitivity of the methodology to uncertainty in the forcing inputs (results are not shown to conserve space). It is also observed that performance deteriorates with increasing error bias. For example, the E_{N-S} of predicted total phosphorus drops from 0.93 to 0.34 when μ is doubled. However, the deterioration of model performance and wider uncertainty bands of model parameters and outputs rather are anticipated with increasing modeling uncertainty. Nevertheless, the computed uncertainty bands were, in general, wide enough to account for modeling uncertainty and encompass the majority of true values of model parameters and outputs.

The Fork Creek fecal coliform TMDL case-study was the simplest among the three applications, but provided yet another benchmark to validate the Bayesian methodology. The Bayesian estimate of the mean (μ) and standard deviation (σ) of the log-bacterial counts were almost similar to those inferred by the maximum likelihood estimates (MLE). The 95% confidence limits of μ and σ were very close to the exact values obtained from theoretical statistical distributions (refer to Table 5). Further, the β-percentile of the 90th-percentile of log-bacterial counts generated using Bayesian inferences closely replicated those computed on the basis of the theoretical relationship (Hahn and Meeker 1991), which is exclusively applicable to Gaussian data. This TMDL case study furnishes an example of how currently practiced TMDLs need to be modified to formally select the MOS and estimate associated risk of failure on the basis of analysis uncertainty. Alternatively, by estimating uncertainty, the risk associated with currently implemented TMDLs could be assessed using the methodology and could be used during adaptive implementation of the TMDL, a process that examines progress toward achieving water quality standards while relying on monitoring and experimentation to reduce uncertainty (NRC 2001). The state of South Carolina recommended a MOS but did not provide a measure of how good the MOS is. The Bayesian methodology yielded 68% as an estimate for the reliability, which corresponded to an estimated 32% risk of violating the standard.

The Shunganunga Creek, Kansas watershed TMDL further validated the BMCML method by comparison with the multivariate linear regression model and provided yet another example of the power of the methodology in obtaining TMDL loadings and associated risks of violating water quality standards. The likelihood-weighted estimates and 95% confidence limits of model parameters compared very well with known theoretical benchmarks. Most importantly, the degree of protection provided by the use of an implicit MOS was quantified rather systematically as having a 58% chance of meeting the standard. Future monitoring and modeling efforts in the watershed should focus on reducing uncertainty in both observational data and the model.

Cross correlations between parameters may have implications on risk management although more solid demonstrations are still needed. As expected, risk estimation in the hypothetical lake-phosphorus model using the formal approach [Eq. (34)] resulted in a MOS-β relationship identical to that computed using the widely used approach of sampling parameter values from marginal posterior distributions (informal approach); this was because true parameter values were sampled independently and thus were statistically uncorrelated. The Shunganunga Creek TMDL showed that neglecting cross-correlations among model parameters tolerates only zero BOD TMDL at low risk levels (e.g., $\beta < 0.4$), which could potentially lead to costlier abatement strategies. Further, reducing BOD to 0 mg/L would still result in greater than 10% violation of the DO standard of 5 mg/L, as Fig. 9(a) shows for the formal approach. From a management perspective, allocation of wasteloads alone may not be sufficient to achieve a minimum of 90% compliance with respect to DO short of reducing total suspended solids (TSS). In other words, a nonpoint source management strategy might be needed in addition to point source control to achieve a TMDL at no greater than 10% risk of noncompliance.

Despite the concerns raised about the inefficiency of the BMC method (Qian et al. 2003), the preceding demonstrations showed that the performance of the method could be improved by using the likelihood-weighted estimate of the variance of error σ_ε^2 [instead of a priori specification thereof by Qian et al. (2003), or using the minimum value of the variance across all parameter sets by Dilks et al. (1992)] and considering likelihood values of parameter sets in the computations as opposed to sampling of parameter values from posterior marginal PDFs. Sensitivity with respect to the width of noninformative priors was observed as the number of model parameters increases (e.g., Application I).

Summary and Conclusions

A probabilistic framework has been presented for integrated model calibration and water quality management, which relies

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

on Bayesian Monte Carlo method and Maximum Likelihood estimation (BMCML). As in the GLUE methodology, the BMCML considers any parameter set—derived from prior information on feasible ranges of parameter values—as a valid simulator of the environmental system prior to conditioning on observational data. Similarly, by shifting the focus from individual parameter values (i.e., posterior marginal PDF) to the importance of parameter sets, potential interactions between parameters are reflected in the value of the likelihood measure associated with each parameter set; these interactions could influence model predictions and risk estimation and management. However, the methodology follows classical Bayesian inferences and requires the normality of the difference between observed data and model prediction (if necessary, after a proper transformation). Without loss of generality, the error is assumed to be biased and autocorrelated (i.e., described by a first-order Markov process). In contrast to the traditional disjunctive approaches of model calibration and cursory integration of model uncertainty to water quality management, the BMCML framework facilitated a lucid synchronization of model calibration to risk-based water quality management, whereby estimation of the risk of violating fuzzy standards and required load reductions are accomplished on firm application of the rules of mathematical probability, taking into consideration uncertainty in modeling and observational data. Most importantly, the analysis furnished a formal methodology for estimating TMDLs and the risk (i.e., probability of failure) associated with informally selected MOS, which was articulated by the Fork Creek bacterial and Shunganunga Creek watershed TMDL case studies.

The BMCML methodology was examined using three water quality applications and was validated by comparison with statistical theory as a benchmark. In the hypothetical lake-sediment phosphorus application, synthesized true model parameters, stochastic error parameters, and corresponding time-series of true system responses provided reference values to validate the BMCML methodology subject to uncertain system excitations. The lake eutrophication example and Shunganunga Creek case-study showed that neglecting covariation among parameters could lead to the overestimation of low-risk compliance load reductions and, consequently, higher MOS.

The statistical bench marks and example applications altogether point toward an effective BMCML methodology which integrates model calibration and uncertainty estimation to risk management. The application of the risk-based approach to the two TMDL case-studies revealed new insights and valuable information to decision making that otherwise would be masked by the deterministic analyses actually implemented in both TMDLs. Although the BMC technique is inherently computationally intensive for complex models and sensitive to uncertainty in the input forcing, the BMCML method nonetheless is relatively simple to implement regardless of model complexity, and rigorous enough to account for all sources of modeling uncertainty as well as uncertainty in water quality numerical targets and fuzzy classification standards.

Appendix. Evaluation of the Integral in Eq. (34)

Normally-distributed Y: $g(Y) = Y$, $Y \sim N(\mu_Y, \sigma_Y)$

The CDF, $F_Y(y)$, is

$$F_Y(y) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{y-\mu_Y}{\sqrt{2}\sigma_Y}\right)\right] \tag{41}$$

where $\text{erf}(x) = 2/\sqrt{\pi}\int_0^x e^{-u^2}du$ is the error function.

Uniform: $p_{Y^*} \sim U(a,b)$

$$\int_0^\infty \text{erf}\left(\frac{y-\mu_Y}{\sqrt{2}\sigma_Y}\right)p_{Y^*}(y)dy \equiv \Phi(a,b;\mu_Y,\sigma_Y), \quad \text{and}$$

$$\begin{aligned}\Phi(a,b;\mu_Y,\sigma_Y) = \frac{1}{b-a}\int_a^b \text{erf}\left(\frac{y-\mu_Y}{\sqrt{2}\sigma_Y}\right)dy &= \left(\frac{b-\mu_Y}{b-a}\right)\\ &\times \text{erf}\left(\frac{b-\mu_Y}{\sqrt{2}\sigma_Y}\right) - \left(\frac{a-\mu_Y}{b-a}\right)\text{erf}\left(\frac{a-\mu_Y}{\sqrt{2}\sigma_Y}\right)\\ &+ \sqrt{\frac{2}{\pi}}\frac{\sigma_Y}{b-a}\left\{\exp\left[-\frac{1}{2}\left(\frac{b-\mu_Y}{\sigma_Y}\right)^2\right]\right.\\ &\left.- \exp\left[-\frac{1}{2}\left(\frac{a-\mu_Y}{\sigma_Y}\right)^2\right]\right\}\end{aligned} \tag{42}$$

Triangular: $p_{Y^*} \sim T(a, z_m, b)$

$$\begin{aligned}\int_0^\infty \text{erf}\left(\frac{y-\mu_Y}{\sqrt{2}\sigma_Y}\right)p_{Y^*}(y)dy &= \frac{f_m}{z_m-a}G(a;z_m,\mu_Y,\sigma_Y)\\ &- \frac{f_m}{z_m-b}G(b;z_m,\mu_Y,\sigma_Y),\\ f_m &= \frac{2}{b-a}\end{aligned} \tag{43}$$

where

$$\begin{aligned}G(y;u,\mu,\sigma) &= \int_y^u (z-y)\text{erf}\left(\frac{z-\mu}{\sqrt{2}\sigma}\right)dz\\ &= \frac{1}{2}\left\{[(u-\mu)^2 + 2(\mu-y)(u-\mu) - \sigma^2]\text{erf}\left(\frac{u-\mu}{\sqrt{2}\sigma}\right)\right.\\ &\left.+ [(y-\mu)^2 + \sigma^2]\text{erf}\left(\frac{y-\mu}{\sqrt{2}\sigma}\right)\right\}\\ &+ \frac{\sigma}{\sqrt{2\pi}}\left\{[(u-\mu) + 2(\mu-y)]\exp\left[-\left(\frac{u-\mu}{\sqrt{2}\sigma}\right)^2\right]\right.\\ &\left.- (\mu-y)\exp\left[-\left(\frac{y-\mu}{\sqrt{2}\sigma}\right)^2\right]\right\}\end{aligned} \tag{44}$$

Log-normally-distributed Y: $g(y) = \ln Y$, $\ln Y \sim N(\lambda, \zeta)$

In this case, the CDF, $F_Y(y)$, is given by

$$F_Y(y) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\ln y - \lambda}{\sqrt{2}\zeta}\right)\right] \tag{45}$$

Uniform: $p_{Y^*} \sim U(a,b)$

$$\int_0^\infty \text{erf}\left(\frac{\ln y-\lambda}{\sqrt{2}\zeta}\right)p_{Y^*}(y)dy \equiv \Gamma(a,b;\lambda,\zeta), \quad \text{where}$$

$$\begin{aligned}\Gamma(a,b;\lambda,\zeta) = \frac{1}{b-a}\int_a^b \text{erf}\left(\frac{\ln y-\lambda}{\sqrt{2}\zeta}\right)dy &= \frac{b}{b-a}\text{erf}\left(\frac{\ln b-\lambda}{\sqrt{2}\zeta}\right)\\ &- \frac{a}{b-a}\text{erf}\left(\frac{\ln a-\lambda}{\sqrt{2}\zeta}\right) - \frac{e^{\lambda+(1/2)\zeta^2}}{b-a}\\ &\times\left[\text{erf}\left(\frac{\ln b-\lambda}{\sqrt{2}\zeta} - \frac{\zeta}{\sqrt{2}}\right) - \text{erf}\left(\frac{\ln a-\lambda}{\sqrt{2}\zeta} - \frac{\zeta}{\sqrt{2}}\right)\right]\end{aligned} \tag{46}$$

Triangular: $p_{Y^*} \sim T(a, z_m, b)$

$$\begin{aligned}\int_0^\infty \text{erf}\left(\frac{\ln y-\lambda}{\sqrt{2}\zeta}\right)p_{Y^*}(y)dy &= \frac{f_m}{z_m-a}\Psi(a;z_m,\lambda,\zeta)\\ &- \frac{f_m}{z_m-b}\Psi(b;z_m,\lambda,\zeta)\end{aligned} \tag{47}$$

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.

where

$$\Psi(y;u,\lambda,\zeta) = \int_y^u (z-y)\mathrm{erf}\left(\frac{\ln z-\lambda}{\sqrt{2}\zeta}\right)dz$$
$$= \frac{1}{2}\left[(u^2-2yu)\mathrm{erf}\left(\frac{\ln u-\lambda}{\sqrt{2}\zeta}\right)+y^2\mathrm{erf}\left(\frac{\ln y-\lambda}{\sqrt{2}\zeta}\right)\right]$$
$$-\frac{1}{2}e^{2(\lambda+\zeta^2)}\left[\mathrm{erf}\left(\frac{\ln u-\lambda}{\sqrt{2}\zeta}-\sqrt{2}\zeta\right)\right.$$
$$\left.-\mathrm{erf}\left(\frac{\ln y-\lambda}{\sqrt{2}\zeta}-\sqrt{2}\zeta\right)\right]$$
$$+ye^{\lambda+(1/2)\zeta^2}\left[\mathrm{erf}\left(\frac{\ln u-\lambda}{\sqrt{2}\zeta}-\frac{\zeta}{\sqrt{2}}\right)\right.$$
$$\left.-\mathrm{erf}\left(\frac{\ln y-\lambda}{\sqrt{2}\zeta}-\frac{\zeta}{\sqrt{2}}\right)\right] \quad (48)$$

Acknowledgments

The U.S. Environmental Protection Agency through its Office of Research and Development partially funded and collaborated in the research described here under contract (EP-C-11-006) with Pegasus Technical Services, Inc. It has not been subject to the Agency review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

References

Beven, K. (2006). "A manifesto for the equifinality thesis." *J. Hydrol.*, 320(1–2), 18–36.

Beven, K., and Binley, A. M. (1992). "The future of distributed models: model calibration and uncertainty prediction." *Hydrol. Proc.*, 6, 279–298.

Beven, K., and Freer, J. (2001). "Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems." *J. Hydrol.*, 249(1–4), 11–29.

Borsuk, M. E., Stow, C. A., and Reckhow, K. H. (2002). "Predicting the frequency of water quality standard violations: A probabilistic approach for TMDL development." *Environ. Sci. and Tech.*, 36(10), 2109–2115.

Chapra, S. C., and Tarapchak, S. J. (1976). "A chlorophyll a model and its relationship to phosphorus loading plots for lakes." *Water Resour. Res.*, 12(6), 1260–1264.

Chin, D. A. (2009). "Risk-based TMDLs in pathogen-impaired waters." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(2009)135:6(521), 521–527.

Dilks, D. W., Canale, R. P., and Meier, P. G. (1992). "Development of Bayesian Monte Carlo techniques for water quality model uncertainty." *Ecol. Model.*, 62(1–3), 149–162.

Dilks, D. W., and Freedman, P. L. (2004). "Improved consideration of the margin of safety in total maximum daily load development." *J. Environ. Eng.*, 10.1061/(ASCE)0733-9372(2004)130:6(690), 690–694.

Faulkner, B. R. (2008). "Bayesian modeling of the assimilative capacity component of nutrient total maximum daily loads." *Water Resour. Res.*, 44(8), W08415.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian data analysis*, 2nd Ed., Chapman and Hall, London.

Hahn, G., and Meeker, W. (1991). *Statistical intervals: A guide for practitioners*, Wiley, New York.

Hantush, M. M. (2009). "Estimation of TMDLs and margin of safety under conditions of uncertainty." *World Environmental and Water Resources Congress 2009: Great Rivers*, ASCE, Reston, VA.

Liu, Y., Yang, P., Hu, C., and Guo, H. (2008). "Water quality modeling for load reduction under uncertainty: A Bayesian approach." *Water Res.*, 42(13), 3305–3314.

Melching, C. S., and Bauwens, W. (2001). "Uncertainty in coupled nonpoint sources and stream water-quality models." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(2001)127:6(403), 403–413.

Neter, J., Wasserman, W., and Kutner, M. H. (1985). *Applied linear regression models*, Irwin, Homewood, IL.

Novotny, V. (2003). *Water quality: Diffuse pollution and watershed management*, 2nd Ed., Wiley, New York.

Novotny, V. (2004). "Simplified databased total maximum daily loads, or the world is log-normal." *J. Environ. Eng.*, 10.1061/(ASCE)0733-9372(2004)130:6(674), 674–683.

National Research Council (NRC). (2001). *Assessing the TMDL approach to water quality management*, National Academy of Science, Washington, DC.

Patil, A., and Deng, Z. Q. (2011). "Bayesian approach to estimating margin of safety for total maximum daily load development." *J. Environ. Manage.*, 92(3), 910–918.

Qian, S., Stow, C. A., and Borsuk, M. E. (2003). "On Monte Carlo methods for Bayesian inference." *Ecol. Model.*, 159(2–3), 269–277.

Reckhow, K. H. (2003). "On the need for uncertainty assessment in TMDL modeling and implementation." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(2003)129:4(245), 245–246.

Shen, J., and Zhao, Y. (2010). "Combined Bayesian statistics and load duration curve method for bacteria nonpoint source loading estimation." *Water Res.*, 44(1), 77–84.

Shirmohammadi, A. I., et al. (2006). "Uncertainty in TMDL models." *Trans. ASABE*, 49(4), 1033–1049.

Sorooshian, S., and Dracup, J. A. (1980). "Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases." *Water Resour. Res.*, 16(2), 430–442.

U.S. Environmental Protection Agency. (1997). *Guidelines for preparation of the comprehensive state water quality assessments*, Office of Water, Washington, DC.

U.S. Environmental Protection Agency. (1999). *Proposed revisions to the water quality planning and management regulations, proposed rule 40 CFR Part 130, Fed. Reg. 64(162)*.

U.S. Environmental Protection Agency. (2007). *Kansas/lower republican basin total maximum daily load*, Waterbody/Assessment Unit: Shunganunga Creek, Water Quality Impairment: Dissolved oxygen.

Vollenweider, R. A., Giovanardi, F., Montanari, G., and Rinaldi, A. (1998). "Characterisation of the trophic conditions of marine coastal waters with special reference to the NW Adriatic Sea: Proposal for a trophic scale, turbidity and generalised water quality index." *Environmetrics*, 9, 329–357.

Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V., and Robinson, B. A. (2009). "Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?" *Stochastic Environ. Res. Risk Assess.*, 23(7), 1011–1026.

Wang, D., Singh, V. P., Zhu, Y. S., and Wu, J. C. (2009). "Stochastic observation error and uncertainty in water quality evaluation." *Adv. Water Resour.*, 32(10), 1526–1534.

Zhang, H. X., and Yu, S. L. (2004). "Applying the first-order error analysis in determining the margin of safety for total maximum daily load computations." *J. Environ. Eng.*, 10.1061/(ASCE)0733-9372(2004)130:6(664), 664–673.

Zheng, Y., and Keller, A. A. (2007). "Uncertainty assessment in watershed-scale water quality modeling and management: 2. Management objectives constrained analysis of uncertainty (MOCAU)." *Water Resour. Res.*, 43(8), W08408.

Zheng, Y., and Keller, A. A. (2008). "Stochastic watershed water quality simulation for TMDL development—A case study in the Newport Bay Watershed." *J. Am. Water Resour. Assoc.*, 44(6), 1397–1410.

Downloaded from ascelibrary.org by Auburn University on 08/29/16. Copyright ASCE. For personal use only; all rights reserved.