Jun 16th, 10:40 AM - 12:00 PM

Kernel Density Independence Sampling based Monte Carlo Scheme (KISMCS) for inverse hydrological modeling

Mojtaba Shafiei
Ferdowsi University of Mashhad

Shervan Gharari
Delft University of Technology

Saket Pande
Delft University of Technology, s.pande@tudelft.nl

Sandjai Bhulai
VU University Amsterdam

International Environmental Modelling and Software Society (iEMSs)
7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA,
Daniel P. Ames, Nigel W.T. Quinn and Andrea E. Rizzoli (Eds.)
http://www.iemss.org/society/index.php/iemss-2014-proceedings

Kernel Density Independence Sampling based Monte Carlo Scheme (KISMCS) for inverse hydrological modeling

Mojtaba Shafiei[1,2], Shervan Gharari[2], Saket Pande[2], Sandjai Bhulai[3]

[1] *Department of Water Engineering, Faculty of Agriculture, Ferdowsi University of Mashhad, Iran.*
[2]*Department of Water Management, Delft University of Technology, Delft Netherlands.*
(*s.pande@tudelft.nl*) [3]*Department of Mathematics, VU University, Amsterdam, Netherlands.*

Abstract: Posterior sampling methods are increasingly being used to describe parameter and model predictive uncertainty in hydrologic modelling. This paper proposes an alternative to random walk chains (such as DREAM-zs). We propose a sampler based on independence chains with an embedded feature of standardized importance weights based on Kernel density estimates. A Markov Chain Monte Carlo sampling algorithm is proposed with Metropolis-Hastings (M-H) updates using an independence sampler. The independence sampler ensures that candidate observations are drawn independently of the current state of a chain, thereby ensuring efficient exploration of the target distribution. The M-H acceptance-rejection criterion is used to sample across 3 chains, which ensures that the chains are well mixed. Kernel density estimation on last 600 samples in a chain is used to calculate standardized importance weights within the independence sampler to ensure fast convergence of sampled points to the target distribution. Its performance is contrasted with a state of the art algorithm, Differential Evolution Adaptive Metropolis (DREAM-zs), based on a toy 10 dimensional bi-modal Gaussian mixture distribution and HYMOD model based synthetic and real world case studies. The comparison of KISMCS and DREAM-zs is done based on their convergence to 'true' posterior parameter distributions in case of synthetics case studies and their convergence to a stationary distribution in case of real world hydrological modeling case studies.

Keywords: Markov Chain Monte Carlo; rainfall runoff modelling; posterior parameter distribution inference; inverse hydrological modelling.

1. INTRODUCTION

In the past decade, a variety of sampling methods have been proposed to infer posterior parameter distributions of hydrological models. These methods are extensively being used to describe parameter and model predictive uncertainty in hydrologic modeling. Monte Carlo methods have been successful in generating samples from posterior parameter distributions but such algorithms are known to be inefficient when encountered with complex and high-dimensional problems. The advent of Markov Chain Monte Carlo (MCMC) methods has helped to address some of these inefficiency and computational difficulties (Kuczera and Parent, 1998; Campbell et al., 1999; Bates and Campbell, 2001; Marshall et al., 2004).
The most general MCMC methods are random walk Metropolis (RWM) algorithm (Metropolis et al., 1953) and Metropolis-Hastings (MH) algorithm (Hastings, 1970) that have become the core of many existing MCMC sampling schemes. However, convergence of such algorithms usually suffer from initialization problems associated with the variance of the proposal being either too large or too small, leading to disturbingly slow convergence rates (Haario et al., 2006; Vrugt et al., 2009). Recently a variety of different approaches have been proposed to improve the efficiency of RWM and MH algorithms in MCMC simulations, including those by self-adaptive single-chain methods such as the adaptive Metropolis (AM) (Haario et al., 2001), delayed rejection adaptive Metropolis (DRAM) algorithms (Haario et al., 2006), a combination of Differential Evolution (DE) evolutionary algorithm (Storn and Price, 1997) with MCMC (DE-MC, see ter Braak, 2006) and a modified version of DE-MC that employs self-adaptive randomized subspace sampling for efficient exploration of a posterior distribution (Vrugt et al., 2008; Vrugt et al., 2009). Although these kinds of samplers are capable of

efficiently sampling from underlying but unknown distributions (Kuczera and Parent, 1998; Braak, 2006; Vrugt et al., 2009), they emphasize on random walk metropolis based algorithms. While efficient, these algorithms may suffer from poor convergence issues especially when the target distribution is non-Gaussian, often requiring long burn-in simulations (Braak, 2006; Vrugt et al., 2009), and remain sensitive to prior specification (Scharnagl et al., 2011).
The main goal of this paper is to suggest an alternate to random walk chains and compare and contrast its performance with state of the art DREAM-zs algorithm (ter Braak and Vrugt, 2008) on a toy problem, a synthetic conceptual rainfall runoff modeling case study and a real world hydrological modeling case study. We propose independence chains (Tierney, 1994) with an embedded feature of standardized importance weights based on Kernel density estimates (Pande, 2013a, 2013b). The sampler is called Kernel Density Independence Sampling based Monte Carlo Scheme (KISMCS).

2. KERNEL DENSITY INDEPENDENCE SAMPLING BASED MONTE CARLO SCHEME (KISMCS)

KISMCS is a Markov Chain Monte Carlo method with Metropolis-Hastings updates (Kuczera and Parent, 1998) using an independence sampler (Brooks, 1998). The independence sampler ensures that candidate observations are drawn independently of the current state of a chain, thereby ensuring efficient exploration of the target distribution (Pasarica and Gelman, 2010). The M-H acceptance-rejection criteria is also used to sample across C number of chains, which ensures that the chains are well mixed. Kernel density estimation (Hardle et al., 2004) on last m samples in a chain is used to calculate standardized importance weights (Givens and Raftery, 1996) within the independence sampler to ensure fast convergence of sampled points to the target distribution. An over-dispersed distribution, a multivariate t-distribution, is used as the kernel (Gelman and Rubin, 1992) to ensure exhaustive exploration of the parameter space. More details about KISMCS can be found in Shafiei et al. (2014). In this study, the algorithm is implemented with m = 600, a Gaussian reference bandwidth for multivariate t-distribution, sampled parameter covariance matrix updates every 600 function evaluations and the number of chains, C, used are 3. We use 600,000 for the number of model evaluations in KISMCS and DREAM-zs.
The technical parameters of DREAM-zs algorithm are now described. The dimension of the Z matrix that contains the current and past states of the chains is (M0, d), where d is the dimension (or the number of parameters) of the problem and M0 is the initial number of rows of Z which is assumed to be 10xd (ter Braak and Vrugt, 2008). We use 10% snooker updates and 90% parallel direction updates with adaptive tuning of crossover values (ter Braak and Vrugt, 2008; Vrugt et al., 2009).

3. CASE STUDIES

3.1 Toy Problem: A 10-dimensional bimodal Gaussian distribution

A 10-dimension bimodal Gaussian mixture distribution is considered as the target distribution for KISMCS (ter Braak, 2006). This target distribution is rather difficult to sample using standard MCMC schemes (Smith and Marshall, 2008). The distribution function $\pi(\mathbf{x})$, as given in the following is used as the likelihood function within KISMCS.

$$\pi(\mathbf{x}) = 1/3 N_{10}(\mathbf{-5},\mathbf{I_{10}}) + 2/3\, N_{10}(\mathbf{5},\mathbf{I_{10}}),$$

where **-5** and **5** are 10-dimensional vectors, $\mathbf{I_{10}}$ is a 10 x 10 dimensional identity matrix and $N_{10}(\mathbf{m},\Sigma)$ represents a 10 dimensional Gaussian distribution with mean **m** and covariance matrix Σ. The KISMCS chains are initially populated by $N_{10}(\mathbf{0},\mathbf{I_{10}})$.
The estimated marginal posterior probability distribution of x1 (first dimension of the 10 dimensional distribution) is depicted in Figure. 1 that shows a good match with the known target distribution (indicated with the red line). Also, its two modes are well-separated.

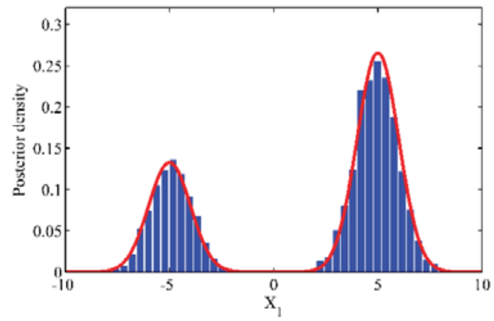

Figure. 1. Results of KISMCS for the 10-dimensional bimodal mixture model. Here the distribution for only the first dimension of the 10-dimensional distribution is presented.

3.2 HYMOD example1: A synthetic case study

We employ the daily rainfall-runoff HYdrologic MODel HYMOD (Wagener, 2001). HYMOD is chosen for its low number of parameters while still maintaining adequate process representation including slow and fast responses together with a non-linear soil moisture component. The model is characterized by 5 reservoirs, including the soil moisture reservoir (S_M), three linear reservoirs in series (S_{F1}, S_{F2}, S_{F3}) mimicking the fast runoff component, and one slow reservoir (S_{S1}). It has 5 parameters representing the maximum soil moisture storage capacity ($S_{M,max}$ [L]), the spatial variability of soil moisture (β[-]), the partitioning between fast reservoirs and slow reservoirs (α[-]), as well as the timescales of the fast and slow reservoirs (R_F, R_S [T^{-1}]). Model equations are solved using the forward explicit Euler method using 12-h resolution time series. The model inputs include mean areal precipitation and potential evaporation. The model has been applied widely in many previous studies to evaluate streamflow simulation and model calibration (Gharari et al., 2013). The upper and lower bounds of the model parameters, which define the *a priori* uncertainty ranges are given in Table 1.

Table 1. Prior ranges of HYMOD model parameters. The last column 'Syn' contains the parameter values used for the synthetic case study.

Parameters	Description	Min	Max.	Syn
$S_{M,max}$ (*mm*)	Maximum storage capacity of the watershed	1.00	500.00	200.00
β (-)	Spatial variability of the soil moisture storage	0.10	2.00	0.10
α (-)	Partitioning of flow between two reservoirs	0.10	0.99	0.90
R_F ($days^{-1}$)	Residence time parameter of quick-flow	0.10	0.99	0.40
R_S ($days^{-1}$)	Residence time parameter of slow-flow	0.00	0.10	0.08

For comparing KISMSC and DREAM-zs we use historical data from the Leaf River watershed (1944 km2) in the USA (Duan et al, 2006). 5 years of data from October 1, 1953 to September 30, 1958 is used to generate synthetic streamflow time series and infer posterior parameter distributions. The time series of daily streamflow data is generated with parameter values provided in Table 1 using observed daily precipitation and potential evaporation over the specified period. This synthetic discharge record is corrupted with Laplacian noise that is autocorrelated and heteroscedastic. The noise has homoscedastic variance of 0.01, heteroscedastic variance of 0.09 and first order autoregressive coefficient of 0.7 (Schoups and Vrugt, 2010).

Figure 2 and 3 present marginal posterior density of the HYMOD model parameters using KISMCS and DREAM-zs algorithms for modelling synthetic Leaf River streamflow. Also, the true values of synthetic case study (Table 1) are shown in red points. Results show that both the algorithms have similar good performance with almost sharp posteriors. However, the mode of the posterior sampled by DREAM-zs better matches the true parameter of R_s.

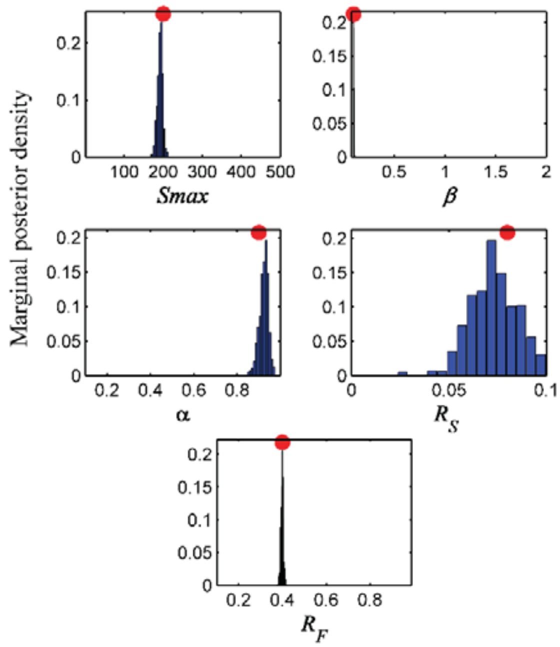

Figure. 2. Posterior distribution of the HYMOD model parameters for synthetic case study using KISMCS.

To further compare the two algorithms, the performance of the mode of parameter distributions that they sample after every 1800 model evaluations is compared in Figure 4. Figure. 4 in effect compares the rates of convergence (in mode of the sampled posterior) of the two algorithms. The results show that both algorithms have almost similar performance in the rate of convergence with DREAM-zs performing slightly better.

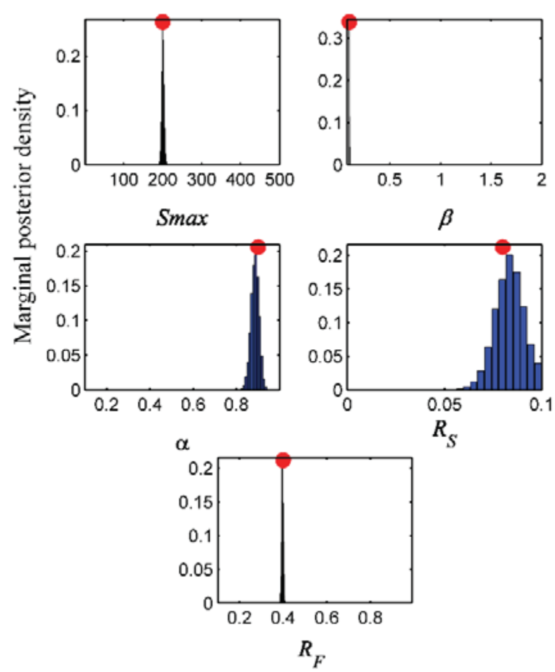

Figure. 3. Posterior distribution of the HYMOD model parameters for synthetic case study using DREAM-zs.

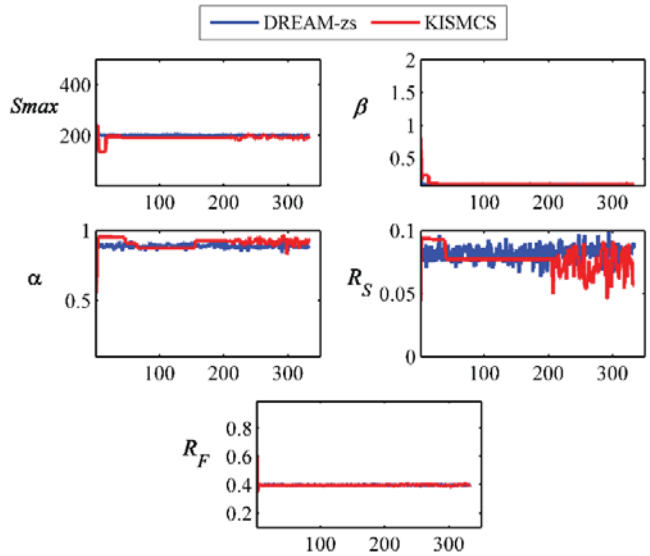

Figure. 4. Rates of convergence (in mode) for the HYMOD parameters for the synthetic case study using KISMCS and DREAM-zs. X-axis is the number of function evaluations/1800.

3.3 HYMOD example2: Leaf River Basin case study

This case study uses the 5 year Leaf River basin dataset and aims to infer posterior HYMOD parameter distributions based on observed discharge data set. Similar to the synthetic case study in section 3.2, results are presented for converged marginal posteriors distributions (Figure. 5 and 6) and rates of convergence (Figure. 7) for the two algorithms. The posterior densities inferred by the two methods are almost the same, however the *Smax* posterior sampled by DREAM-zs is narrower than the one sampled by KISMCS. The rate of convergence of DREAM-zs for parameter R_s is better than that of KISMCS.

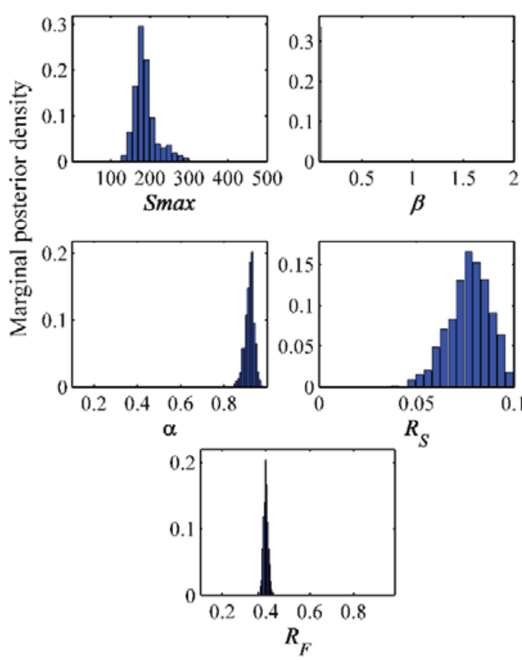

Figure. 5. Posterior distributions of the HYMOD model parameters for real case study sampled by KISMCS.

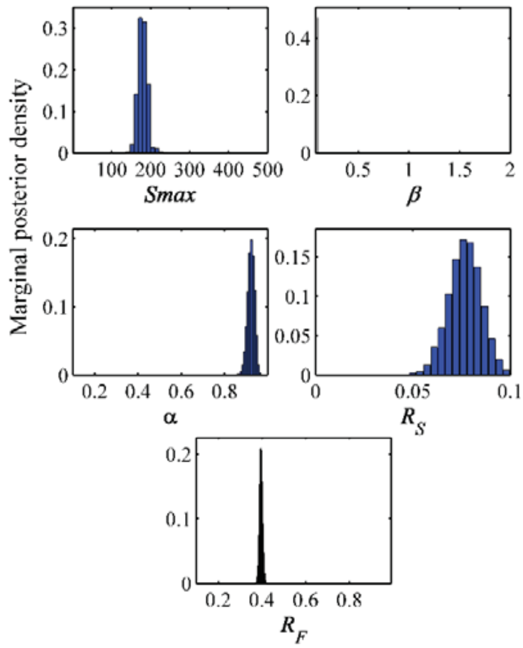

Figure. 6. Posterior distributions of the HYMOD model parameters for real case study sampled by DREAM-zs.

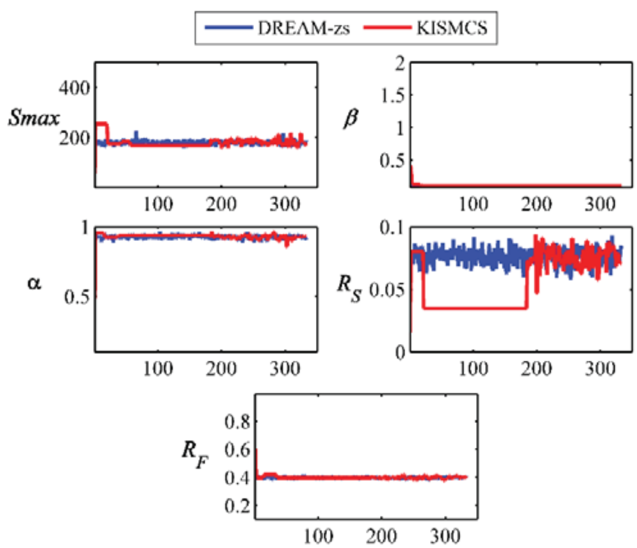

Figure. 7. Rates of convergence for the HYMOD parameters for real case study sampled by KISMCS and DREAM-zs. X-axis is the number of function evaluations/1800

4. CONCLUSIONS

The goal of this paper was to introduce the Kernel Density Independence Sampling based Monte Carlo Scheme (KISMCS) to infer the posterior sampling problems. Several examples including a toy case study and two hydrological case studies were used to demonstrate its robust performance. Its performance was compared with with the performance of state of the art DREAM-zs algorithm. The sampler used within KISMCS is based on independence sampling while DREAM-ZS is based on random walk sampling with differential evolution.

KISMCS advanced the knowledge of underlying search mechanisms of evolutionary algorithms by introducing two an independence sampler and the use of kernel density estimation of a proposal distribution. The motivation was to strike a good balance between two competing demands of a good sampler: to efficiently explore the parameter space and to increase the rate of convergence of a sampler. Results showed the good performance of KISMCS that was comparable to DREAM-zs. Further investigation of KISMCS is envisaged in future studies that aim to improve its computational efficiency, in particular with regards to the computational demand of kernel density estimation.

5. REFERENCES

Bates, B.C., Campbell, E.P., (2001). A Markov Chain Monte Carlo Scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. Water Resour. Res. 37, 937–947.

Brooks, S., (1998). Markov chain Monte Carlo method and its application. J. R. Stat. Soc. Ser. D (the Stat. 47, 69–100.

Campbell, E.P., Fox, D.R., Bates, B.C., (1999). A Bayesian Approach to parameter estimation and pooling in nonlinear flood event models. Water Resour. Res. 35, 211–220.

Duan, Q., Schaake, J., Andreassian, V., Franks, S., Goteti, G., Gupta, H. V, Gusev, Y.M., Habets, F., Hall, A., Hay, L., (2006). Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. J. Hydrol. 320, 3–17.

Gelman, A., Rubin, D.B., (1992). Inference from iterative simulation using multiple sequences. Stat. Sci. 457–472.

Gharari, S., Hrachowitz, M., Fenicia, F., Savenije, H.H.G., (2013). An approach to identify time consistent model parameters: sub-period calibration. Hydrol. Earth Syst. Sci. 17, 149–161.

Givens, G.H., Raftery, A.E., (1996). Local adaptive importance sampling for multivariate densities with strong nonlinear relationships. J. Am. Stat. Assoc. 91, 132–141.

Haario, H., Laine, M., Mira, A., Saksman, E., (2006). DRAM: efficient adaptive MCMC. Stat. Comput. 16, 339–354.

Haario, H., Saksman, E., Tamminen, J., (2001). An adaptive Metropolis algorithm. Bernoulli 7, 223–242.

Hardle, W., Muller, M., Sperlich, S., Werwatz, A., (2004). Nonparametric and semiparametric models. Berlin: Springer.

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

Kuczera, G., Parent, E., (1998). Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. J. Hydrol. 211, 69–85.

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

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., (1953). Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092.

Pande, S., (2013a). Quantile hydrologic model selection and model structure deficiency assessment: 1. Theory. Water Resour. Res. 49, 5631–5657.

Pande, S., (2013b). Quantile hydrologic model selection and model structure deficiency assessment: 2. Applications. Water Resour. Res. 49, 5658–5673.

Pasarica, C., Gelman, A., (2010). Adaptively scaling the Metropolis algorithm using expected squared jumped distance. Stat. Sin. 20, 343.

Scharnagl, B., Vrugt, J. a., Vereecken, H., Herbst, M., (2011). Inverse modelling of in situ soil water dynamics: investigating the effect of different prior distributions of the soil hydraulic parameters. Hydrol. Earth Syst. Sci. 15, 3043–3059.

Schoups, G., Vrugt, J. a., (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. Water Resour. Res. 46, W10531.

Shafiei, M., Gharari, S., Pande, S., Bhulai, S., (2014). Kernel Density Independence Sampling based Monte Carlo Scheme (KISMCS): A new algorithm to infer posterior densities of hydrologic model parameters. (In preparation)

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

Storn, R., Price, K., (1997). Differential evolution--a simple and efficient heuristic for global optimization over continuous spaces. J. Glob. Optim. 11, 341–359.
Ter Braak, C.J.F., (2006). A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. Stat. Comput. 16, 239–249.
Ter Braak, C.J.F., Vrugt, J. a., (2008). Differential Evolution Markov Chain with snooker updater and fewer chains. Stat. Comput. 18, 435–446.
Tierney, L., (1994). Markov chains for exploring posterior distributions. Ann. Stat. 1701–1728.
Vrugt, J.A., Braak, C.J.F., Diks, C.G.H., Robinson, B.A., Hyman, J.M., Higdon, D., Modeling, M., T-, A., (2009). Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling 10, 271–288.
Vrugt, J.A., ter Braak, C.J.F., Gupta, H. V., Robinson, B. a., Braak, C.J.F., (2008). Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stoch. Environ. Res. Risk Assess. 23, 1061–1062.