ELSEVIER

Ecological Modelling 159 (2003) 269–277

ECOLOGICAL MODELLING

www.elsevier.com/locate/ecolmodel

On Monte Carlo methods for Bayesian inference

Song S. Qian [a,*], Craig A. Stow [b], Mark E. Borsuk [b,1]

[a] *The Cadmus Group, Inc., 6330 Quadrangle Drive, Suite 180, Chapel Hill, NC 27517, USA*
[b] *Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708, USA*

Received 10 December 2001; received in revised form 22 July 2002; accepted 31 July 2002

Abstract

Bayesian methods are experiencing increased use for probabilistic ecological modelling. Most Bayesian inference requires the numerical approximation of analytically intractable integrals. Two methods based on Monte Carlo simulation have appeared in the ecological/environmental modelling literature. Though they sound similar, the Bayesian Monte Carlo (BMC) and Markov Chain Monte Carlo (MCMC) methods are very different in their efficiency and effectiveness in providing useful approximations for accurate inference in Bayesian applications. We compare these two methods using a low-dimensional biochemical oxygen demand decay model as an example. We demonstrate that the BMC is extremely inefficient because the prior parameter distribution, from which the Monte Carlo sample is drawn, is often a poor surrogate for the posterior parameter distribution, particularly if the parameters are highly correlated. In contrast, MCMC generates a chain that converges, in distribution, on the posterior parameter distribution, that can be regarded as a sample from the posterior distribution. The inefficiency of the BMC can lead to marginal posterior parameter distributions that appear irregular and may be highly misleading because the important region of the posterior distribution may never be sampled. We also point out that *a priori* specification of the model error variance can strongly influence the estimation of the principal model parameters. Although the BMC does not require that the model error variance be specified, most published applications have treated this variance as a known constant. Finally, we note that most published BMC applications have chosen a uniform prior distribution, making the BMC more similar to a likelihood-based inference rather than a Bayesian method because the posterior is unaffected by the prior. Though other prior distributions could be applied, the treatment of Monte Carlo samples with any other choice of prior distribution has not been discussed in the BMC literature.
© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Biochemical oxygen demand; MCMC; Uncertainty analysis; WinBUGS

1. Introduction

Bayesian methods are currently experiencing an increase in popularity in the sciences as a means of probabilistic inference (Malakoff, 1999). Among their advantages are the ability to include prior information, the ease of incorporation into a

* Corresponding author. Tel.: +1-919-403-5105; fax: +1-919-401-5867
E-mail address: sqian@cadmusgroup.com (S.S. Qian).
[1] Current address: Department of Systems Analysis, Integrated Assessment, and Modelling (SIAM), Swiss Federal Institute for Environmental Science and Technology (EAWAG), P.O. Box 611, 8600 Dübendorf, Switzerland.

0304-3800/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0304-3800(02)00299-5

formal decision analytic context, the explicit handling of uncertainty, and the straightforward ability to assimilate new information in contexts such as adaptive management. The Bayesian approach has been shown to be particularly useful for ecological models with poor parameter identifiability (Reichert and Omlin, 1997).

In a modelling application, Bayesian inference concerns the estimation of the values of p unknown model parameters: θ:(θ_1,..., θ_p) about which there may be some prior beliefs. These prior beliefs can be expressed as a probability density function, $\pi(\theta)$, and may be interpreted as the probability placed on all possible parameter values before collecting any new data. The dependence of observations $D=(d_1,\ldots, d_m)$ on the p parameters θ can be expressed as the probability density function, $L(D|\theta)$. This p.d.f. is often referred to as the *likelihood function* and is used to update the prior beliefs on θ to account for the new data, D. This updating is performed using Bayes' theorem which can be expressed:

$$\pi(\theta|D)=\frac{\pi(\theta)L(D|\theta)}{\int_{\theta}\pi(\theta)L(D|\theta)\mathrm{d}\theta} \tag{1}$$

where $\pi(D|\theta)$ is called the *posterior* distribution and expresses the probability of the parameter values after observing the new data. Because the denominator in Eq. (1) is a normalizing constant, Bayes' theorem is often expressed as:

$$\pi(\theta|D)\propto\pi(\theta)L(D|\theta) \tag{2}$$

indicating that the prior expectations are modified by the likelihood function to yield the posterior belief.

Once the posterior distribution is available, any features of θ, such as the marginal distributions or means and variances of the individual θ_i, as well as the predictive distribution of future observations, require integrating over the posterior distribution. For example, the marginal posterior distribution of an individual θ_i can be calculated as

$$\pi(\theta_i|D)=\int_{\theta_{-i}}\pi(\theta|D)\mathrm{d}\theta_{-i} \tag{3}$$

where θ_{-i} represents all θ's except θ_i.

Most Bayesian inference problems can be succinctly expressed as the expectation of a function of interest, $g(\theta)$, evaluated over the posterior distribution:

$$E(g(\theta)|D)=\int_{\theta}\pi(\theta|D)g(\theta)\mathrm{d}\theta \tag{4}$$

where E denotes the expectation operator.

A major difference between Bayesian methods and more familiar parameter estimation methods such as least squares (ordinary and nonlinear), or maximum likelihood, is that inference using Bayes theorem is typically made over the whole support of $\pi(\theta|D)$, not just at single values of θ that optimize a designated objective function (such as the likelihood function, $L(D|\theta)$. This will make little difference in the conclusion for some simple models using an appropriate choice for the prior distribution. However, in some problems a Bayesian approach has been shown to lead to very different conclusions than a classical approach (Ludwig, 1996; Al-Khatib et al., 2001).

An historical limitation of using Bayesian methods for scientific inference was that analytical solutions for the required integrations were available for fairly limited combinations of model forms and probability functions (such as the normal linear model). For most nonlinear models, or models where θ was high-dimensional, the inability to solve such integrals made the implementation of Bayes theorem prohibitively difficult. However, the advent of fast and inexpensive computing has promoted the development of several methods for performing Bayesian inference. We discuss two of the methods that are based on Monte Carlo sampling in the following sections. While the two techniques have similar names, they have some important differences that we will illustrate with a simple example.

Monte Carlo sampling has been widely used in ecological modelling for uncertainty analysis (e.g., van Horssen et al., 2002; Carroll and Warwick, 2001; Hakanson, 2000; Phillips and Marks, 1996; Yool, 1999; Annan, 1997), parameter estimation (e.g., Dilks et al., 1992; Gertner et al., 1999), and model evaluation (e.g., Hakanson, 1995). The Markov chain Monte Carlo sampling method

discussed in this paper differs from the traditional applications of Monte Carlo in that the parent distributions of uncertain parameters do not have to be known explicitly.

2. Markov Chain Monte Carlo

One technique for Bayesian inference that is commonly used among statisticians is called Markov chain Monte Carlo (MCMC). MCMC is a general methodology that provides a solution to the difficult problem of sampling from a high-dimensional distribution for the purpose of numerical integration. The idea behind MCMC for Bayesian inference is to create a random walk, or Markov process, that has $\pi(\theta|D)$ as its stationary distribution and then to run the process long enough so that the resulting sample closely approximates a sample from $\pi(\theta|D)$ (Gelman et al., 1995). These samples can be used directly for parameter inference and prediction. For example, Monte Carlo integration estimates $E(g(\theta|D))$ by drawing n samples, θ^i $(i=1,\ldots, n)$, from $\pi(\theta|D)$ and calculating the mean:

$$E(g(\theta|D)) \approx \frac{1}{n}\sum_{i=1}^{n} g(\theta^i). \quad (5)$$

With independent samples, the law of large numbers ensures that the approximation can be made increasingly accurate by increasing the sample size, n. The result still holds when samples θ^i are not independent, as long as the samples are drawn throughout the support of $\pi(\theta|D)$ in the correct proportions.

There are many ways of constructing the appropriate Markov chain, but all of them, including the Gibbs sampler (Gelfand et al., 1990; Gelfand and Smith, 1990; Smith and Roberts, 1993; Casella and George, 1992), are special cases of the general framework of Metropolis et al. (1953) and Hastings (1970), or the Metropolis–Hastings algorithm. Details of the MCMC method can be found in Gilks et al. (1996). An application of MCMC in ecological modelling can be found in Borsuk et al. (2001).

It is worthwhile to note that we only need to know the the posterior distribution up to a proportional constant to generate random samples using the Metropolis–Hastings algorithm. In other words, we can use the right-hand-side of Eq. (2) for generating posterior samples, without the normalizing constant (the denominator in Eq. (1)). Additionally, using MCMC we can sample from the joint posterior distribution of all parameters, rather then from the marginal parameter distributions. As Reichert and Omlin (1997) and Omlin and Reichert (1999) indicated, multidimensional parameter distributions are important in model uncertainty analysis and may be crucial in locating narrow and 'hard to locate' highest probability density regions.

3. Bayes Monte Carlo

A second technique for Bayesian inference has been presented in the literature under the name 'Bayesian Monte Carlo' or 'Bayes Monte Carlo' (BMC) (Dilks et al., 1992). Applications to date using BMC have included ground water flow modelling (Sohn et al., 2000), air quality prediction (Bergin and Milford, 2000), water resource management (Venkatesh and Hobbs, 1999), soil remediation (Dakins et al., 1996), and health risk assessment (Brand and Small, 1995).

Given a model relating a set of input or predictor variables, X, to a set of output or response variables, Y, or $Y=f(X, \theta)$, the BMC method begins by identifying probability distributions for the model parameters θ that represent the best estimate of the modeler before comparing model results to actual observations. In most applications this has consisted of a uniform distribution for each θ_k that covers what is believed to be the range of plausible values. Multiple sets of model predictions are then generated by randomly drawing $i=1,2,\ldots, n$ sets of samples (θ^i) from the parameter distributions using Monte Carlo methods and generating a set of outputs for each set of parameter draws $[Y^i = f(X, \theta^i)]$. Each of these sets of predictions are then quantitatively compared with actual observations D using a likelihood function expressing the

known or expected error structure. For example, with $j = 1,\ldots, m$ observations and corresponding predictions of a response variable and assuming independent and normally distributed errors with mean 0, and variance σ^2, the likelihood of the ith Monte Carlo prediction, Y^i, is calculated as

$$L(D|Y^i) = \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(D_j - Y_j^i)^2}{2\sigma^2}\right). \quad (6)$$

This likelihood value is then combined with the prior probability of obtaining that prediction (the joint probability of the set of parameters θ^i that generated it) using Bayes' theorem:

$$\pi(Y^i|D) = \frac{\pi(Y^i)L(D|Y^i)}{\sum_{j=1}^{n} \pi(Y^j)L(D|Y^j)} \quad (7)$$

where $\pi(Y^i|D)$ is the posterior probability of the model prediction Y^i. This calculation is repeated for all n Monte Carlo samples with the results interpreted to represent the values of the density function for the corresponding sampled parameter values. The posterior mean, variance, and correlation coefficient can be calculated by the usual methods using the posterior probabilities as weights (Sohn et al., 2000). The posterior statistics summarize uncertainty about parameter values after comparison with the observed data.

While the BMC method has been presented as a 'straightforward', 'objective and statistically rigorous' method for performing Bayesian analysis (Dilks et al., 1992), we believe that BMC is not an effective technique for most applications and may result in incorrect inference. We make this claim for several reasons.

First, according to BMC, all Monte Carlo samples are drawn from the prior parameter distributions, not directly from the posterior. If the individual $\theta_1,\ldots, \theta_p$ are highly correlated or the prior distributions are very wide, the space occupied by the most probable region of $\pi(\theta|D)$, in other words the space we most want to sample, may be an extremely small proportion of the space represented by $\pi(\theta)$. At best, this is likely to be extremely inefficient, resulting in a large number of wasted model runs. At worst, it is possible that few or even no samples will be drawn from the most probable region of $\pi(\theta|D)$. Inferences from such results will be either uninformative or misleading.

Second, the BMC method presents only the joint posterior probability and the marginal posterior probabilities of the p individual parameters, $\pi(\theta_1),\ldots, \pi(\theta_p)$, are usually not estimated. As a result, it is not possible to use the Bayesian credible interval as a means of summarizing uncertainty.

Third, BMC, as generally applied, involves *a priori* specification of the error variance, σ^2, in Eq. (6). Dilks et al. (1992) have suggested that the model error term can be equated with measurement or observation error, and that a value for σ^2 can be derived from statistical analysis of laboratory or field data. However, the assumption that model predictions deviate from observations only because of error in the observations implies an excessive faith in model validity. In reality, model error can result from imperfect system representation, inherent randomness, and sub-scale variability. Specifying an appropriate value for σ^2 in advance is difficult and can strongly influence the estimation of model parameters.

4. Example application

4.1. Model description

To compare the results and efficiency of the two methods of Bayesian inference discussed earlier, we consider a simple example. Borsuk and Stow (2000) present a mixed-order biochemical oxygen demand (BOD) decay model of the form

$$L_t = \begin{cases} L_0 - \{L_0^{1-N} - k_n t(1-N)\}^{\frac{1}{1-N}} + \varepsilon, & N \neq 1 \\ L_0 + \varepsilon, (1 - \exp(-k_n t)) & N = 1 \end{cases} \quad (8)$$

where L_t is the oxygen consumed (BOD exerted, mg/l) at time t, L_0 is the ultimate BOD (mg/l), k_n is the reaction rate constant ((mg/l)$^{(1-N)}$/days), and N is the 'pseudo-order' parameter. This model is a generalization of the first-order decay model, with the order of the reaction, N, left as a free parameter to be estimated from data (rather than

set *a priori* to a value of one, or sometimes two). In this example θ is low-dimensional, consisting of three model parameters, L_0, N, and k_n, as well as the model error variance σ^2. The low-dimensionality makes it possible to use plots in two and three dimensions to illustrate some of the problems encountered using BMC. The example also demonstrates the difficulties in using BMC method with highly correlated parameters. We used the data for wastewater sample 1 in Borsuk and Stow (2000).

4.2. Sampling methods

MCMC sampling was done using WinBUGS, a publicly available program for conducting Bayesian inference using Gibbs sampling (Spiegelhalter et al., 1996). We used a uniform (non-informative) prior distribution for all model parameters so that the results would be comparable to the BMC results. The model error variance, σ^2, was estimated from the data as a free parameter, using a flat prior. Five thousand sets of the four parameters were randomly sampled from the joint posterior distribution.

The BMC results are based on 10 000 sets of uniformly distributed random samples of the three principal model parameters. Prior ranges were 60–100, 10^{-6}–1.0, and 1.0–4 for L_0, k_n, and N, respectively. We set σ^2 to 100, based on reported measurement error values (APHA, 1998), and also to 0.24, the marginal mode from the MCMC sampling, to evaluate the effect of alternative fixed values for σ^2. Marginal distributions of the three model parameters are evaluated with the conventional histogram method used in importance sampling (Tanner, 1993), where the prior range of each parameter is divided into many equally spaced bins (we used 20) and the sum of posterior probability values within each bin is the marginal posterior probability for the parameter value represented by the center of the bin.

4.3. Results

The MCMC results show that three principal model parameters are highly correlated, with only a small proportion of the prior parameter space supported by the data (Fig. 1). Bivariate plots highlight this correlation (Fig. 2a–c). Marginal posterior densities indicate the most likely values and plausible ranges for each model parameter (Fig. 2d–f). The marginal posterior density of σ^2 (Fig. 3) indicates that, as a free parameter, the error variance is well-determined from the data, with a plausible range of approximately 0.1–0.5, and a mode of 0.24.

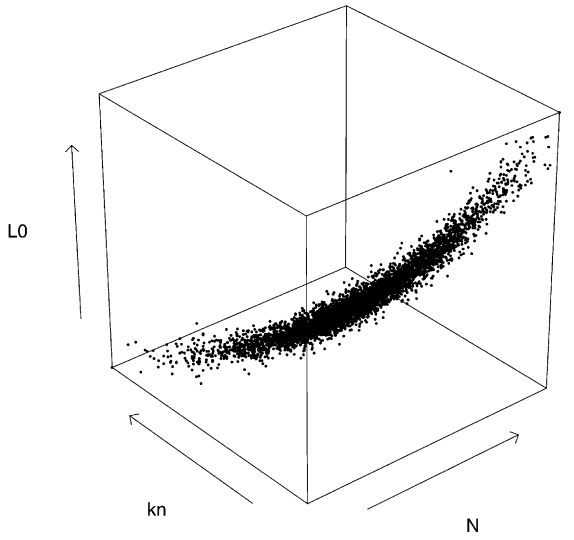

Fig. 1. Three-dimensional plot of the 5000 sets of samples from the joint posterior distribution of the three parameters. Samples are concentrated around a narrow banana-shaped region, representing a fraction of the prior parameter space. Each of the three axes cover only fraction of the their respective prior range.

Using the BMC method, estimated posterior probability values for each of the three model parameters (Fig. 4a–c) suggest irregular marginal probability distributions, with adjacent parameter values having very different probabilities. This result is not possible for a smooth continuous model. It is most likely caused by insufficient number of samples (although 10 000 were used), because the posterior parameter space as estimated by MCMC is very narrow compared to the prior parameter space. The BMC estimated means and S.D. are not consistent with those from MCMC (Table 1). With σ^2 set to 0.24, based on the MCMC mode (Fig. 3), most BMC samples had a posterior probability of effectively zero (Fig. 4d–

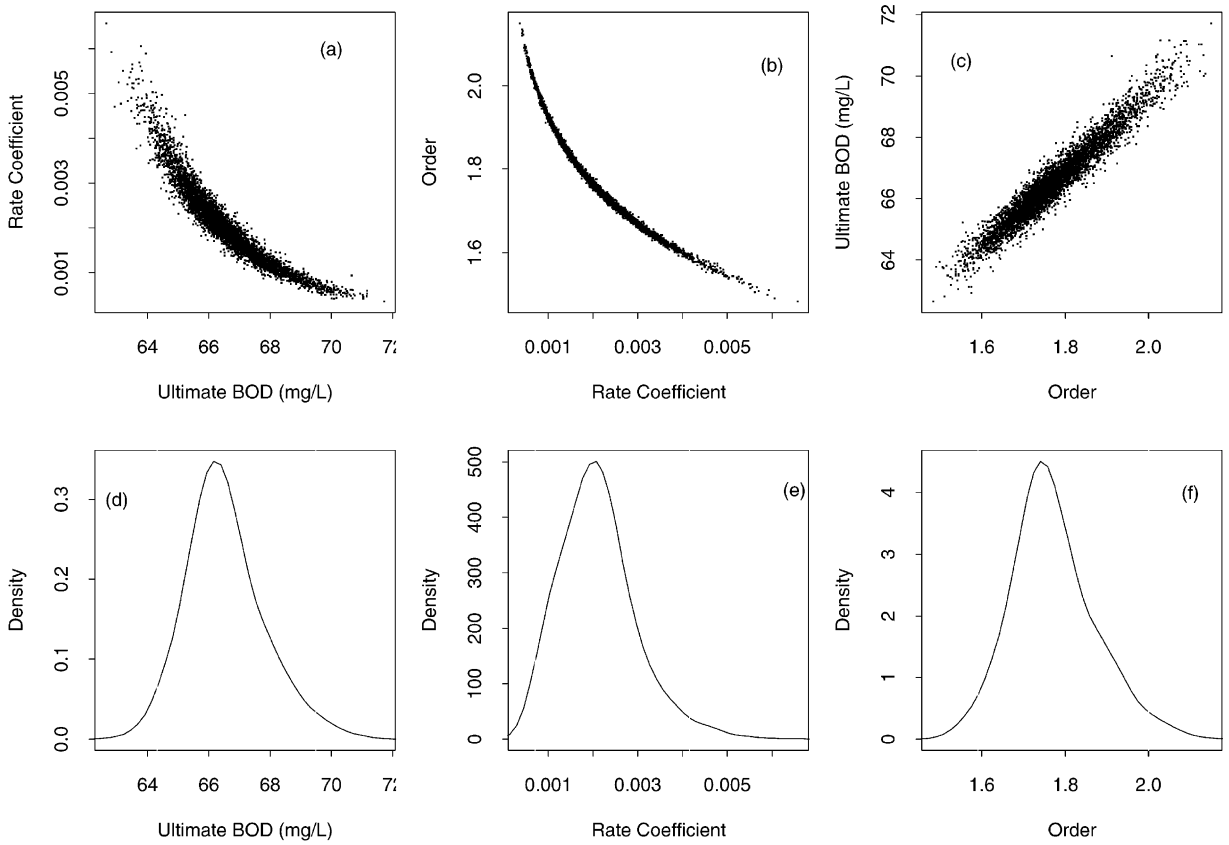

Fig. 2. Bivariate scatter plots (a–c) of the 5000 sets of posterior samples and the marginal posterior distributions (d–f) for the three model parameters. The shaded vertical lines show the 95% credible intervals. Each x-axis covers only fraction of the respective prior range.

f)). Only one sample was chosen from the plausible region of the posterior, in 10 000 samples.

5. Discussion

A major problem in the sampling performance of the BMC is that it does not converge toward the most probable region of the posterior distribution. Consequently, it can be extremely inefficient, rarely sampling from the most probable region (Fig. 1). This problem is exacerbated when the parameter values are not well known, and the prior ranges are made very wide to include all possible values. When this is the case, the chance of sampling the important region becomes increasingly small. Conversely, if, for efficiency reasons, the prior parameter range is kept small, there is a risk of cutting off important regions. The difficulty imposed by these conflicting constraints increases as the dimension of θ increases.

MCMC is specifically designed to sample from the posterior distribution, to eliminate these problems. The procedure is adaptive, so that it will converge on the posterior distribution from an initial starting point. The rate of convergence is case specific. However, techniques are available to evaluate whether the MCMC sample chain has appropriately converged (Brooks and Gelman, 1998).

The choice of an *a priori* value for σ^2 has a clear impact on the sample obtained (Fig. 4). Our choice of two very different values, 0.24 and 100, was done not only to provide a contrast, but also to

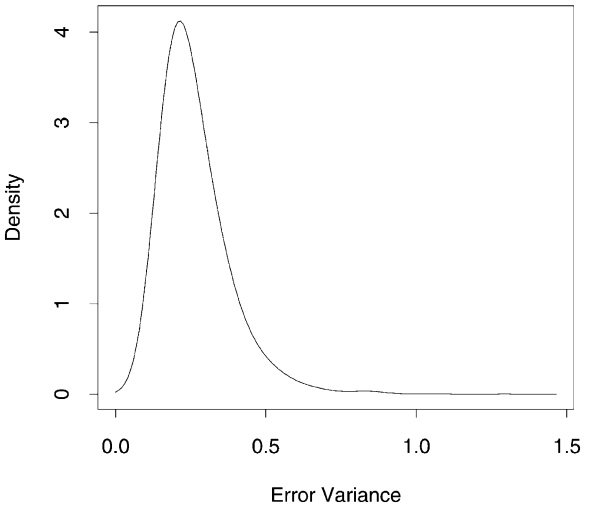

Fig. 3. Posterior marginal distribution of the model error variance σ^2.

indicate the difficulty of picking an *a priori* value. Previous works have based the choice for σ^2 on estimates of measurement error, and 100 was based on reported BOD observation error data (APHA, 1998). However, 0.24, a value well-determined by the data, is obviously very different. We recommend that if prior information is available, it should be expressed, not as a precise value, but as a prior distribution, so that it is subject to modification via Bayes theorem.

The most serious problem with the BMC is that there is no guarantee that the posterior parameter space is sufficiently covered by the Monte Carlo sample. As a result, marginal distributions of model parameters are not guaranteed to be accurate. The volume of the important region of the posterior parameter space can be small compared to the volume of the sampled space, particularly with correlated parameters. In our example this problem is easy to depict, because the model is only four-dimensional (see Omlin and Reichert (1999), for more examples). The 'curse of dimensionality' both exacerbates and obscures this difficulty in more complex, higher dimensional models. In principle, the BMC could produce good marginal estimates, but extremely large Monte Carlo samples would be required. Even if the samples cover the posterior parameter space well, the marginal distributions may still be inadequate if the sample size is not large enough. For example, if there are 1000 samples (as commonly used) and 20 histogram bars are used for estimating the marginal distribution, on average, there are about 50 samples for each marginal probability value. This is probably not enough. This is the reason that the marginal parameter distributions appear irregular. In principle, the BMC could be used to provide a sample for numerically estimated integration, but in many instances it would still be very inefficient, often missing the important region of the posterior distribution.

The Bayes Monte Carlo method will provide sound inference under some very limited conditions. If the parameter space is one-dimensional and the prior range is similar to the plausible posterior range of the parameter, and σ^2 is well-chosen, this approach may work. In one-dimension each value of the estimated parameter is associated with a unique posterior probability value, so the problem of numerical integral estimation disappears. Otherwise, unless one is extremely lucky, the BMC will provide poor estimates.

Finally, we note that the BMC, as implemented in the literature, is more a likelihood-based inference than a true Bayesian procedure. Bayes theorem indicates that the posterior distribution is proportional to the product of the prior distribution and the likelihood function (Eq. (2)). All BMC applications that we could identify in the literature have used uniform priors, with wide ranges chosen for each unknown parameter. Because uniform priors were used, the posterior parameter distribution was proportional to the likelihood function. Thus, evaluation of the posterior could be accomplished by just substituting Monte Carlo samples into the likelihood function. For any other choice of prior this would be incorrect. Additionally, the relative weights applied to the Monte Carlo samples would require modification if any other distribution besides a uniform was used to draw the Monte Carlo samples. While the treatment of Monte Carlo samples has been well explored elsewhere (Geweke, 1989), it has not been addressed in the BMC literature.

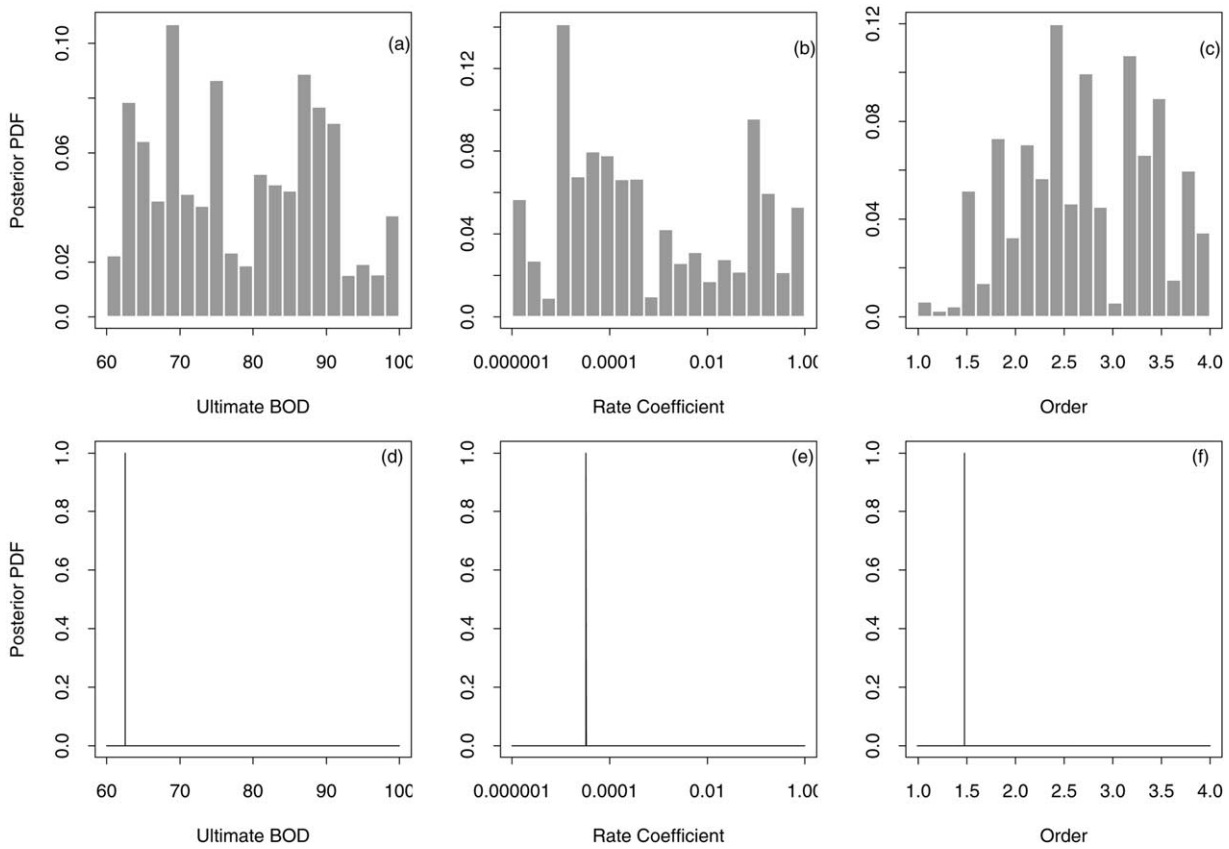

Fig. 4. Marginal posterior probability distributions of the three model parameters calculated using the BMC procedure. (a)–(c) are based on an error variance of 100 and (d)–(f) are based on an error variance of 0.24. Note that each x-axis covers the entire respective prior range.

We encourage the continued implementation of Bayesian techniques for inference in the ecological/environmental sciences. The incorporation of uncertainty will become increasingly important as the necessity to evaluate management alternatives drives our need for ecological prediction (Clark et al., 2001). Bayesian methods provide an explicit, straightforward means of incorporating uncertainty into modelling and forecasting. However, incorrect numerical implementation can negate the advantages of a Bayesian approach and result in misleading inference.

Table 1
Comparison of estimated means and S.D.

	Ultimate BOD		Order		Rate coefficient	
	Mean	S.D.	Mean	S.D.	Mean	S.D
MCMC	66.53	1.27	1.77	0.10	0.0021	0.00085
BMC ($\sigma^2 = 100$)	78.27	10.67	2.69	75.57	0.067	78.2
BMC ($\sigma^2 = 0.24$)	62.52	–	1.47	–	3.29×10^{-5}	–

References

Al-Khatib, S., Califf, R.M., Hasselblad, V., Alexander, J.H., McCrory, D.C., Sugarman, J., 2001. Placebo-controls in short-term clinical trials of hypertension. Science 292, 2013–2015.

Annan, J.D., 1997. On repeated parameter sampling in Monte Carlo simulations. Ecological Modelling 97, 111–115.

APHA, 1998. Standard Methods for the Examinantion of Water and Wastewater, 20th Edition. American Public Health Association, Washington, DC.

Bergin, M.S., Milford, J.B., 2000. Application of Bayesian Monte Carlo analysis to a Lagrangian photochemical air quality model. Atmospheric Environment 34, 781–792.

Borsuk, M.E., Stow, C.A., 2000. Bayesian parameter estimation in a mixed-order model of BOD decay. Water Research 34, 1830–1836.

Borsuk, M.E., Higdon, D., Stow, C.A., Reckhow, K.H., 2001. A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecological Modelling 143, 165–181.

Brand, K.P., Small, M.J., 1995. Updating uncertainty in an integrated risk assessment: conceptual framework and methods. Risk Analysis 15, 719–731.

Brooks, S.P., Gelman, A., 1998. Alternative methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7, 434–455.

Carroll, R.W.H., Warwick, J.J., 2001. Uncertainty analysis of the Carson River mercury transport model. Ecological Modelling 137, 211–224.

Casella, G., George, E., 1992. Explaining the Gibbs sampler. The American Statistician 46, 167–174.

Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A., Lodge, D.M., Pascual, M., Pielke, J.R., Pizer, W., Pringle, C., Reid, W.V., Rose, K.A., Sala, O., Schlesinger, W.H., Wall, D.H., Wear, D., 2001. Ecological Forecasts: an emerging imperitive. Science 293, 660–677.

Dakins, M.E., Toll, J.E., Small, M.J., Brand, K.P., 1996. Risk-based environmental redediation: Bayesian Monte Carlo analysis and the expecte value of sample information. Risk Analysis 16, 67–79.

Dilks, D.W., Canale, R.P., Meier, P.G., 1992. Development of Bayesian Monte-Carlo techniques for water-quality model uncertainty. Ecological Modelling 62, 149–162.

Gelfand, A.E., Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 85, 398–409.

Gelfand, A.E., Hills, S.E., Racin-poon, A., Smith, A.F.M., 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. Journal of the American Statistical Association 85, 972–985.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian Data Analysis. Chapman and Hall, London.

Geweke, J., 1989. Bayesian inference in econometric models using Monte Carlo integration. Econometrica 57, 1317–1339.

Gertner, G.Z., Fang, S., Skovsgaard, J.P., 1999. A Bayesian approach for estimating the parameters of a forest process model based on long-term growth data. Ecological Modelling 119, 249–265.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (Eds.), Markov Chain Monte Carlo in Practice. Chapman and Hall, London 1996.

Hakanson, L., 1995. Optimal size of predictive models. Ecological Modelling 78, 195–204.

Hakanson, L., 2000. The role of characteristic coefficients of variation in uncertainty and sensitivity analyses, with examples related to the structuring of lake eutrophication models. Ecological Modelling 131, 1–20.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109.

Ludwig, D., 1996. Uncertainty and the assessment of extinction probabilities. Ecological Applications 6, 1067–1076.

Malakoff, D.M., 1999. Bayes offers 'new' way to make sense of numbers. Science 286, 1460–1464.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equations of state calculations by fast computing machine. Journal of Chemistry and Physics 21, 1087–1091.

Omlin, M., Reichert, P., 1999. A comparison of techniques for the estimation of model parameter uncertainty. Ecological Modelling 115, 45–59.

Phillips, D.L., Marks, D.G., 1996. Spatial uncertainty analysis: propagation of interpolation errors in spatially distributed models. Ecological Modelling 91, 213–229.

Reichert, P., Omlin, M., 1997. On the usefulness of over-parameterized ecological models. Ecological Modelling 95, 289–299.

Smith, A.F.M., Roberts, G.O., 1993. Bayesian computation via Gibbs sampling and related Markov chain Monte Carlo methods (with discussion). Journal of the Royal Statistical Society, Series B 55, 3–24.

Sohn, M.D., Small, M.J., Pantazidou, M., 2000. Reducing uncertainty in site characterization using Bayes Monte Carlo methods. Journal of Environmental Engineering, ASCE 126, 893–902.

Spiegelhalter, D., Thomas, A., Best, N., Gilks, W., 1996. Bayesian inference Using Gibbs Sampling Manual. MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK.

Tanner, M.A., 1993. Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions, second ed.. Springer, New York.

van Horssen, P.W., Pebesma, E.J., Schot, P.P., 2002. Uncertainties in spatially aggregated predictions from a logistic regression model. Ecological Modelling 154, 93–101.

Venkatesh, B.N., Hobbs, B.F., 1999. Analyzing investments for managing Lake Erie levels under climate change uncertainty. Water Resources Research 35, 1671–1683.

Yool, A., 1999. Comments on the paper: on repeated parameter sampling in Monte Carlo simulations. Ecological Modelling 115, 95–98.