

Improving the efficiency of Monte Carlo Bayesian calibration of hydrologic models via model pre-emption

Mahyar Shafii, Bryan Tolson and L. Shawn Matott

ABSTRACT

Bayesian inference via Markov Chain Monte Carlo (MCMC) sampling and sequential Monte Carlo (SMC) sampling are popular methods for uncertainty analysis in hydrological modelling. However, application of these methodologies can incur significant computational costs. This study investigated using model pre-emption for improving the computational efficiency of MCMC and SMC samplers in the context of hydrological modelling. The proposed pre-emption strategy facilitates early termination of low-likelihood simulations and results in reduction of unnecessary simulation time steps. The proposed approach is incorporated into two samplers and applied to the calibration of three rainfall–runoff models. Results show that overall pre-emption savings range from 5 to 21%. Furthermore, results indicate that pre-emption savings are greatest during the pre-convergence ‘burn-in’ period (i.e., between 8 and 39%) and decrease as the algorithms converge towards high likelihood regions of parameter space. The observed savings are achieved with absolutely no change in the posterior set of parameters.

Key words | AR, calibration, DREAM, pre-emption, SMC, uncertainty analysis

Mahyar Shafii (corresponding author)
Department of Earth and Environmental Sciences,
University of Waterloo,
200 University Avenue West,
Waterloo, Ontario N2L 3G1,
Canada
E-mail: mshafii@uwaterloo.ca

Bryan Tolson
Department of Civil and Environmental
Engineering,
University of Waterloo,
200 University Avenue West,
Waterloo, Ontario N2L 3G1,
Canada

L. Shawn Matott
Center for Computational Research,
University at Buffalo,
701 Ellicott Street, Buffalo,
New York 14203,
USA

INTRODUCTION

This paper focuses on improving the computational efficiency of calibration and uncertainty analysis – two essential components of model assessment, defined as the use of robust procedures to determine the suitability of a given model for a given purpose (Matott *et al.* 2009). Investigations of uncertainty in hydrological modelling have emphasized the use of automatic calibration methods, which develop expressions for parameter uncertainty, ranging from simple Monte Carlo simulations such as GLUE (Beven & Binley 1992) to statistical approaches based on Bayesian inference (Box & Tiao 1973; Kuczera 1983). Owing to the complexity of large-scale hydrological models, Bayesian inference is facilitated through Markov Chain Monte Carlo (MCMC) sampling from parameters’ posterior distributions (e.g., Kuczera & Parent 1998; Haario *et al.* 2001; Vrugt *et al.* 2003, 2009; Kavetski *et al.* 2006).

Sequential Monte Carlo (SMC) simulations have also become very attractive in hydrological modelling in recent

years (Moradkhani *et al.* 2005; Hsu *et al.* 2009; Salamon & Feyen 2010; Jeremiah *et al.* 2011). SMC samplers combine data assimilation principles with a particle filtering strategy (e.g., Moradkhani *et al.* 2005; Smith *et al.* 2008), and generally resemble previous developed sampling importance resampling approaches (e.g., Del Moral *et al.* 2006). More recently, Jeremiah *et al.* (2011) compared an example SMC sampler with an adaptive MCMC sampler and found that both methods displayed robustness and convergence.

Despite the common use of MCMC and SMC approaches, their application can incur high computational costs. Therefore, strategies for improving the efficiency of such samplers are an ongoing area of research. In MCMC sampling, efforts to improve efficiency include utilizing prior information (Mertens *et al.* 2004; Sikorska *et al.* 2012), developing adaptive algorithms (e.g., Haario *et al.* 2001; Vrugt *et al.* 2003, 2009; Craiu *et al.* 2009), and using ‘limited-memory’ sampling (Kuczera *et al.* 2010). Efforts for overcoming practical SMC issues include using importance

sampling (Cheng & Druzdzel 2000) and seeding initial solutions using empirical Bayes (Chen *et al.* 2004).

As a complementary approach to the aforementioned efforts, this study explores the use of the deterministic model ‘pre-emption’ concept, first introduced in Razavi *et al.* (2010), to improve the computational efficiency of MCMC and SMC samplers in the context of hydrological modelling. Model pre-emption is a relatively simple strategy for identifying low-quality simulations and terminating them early before the entire simulation run time completes. Model pre-emption requires the continuous computation of a goodness-of-fit criterion as the simulation proceeds and deterministic model pre-emption is the application of this concept within a sampling algorithm for which a threshold criterion value can be defined such that candidate solutions with goodness-of-fit measures beyond this value are known to have no impact on the search algorithm results. Previous research by Razavi *et al.* (2010) establishes that model pre-emption can yield substantial computational savings when applied to various optimization-based calibration strategies (DDS and PSO) and various ‘informal’ uncertainty-based calibration strategies, e.g., GLUE (Beven & Binley 1992) and DDS-AU (Tolson & Shoemaker 2007). In contrast, this study investigated model pre-emption for use within formal likelihood functions embedded within the MCMC and SMC sampling algorithms. These pre-emption-enabled formal samplers were then applied to the calibration and uncertainty analysis of three rainfall–runoff models. To the best of our knowledge, such an implementation has not been considered in previous studies on the use of MCMC and SMC sampling in hydrological modelling.

METHODS

Model pre-emption was applied to two algorithms, i.e., an MCMC implementation known as differential evolution adaptive metropolis (DREAM) (Vrugt *et al.* 2009) and an SMC implementation described by Jeremiah *et al.* (2011) and referred to herein as JSMC. DREAM runs multiple Markov chains simultaneously to facilitate efficient global exploration of the parameter space and its convergence is monitored using the Gelman–Rubin metric (\hat{R}) (Gelman &

Rubin 1992), where values less than 1.2 indicate convergence. The JSMC sampling process propagates a population of parameter vectors (or particles) of size N from an initial sampling distribution to the desired posterior distribution. For more information on the JSMC sampler refer to Jeremiah *et al.* (2011).

Both DREAM and JSMC are designed to take samples from the Bayesian posterior distributions of model parameters. Two Bayesian formulations were investigated in this study, as described below. Consider a time series of N stream flow observations, Y_t , $t = 1, \dots, N$ (or Y in vector notation) used to calibrate hydrologic model $h(\theta)$ given its parameter vector (θ). Assuming the model errors are uncorrelated and Gaussian distributed with zero mean and variance σ_ε^2 , the posterior probability density function $p(\theta|Y)$ has the following form (after integrating out σ_ε^2):

$$p(\theta|Y) \propto \left(\sum_{t=1}^N \varepsilon_t^2 \right)^{-(N/2)} \quad (1)$$

where $\varepsilon_t = Y_t - h(\theta)$ is a vector of residuals. Equation (1) assumes that errors are uncorrelated, but this is not a very realistic assumption in the context of hydrologic modelling. One approach to account for auto-correlation is to use a first-order auto-regressive (AR) scheme for the error series (Sorooshian & Dracup 1980)

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t \quad t = 1, \dots, N \quad (2)$$

where $\varepsilon_0 = 0$, ρ is the first-order correlation coefficient, and v_t is the remaining error prescribed to have a zero mean and constant variance σ_v^2 . The resulting joint posterior distribution of θ and ρ in this case would be

$$p(\theta, \rho|Y) \propto \left(\varepsilon_1^2(1-\rho)^2 + \sum_{t=2}^N \delta_t^2 \right)^{-(N/2)} \quad (3)$$

where $\delta_t(\theta, \rho|Y) = \varepsilon_t(\theta|Y) - \rho\varepsilon_{t-1}(\theta|Y)$ $t = 1, \dots, N$. It is observed in both Equations (1) and (3) that the posterior density will monotonically decrease when residuals are incorporated time step by time step into the equations. Since the posterior densities calculated with Equations (1) and (3) monotonically degrade as the simulation proceeds

through time, both formulations are suitable for adopting a model pre-emption approach (Razavi *et al.* 2010).

Pre-emption-enabled DREAM and JSMC sampling was applied to the calibration and uncertainty analysis of three different rainfall–runoff models. Table 1 provides summary information on these case studies and lists corresponding case study reference papers containing complete descriptions.

Model pre-emption

In deterministic model pre-emption (Razavi *et al.* 2010), model performance (in terms of some monotonically degrading calibration objective function) is monitored during simulation, and a given simulation is terminated early if it is recognized to be so poor that it will not contribute to guiding the search strategy. In the present study, the DREAM and JSMC sampling algorithms were modified to support deterministic model pre-emption. The first step in implementing pre-emption is to select an appropriate objective function. As noted previously, both Equations (1) and (3) are suitable choices for model pre-emption.

Another important factor in pre-emption implementation is the pre-emption threshold, i.e., a likelihood value where the solutions resulting in likelihood values worse than this threshold would be rejected even if the simulation is carried out completely. Both DREAM and JSMC decide to jump from a current state (θ_n) to a candidate state (θ^*) based on the ratio of the posterior densities of the two states, i.e., $p(\theta^*|Y)/p(\theta_n|Y)$. The candidate state is accepted if $p(\theta^*|Y)/p(\theta_n|Y) > Z$, where Z is a random number uniformly distributed between 0 and 1; otherwise the sampler

remains at the current state. A move to θ^* is accepted only if $p(\theta^*|Y) > Z \times p(\theta_n|Y)$. Thus, the posterior density value of $p(\theta^*|Y)_{\min} = Z \times p(\theta_n|Y)$ can be considered as the pre-emption threshold so long as the random number Z is generated ‘prior’ to evaluating a given candidate solution. Algorithms can then determine, a priori, the minimum acceptable value of the candidate posterior density ($p(\theta^*|Y)_{\min}$) as the pre-emption threshold.

Defining a pre-emption-enabled version of DREAM and JSMC requires slight adjustment of the acceptance/rejection step in the Metropolis–Hastings algorithm, as illustrated in Figure 1. When a given parameter set θ_n is evaluated (box 1 in Figure 1), Z is generated and the pre-emption threshold for candidate θ^* or $p(\theta^*|Y)_{\min}$ is identified (box 2). At any time step (t) of the model simulation, the likelihood can be calculated as $p_t(\theta^*|Y)$ and evaluated against $p(\theta^*|Y)_{\min}$ (boxes 3–6). If the evaluated density of any candidate solution becomes lower than $p(\theta^*|Y)_{\min}$ at any point through the simulation, it is pre-empted (box 7); otherwise, the evaluation of θ^* terminates without any time saving (box 8). Note that a pre-empted candidate would never be accepted by the Metropolis–Hastings algorithm, even if the simulation had not been pre-empted. As such, the pre-emption strategy employed here is deterministic in that it has absolutely no influence, other than computational savings, on the behaviour of the algorithm.

For this study, pre-emption was implemented using a simple file-based hand-shaking mechanism between the MCMC or JSMC algorithm and the simulation model. Prior to invoking a simulation run, the algorithm would write the calculated pre-emption threshold ($p(\theta^*|Y)_{\min}$) for that run to a file and the simulation model would read the

Table 1 | Three rainfall–runoff models and the catchments studied in this paper

Simulation model (type)	No. of para.	Catchment (area km ²)	Forcing data	Calibration period	Refs for more information
HYMOD (lumped)	5	Leaf (1994)	Precipitation temperature	1953–1954	Boyle (2000); Vrugt <i>et al.</i> (2003)
WetSpa (semi-distributed)	6	Baron (965)	PET ^a	1995–2000	Safari <i>et al.</i> (2009)
SWAT (semi-distributed)	26	Cannonsville (37)	Precipitation temperature PET ^a ; Sol. radiation; Rel. humidity	1996–1998	Tolson (2005); Tolson & Shoemaker (2007)

^aPotential evapotranspiration.

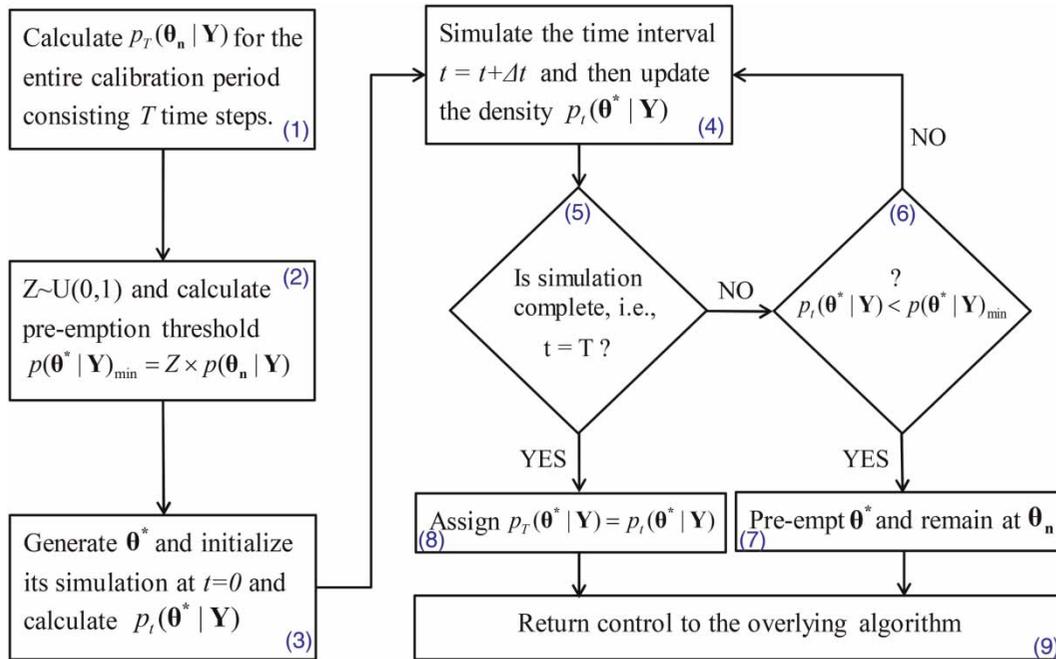


Figure 1 | Modified acceptance/rejection step in Metropolis–Hastings component implemented in pre-emption.

value from the file as part of the model initialization process. During simulation, the objective function was compared against this threshold each time the objective function was updated (e.g., after each time step). Violation of the threshold would then trigger the model to gracefully terminate the simulation.

Assuming computational cost is the same for all model time steps (i.e., the simulation model takes an identical amount of time during different time steps), the associated computational savings for a given application of DREAM or JSMC can be estimated as follows, according to Razavi et al. (2010)

$$S = 100 \times \frac{n - n_p}{n} \quad (4)$$

where S is the computation savings (in per cent), n is the total number of time steps in the calibration period, and n_p is the number of time steps simulated before the simulation is terminated by pre-emption. Note that the pre-emption approach outlined in this section is applicable to any other MCMC or SMC samplers that utilize the Metropolis–Hastings acceptance/rejection approach for evaluating candidate moves.

RESULTS

Non-pre-emptive experiments

A ‘standard’ (i.e., non-pre-emptive) DREAM implementation was applied to three case studies, thereby establishing baseline computational costs for the algorithm. Preliminary investigation of model residuals indicated the standard Bayesian formulation of Equation (1) was sufficient for calibrating the HYMOD case study. Conversely, the AR-based formulation of Equation (3) was required for the WetSpa and SWAT case studies to accommodate correlation among the residuals. The Gelman–Rubin convergence metric (\hat{R} -statistic) indicated that DREAM converged after 7,800, 4,000 and 161,000 simulations in HYMOD, WetSpa and SWAT case studies, respectively. After convergence, 10,000 more samples were taken to form the posterior distribution.

A baseline set of non-pre-emptive JSMC sampling experiments were also applied to the three case studies. Although JSMC convergence and model residuals are treated differently, the same likelihood formulations as the DREAM experiments were used, and the same computational

budgets were considered. Note that the relative efficiencies of DREAM and JSMC were compared previously by [Jeremiah *et al.* \(2011\)](#) and such inter-algorithm comparisons were not pursued in the present study. Instead, the numerical experiments focussed on the application of pre-emption to reduce the computational burden of these methods.

Application of model pre-emption

Pre-emption-enabled versions of DREAM and JSMC were applied to the same calibration problems as mentioned in the section Non-pre-emptive experiments. The pre-emptive DREAM and JSMC experiments were performed using the same sequence of random numbers (generated by a random number generator) applied in previous experiments. Moreover, the same computational budgets were considered for corresponding pre-emptive and baseline experiments. These identical settings ensured that a given sampler's search behaviour was the same with and without pre-emption. As expected, the pre-emption-based DREAM and JSMC samplers yielded the same sets of posterior parameter values as those obtained in the corresponding baseline (i.e., non-pre-emptive) experiments. These identical parameter values are not shown here. In other words, use of model pre-emption did not change the calibration results, and the only effect of using pre-emption was a reduction in the required amount of computation.

[Table 2](#) provides average computational savings (in per cent) for the pre-emption-based DREAM and JSMC experiments. The total average savings ranged from 5 to 21% in DREAM, and from 16 to 18% in JSMC. Extrapolating based on average simulation model runtimes and the

percentage savings yields estimated wall-clock savings of up to 38 hours for our case studies. For more computationally demanding hydrologic models, such as fully distributed models requiring hours of simulation time, the wall-time savings afforded by pre-emption would be even more significant.

For the selected algorithms (i.e., DREAM and JSMC), most of the pre-emption savings occurred during the initial sampling or 'burn-in' period, defined as the period before the Gelman–Rubin metric indicates convergence. As the samplers converge, candidate parameter sets (θ^*) increasingly differ from the current parameter sets (θ_n). This in turn increases the likelihood ratio acceptance criterion, $p(\theta^*|Y)/p(\theta_n|Y)$, and reduces the probability of pre-emption. To quantify this behaviour, the DREAM pre-emption savings were separated into burn-in and post-burn-in periods, and the JSMC results were likewise divided into two halves. The results are shown in [Table 2](#).

[Figure 2](#) illustrates the empirical cumulative distribution function of the simulation time at which model pre-emption terminated a given simulation. For DREAM in the post-burn-in period, almost all pre-emption occurred after 85% of the simulation was completed. This explains why the overall cumulative savings reported for DREAM in [Table 2](#) are relatively low. However, unacceptable simulations were terminated much earlier during the burn-in period and there was considerable computational savings in this stage. Fairly similar pre-emption behaviour was observed for the JSMC sampler (lower panel in [Figure 2](#)).

Sensitivity of pre-emption savings to calibration period

The effectiveness of model pre-emption can be sensitive to the location of large storms within the calibration period ([Razavi *et al.* 2010](#)). For example, inferior parameter sets will generally trigger early exceedance of the pre-emption threshold if major events happen early in the calibration period. However, pre-emption will not help as much if a major storm occurs at the end of a calibration period. This is because the simulation will need to cover most of the calibration period before a pre-emption judgement can be made. As identified by [Razavi *et al.* \(2010\)](#), one approach to enhance efficiency gains due to pre-emption is to choose the calibration period such that the largest magnitude

Table 2 | Average computation saving (in per cent) obtained from pre-emption-enabled MCMC and JSMC in different case studies

Method	Time window	Case study		
		HYMOD	WetSpa	SWAT
MCMC	Burn-in period	17	39	8
	Post-burn-in period	12	13	4
	Entire simulations	14	21	5
JSMC	First half of the simulations	21	28	25
	Second half of the simulations	11	8	9
	Entire simulations	16	18	17

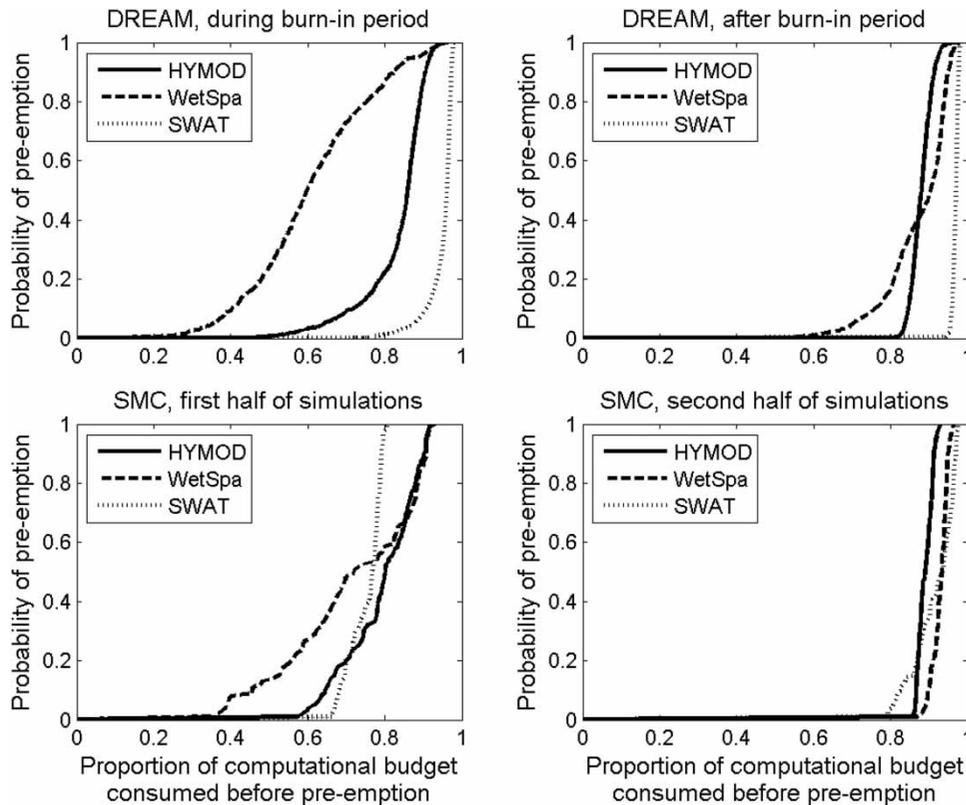


Figure 2 | Empirical cumulative distribution function of the simulation time at which model pre-emption is applied in different calibration problems.

events/responses occur near the beginning of the calibration period.

To explore the sensitivity of model pre-emption to the calibration period, the HYMOD case study was calibrated using pre-emption-enabled DREAM considering four different years from the observation period. Results showed the pre-emption savings varied according to the selected calibration period, and in some cases considerable savings were achieved. Overall pre-emption savings in these experiments ranged from 8 to 35% during the entire simulation and 10 to 39% during the burn-in period.

DISCUSSION AND CONCLUSIONS

In view of the computational burden associated with samplers employed for Bayesian inference (e.g., DREAM or JSMC), a model pre-emption approach was investigated for saving computational time. The proposed approach (i.e., avoiding unnecessary simulations) yielded on average between 5 and

21% computational savings in the three selected case studies. In one of the case studies, it was shown that savings could reach as high as 39% depending on the selected calibration period. The time savings were larger during the initial stage of sampling, and ranged from 8 to 39%. Such savings are considerable for simulation models that require several minutes or hours to complete. Moreover, the pre-emption savings varied according to the selected calibration period, and in some cases considerable savings were achieved. Implementing pre-emption did not change the calibration results compared to when calibrating without pre-emption. Moreover, implementation was straightforward and our approach is generally applicable to any samplers that utilize the Metropolis–Hastings acceptance/rejection approach for evaluating candidate moves in the search space.

The case study results presented here provide strong empirical evidence that a model pre-emption approach is a good choice for application to other case studies involving formal Bayesian calibration. Pre-emption will be most useful in calibration problems where it is very hard to find

good solutions and a great deal of time is wasted fully evaluating bad solutions long after it is known that they will have no influence whatsoever on the sampling algorithm. Moreover, our results suggest that pre-emption savings are most significant in cases where Bayesian samplers do not converge. In practice, convergence failure is relatively common during the initial phases of calibrating complex hydrological models where multiple applications of a Bayesian sampler can be required. For example, refinement of the model, model input forcings and/or likelihood function is often required before a satisfactory calibration result is obtained. The burn-in period of the selected case studies are representative of these no-converged situations and corresponding results suggest that savings of up to 39% can be achieved. In this way, pre-emption can accelerate model development by helping modellers more quickly determine when there is an issue preventing MCMC or SMC algorithm convergence.

ACKNOWLEDGEMENTS

This research was supported with funding from Bryan Tolson's NSERC Discovery Grant (50%). The authors thank Dr De Smedt (Vrije Universiteit Brussel – VUB) and Dr Ali Safari (VUB) for providing the WetSpa case study (and associated input and data files), and Dr Vrugt (University of California at Irvine) for their DREAM source code.

REFERENCES

- Beven, K. & Binley, A. 1992 *The future of distributed models: model calibration and uncertainty prediction*. *Hydrol. Process.* **6** (3), 279–298.
- Box, G. E. P. & Tiao, G. C. 1973 *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Boston, MA.
- Boyle, D. P. 2000 *Multicriteria calibration of hydrological models*. Ph.D. Dissertation thesis, University of Arizona Tucson, AZ.
- Chen, W., Bakshi, B. R., Goel, P. K. & Ungarala, S. 2004 *Bayesian estimation via sequential Monte Carlo sampling: unconstrained nonlinear dynamic systems*. *Ind. Eng. Chem. Res.* **43**, 4012–4025.
- Cheng, J. & Druzdzel, M. J. 2000 AIS-BN: an adaptive importance sampling algorithm for evidential reasoning in large Bayesian networks. *J. Artif. Intell. Res.* **13**, 155–188.
- Craiu, R., Rosenthal, J. & Yang, C. 2009 *Learn from thy neighbor: parallel-chain and regional adaptive MCMC*. *J. Am. Stat. Assoc.* **104** (488), 1454–1466.
- Del Moral, P., Doucet, A. & Jasra, A. 2006 *Sequential Monte Carlo samplers*. *J. R. Stat. Soc., Ser. C Appl. Stat. Ser. B* **68**, 411–436.
- Gelman, A. & Rubin, D. B. 1992 *Inference from iterative simulation using multiple sequences*. *Stat. Sci.* **7** (4), 457–472.
- Haario, H., Saksman, E. & Tamminen, J. 2001 *An adaptive Metropolis algorithm*. *Bernoulli* **7** (2), 223–242.
- Hsu, K., Moradkhani, H. & Sorooshian, S. 2009 *A sequential Bayesian approach for hydrologic model selection and prediction*. *Water Resour. Res.* **45**, W00B12.
- Jeremiah, E., Sisson, S., Marshall, L., Mehrotra, R. & Sharma, A. 2011 *Bayesian calibration and uncertainty analysis of hydrological models: a comparison of adaptive Metropolis and sequential Monte Carlo samplers*. *Water Resour. Res.* **47**, W07547.
- Kavetski, D., Kuczera, G. & Franks, S. W. 2006 *Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory*. *Water Resour. Res.* **42**, W03407.
- Kuczera, G. 1983 *Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty*. *Water Resour. Res.* **19** (5), 1151–1162.
- Kuczera, G. & Parent, E. 1998 *Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm*. *J. Hydrol.* **211** (1–4), 69–85.
- Kuczera, G., Kavetski, D., Renard, B. & Thyer, M. 2010 *A limited-memory acceleration strategy for MCMC sampling in hierarchical Bayesian calibration of hydrological models*. *Water Resour. Res.* **46** (7), W07602.
- Matott, L. S., Babendreier, J. E. & Purucker, S. T. 2009 *Evaluating uncertainty in integrated environmental models: a review of concepts and tools*. *Water Resour. Res.* **45**, W06421.
- Mertens, J., Madsen, H., Feyen, L., Jacques, D. & Feyen, J. 2004 *Including prior information in the estimation of effective soil parameters in unsaturated zone modelling*. *J. Hydrol.* **294** (4), 251–269.
- Moradkhani, H., Hsu, K.-L., Gupta, H. & Sorooshian, S. 2005 *Uncertainty assessment of hydrologic model states and parameters: sequential data assimilation using the particle filter*. *Water Resour. Res.* **41**, W05012.
- Razavi, S., Tolson, B., Matott, L. S., Thomson, N., MacLean, A. & Seglinieks, F. 2010 *Reducing the computational cost of automatic calibration through model pre-emption*. *Water Resour. Res.* **46** (11), W11523.
- Safari, A., De Smedt, F. & Moreda, F. 2009 *Wetspa model application in the distributed model intercomparison project (DMIP2)*. *J. Hydrol.* **418–419**, 78–89.
- Salamon, P. & Feyen, L. 2010 *Disentangling uncertainties in distributed hydrological modeling using multiplicative error models and sequential data assimilation*. *Water Resour. Res.* **46**, W12501.
- Sikorska, A. E., Scheidegger, A., Banasik, K. & Rieckermann, J. 2012 *Bayesian uncertainty assessment of flood predictions in*

- ungauged urban basins for conceptual rainfall-runoff models. *Hydrol. Earth Syst. Sci.* **16**, 1221–1236.
- Smith, P. J., Beven, K. J. & Tawn, J. A. 2008 Detection of structural inadequacy in process-based hydrological models: a particle-filtering approach. *Water Resour. Res.* **44**, W01410.
- Sorooshian, S. & Dracup, J. A. 1980 Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: correlated and heteroscedastic error cases. *Water Resour. Res.* **16** (2), 430–442.
- Tolson, B. A. 2005 *Automatic Calibration, Management and Uncertainty Analysis: Phosphorus Transport in the Cannonsville Watershed*. Cornell University, Ithaca, NY.
- Tolson, B. A. & Shoemaker, C. A. 2007 Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resour. Res.* **43**, W01413.
- Vrugt, J. A., Gupta, H. V., Bouten, W. & Sorooshian, S. 2003 A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.* **39** (8), 1201.
- Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., Higdon, D., Robinson, B. A. & Hyman, J. A. 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.

First received 21 March 2014; accepted in revised form 5 February 2015. Available online 23 February 2015