Bayesian calibration and uncertainty analysis of hydrological models: A comparison of adaptive Metropolis and sequential Monte Carlo samplers

Erwin Jeremiah,[1] Scott Sisson,[2] Lucy Marshall,[3] Rajeshwar Mehrotra,[1] and Ashish Sharma[1]

Received 6 November 2010; revised 9 March 2011; accepted 1 April 2011; published 27 July 2011.

[1] Bayesian statistical inference implemented by stochastic algorithms such as Markov chain Monte Carlo (MCMC) provides a flexible probabilistic framework for model calibration that accounts for both model and parameter uncertainties. The effectiveness of such Monte Carlo algorithms depends strongly on the user-specified proposal or sampling distribution. In this article, a sequential Monte Carlo (SMC) approach is used to obtain posterior parameter estimates of a conceptual hydrologic model using data from selected catchments in eastern Australia. The results are evaluated against the popular adaptive Metropolis MCMC sampling approach. Both methods display robustness and convergence, but the SMC displays greater efficiency in exploring the parameter space in catchments where the optimal solutions lie in the tails of the prescribed prior distribution. The SMC method is also able to identify a different set of parameters with an overall improvement in likelihood and Nash-Sutcliffe efficiency for selected catchments. As a result of its population-based sampling mechanism, the SMC method is shown to offer improved efficiency in identifying parameter optimization and to provide sampling robustness, in particular in identifying global posterior modes.

Citation: Jeremiah, E., S. Sisson, L. Marshall, R. Mehrotra, and A. Sharma (2011), Bayesian calibration and uncertainty analysis of hydrological models: A comparison of adaptive Metropolis and sequential Monte Carlo samplers, *Water Resour. Res.*, *47*, W07547, doi:10.1029/2010WR010217.

1. Introduction

[2] Hydrological models are routinely used in applications such as urban planning, policy making and water resource management. Because many of these models are intended to be applicable to different catchments, they typically require parameters that need to be optimized (calibrated) to closely match the input-output behavior of the model to the real system it represents. The usefulness of hydrologic models for the purpose of operational predictions depends on how well the model is optimized. Therefore, the calibration must be conducted carefully to maximize the reliability of the model. In this paper, the models are represented as $Q_t = f(x_t;\phi) + \varepsilon_t$ for $t = 1\ldots T$, where Q_t is the observed flow for a catchment at time t, $f(x_t;\phi)$ is the corresponding model output (modeled flow), x_t is the model input (precipitation and evapotranspiration at time t), ϕ is a vector of unknown variables or model parameters, and ε_t is an error term.

[3] Traditional calibration procedures, which involve "manual" adjustment of the parameter values, are labor intensive, and their success is strongly dependent on the experience of the modeler. Alternatively, automatic calibration approaches that seek to take advantage of the speed and power of computers, while being objective and relatively easy to implement, are becoming increasingly popular [*Boyle et al.*, 2000; *Duan et al.*, 1992; *Sisson and Fan*, 2009; *Sorooshian and Gupta*, 1983]. These automatic calibration methods are primarily developed with the aim of successfully finding a single best set of parameter values. However, in hydrologic modeling, the data used for calibration is of limited length, susceptible to measurement errors, and the model structure assumed offers only an approximate representation of the real system. As such, it is generally impossible to find a unique parameter set leading to simulations consistent with a validation record. For this reason, a variety of methods for assessing parameter uncertainty in hydrologic models have been proposed [e.g., *Kuczera and Mroczkowski*, 1998; *Beven and Binley*, 1992; *Tarantola*, 1987; *Kuczera and Parent*, 1998; *Clark et al.*, 2008]. Out of these approaches, Bayesian inference, implemented by stochastic algorithms such as Markov chain Monte Carlo (MCMC), is popular and effective in providing a natural and general probabilistic framework that simultaneously accounts for both model and parameter uncertainties [*Bulygina and Gupta*, 2009; *Marshall et al.*, 2004; *Micevski and Kuczera*, 2009; *Sisson et al.*, 2006; *Smith and Marshall*, 2008].

[1] School of Civil and Environmental Engineering, University of New South Wales, Sydney, New South Wales, Australia.
[2] School of Mathematics and Statistics, University of New South Wales, Sydney, New South Wales, Australia.
[3] Land Resources and Environmental Science, Montana State University, Bozeman, Montana, USA.

0043-1397/11/2010WR010217

[4] Desirable properties for MCMC sampling methods are high efficiency and good chain mixing, which in combination can lead to rapid chain convergence to the posterior distribution of the model parameters [*Fearnhead*, 2008]. In practice, however, convergence may be disturbingly slow, often as a result of an inappropriate selection of the proposal distribution used to generate candidate moves in the Markov chain. Further, in spite of the stochastic nature of MCMC algorithms, there is a risk that the sampler becomes trapped in local posterior modes when dealing with complex models with many parameters. The issue of efficient implementation of MCMC methods on complex problems has been discussed in a number of recent publications: for example, *Papaspilopoulos et al.* [2003] discuss MCMC algorithm design and implementation, *Jasra et al.* [2007] and *Kou et al.* [2006] investigate the use of population-based MCMC ideas, and *Redelings and Suchard* [2005] highlight the need for tailored algorithms for many applications.

[5] The popular adaptive Metropolis (AM) MCMC algorithm [*Haario et al.*, 1999], which continually determines and updates the covariance matrix of a Gaussian random walk proposal distribution, works well for relatively simple inference problems. However, it can be inefficient and unreliable when the posterior exhibits heavy tails, complex relationships between the parameters, or multiple posterior modes. A single Markov chain with a random walk proposal is typically unable to efficiently surmount such difficulties, and accordingly often requires very long algorithm runs to generate a representative posterior sample. Significant improvements to the efficiency of MCMC simulation can be made by running multiple interacting chains simultaneously [*Vrugt et al.*, 2003]. The use of parallel chains with different starting points enables the sampler to simultaneously explore multiple regions of the parameter space, thereby improving the posterior coverage over that of a single chain. Sampler mixing problems involving high dimensionality of the parameter vector, posterior multimodality, and nonlinear parameter relationships can be efficiently handled by using these approaches [*Vrugt et al.*, 2009].

[6] Despite the variety of data assimilation approaches and their proven benefits for uncertainty analysis and model prediction, the application of these techniques to rainfall-runoff models, in particular with the use of particle filters, has only recently been considered [e.g., *Moradkhani et al.*, 2005b; *Smith et al.*, 2008; *Weerts and El Serafy*, 2006]. Recent publications [e.g., *Doucet et al.*, 2001; *Moradkhani et al.*, 2005a; *Vrugt et al.*, 2005; *Sisson et al.*, 2007] have identified the positive impact on parameter estimation and model predictions when employing sequential data assimilation techniques for parameter estimation. The sequential Monte Carlo (SMC) approach is known for its capability in tracking the states and parameters of a nonlinear dynamical process based on Bayes' rule and is constructed through various sampling and resampling strategies. An SMC sampler can also retain the benefits of MCMC algorithms by adopting an MCMC kernel for the particle mutation (proposal) steps [e.g., *Del Moral et al.*, 2006; *Yang et al.*, 2007]. However, the method does not rely on the convergence of a Markov chain as in the MCMC sampler [*Fan et al.*, 2008], thereby avoiding the difficulty of ensuring (possibly multiple) chain convergence to the posterior distribution when working with complex models.

[7] In this article, we consider the issues of data and parameter uncertainty and the presence of "global" and "local" posterior modes within the Bayesian framework as a result of employing different Monte Carlo alternatives for the calibration of hydrologic models. In particular we consider the sequential Monte Carlo and adaptive Metropolis algorithms. We will demonstrate that significant improvements to sampler efficiency and robustness to the presence of multiple posterior modes can be achieved through a properly designed SMC method for parameter estimation of a hydrological model, when compared to the more popular AM algorithm.

[8] The paper is organized as follows. Section 2 presents a summary of the Bayesian approach and its importance in hydrological modeling, followed by the description of the AM and SMC algorithms. Section 3 provides detail on the hydrological model used, and the background for the case study. The results of the case study are presented, interpreted and discussed in section 4, including a comparison of the performance of each sampler, and a summary of the advantages and disadvantages of the SMC approach. Finally, section 5 presents the main conclusions to the study and the scope for future work.

2. Bayesian Inference and Monte Carlo Sampling

[9] Bayesian inference is based on the posterior distribution, $\pi(\theta|y)$, which characterizes the joint probability distribution of the vector of model parameters, θ given the observed data, y. The posterior distribution is generated by combining the prior distribution, $p(\theta)$, (representing preexisting knowledge) with the likelihood function, $p(y|\theta)$. The attraction of Bayesian methodology is the ability of Bayesian inference to adopt preexisting knowledge through the prior distribution, and update the posterior as additional data becomes available [*Marshall et al.*, 2004]. Bayes' theorem can be expressed as

$$\pi(\theta|y) \propto p(y|\theta)p(\theta). \tag{1}$$

[10] Unfortunately, the nonlinear, multiple parameter nature of most hydrological models, coupled with multiple local posterior modes and discontinuous derivatives, means that there is often no simple analytic form of the posterior distribution [*Bates and Campbell*, 2001]. As such, Monte Carlo methods are routinely employed to estimate posterior quantities of interest [*Gelman et al.*, 1995]. In particular, MCMC samplers are frequently used for this task. These algorithms generate a random, correlated sample (that is, the realization of a Markov chain) from the posterior distribution of parameter values under the hydrological model given the observed data [*Bates and Campbell*, 2001; *Micevski and Kuczera*, 2009]. Numerous MCMC variations exist such as those described by *Kuczera and Parent* [1998], *Marshall et al.* [2004], *Haario et al.* [2006], and *Sisson and Fan* [2010].

[11] The underpinning idea behind MCMC samplers is to construct a Markov chain such that its resulting stationary distribution is the posterior distribution of interest. A realization of the chain after convergence thereby represents a sample from the posterior. Mechanistically, each state in the Markov chain represents a particular configuration of the model parameter values. The next state in the chain is obtained by proposing a new parameter state from a proposal distribution, possibly depending on the current state (for example, the proposal could result from a multivariate Gaussian random walk centered on the current state of the

chain). This proposal is then accepted as the next state of the chain according to an acceptance probability designed to ensure convergence to the desired posterior, or it is rejected, in which case the next state of the chain is set equal to the current state. See, for example, *Gelman et al.* [1995] for further information on MCMC samplers.

[12] Sequential Monte Carlo samplers offer an alternative mechanism for posterior simulation. In essence, SMC samplers propagate a population of parameter vectors (termed "particles") through a specified sequence of related distributions, using importance sampling, particle mutation (movement) and resampling strategies. At each stage of the algorithm, the particle population represents a sample from one of the intermediate distributions in the sequence, with the final distribution in the sequence being the desired posterior distribution. Since it is typically difficult to sample from the final distribution directly, SMC provides a useful platform to tackle the problem in sequential increments, by sampling from each of the intermediate distributions in turn [*Del Moral et al.*, 2006, 2007]. The population-based mechanism of the sampler means that it is sometimes more suitable for posterior simulation than MCMC algorithms and may capture modes or tails of the distribution with greater accuracy and computational efficiency that MCMC alternatives can [see, e.g., *Sisson et al.*, 2007]. For detailed information on SMC samplers, see, e.g., *Del Moral et al.* [2006] and *Doucet et al.* [2001].

[13] Monte Carlo algorithms implemented in hydrologic applications attempt to balance the efficiency or flexibility of the method with the complexity of the hydrologic model parameter space. We now provide further detail on the two Monte Carlo algorithms used in this study: adaptive Metropolis MCMC and sequential Monte Carlo.

2.1. Adaptive Metropolis Markov Chain Monte Carlo

[14] The adaptive Metropolis algorithm [*Haario et al.*, 2001] is a particular instance of the traditional Metropolis algorithm [*Metropolis et al.*, 1953] for MCMC sampling. With the Metropolis method, it is crucial to have a good proposal distribution that induces a well mixing Markov chain, that is, one that quickly and efficiently explores the posterior distribution. This can be difficult to achieve in high-dimensional, multimodal problems with strong parameter dependencies. The AM algorithm implements a Metropolis sampler using a (symmetric) multivariate Gaussian proposal distribution, centered at the current state of the Markov chain. However, rather than requiring the user to provide the associated covariance matrix, which may have a large impact on the performance of the sampler if misspecified, the AM algorithm determines this as proportional to the sample covariance matrix of the past history of the chain. As such, the scale and correlations of the proposal distribution are continually and adaptively determined to relate to the complexity of the underlying posterior distribution [*Haario et al.*, 2001; *Marshall et al.*, 2004]. While the underlying sampler is strictly no longer a first-order Markov chain (as a result of estimating the proposal covariance from the past history of the chain), expectations computed under the resulting stationary distribution coincide with those under the desired posterior [*Haario et al.*, 2001]. Further discussions on the AM algorithm are available in work by *Marshall et al.* [2004], *Craiu et al.* [2009], and P. H. Garthwaite et al. (Adaptive optimal scaling of Metropolis-Hastings algorithms using the Robbins-Monro process, 2010, http://arxiv.org/abs/1006.3690, hereinafter referred to as Garthwaite et al., unpublished manuscript, 2010).

2.2. Sequential Monte Carlo

[15] Sequential Monte Carlo samplers are based on the propagation of a population of parameter vectors (particles) from an initial sampling distribution to the desired posterior distribution. They repeatedly employ three processes: reweighting, mutation (move) and resampling [*Del Moral et al.*, 2006; *Doucet et al.*, 2001]. Suppose that an initial population of N particles (parameter vectors), $\theta_1^0,\ldots,\theta_N^0$, are available, randomly generated from an initial sampling distribution $p_0(\theta)$, where each θ_j^0, for $j = 1,\ldots,N$, represents a possible model configuration. The initial sampling distribution is commonly the prior distribution, $p_0(\theta) = p(\theta)$, although this is not required. As the SMC sampler is constructed using importance sampling concepts, the particle population is represented as a weighted sample $\{\theta_j^0, w_j^0\}$, where the weight $w_j^0 = 1/N$, for $j = 1,\ldots,N$.

[16] The aim of an SMC sampler is to propagate the particle population through a sequence of related distributions, with the final distribution given by the desired posterior. There are a number of ways to achieve this [*Doucet et al.*, 2001]. A common approach is through a geometric bridge [*Fan et al.*, 2008], so that

$$\pi_s(\theta) \propto p_0(\theta)^{1-\beta s}\pi(\theta|y)^{\beta s}, \tag{2}$$

where $\pi_s(\theta)$ denotes the sth distribution in the sequence, $s = 0,\ldots,S$. By specifying some sequence, $\{\beta_s\}$, such that $0 = \beta_0 \le \beta_1 \le \ldots \le \beta_S = 1$, the distribution $\pi_s(\theta)$ will smoothly transition from the initial sampling distribution, $\pi_0(\theta) \propto p_0(\theta)$, when $s = 0$, to the posterior, $\pi_S(\theta) \propto \pi(\theta|y)$, when $s = S$.

[17] A simple importance sampling strategy can be employed to transition the particle population between elements of the distributional sequence. If $\{\theta_j^{s-1}, w_j^{s-1}\}$ denotes the weighted particle population from distribution $\pi_{s-1}(\theta)$ at step $s-1$, then by setting $\theta_j^s = \theta_j^{s-1}$ and

$$w_j^s = w_j^{s-1}\frac{\pi_s\left(\theta_j^s\right)}{\pi_{s-1}\left(\theta_j^{s-1}\right)}, \tag{3}$$

the sample $\{\theta_j^s, w_j^s\}$ represents a sample from distribution $\pi_s(\theta)$ at step s. Note that (3) will increase the weight of particles that are in regions of higher density under $\pi_s(\theta)$ than under $\pi_{s-1}(\theta)$, and decrease the weight of particles that are in regions of lower density under $\pi_s(\theta)$ than under $\pi_{s-1}(\theta)$ Implementing this procedure for $s = 1,\ldots,S$, results in importance weights of $w_j^s = \frac{\pi\left(\theta_j^s\right)}{p_0\left(\theta_j^0\right)}$, which are identical to one-step importance sampling weights from $p_0(\theta)$ to the target $\pi_S(\theta)$. However, as $\theta_j^0 = \theta_j^s$ in such a one-step importance sampling procedure, efficiency of the method is highly dependent on the initial sampling distribution $p_0(\theta)$.

[18] Accordingly, at each stage, s, in the sampler, each particle is mutated (i.e., moved) according to a Markov transition kernel, K_s, so that $\theta_j^s \sim K_s(\theta_j^{s-1}, \theta)$. This permits increased diversity in the particle population. The form of K_s is essentially arbitrary provided the appropriate particle

weight is computed [*Del Moral et al.*, 2006]. However, in this article we will follow [*Fan et al.*, 2008] and define K_s to be an MCMC transition kernel that admits $\pi_s(\theta)$ as the stationary distribution [*Del Moral et al.*, 2006; *Fan et al.*, 2008; *Yoon et al.*, 2009]. In this situation, the particle weight remains unchanged following the MCMC particle mutation step.

[19] Through the process of reweighting and mutation, the quality of the particle approximation to $\pi_s(\theta)$ will deteriorate as s increases. This occurs as the variability of the weights, $\{w_j^s\}$, increases as a function of s: a few particles with large weights will dominate computations performed on the posterior, whereas many particles with small weight will have little impact. A dynamic resampling strategy is commonly employed to avoid this problem. A proxy for the variability of the weights, the effective sample size (ESS), is computed after each reweighting stage of the sampler. The ESS corresponds to the effective number of independent particles in the sample, and satisfies $1 \le ESS \le N$. If the ESS falls below a predetermined level (say, $N/2$), then the particles are resampled from their empirical distribution, $\{\theta_j^s, w_j^s\}$, and the weights are reset to $w_j^s = 1/N$ for $j = 1,\ldots,N$. This allows those particles with very small weights to be discarded in favor of particles with larger weights, which become replicated. See, e.g., *Gordon et al.* [1993], *Liu and Chen* [1998], *Kitagawa* [1996], and *Liu* [2001] for more discussion on resampling techniques.

[20] The full SMC sampler algorithm is presented below. See *Fan et al.* [2008] for more details on this algorithm, while for further reading refer to *Doucet et al.* [2001] and *Del Moral et al.* [2006] for additional details on how this relates to more general SMC algorithms.

Initialization

Generate particles $\theta_1^0,\ldots,\theta_N^0$ from initial sampling distribution $p_0(\theta)$.

Set weights $w_j^0 = 1/N$ for $j = 1,\ldots,N$.

Determine the sequence $\beta_1,\ldots,\beta_S$ and set population indicator $s = 1$.

Reweight

For each particle, $j = 1,\ldots,N$:

Set $\theta_j^s = \theta_j^{s-1}$.

Calculate the new particle weight as $w_j^s = w_j^{s-1} \frac{\pi_s(\theta_j^s)}{\pi_{s-1}(\theta_j^{s-1})}$.

Normalize the weights so that $\sum_{j=1}^{N} w_j^s = 1$.

Resample

Calculate the effective sample size, $\mathrm{ESS} = \left(\sum_{j=1}^{N} w_j^s\right)^2 / \sum_{j=1}^{N} (w_j^s)^2$.

If $ESS < N/2$ then resample the particles from the empirical distribution $\{\theta_j^s, w_j^s\}$, and set $w_j^s = 1/N$ for $j = 1,\ldots,N$.

Mutation

For each particle θ_j^s, $j = 1,\ldots,N$:

Generate a candidate particle $\theta^* \sim q_s(\theta|\theta_j^s)$ from proposal distribution q_s.

Calculate the acceptance probability

$$\alpha = \min\left\{1 \frac{\pi_s\left(\theta^*\right) q_s\left(\theta_j^s|\theta^*\right)}{\pi_s\left(\theta_j^s\right) q_s\left(\theta^*|\theta_j^s\right)}\right\}.$$

With probability α set $\theta_j^s = \theta^*$, else leave θ_j^s unchanged.

If $s < S$ then increment $s = s + 1$ and go to the reweight step.

[21] The sequence $\beta_1,\ldots,\beta_S$ may be determined during algorithm initialization, as described above, or dynamically during algorithm implementation. We implement a dynamic sequence in section 4, the rationale for which is described in section 4.2. Unlike standard MCMC, the proposal distribution, q_s, for the MCMC particle mutation step may freely depend on any particles in previous populations [e.g., *Chopin*, 2002]. For the remainder of this article we specify $q_s(\theta|\theta_j^s) = N(\theta_j^s, \hat{\sum}_{s-1})$, similar to an MCMC independence sampler scheme, where $\hat{\sum}_{s-1}$ denotes the (weighted) sample covariance matrix of $\{\theta_j^{s-1}, w_j^{s-1}\}$. Finally, Monte Carlo estimates of expectations converge to the true expectation under the posterior as the number of particles, N, approaches infinity. As such, N should be large enough to permit a credible representation of each distribution in the sequence $\pi_s(\theta)$. Monte Carlo error and sampler stability will improve with increasing N. For this article we adopt $N = 1000$ particles for the eight-parameter model described in section 3, following *Fan et al.* [2008]. We found that this choice provided good overall performance of the sampler.

[22] We now describe the hydrologic model that will be used to assess the performance of the AM and SMC samplers.

3. Case Study

[23] The case study builds on previous work by *Bates and Campbell* [2001] and *Marshall et al.* [2004]. To provide a fair comparison of our results with these earlier studies, we adopt the same data set, rainfall-runoff model, initial model conditions, and likelihood function. The catchment selected is the Bass River, a 52 km^2 catchment located at Loch in the South Gippsland Basin in Victoria, on the western slope of the Strzelecki Ranges. A total of 11 years of daily rainfall, evapotranspiration, and runoff data is available for this catchment. The rainfall-runoff model used is the Australian Water Balance Model (AWBM) with 8 parameters [*Boughton*, 1993]. The model requires daily rainfall and evapotranspiration data as inputs to provide the predicted runoffs based on water balance (Figure 1).

[24] The AWBM consists of 3 different surface storages, C1, C2, and C3 with the condition that C1 < C2 < C3. These storages are associated with fractional areas, A1, A2, and A3. The combination of the associated surface storage and fractional area determines the soil saturation limits before contributing to the surface overflow. The base flow index (BFI) is defined as the ratio of the amount of base flow to the total amount of surface runoff, and determines the contribution to groundwater storage. In this article we hold the BFI constant with a value of 0.4, following *Bates and Campbell* [2001] and *Marshall et al.* [2004]. The final model parameter is the daily recession constant (K) which determines the base flow from the groundwater storage.

[25] Following *Bates and Campbell* [2001] and *Marshall et al.* [2004], the likelihood function is formulated assuming Gaussian homoscedastic, uncorrelated (independent) error terms, so that

$$p(Q|\theta) = \left(2\pi\sigma^2\right)^{-T/2} \prod_{t=1}^{T} \exp\left\{-\frac{[Q_t - f(x_t;\phi)]^2}{2\sigma^2}\right\}, \quad (4)$$

where ϕ = (K, A1, A2, A3, C1, C2, C3) is the vector of unknown hydrologic model parameters, and $\theta = (\phi, \sigma)$ is the

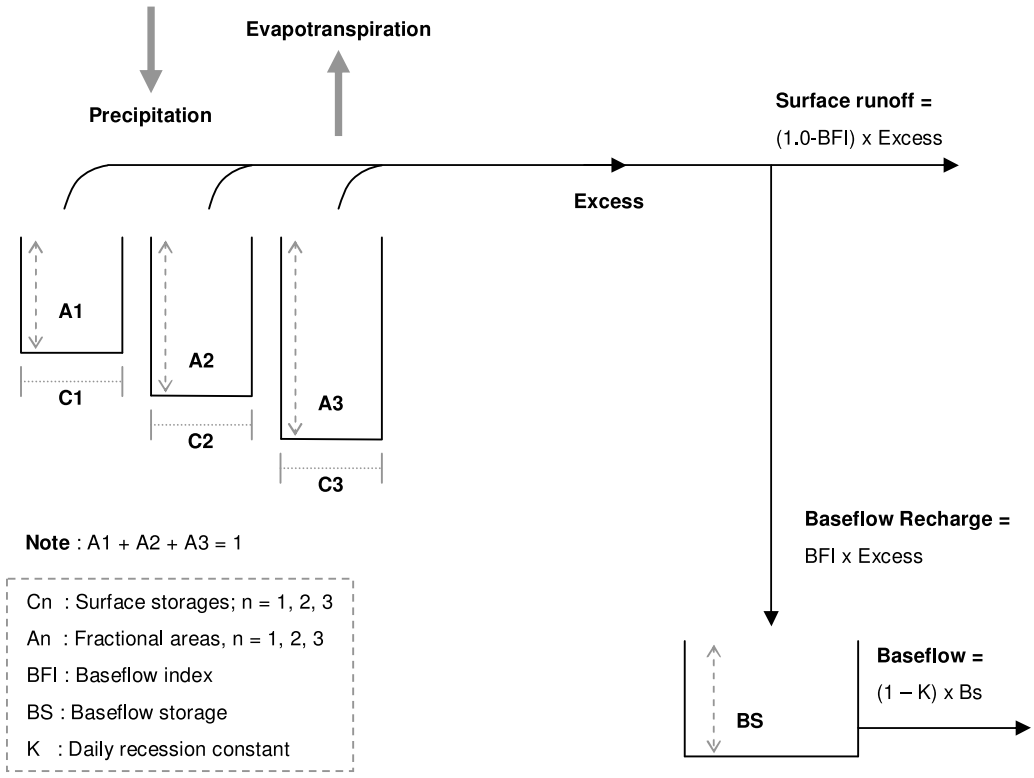

Figure 1. Structure of the AWBM [*Boughton*, 1993].

complete vector of model parameters about which inference is required. Following *Bates and Campbell* [2001], the prior distributions for the parameters are specified as A1 ~ Beta(1.4, 2.6) and A2, A3 ~ Beta(2.0, 2.5) for the fractional areas and $Ck \sim$ Weibull(2.16, $136c_k$) for the surface storage capacities for $k = 1, 2, 3$, where c_1, c_2, c_3 are given by 0.5, 0.75, and 1.5, respectively. The prior for the daily recession constant is given by a mixture of Beta distributions so that

$$K \sim 0.271[\text{Beta}(51.9, 4.17)] + (1 - 0.271)[\text{Beta}(255, 9.6)] \quad (5)$$

and the prior for σ^2 is a scaled inverse chi-square distribution with 8.6 degrees of freedom and scale parameter 46. Readers are referred to *Bates and Campbell* [2001] for additional justifications on the prior distributions used in the AWBM for the case study presented here. Each model run was initialized with all stores empty, with the first 150 days of model simulations discarded to reduce the impact of initial condition assumptions. The SMC sampler in this article uses the joint prior as the initial sampling distribution so that $p_0(\theta) = p(\theta)$.

4. Results and Discussion

[26] We now compare the performance of the AM and SMC samplers as applied to the posterior distribution of model parameters described in section 3.

4.1. Overview of AM Results and Comparisons to SMC

[27] As per *Gilks et al.* [1996], ten Markov chains, each 100,000 iterations long, were generated using the AM sampler to assess convergence of the results. The first 10,000 iterations were discarded as burn-in to reduce the impact of the initial parameter values (randomly sampled from the prior) and the adaptation of the covariance matrix. While automated methods exist for determining the scaling factor of the AM algorithm (see Garthwaite et al., unpublished manuscript, 2010), for simplicity a manual approach was used that ensured acceptance rates above 30%. Sampler convergence was assessed for each parameter independently via the $\hat{R}$ statistic [*Gelman and Rubin*, 1992], defined as the ratio of between-chain variance to within-chain variance. This statistic approximates to a value of one if all the chains have converged, although the converse does not necessarily hold [*Gilks et al.*, 1996]. Similarly, ten SMC sampler replicates were generated using $N = 1000$ particles, with initial values randomly sampled from the joint prior distribution. The replicate runs provide a good indication of the variability of the posterior results obtained [*Chopin*, 2002].

[28] Summaries of the posterior distribution determined by both algorithms are presented in Table 1. The values represent the posterior mean and median of each parameter averaged over the ten sampler replicates, along with the log likelihood evaluated at the maximum a posteriori (MAP) parameter estimate (determined as the maximum log likelihood across all ten replicates). The values in parentheses represent the standard errors of these quantities. Also presented are the marginal $\hat{R}$ estimates for each parameter sampled based on the ten AM replicate chains. These values are approximately one, which is consistent with convergence of the generated chains. Visual inspection of parameter trace plots also supported chain convergence. The results published by *Marshall et al.* [2004] using the AM method are also presented for comparison.

[29] While the AM results are similar to those reported by *Marshall et al.* [2004], the SMC method provides some

Table 1. Summary of Posterior Marginal (Parameter) Distributions for the AWBM Under AM and SMC Samplers[a]

Parameter	*Marshall et al.* [2004] Mean	AM (log(MAP) = −6823.6) Mean	AM (log(MAP) = −6823.6) Median	AM (log(MAP) = −6823.6) $\hat{R}$ Factor	SMC (log(MAP) = −6784.2) Mean	SMC (log(MAP) = −6784.2) Median
K	0.842	0.845 (3×10^{-5})	0.845 (4×10^{-5})	1.003	0.735 (1×10^{-4})	0.735 (1×10^{-4})
A1	0.276	0.278 (4×10^{-4})	0.277 (4×10^{-4})	1.013	0.289 (3×10^{-4})	0.288 (3×10^{-4})
A2	0.489	0.494 (5×10^{-4})	0.494 (5×10^{-4})	1.013	0.468 (3×10^{-4})	0.468 (3×10^{-4})
A3	0.235	0.228 (5×10^{-4})	0.228 (4×10^{-4})	1.017	0.243 (9×10^{-5})	0.243 (9×10^{-5})
C1	107	106.86 (6×10^{-2})	106.71 (5×10^{-2})	1.004	109.35 (6×10^{-2})	108.83 (9×10^{-2})
C2	187	187.70 (6×10^{-2})	187.07 (2×10^{-2})	1.030	195.93 (4×10^{-2})	196.05 (2×10^{-2})
C3	541	421.58 (107×10^{-2})	417.02 (15×10^{-2})	1.015	3108.69 (12×10^{-2})	3109.42 (4×10^{-2})
σ^2	1.378	1.467 (3×10^{-4})	1.466 (5×10^{-4})	1.006	1.215 (5×10^{-4})	1.215 (5×10^{-4})

[a]Posterior means and medians are based on 10 replicate samplers. Standard errors are given in parentheses. Posterior means based on the AM sampler of *Marshall et al.* [2004] are also shown.

notably differing estimates, in particular for parameters K and C3. We now examine these differences and their ensuing impacts on flow estimates in more detail.

[30] Figure 2 illustrates hydrograph comparisons of predicted flows (black lines) to observed flows (gray lines), based on maximum a posteriori (MAP) estimates obtained under both AM and SMC samplers (ascertained as the means of the 10 MAP estimate replicates using each method). The detailed visual inspection indicates that the predicted runoff under the SMC-derived model parameters (Figure 2d) provides a better fit to the low-flow components of the hydrographs, in comparison to the calculated runoff with the

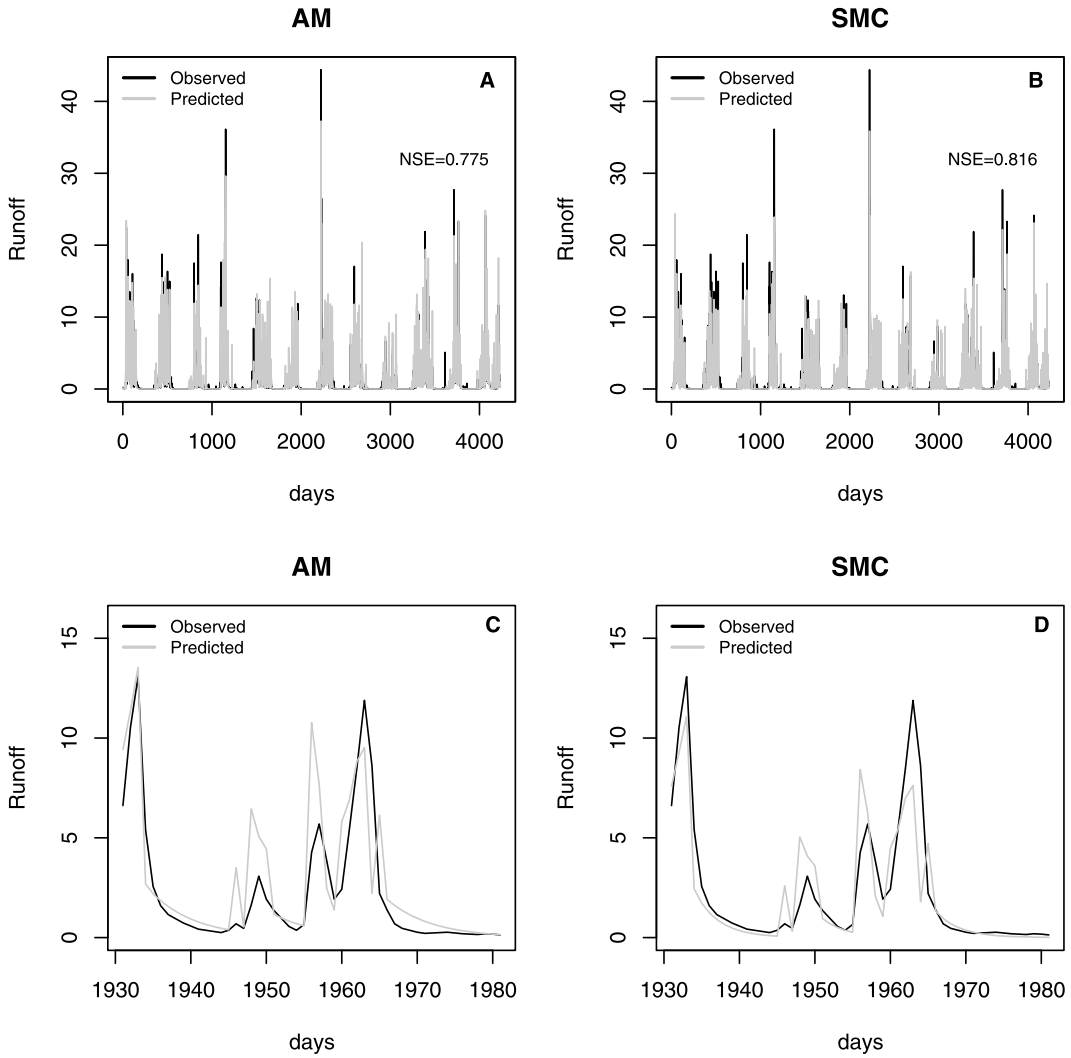

Figure 2. Comparison of observed and calculated runoffs under AM and SMC samplers. (a, b) Comparison of overall runoffs, where black and gray lines denote the observed and calculated runoffs, respectively. (c, d) Runoffs for period from 1930 to 1980 days.

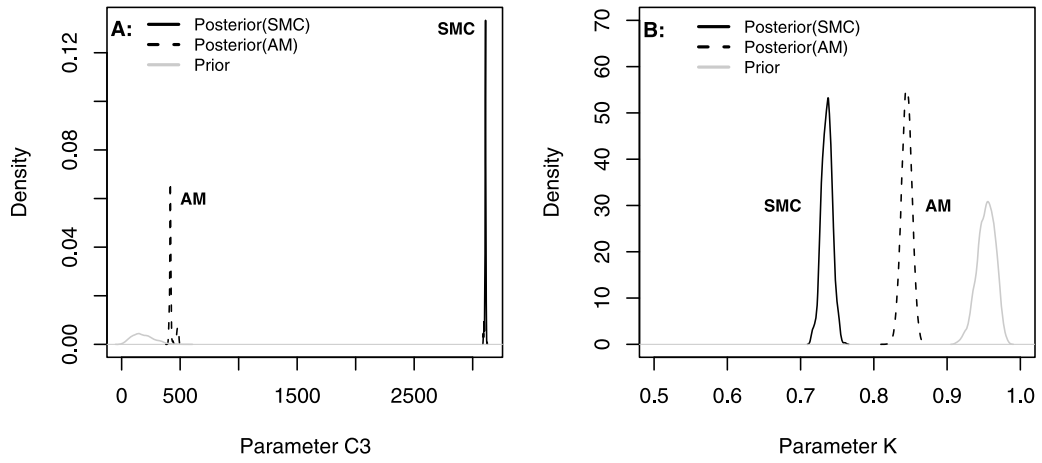

Figure 3. Posterior marginal densities for parameters C3 and *K* with the respective priors under the different samplers. Solid and dashed black lines indicate the marginal posteriors for each parameter under the SMC and AM samplers, respectively. The gray line denotes the prior distribution: (a) Weibull distribution for C3 and (b) a mixture of Beta distributions for *K*.

AM-derived parameters (Figure 2c). The log likelihood evaluated at the MAP estimate is also higher under the SMC sampler (see Table 1), indicating a better fit to the observed data.

[31] The predictive power of each model parameterization can be quantified through the Nash-Sutcliffe model efficiency coefficient (NSE) [*Hall*, 2001; *McMillan and Clark*, 2009]. Here $0 < \text{NSE} \leq 1$ indicates that the model has greater predictive ability than the mean of the observed data, with NSE = 1 indicating perfect prediction. The AM-derived model parameterization gives NSE = 0.775, whereas the SMC-derived parameterization gives NSE = 0.816. As such, by this measure the SMC algorithm produces parameter estimates with better predictive ability.

[32] Figure 3 illustrates the prior and posterior marginal densities of C3 and *K* obtained under both AM and SMC samplers using one sampler replicate. The prior distribution

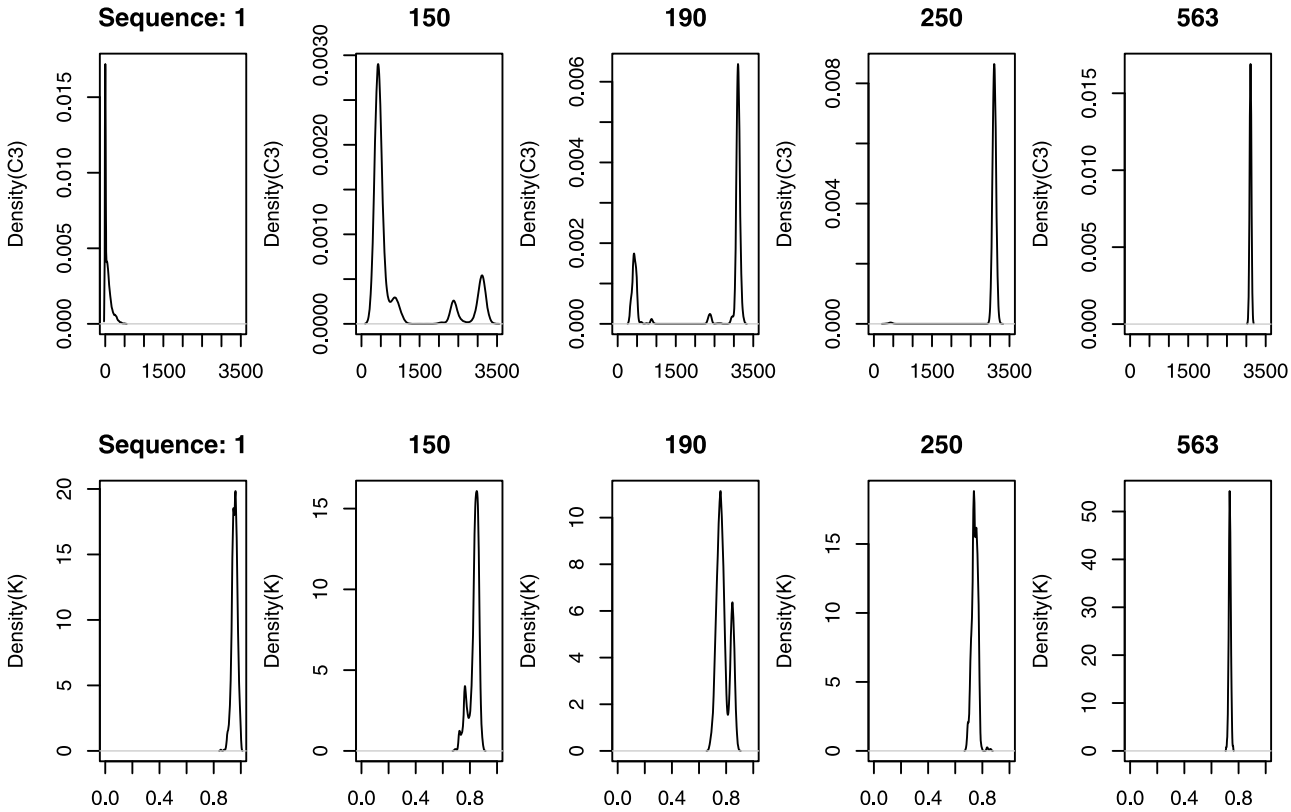

Figure 4. Marginal posterior densities of (top) C3 and (bottom) *K* at various stages of the SMC algorithm. Densities are illustrated at sampler stages *s* = 1 (close to the prior), 150, 190, 250, and 563 (the posterior).

Table 2. Estimates of the Marginal Parameter Modal Values of the Distribution $\pi_{170}(\theta)$ Using the SMC Sampler[a]

Parameter	Mode 1	Mode 2	Mean
C1	108.06	107.171	107.557
C2	189.093	191.988	190.478
C3	448.167	3011.629	1673.502
A1	0.286	0.288	0.286
A2	0.488	0.474	0.482
A3	0.226	0.238	0.232
K	0.846	0.760	0.805
σ^2	1.488	1.295	1.396

[a]The distribution $\pi_{170}(\theta)$ is midway between stages 150 and 190, as shown in Figure 4. Parameter estimates are computed as the mean modal value over the 10 replicate samplers.

of a hydrological model is typically based on regional records or past studies [*Bates and Campbell*, 2001]. Though these can often conflict with the observed data, the impact on the posterior will not be significant given enough data [*Marshall et al.*, 2004]. From Figure 3, the posteriors for each parameter are on similar scales, although there are clear differences in location. This is particularly the case for C3, whereby the SMC sampler identifies a substantially larger surface storage. The location shift for K under the SMC sampler indicates a decrease in discharge from the base flow storage in order to better fit the recession curve of the hydrograph. It seems evident that one of the samplers has become trapped in a local mode of the posterior distribution. Given the degree of conflict in the resulting inferential interpretations, it is of importance to determine which sampler is producing the correct output.

[33] It is conceivably simple for the AM sampler to become trapped in a local mode, although the same argument applies to many MCMC samplers: Consider a bimodal posterior distribution with well-separated modes, and suppose that the Markov chain is initialized near one of these modes. During the initial iterations, the AM sampler will explore the closest mode, and be unable to move to the other modes with the random walk proposal without realizing very low sampler acceptance rates (obtained by increasing the proposal variance, for example). When the adaptive phase of the AM sampler begins, the random walk proposal variance will be determined by the scale of the local mode. If the scale of separation of the posterior modes exceeds the range of this proposal, then the sampler will have little opportunity to escape the local mode. As such, there is a strong possibility that the AM sampler will become trapped in the first mode that the Markov chain explores [*Vrugt et al.*, 2009]. Markov chain mixing between multiple modes is a well known problem [e.g., *Celeux et al.*, 2000]. In the current situation, the marginal posterior distributions (Figure 3) for C3 and K, for the AM sampler, are both closer to the prior high-density regions than the posterior obtained through the SMC sampler. This observation is consistent with the idea of the AM sampler being initialized using random draws from the prior distribution, and then "converging" to the closest mode of the posterior. To confirm this, of the 10 replicate AM sampler runs, each randomly initialized from the prior, all failed to converge to the posterior distribution identified using the SMC method. A further implication of this analysis is that a multiple-chain MCMC sampler, similarly initialized, could reasonably be inferred to also have the same convergence problems as the AM method.

[34] In some sense, the SMC sampler utilizing an MCMC mutation kernel is similar to the AM scheme: an MCMC update step is performed on each particle using an MCMC independence sampler with covariance matrix adaptively determined by the sample covariance of the previous (particle) population, with the initial sampling distribution usually taken as the prior. Quite naturally, the SMC sampler will initially locate modes closest to the prior (initial sampling distribution). However, two aspects of the SMC sampler make it less susceptible to becoming stuck in local modes than the AM sampler. First, particle mutations are performed on the whole population of particles at each stage, rather than on single vector AM sampler updates. As such, the parameter space coverage of the proposed MCMC moves under the SMC sampler is larger than that under the AM sampler, even when comparing a single update of the N particles under the SMC sampler to N repeated updates of the AM Markov chain. This argument has some similarities to the mixing benefits obtained when combining multiple independent MCMC samplers, over a single Markov chain sampler [*Haario et al.*, 2006; *Smith and Marshall*, 2008; *Vrugt et al.*, 2003].

[35] The second benefit of the SMC sampler is that it is implemented on a sequence of related distributions (e.g., equation (2)) that gradually evolve from the initial sampling distribution (e.g., the prior) to the posterior distribution. The standard AM/MCMC sampler explores the posterior distribution only (that is, the final distribution in the SMC distributional sequence), which is potentially highly peaked and multimodal, and is therefore difficult to traverse. Because the distributional sequences of the SMC sampler tend to include many more diffuse distributions than the final posterior, it is far simpler for the sampler to move around the parameter space, and not become trapped in local modes. This idea could also be incorporated in a standard MCMC sampler through simulated annealing or tempering arguments [e.g., *Geyer and Thompson*, 1995], although incorporating the adaptive proposal mechanism of the AM sampler in this setting could be problematic.

[36] The outcome of these advantages can be observed in Figure 4, which displays estimates of the marginal densities of C3 and K at various stages of the SMC algorithm. The marginal densities at iteration $s = 0$ correspond to the prior $\pi_0(\theta) \propto p_0(\theta)$, whereas iteration $S = 563$ corresponds to the target posterior. For both C3 and K, the SMC sampler initially ($s = 150$) identifies the closest mode to the prior, although allowing some mass on alternative modes. The dominating mode corresponds to the posterior identified by the AM algorithm. Between iterations 150 and 190 there is a transition whereby the posterior mass is transferred from one mode to the other. This occurs as $\pi_s(\theta)$ contains a greater likelihood contribution for increasing s, and so the data begin to favor one mode over the other. Finally by iteration $s = 563$, the initial mode has been ruled out as effectively containing zero mass (to the limits of the accuracy of the particle approximation).

[37] To demonstrate that the AM sampler offers support for the posterior distribution obtained by the SMC sampler, we reimplemented the AM algorithm with the modification that the Markov chain was initialized midway between the two competing posterior modes. More precisely,

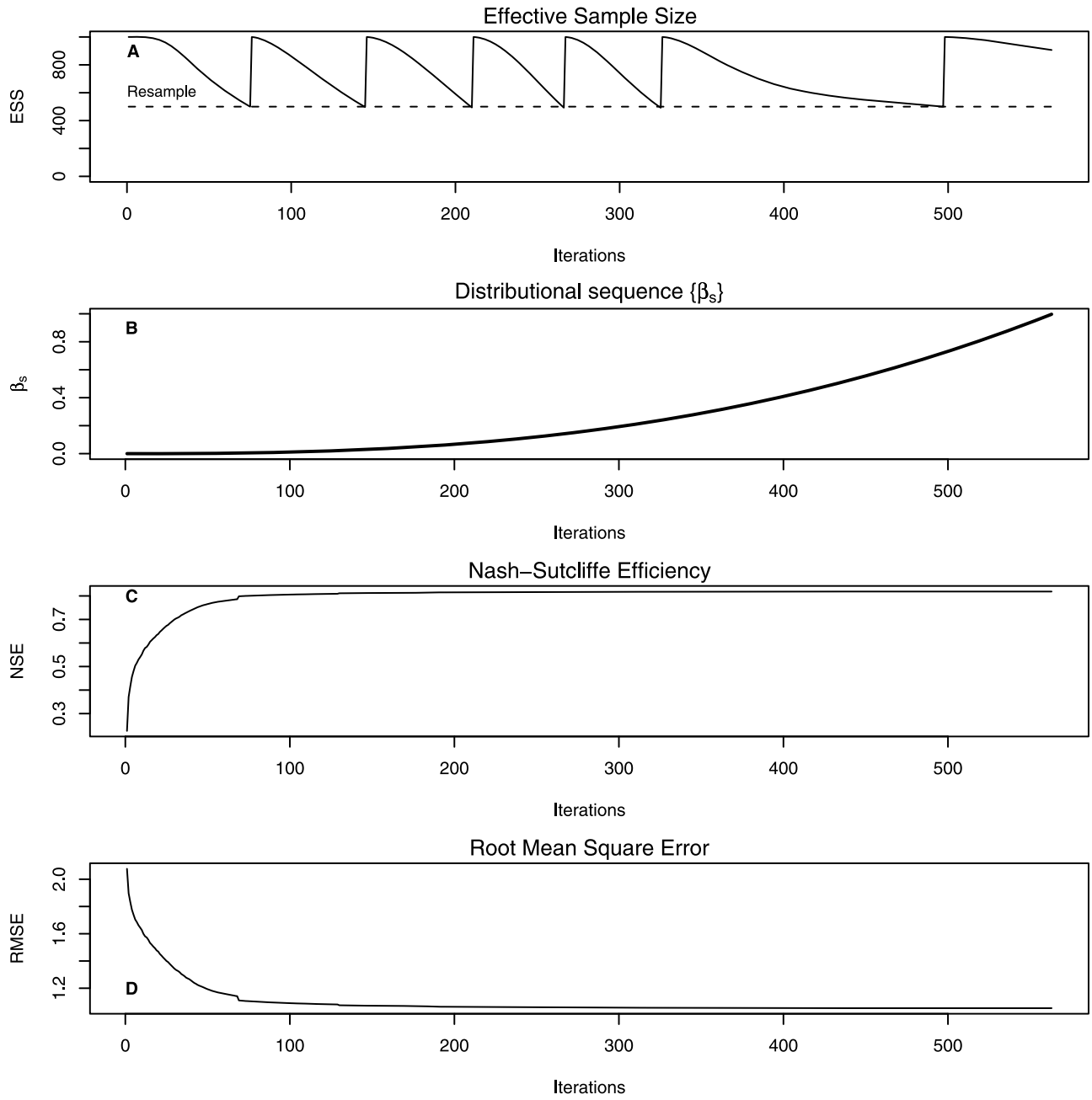

Figure 5. The dynamic evolution of the SMC sampler and the AWBM model fit as a function of algorithm iteration. (a) The effective sample size, with the resample threshold set at 0.5*N*. (b) The transition path of β_S (c) model's predictive performance through the Nash-Sutcliffe coefficient of efficiency and (d) the model's root-mean-square error (based on the MAP estimate).

Table 2 presents marginal posterior mode estimates for each parameter based on the $\pi_{170}(\theta)$ distribution. This distribution was chosen as it is midway between $\pi_{150}(\theta)$ and $\pi_{190}(\theta)$ (see Figure 4), and the two modes are particularly well defined for $s = 170$. The AM sampler was initialized at the componentwise mean of the two modes (last column of Table 2). In each case of ten sampler replicates, the AM rapidly converged to the posterior mode identified by the SMC algorithm, and produced a posterior sample indistinguishable from that obtained by the SMC sampler. This clearly illustrates that the SMC sampler had previously identified the correct posterior distribution, and that appropriate initialization of the AM algorithm is critical to produce correct sampler output.

4.2. Determination of the $\beta_0,\ldots,\beta_S$ Schedule

[38] In practice, SMC algorithms are susceptible to particle degeneracy, whereby the variability of the particle weights increases at each sampler iteration. Left unchecked, this results in a relatively few particles possessing almost all the posterior mass, resulting in poor and highly variable

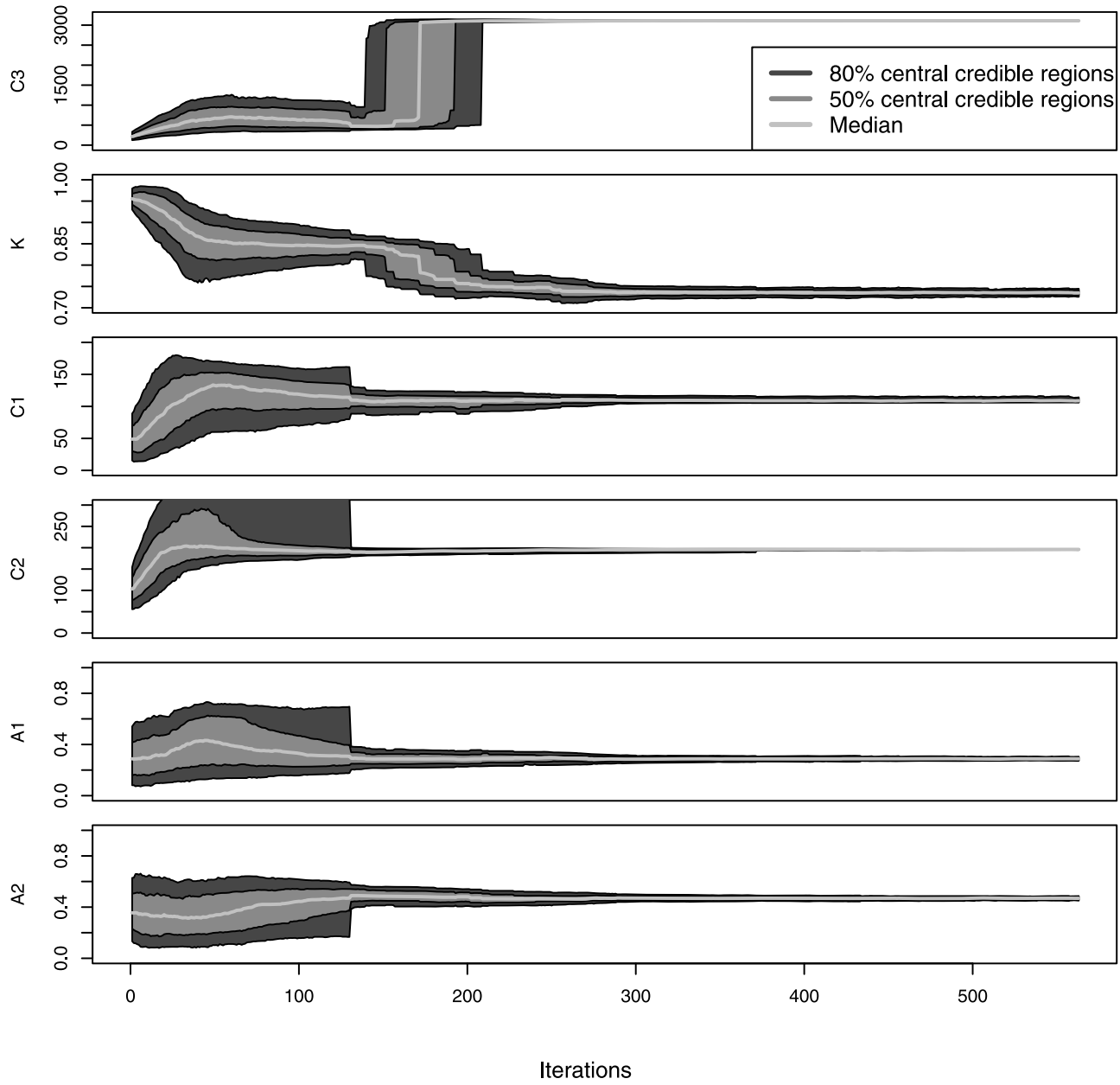

Figure 6. Posterior marginal parameter uncertainty against iteration of the SMC algorithm. In each plot, the white line corresponds to the distributional median, while the light and darker shaded sections represent pointwise 50% and 80% central credible regions, respectively.

posterior inference [*Arulampalam et al.*, 2002; *Moradkhani et al.*, 2005b]. One way this is controlled in the SMC algorithm is through the resample step, which dynamically occurs if the effective sample size (a proxy for the variance of the weights) drops below a predetermined level, here $N/2$. After resampling, the empirical weights are equal (i.e., no variability), although at the price of particle replication. The primary driver of weight deterioration is the distributional sequence $\pi_0(\theta)$ to $\pi_S(\theta)$ If the change between any $\pi_{s-1}(\theta)$ and $\pi_s(\theta)$ is too great, the variability of the importance weights, which is determined by (3), will rapidly increase. A robust sampler will reduce the chance of this occurring. A typical approach to determine an appropriate distributional sequence is to perform a preliminary analysis, such as a fixed linear specification, resulting in a predetermined number of sequential steps. Naturally, this approach may cause a too large between-distribution change, thereby causing the particle system to collapse. Our initial investigations used this approach, but soon we realized that an exponential change specification gave a more realistic means of achieving the targeted posterior distribution. However, given that this approach can also possibly collapse, we developed a method to automate the distributional sequence and allow it to adapt to the problem at hand.

[39] In this article, for the distributional sequence (2), we adopt the exponential relationship $\beta_s = P_0 s^{P_1}$ with P_0, $P_1 > 0$, for which β_s grows initially slowly, but at an increasing rate with s. However, at some stage in the algorithm, if the value of β_s is too large, the SMC particle approximation may rapidly deteriorate. To prevent this, we

Table 3. Proposed Number of Particles to Be Used With the Hydrological Model Dimensions[a]

Model Dimension	Proposed Number of Particles
6 parameters	600
8 parameters	900
12 parameters	2000
16 parameters	>6500

[a]It should be noted that the 16 parameter model was assessed with a maximum of 6500 particles.

introduce an adjustment to the rate of change in β_s, if at any stage the effective sample size falls below a prespecified amount; here we set this value to be $0.2N$. In this case, the value of P_1 is reduced by some amount (while maintaining $P_1 > 0$), and the sampler then returns to the mutation step at the previous stage of the algorithm (i.e., before the deterioration of the particle approximation to the posterior) using the new value of P_1. As such, the rate of change of the distributional sequence $\pi_s(\theta)$, and accordingly the quality of the SMC particle approximation to the posterior, can be dynamically controlled.

[40] Additionally, this approach provides an improved degree of computational efficiency, as further computations (arising from a slower rate of increase in β_s, and therefore more steps S before $\beta_s = 1$ is achieved) are introduced only when they are required. This idea could be optionally extended to increase the rate of change of β_s (by increasing P_1) if, for example, the effective sample size has remained above some high level for a certain number of consecutive steps (that is, when the rate of change is too slow). See, e.g., *Drovandi and Pettitt* [2011] for an alternative example of dynamic determination of the distributional sequence, $\pi_s(\theta)$ For the present analyses we specify $P_0 = 7 \times 10^{-8}$ (recall that $\beta_1 = P_0$ should be small), $P_1 = 2.6$ (which in combination with $P_0 = 7 \times 10^{-8}$ determines a slowly changing β_s for the initial iterations), and dynamically reduce $P_1 = P_1 - 0.15 \times e^{(-0.2r)}$ thereby ensuring that $P_1 > 0$ decreases by increasingly small amounts for $r = 3,4,\ldots$.

[41] With this specification, Figure 5 illustrates the dynamic evolution of the SMC sampler, and the performance of the AWBM fit, as a function of algorithm iteration $s = 0,\ldots,S$. Figure 5a displays the changes in the effective sample size: the algorithm performs a resample step when the ESS drops below $0.5N$ (dashed line). The resulting sequence of β_S is shown in Figure 5b. The implementation of an exponential curve for β_S has proven to be effective. However, with a slower rate of change between each β_S especially in the initial stages of the simulation, the ESS experienced rapid decline until iteration 300, with gradually smoothening out without collapsing. The fast decline of the ESS coincides with the mass of particle population shift as seen in Figure 4.

[42] Figures 5c and 5d illustrate the evolution in the model's predictive performance, through the NSE and the root-mean-square error (RMSE), based on the posterior MAP estimate. The RMSE enumerates the uncertainties that exist between the predicted runoffs under the model, and the observed data [*Johansson and Lehmann*, 2009]. Aside from the initial rapid improvements in model performance (following algorithm initialization in a region of low posterior density), the region of greatest improvement in both NSE and RMSE is again in the region of greatest change in $\pi_s(\theta)$: at around 150–250 iterations. This is further illustrated by Figure 6, which displays the marginal median, and 50% and 80% central credible regions of several model parameters as a function of sampler iteration. The between-mode transition around 150–190 iterations for C3 and K is especially evident (Figure 6 displays a "jump" transition once more than 10% of marginal posterior mass is contained within the second mode). Beyond this point, the credible regions of all parameters, and the RMSE and NSE become more stable, and slowly improve as the SMC sampler gradually incorporates more of the target posterior distribution with increasing s.

4.3. Sensitivity Analysis

[43] The effectiveness of the SMC sampler in exploring the parameter space also depends on the concentration of the particle population: the higher number of particles (N), the better is the representation of the posterior distribution. However, computational overheads also increase with the number of particles. A sensitivity analysis was performed to determine the minimum number of particles required to achieve a prespecified precision, as a function of the model dimension. We replicate the AWBM hydrological model analysis but vary the number of surface storages, C and fractional areas, A. Altogether, 4 AWBMs with 6, 8, 12, and 16 parameters were used over a range of particle population sizes. Each combination was replicated 100 times with the aim of obtaining the Monte Carlo variability of posterior estimates. The optimal number of particles to be used for each model configuration was then assessed as the number that led to the variability in the MSE (across the 100 replicates) being equivalent to that obtained using the most complex formulation (16 parameter AWBM using 6500 particles). Based on this assessment, an optimal number of particles

Table 4. Nash-Sutcliffe Efficiency (NSE) for the AWBM Hydrological Model, Based on MAP Parameter Estimates Obtained Under AM and SMC Samplers, Using Data From Nine Other Catchments Located Within Australia

Catchment	Size (km^2)	Data Length (days)	Rainfall/Runoff Ratio	NSE – AM	NSE – SMC
Black River	260	8766	0.26	0.66	0.66
Barambah Creek	640	28,489	0.09	0.70	0.70
Orara River	135	26,663	0.42	0.62	0.62
Murrumbidgee River	1891	26,298	0.15	0.48	0.48
Mosquito Creek	1130	9,862	0.04	0.77	0.78
Cotter	148	12,419	0.28	0.60	0.61
Magela Creek	260	7,305	0.3	0.45	0.45
Leven River	500	12,784	0.57	0.78	0.78
Never Never River	51	5,844	0.55	0.65	0.65

for each of the models was ascertained (Table 3). Note that the results here are sensitive to the specified degree of variability in the MSE. Alternatively precisions and statistics may be a more appropriate in different modeling scenarios. As such, the 1000 particles used in this article (for an 8 parameter AWBM) appear appropriate, and are consistent with other applications [e.g., *Fan et al.*, 2008; *Del Moral et al.*, 2007].

5. Conclusion

[44] The aim of this study was to evaluate the utility of two Monte Carlo samplers, the AM and the SMC algorithm, in estimating the posterior distributions of AWBM parameters. We have demonstrated that the SMC is more reliable than the AM sampler in this respect: the AM algorithm, if initialized poorly, can become trapped in a local posterior mode, and as a result, provide suboptimal parameter estimates for the hydrological model.

[45] The case study in our analysis focused on previous work by *Bates and Campbell* [2001] and *Marshall et al.* [2004]. We also examined the effectiveness of each sampler for other data sets across a range of catchment conditions, such as catchment size and data set length. The results shown in Table 4, are based on an identical sampler setup to that used previously. Both samplers produce near-identical results in terms of NSE performance, for all catchments (with trivial improvement for SMC for Cotter and Mosquito Creek catchments). These results lend support for the ability of both samplers to achieve similar outcomes, especially when the optimal solutions are not in the tails of the specified prior distribution. The strength of the SMC method lies in the dispersion of N particles (compared to a single point in an AM sampler), combined with a gradual smooth transition through the distributional sequence to the target posterior.

[46] It is also important to note that the results for both use equivalent computation times as measured by the number of times the rainfall-runoff model has to be simulated using proposed parameter values. However, in practice, for a given data set, it is difficult to determine in advance whether the AM sampler output is likely to be reliable. While the AM and SMC algorithms in general require similar computational overheads and user tuning requirements, the need to implement the AM sampler multiple times, from dispersed starting points in order to ensure correct sampler convergence [*Gelman and Rubin*, 1992] means that the SMC sampler provides a more efficient way of exploring the posterior distribution.

[47] In closing, we emphasize the need for more research to allow use of the SMC sampler in a broader context. Specifically, we see the need for applications using alternate model configurations and complexities, the need for assessing the generality of the distributional sequence proposed in section 4.2 for higher-dimensional settings, and the need for efficient coding of the SMC algorithm to take advantage of the parallel processing opportunities that can be exploited and will be especially useful with assessments of more complex and heavily parameterized rainfall-runoff models.

[48] **Acknowledgments.** Funding for this research came from an Australian Research Council Discovery project. The streamflow data sets used in generating our results were provided to us by Francis Chiew.

References

Arulampalam, M. S., S. Maskell, N. Gordon, and T. Clapp (2002), A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Trans. Signal Process.*, *50*(2), 174–188, doi:10.1109/78.978374.

Bates, B. C., and E. P. Campbell (2001), A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling, *Water Resour. Res.*, *37*(4), 937–947, doi:10.1029/2000WR900363.

Beven, K., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*(3), 279–298, doi:10.1002/hyp.3360060305.

Boughton, W. C. (1993), A hydrograph-based model for estimating the water yield of ungauged catchments, Inst. of Eng. Aust., Canberra.

Boyle, D. P., H. V. Gupta, and S. Sorooshian (2000), Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resour. Res.*, *36*(12), 3663–3674, doi:10.1029/2000WR900207.

Bulygina, N., and H. Gupta (2009), Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation, *Water Resour. Res.*, *45*, W00B13, doi:10.1029/2007WR006749.

Celeux, G., M. Hurn, and C. P. Robert (2000), Computational and inferential difficulties with mixture posterior distributions, *J. Am. Stat. Assoc.*, *95*(451), 957–970, doi:10.2307/2669477.

Chopin, N. (2002), A sequential particle filter method for static models, *Biometrika*, *89*(3), 539–552, doi:10.1093/biomet/89.3.539.

Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, *44*, W00B02, doi:10.1029/2007WR006735.

Craiu, R. V., J. Rosenthal, and C. Yang (2009), Learn from thy neighbor: Parallel-chain and regional adaptive MCMC, *J. Am. Stat. Assoc.*, *104*(488), 1454–1466, doi:10.1198/jasa.2009.tm08393.

Del Moral, P., A. Doucet, and A. Jasra (2006), Sequential Monte Carlo samplers, *J. R. Stat. Soc., Ser. B*, *68*, 411–436, doi:10.1111/j.1467-9868.2006.00553.x.

Del Moral, P., A. Doucet, and A. Jasra (2007), Sequential Monte Carlo for Bayesian computation, in *Bayesian Statistics*, vol. 8, edited by J. M. Bernardo et al., pp. 1–34, Oxford Univ. Press, Oxford, U. K.

Doucet, A., N. De Freitas, and N. Gordon (2001), *Sequential Monte Carlo Methods in Practice*, 581 pp., Springer, New York.

Drovandi, C. C., and A. N. Pettitt (2011), Estimation of parameters for microparisite population evolution using approximate Bayesian computation, *Biometrics*, *67*(1), 225–233.

Duan, Q. Y., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*(4), 1015–1031, doi:10.1029/91WR02985.

Fan, Y., D. S. Leslie, and M. P. Wand (2008), Generalised linear mixed model analysis via sequential Monte Carlo sampling, *Electron. J. Stat.*, *2*, 916–938.

Fearnhead, P. (2008), Computational methods for complex stochastic systems: A review of some alternatives to MCMC, *Stat. Comput.*, *18*(2), 151–171, doi:10.1007/s11222-007-9045-8.

Gelman, A., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Stat. Sci.*, *7*, 457–472, doi:10.1214/ss/1177011136.

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

Geyer, C. J., and E. A. Thompson (1995), Annealing Markov chain Monte Carlo with applications to ancestral inference, *J. Am. Stat. Assoc.*, *90*(431), 909–920, doi:10.2307/2291325.

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

Gordon, N. J., D. J. Salmond, and A. F. M. Smith (1993), Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proc. F Radar Signal Process.*, *140*(2), 107–113, doi:10.1049/ip-f-2.1993.0015.

Haario, H., E. Saksman, and J. Tamminen (1999), Adaptive proposal distribution for random walk Metropolis algorithm, *Comput. Stat.*, *14*(3), 375–395, doi:10.1007/s001800050022.

Haario, H., E. Saksman, and J. Tamminen (2001), An adaptive Metropolis algorithm, *Bernoulli*, *7*(2), 223–242, doi:10.2307/3318737.

Haario, H., M. Laine, A. Mira, and E. Saksman (2006), DRAM: Efficient adaptive MCMC, *Stat. Comput.*, *16*(4), 339–354, doi:10.1007/s11222-006-9438-0.

Hall, M. J. (2001), How well does your model fit the data?, *J. Hydroinf.*, *3*(1), 49–55.

Jasra, A., D. Stephens, and C. Holmes (2007), On population-based simulation for static inference, *Stat. Comput.*, *17*(3), 263–279, doi:10.1007/s11222-007-9028-9.

Johansson, A. M., and E. A. Lehmann (2009), Evolutionary optimization of dynamics models in sequential Monte Carlo target tracking, *IEEE Trans. Evol. Comput.*, *13*(4), 879–894, doi:10.1109/TEVC.2009.2017518.

Kitagawa, G. (1996), Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, *J. Comput. Graphical Stat.*, *5*(1), 1–25, doi:10.2307/1390750.

Kou, S. C., Q. Zhou, and W. H. Wong (2006), Equi-energy sampler with applications in statistical inference and statistical mechanics, *Ann. Stat.*, *34*(4), 1581–1619, doi:10.1214/009053606000000515.

Kuczera, G., and M. Mroczkowski (1998), Assessment of hydrologic parameter uncertainty and the worth of multiresponse data, *Water Resour. Res.*, *34*(6), 1481–1489, doi:10.1029/98WR00496.

Kuczera, G., and E. Parent (1998), Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm, *J. Hydrol.*, *211*(1–4), 69–85, doi:10.1016/S0022-1694(98)00198-X.

Liu, J. S. (2001), *Monte Carlo Strategies in Scientific Computing*, Springer, New York.

Liu, J. S., and R. Chen (1998), Sequential Monte Carlo methods for dynamic systems, *J. Am. Stat. Assoc.*, *93*(443), 1032–1044, doi:10.2307/2669847.

Marshall, L., D. Nott, and A. Sharma (2004), A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling, *Water Resour. Res.*, *40*, W02501, doi:10.1029/2003WR002378.

McMillan, H., and M. Clark (2009), Rainfall-runoff model calibration using informal likelihood measures within a Markov chain Monte Carlo sampling scheme, *Water Resour. Res.*, *45*, W04418, doi:10.1029/2008WR007288.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, *21*(6), 1087–1092, doi:10.1063/1.1699114.

Micevski, T., and G. Kuczera (2009), Combining site and regional flood information using a Bayesian Monte Carlo approach, *Water Resour. Res.*, *45*, W04405, doi:10.1029/2008WR007173.

Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. R. Houser (2005a), Dual state-parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.*, *28*(2), 135–147, doi:10.1016/j.advwatres.2004.09.002.

Moradkhani, H., K. L. Hsu, H. Gupta, and S. Sorooshian (2005b), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, *41*, W05012, doi:10.1029/2004WR003604.

Papaspilopoulos, O., G. O. Roberts, and M. Sköld (2003), *Non-centred Parameterisations for Hierarchical Models and Data Augmentation (With Discussion)*, Clarendon, London.

Redelings, B. D., and M. A. Suchard (2005), Joint Bayesian estimation of alignment and phylogeny, *Syst. Biol.*, *54*(3), 401–418, doi:10.1080/10635150590947041.

Sisson, S. A., and Y. Fan (2009), Towards automating model selection for a mark-recapture-recovery analysis, *J. R. Stat. Soc., Ser. C Appl. Stat.*, *58*(2), 247–266, doi:10.1111/j.1467-9876.2008.00656.x.

Sisson, S. A., and Y. Fan (2010), *Likelihood-Free Markov Chain Monte Carlo*, Chapman and Hall, Boca Raton, Fla.

Sisson, S., L. Pericchi, and S. Coles (2006), A case for a reassessment of the risks of extreme hydrological hazards in the Caribbean, *Stochastic Environ. Res. Risk Assess.*, *20*(4), 296–306, doi:10.1007/s00477-005-0246-4.

Sisson, S. A., Y. Fan, and M. M. Tanaka (2007), Sequential Monte Carlo without likelihoods, *Proc. Natl. Acad. Sci. U. S. A.*, *104*(6), 1760–1765, doi:10.1073/pnas.0607208104. [Correction, *Proc. Natl. Acad. Sci. U. S. A.*, *106*(39), 16,889–16,889, doi:10.1073/pnas.0908847106, 2009.]

Smith, T. J., and L. A. Marshall (2008), Bayesian methods in hydrologic modeling: A study of recent advancements in Markov chain Monte Carlo techniques, *Water Resour. Res.*, *44*, W00B05, doi:10.1029/2007WR006705.

Smith, P. J., K. J. Beven, and J. A. Tawn (2008), Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach, *Water Resour. Res.*, *44*, W01410, doi:10.1029/2006WR005205.

Sorooshian, S., and V. K. Gupta (1983), Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness, *Water Resour. Res.*, *19*(1), 260–268, doi:10.1029/WR019i001p00260.

Tarantola, A. (1987), *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier, Amsterdam.

Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.

Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, *41*, W01017, doi:10.1029/2004WR003059.

Vrugt, J. A., C. J. F. ter Braak, C. G. H. Diks, B. A. Robinson, J. M. Hyman, and D. Higdon (2009), Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. Nonlinear Sci. Numer. Simul.*, *10*(3), 273–290, doi:10.1515/IJNSNS.2009.10.3.273.

Weerts, A. H., and G. Y. H. El Serafy (2006), Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, *Water Resour. Res.*, *42*, W09403, doi:10.1029/2005WR004093.

Yang, X. J., K. L. Shi, T. Huang, and K. Y. Xing (2007), *Combined Parameter and State Estimation in Particle Filtering*, pp. 1614–1617, Inst. of Electr. and Electron. Eng, New York.

Yoon, J. W., S. J. Roberts, M. Dyson, and J. Q. Gan (2009), Adaptive classification for brain computer interface systems using sequential Monte Carlo sampling, *Neural Networks*, *22*(9), 1286–1294, doi:10.1016/j.neunet.2009.06.005.

E. Jeremiah, R. Mehrotra, and A. Sharma, School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW 2052, Australia.

L. Marshall, Land Resources and Environmental Science, Montana State University, Bozeman, MT 59717, USA.

S. Sisson, School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia.