Environmental Modelling & Software 82 (2016) 21–30

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

A comparison of two Bayesian approaches for uncertainty quantification

Thierry A. Mara [a,*], Frederick Delay [b], François Lehmann [b], Anis Younes [b,c]

[a] PIMENT, EA 4518, Université de La Réunion, FST, 15 Avenue René Cassin, 97715 Saint-Denis, Reunion
[b] LHyGeS, UMR-CNRS 7517, Université de Strasbourg/EOST, 1 rue Blessig, 67084 Strasbourg, France
[c] IRD UMR LISAH, F-92761 Montpellier, France

ARTICLE INFO

Article history:
Received 22 September 2015
Received in revised form
8 April 2016
Accepted 8 April 2016

Keywords:
Bayesian parameter estimation
Parameter uncertainty
Predictive uncertainty
MCPD sampler
$\text{DREAM}_{(ZS)}$
MCMC
Soil hydraulic parameter identification

ABSTRACT

Statistical calibration of model parameters conditioned on observations is performed in a Bayesian framework by evaluating the joint posterior probability density function (pdf) of the parameters. The posterior pdf is very often inferred by sampling the parameters with Markov Chain Monte Carlo (MCMC) algorithms. Recently, an alternative technique to calculate the so-called Maximal Conditional Posterior Distribution (MCPD) appeared. This technique infers the individual probability distribution of a given parameter under the condition that the other parameters of the model are optimal. Whereas the MCMC approach samples probable draws of the parameters, the MCPD samples the most probable draws when one of the parameters is set at various prescribed values. In this study, the results of a user-friendly MCMC sampler called $\text{DREAM}_{(ZS)}$ and those of the MCPD sampler are compared. The differences between the two approaches are highlighted before running a comparison inferring two analytical distributions with collinearity and multimodality. Then, the performances of both samplers are compared on an artificial multistep outflow experiment from which the soil hydraulic parameters are inferred. The results show that parameter and predictive uncertainties can be accurately assessed with both the MCMC and MCPD approaches.

1. Introduction

The validation of computer models is an essential task to increase their credibility. One of the most important exercises in the validation framework is to check whether the computer model adequately represents reality (Bayarri et al., 2007). This is achieved by comparing model predictions to observation data. This exercise generally leads to model calibration because the model parameters are usually poorly known a priori (*i.e*. before collecting data). Good practice in calibration of computer models consists of searching for all parameter values that satisfactorily fit the data, thus determining their plausible range of uncertainty. This can be achieved in a Bayesian framework in which the prior knowledge about the model and the observed data are merged to define the joint posterior probability distribution function (pdf) of the parameters. The issue is then to assess the joint posterior pdf.

The inference of model parameter posterior pdf by means of Markov chain Monte Carlo (MCMC) sampling techniques (Metropolis et al., 1953; Hastings, 1970) has received much attention in the last two decades. MCMC explores the region of plausible values in the parameter space and provides successive parameter draws directly sampled from the target joint pdf. Some selection criteria are used to ensure that the successive draws in the chain improve. This means that, throughout the sampling process, probable draws with respect to the target distribution are more likely drawn. Many developments and improvements have been proposed to accelerate MCMC convergence.

Grenander and Miller (1994) developed the Langevin MCMC, which accelerates the convergence of the chains by exploiting the Jacobian of the target distribution. This MCMC sampler may require that the computer model provide the local sensitivities to compute the Jacobian of the target distribution. In practice, modelers generally estimate the gradient by finite differences via a surrogate or coarse-scale model to alleviate the computational burden (see for instance, Dostert et al., 2009; Angelikopoulos et al., 2015). Haario et al. (2006) developed the Delayed Rejection Adaptive Metropolis (DRAM), an algorithm that increases the rate of acceptance of MCMC draws by exploiting the delayed rejection trick

* Corresponding author.
E-mail address: mara@univ-reunion.fr (T.A. Mara).

http://dx.doi.org/10.1016/j.envsoft.2016.04.010

proposed in Tierney and Mira (1999) and the adaptive Metropolis algorithm of Haario et al. (2001). ter Braak and Vrugt (2008) developed the Differential Evolution-Markov Chain (DE-MC) algorithm, which merges the differential evolution method of ter Braak (2006) and the Shuffled Complex Evolution Metropolis (SCEM) method proposed by Vrugt et al. (2003). DREAM improves the efficiency of MCMC by running multiple chains in parallel for a wider and quicker exploration of the parameter space in addition to a self-adaptive randomized subspace sampling (Vrugt et al., 2009). Recently, the algorithm of DREAM has been embedded in UCODE_2014, dedicated to inverse modeling (Lu et al., 2014). Laloy and Vrugt (2012) then developed DREAM$_{(ZS)}$, that ensures convergence with fewer chains in parallel than DREAM.

Recently, Mara et al. (2015) proposed a new probabilistic approach to the inverse problem whose main idea is to maximize the joint posterior pdf of a parameter set with one selected parameter sampling successive prescribed values. This provides the so-called Maximal Conditional Posterior Distribution (MCPD) of the selected parameter. The main advantage of the recent MCPD technique is that parameter distributions can be inferred independently. Therefore, the MCPDs can be simultaneously evaluated on multicore computers (or on multiple computers). This drastically reduces the computational effort in terms of computational time units (CTU).

The MCPD and MCMC samplers assess the same target distribution, namely, the parameter joint posterior pdf. Nevertheless, the two samplers do not provide the same results. In general, the MCPD of a single parameter does not correspond to its marginal posterior distribution. In addition, the MCPD sampler only provides a few sets of probable draws while MCMC generates a large number of draws sampled in agreement with the target distribution. Nevertheless and as advocated in this study, both samplers are valuable Bayesian methods for statistical inverse problems. Hence, the main objective of the present work is to compare the ability of MCPD and DREAM$_{(ZS)}$ MCMC samplers to quantify model output and model parameter uncertainties.

The paper is organized as follows: Section 2 summarizes the inversion in a Bayesian framework and recalls the principles of the recent MCPD technique. The general algorithms ruling the DREAM$_{(ZS)}$ MCMC and MCPD samplers are introduced in Section 3. In Section 4, we discuss on the analogy and the differences between MCPD and MCMC draws. Section 5, emphasizes the comparison between MCMC and MCPD samplings: 1) for the inversion of multimodal and correlated functions, and 2) for the evaluation of soil hydraulic properties from a synthetic one-dimensional drainage experiment. Finally, a summary with conclusions is presented in Section 6.

2. Inverse problem

2.1. Bayesian inference

In probabilistic inverse modeling, the parameter set $\mathbf{x} = (x_1, \ldots, x_d)$ of a computer model is inferred from a set of observation data $\mathbf{y}$ using the Bayesian inference, which defines the conditional joint posterior pdf as follows:

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x}), \tag{1}$$

where $p(\mathbf{x})$ is the prior density that characterizes the investigator's beliefs about the parameters before collecting the new observations, and $p(\mathbf{y}|\mathbf{x})$ is the likelihood function, which measures how well the model fits the data. The parameter set that maximizes Eq. (1), namely:

$$\mathbf{x}^{\mathrm{MAP}} = \underset{\mathbf{x}}{\mathrm{argmax}}\ \mathrm{p}(\mathbf{x}|\mathbf{y}), \tag{2}$$

is called the Maximum A Posteriori (MAP) estimate of the parameters. It is the most probable parameter set given the data and can be inferred via an optimization technique. The marginal posterior pdf that characterizes the uncertainty of a single parameter is defined by the following integral:

$$p(x_i|\mathbf{y}) = \int p(\mathbf{x}|\mathbf{y})\mathrm{d}\mathbf{x}_{-\mathrm{i}}, \quad \forall i = 1, \ldots, d \tag{3}$$

where $\mathbf{x}_{-i}$ represents all the parameters except x_i. Usually, the integral in Eq. (3) is evaluated by a multidimensional quadrature method or by direct summations in a large sample of $p(x_i|\mathbf{y})$ obtained, for instance, via an MCMC technique.

2.2. Maximal conditional posterior distribution

Mara et al. (2015) define the maximal conditional posterior distribution of x_i as follows:

$$\mathcal{P}(x_i) = \max_{\mathbf{x}_{-i}}(p(\mathbf{x}_{-i}|\mathbf{y}, x_i)) \times p(x_i|\mathbf{y}). \tag{4}$$

An informal definition can be given by stating that a point estimate of the MCPD is the maximal value reached by the joint pdf Eq. (1) for a given (prescribed) value of one parameter (*i.e.* x_i). This maximal value, in the context of model inversion, assumes that the set $\mathbf{x}_{-i}$ maximizes Eq. (1), knowing that x_i is prescribed. By applying the axiom of conditional probabilities to Eq. (4), it can be stated that $\max\{p(\mathbf{x}_{-i}|\mathbf{y}, x_i)\} \times p(x_i|\mathbf{y}) = \max_{\mathbf{x}_{-i}}\{p(\mathbf{x}_{-i}, x_i|\mathbf{y})\}$. Therefore, the MAP estimate (when it exists) belongs to the MCPD of all parameters.

In view of the MCPD definition, especially its interpretation in terms of the x_i draws for the other parameters at their optimal values, the MCPD can provide information on the uncertainty attached to a single parameter. Obtaining uncertainties for all parameters is simply achieved by calculating the individual MCPD of all parameters.

3. Parameter uncertainty assessment

3.1. The DREAM$_{(ZS)}$ MCMC sampler

The MCMC samplers generate successive draws of parameter sets that converge toward the posterior density $p(\mathbf{x}|\mathbf{y})$. Several methods are reported in the literature (*e.g.* Grenander and Miller, 1994; Haario et al., 2006; Vrugt et al., 2009; Laloy and Vrugt, 2012), but they all rely on the Metropolis-Hasting algorithm, which proceeds according to the following schedule:

(i) Choose an initial estimate of the parameter set $\mathbf{x}^0$ and a proposal distribution $q(\mathbf{a}, \mathbf{b})$ that randomly derives the parameter set $\mathbf{a}$ from an input $\mathbf{b}$.
(ii) From the current set $\mathbf{x}^k$, generate a new candidate $\mathbf{x}^*$ with the generator $q(\mathbf{x}^*, \mathbf{x}^k)$.
(iii) Compute $\alpha = p(\mathbf{x}^*|\mathbf{y})p(\mathbf{x}^* \leftarrow \mathbf{x}^k)/p(\mathbf{x}^k|\mathbf{y})p(\mathbf{x}^k \leftarrow \mathbf{x}^*)$, where $p(\mathbf{b} \leftarrow \mathbf{a})$ is the transition probability from individual $\mathbf{a}$ to individual $\mathbf{b}$ associated with the generator q. Additionally, draw a random number $u \in [0, 1]$ from a uniform distribution.
(iv) If $\alpha \geq u$, set $\mathbf{x}^{k+1} = \mathbf{x}^*$, otherwise, set $\mathbf{x}^{k+1} = \mathbf{x}^k$.
(v) Resume from (ii) until the chain $\{\mathbf{x}^0, \ldots, \mathbf{x}^k\}$ converges or a prescribed number of iterations k_{max} is reached.

The calculation of $p(\mathbf{x}^*|\mathbf{y})$ in (iii) requires that the forward model