Geophys. J. Int. (2015) **202,** 961–975
GJI Marine geosciences and applied geophysics
doi: 10.1093/gji/ggv196

Systematic evaluation of sequential geostatistical resampling within MCMC for posterior sampling of near-surface geophysical inverse problems

Paolo Ruggeri, James Irving and Klaus Holliger

Applied and Environmental Geophysics Group, Institute of Earth Sciences, University of Lausanne, Lausanne, Switzerland. E-mail: james.irving@unil.ch

Accepted 2015 May 11. Received 2015 May 8; in original form 2014 June 18

SUMMARY

We critically examine the performance of sequential geostatistical resampling (SGR) as a model proposal mechanism for Bayesian Markov-chain-Monte-Carlo (MCMC) solutions to near-surface geophysical inverse problems. Focusing on a series of simple yet realistic synthetic crosshole georadar tomographic examples characterized by different numbers of data, levels of data error and degrees of model parameter spatial correlation, we investigate the efficiency of three different resampling strategies with regard to their ability to generate statistically independent realizations from the Bayesian posterior distribution. Quite importantly, our results show that, no matter what resampling strategy is employed, many of the examined test cases require an unreasonably high number of forward model runs to produce independent posterior samples, meaning that the SGR approach as currently implemented will not be computationally feasible for a wide range of problems. Although use of a novel gradual-deformation-based proposal method can help to alleviate these issues, it does not offer a full solution. Further, we find that the nature of the SGR is found to strongly influence MCMC performance; however no clear rule exists as to what set of inversion parameters and/or overall proposal acceptance rate will allow for the most efficient implementation. We conclude that although the SGR methodology is highly attractive as it allows for the consideration of complex geostatistical priors as well as conditioning to hard and soft data, further developments are necessary in the context of novel or hybrid MCMC approaches for it to be considered generally suitable for near-surface geophysical inversions.

Key words: Inverse theory; Tomography; Probability distributions; Ground penetrating radar.

1 INTRODUCTION

Applied near-surface geophysical methods have gained significant and increasing interest in the field of hydrology because of their ability to provide valuable information on the distribution of subsurface physical properties at spatial and temporal scales that are rarely attainable with traditional hydrological measurements. However, to make meaningful hydrological predictions and sound environmental decisions based on geophysical data, knowledge regarding the uncertainty in the estimated parameters is essential. While in many cases such uncertainty is ignored, this is a dangerous practice that can easily lead to misconceptions regarding the utility and reliability of geophysical methods.

A variety of possibilities exist for assessing the uncertainty of an estimated set of spatially distributed subsurface parameters given the corresponding set of geophysical measurements. Arguably the most common of these involves the use of a single solution to the inverse problem, generally obtained through gradient-based optimization approaches, around which a Taylor-series expansion is made in order to approximate the model parameter covariance and/or resolution matrices (e.g. Menke 1989; Alumbaugh & Newman 2000; Tarantola 2005). Although such linearized estimates provide highly useful local information and can be calculated in a computationally efficient manner, it is well understood that they may significantly underrepresent model parameter uncertainty in cases where there exists substantial nonlinearity in the geophysical forward problem (e.g. Sambridge & Mosegaard 2002; Tarantola 2005; Trainor-Guitton & Hoversten 2011). This in turn leads to overly narrow ranges of predictions based on the geophysical data, a situation which in all practical applications should be avoided (Kaipio & Somersalo 2005).

To address the above limitation, much recent interest has been expressed in the use of stochastic inverse methods for near-surface geophysical parameter estimation and uncertainty analysis. In particular, continuing advances in computational power have led to an increased use of Bayes' Theorem combined with Markov-chain-Monte-Carlo (MCMC) sampling to generate stochastic realizations from the posterior distribution of model parameters (e.g. Bosch

1999; Mosegaard & Tarantola 1995; Ramirez *et al.* 2005; Tarantola 2005). The resulting ensemble can then be used to calculate a variety of posterior model parameter statistics as well as to bound predictions. Bayesian-MCMC methods have the advantage of being able to provide more comprehensive estimates of model parameter uncertainty than linearized approaches. They are also extremely flexible in that they can incorporate any information that can be expressed in terms of probabilities into the inverse problem. These methods do, however, have the notable disadvantage of being limited by their high computational cost, which results from the typically large numbers of model parameters and data in geophysical problems combined with the need for small model perturbations along the Markov chain in order to ensure reasonable rates of proposal acceptance.

One way to reduce the computational cost of MCMC-based inversions is to decrease the number of estimated model parameters through an appropriate reparameterization of the subsurface domain of interest. In particular, model reduction through the use of a small number of well-chosen basis functions can exploit the spatial correlation naturally found in subsurface properties, thus allowing them to be effectively represented using only a few weighting coefficients as opposed to high-dimensional pixel-based parameterizations. To date, the most widely used model reduction approaches for geophysical and hydrological inverse problems involve either the discrete cosine transform (e.g. Jafarpour *et al.* 2009, 2010; Linde & Vrugt 2013), the wavelet transform (e.g. Davis & Li 2011; Jafarpour 2011) or the singular value decomposition (e.g. Oware *et al.* 2013). While use of such methods has proven quite effective, it is particularly difficult in a Bayesian context to know the bias on posterior uncertainty estimates imposed by seemingly benign prior assumptions on the reduced set of model parameters. For example, the assumption of simple uniform or uncorrelated Gaussian priors for the basis weighting coefficients may have strong implications for the prior distribution in the original pixel-based discretized space. In addition, conditioning inversions to local hard or soft measurements is not trivial with these approaches.

Another means of lowering the cost of MCMC-based geophysical inversions, and the focus of our work in this paper, retains the full pixel-based parameterization of the model space and instead concentrates on incorporating as much prior information as possible into the MCMC proposal distribution. In this way, the size (but not the dimension) of the parameter space to be explored is reduced because the number of subsurface configurations that may be tested is limited to a small subset of the total number of possibilities. Geostatistical simulation methods are highly attractive in this regard because of the inherent flexibility and ease with which they can represent complex geological information as well as the fact that they can be conditioned to a wide variety of measured and previously simulated data. Indeed, subsets of the model parameters can be geostatistically resampled at each MCMC iteration using sequential simulation, thereby allowing the proposition of new models along the Markov chain that are consistent with a set of prior statistical constraints yet represent small perturbations with respect to previously accepted parameter configurations. Recent results involving the application of the latter methodology, which we refer to in this paper as sequential geostatistical resampling (SGR), have appeared promising for both geophysical and hydrological inverse problems (Fu & Gómez-Hernández 2009; Irving & Singha 2010; Mariethoz *et al.* 2010; Cordua *et al.* 2012; Hansen *et al.* 2012). However, what is critically missing from the existing literature is a systematic examination of the methodology under a variety of realistic data and model scenarios in order to determine (i) what algorithmic choices lead to the most efficient MCMC implementations and (ii) more importantly, whether the approach as currently formulated can be considered to be computationally feasible and thus generally suitable for a wide range practical applications.

Here, through a set of relatively simple synthetic crosshole georadar tomographic examples characterized by different prescribed numbers of data, levels of data error and degrees of model parameter spatial correlation, we explore the performance of SGR within MCMC with regard to its ability to generate independent realizations from the Bayesian posterior distribution. The latter serves as our metric for algorithmic efficiency, as clearly the most efficient MCMC implementations will be capable of generating independent posterior realizations in the fewest number of iterations. We begin with a review of the Bayesian approach to geophysical inversion, MCMC methods for posterior sampling and a description of the overall SGR approach. We then examine the performance of three different SGR implementations (random cells resampling, block resampling and a new gradual-deformation-based resampling method) for the set of tomographic example cases. Based on our findings, we draw some general conclusions regarding the overall computational feasibility of the approach, and we outline some critical needs for future work.

2 METHODOLOGICAL BACKGROUND

2.1 Bayesian geophysical inversion

Consider to begin the general geophysical forward problem linking a set of spatially distributed subsurface model parameters $\mathbf{m}$ to a set of measured data $\mathbf{d}$:

$$\mathbf{d} = g(\mathbf{m}), \tag{1}$$

where $g(\cdot)$ is the forward operator containing the underlying physics and geometry of each measurement. The corresponding inverse problem, involving the estimation of $\mathbf{m}$ given $\mathbf{d}$, generally requires knowledge of $g(\cdot)$ along with some degree of prior information about the model parameters. Amongst a wide variety of methods for solving geophysical inverse problems, the probabilistic approach originally proposed by Tarantola & Valette (1982) has gained widespread popularity because it provides a flexible and intuitive framework that naturally allows for uncertainty quantification as well as the integration of different sources of data. Within this framework, the solution to the problem can be formulated using Bayes' theorem, whereby an initial prior state of information for the model parameters $\rho(\mathbf{m})$ is updated into a more refined posterior state of knowledge $\sigma(\mathbf{m})$ based on the available data. That is,

$$\sigma(\mathbf{m}) = k\, L(\mathbf{m})\, \rho(\mathbf{m}), \tag{2}$$

where k is a normalization constant that ensures that the posterior pdf integrates to unity, and $L(\mathbf{m})$ is the likelihood function which measures the degree of fit between the observed geophysical measurements and the data predicted for a particular model configuration using $g(\mathbf{m})$. Assuming independent, identically normally distributed data errors as well as perfect knowledge of the forward operator, we obtain the following expression for $L(\mathbf{m})$ (e.g. Mosegaard & Tarantola 1995):

$$L(\mathbf{m}) = \frac{1}{(2\pi\sigma_d^2)^{N/2}} \exp\left[-\frac{[g(\mathbf{m})-\mathbf{d}]^T[g(\mathbf{m})-\mathbf{d}]}{2\sigma_d^2}\right], \tag{3}$$

where σ_d is the estimated error standard deviation and N is the number of data.

Eqs (2) and (3) together provide a means of calculating the posterior probability of a particular configuration of model parameters. In an approach conceptually similar to deterministic inversion, $\sigma(\mathbf{m})$ can be optimized to find the model having the maximum *a posteriori* (MAP) probability, which represents a single 'most probable' solution to the inverse problem (e.g. Tarantola 2005). However, properly quantifying the posterior resolution or uncertainty of the model parameters is a much greater challenge. In all but the simplest of cases, linearized uncertainty estimates made about a single solution to the inverse problem will tend to underrepresent our lack of knowledge (e.g. Sambridge & Mosegaard 2002; Tarantola 2005). On the other hand, performing the multidimensional integrations necessary to obtain the true statistical moments of $\sigma(\mathbf{m})$ is generally not feasible because of the high dimension of the model space and the often complex nonlinear nature of the forward operator. Nevertheless, numerical algorithms such as MCMC can be used to generate random samples from the posterior pdf in eq. (2), which in turn can be used to calculate uncertainties and bound predictions.

2.2 MCMC and the extended Metropolis algorithm

MCMC defines a class of statistical methods in which a Markov chain is constructed over the state space of model parameters in order to generate samples from a target distribution of interest. Transition probability rules are defined for the chain such that, with enough iterations, it is guaranteed to equilibrate to the target distribution no matter what the starting state. Although an infinite number of possibilities exist with regard to how such transition probabilities can be prescribed, the Metropolis–Hastings algorithm (Metropolis *et al.* 1953; Hastings 1970) has become one of the more popular MCMC approaches because of its generality and simplicity. With this method, model perturbations along the chain are generated through a proposal distribution $Q(\mathbf{m}'|\mathbf{m}_i)$, where $\mathbf{m}_i$ is the current state of the chain at iteration i and $\mathbf{m}'$ is the proposed transition. To assure convergence to the target distribution $\sigma(\mathbf{m})$, the proposals are stochastically accepted at each iteration with probability

$$P_{\text{acc}} = \min\left[1, \frac{\sigma(\mathbf{m}')\, Q(\mathbf{m}_i|\mathbf{m}')}{\sigma(\mathbf{m}_i)\, Q(\mathbf{m}'|\mathbf{m}_i)}\right]. \tag{4}$$

If the proposed transition is accepted, then $\mathbf{m}_{i+1} = \mathbf{m}'$. Conversely, if the proposal is rejected, then the chain remains in its current state and $\mathbf{m}_{i+1} = \mathbf{m}_i$. In the case where the proposal distribution is symmetric, that is where $Q(\mathbf{m}'|\mathbf{m}_i) = Q(\mathbf{m}_i|\mathbf{m}')$, the Metropolis–Hastings algorithm reduces to the original algorithm of Metropolis *et al.* (1953) having acceptance probability

$$P_{\text{acc}} = \min\left[1, \frac{\sigma(\mathbf{m}')}{\sigma(\mathbf{m}_i)}\right]. \tag{5}$$

Although both the Metropolis–Hastings and Metropolis samplers have been used in a variety of geophysical and hydrological inverse problems to date (e.g. Bosch *et al.* 2006; Buland & Kolbjørnsen 2012; Scholer *et al.* 2012), one drawback of these MCMC formulations for Bayesian problems is that they require a means of evaluating the prior probability of a particular model configuration $\rho(\mathbf{m})$ at each iteration in order to obtain $\sigma(\mathbf{m})$. This may be impractical when the prior distribution for the model parameters is complex and/or defined by a series of simulation steps rather than an analytical pdf. Further, the algorithms may prove computationally inefficient for complex priors as they typically rely upon relatively simple uniform proposal distributions (Hansen *et al.* 2012). To address these issues, Mosegaard & Tarantola (1995) noted that the proposal distribution in the MCMC procedure can be chosen so as to simulate directly from the prior distribution. That is, they proposed that Q can be defined such that a Markov chain with Q as its transition probability equilibrates to $\rho(\mathbf{m})$. In this case, the following detailed balance equation must hold

$$\frac{Q(\mathbf{m}_i|\mathbf{m}')}{Q(\mathbf{m}'|\mathbf{m}_i)} = \frac{\rho(\mathbf{m}_i)}{\rho(\mathbf{m}')}, \tag{6}$$

which in turn implies that the acceptance probability in eq. (4) involves a simple comparison of the likelihoods of the proposed and current states of the Markov chain as follows:

$$P_{\text{acc}} = \min\left[1, \frac{L(\mathbf{m}')}{L(\mathbf{m}_i)}\right]. \tag{7}$$

For Bayesian problems, MCMC algorithms that are based on the above procedure, which has been named the extended Metropolis algorithm by Hansen *et al.* (2012) (Fig. 1), can be significantly more efficient than algorithms based on the standard Metropolis or Metropolis–Hastings samplers (Mosegaard & Tarantola 1995). More importantly, the procedure offers great flexibility with regard to how the prior is defined, in the sense that all that is needed is a means of generating prior samples rather than a formal prior probability distribution. Critical to the algorithm's effectiveness, however, is how each new set of model parameters is generated from the prior distribution conditional to the current point in the Markov chain. Clearly, if the change between the proposed and previously accepted models is too large from one iteration to the next, then the probability of rejecting the proposal will be high because of the large difference in likelihood (eq. 7). Conversely, if the change is too small then the MCMC algorithm will be slow to equilibrate and produce independent posterior samples because models along the Markov chain will exhibit a high degree of correlation. We thus require a proposal mechanism that is able to effectively generate prior samples, yet at the same time ensure that such samples represent reasonable perturbations from one model to the next.

2.3 Sequential geostatistical resampling (SGR)

Sequential simulation aims to generate stochastic realizations of some subsurface parameter of interest conditional to (i) known values of the parameter at certain locations throughout the model space, and (ii) knowledge regarding the overall spatial statistical structure of the parameter under the assumption of stationarity. Originally developed for the simulation of Gaussian random fields based on two-point geostatistics (e.g. Goovaerts 1997; Deutsch 2002), the method is now commonly used for the generation of complex multipoint statistical structures (e.g. Strebelle 2002) and has gained much corresponding popularity in reservoir characterization and groundwater studies (e.g. Caers & Zhang 2004; Huysmans & Dassargues 2009). The technique is iterative and essentially involves the generation of parameter values along a randomly chosen path through the model space. At each iteration, the conditional pdf for the model cell of interest is calculated based on known data and previously simulated points, which is then used to simulate a random value at that location. Multiple realizations are easily generated by repeating the procedure using different random paths.

One of the most promising recent applications of the extended Metropolis approach outlined in Section 2.2 involves incorporating detailed geostatistical prior information into the inverse problem through the use of sequential simulation. Specifically, recent efforts have explored the idea of using sequential simulation to iteratively perturb models along the Markov chain by resampling randomly

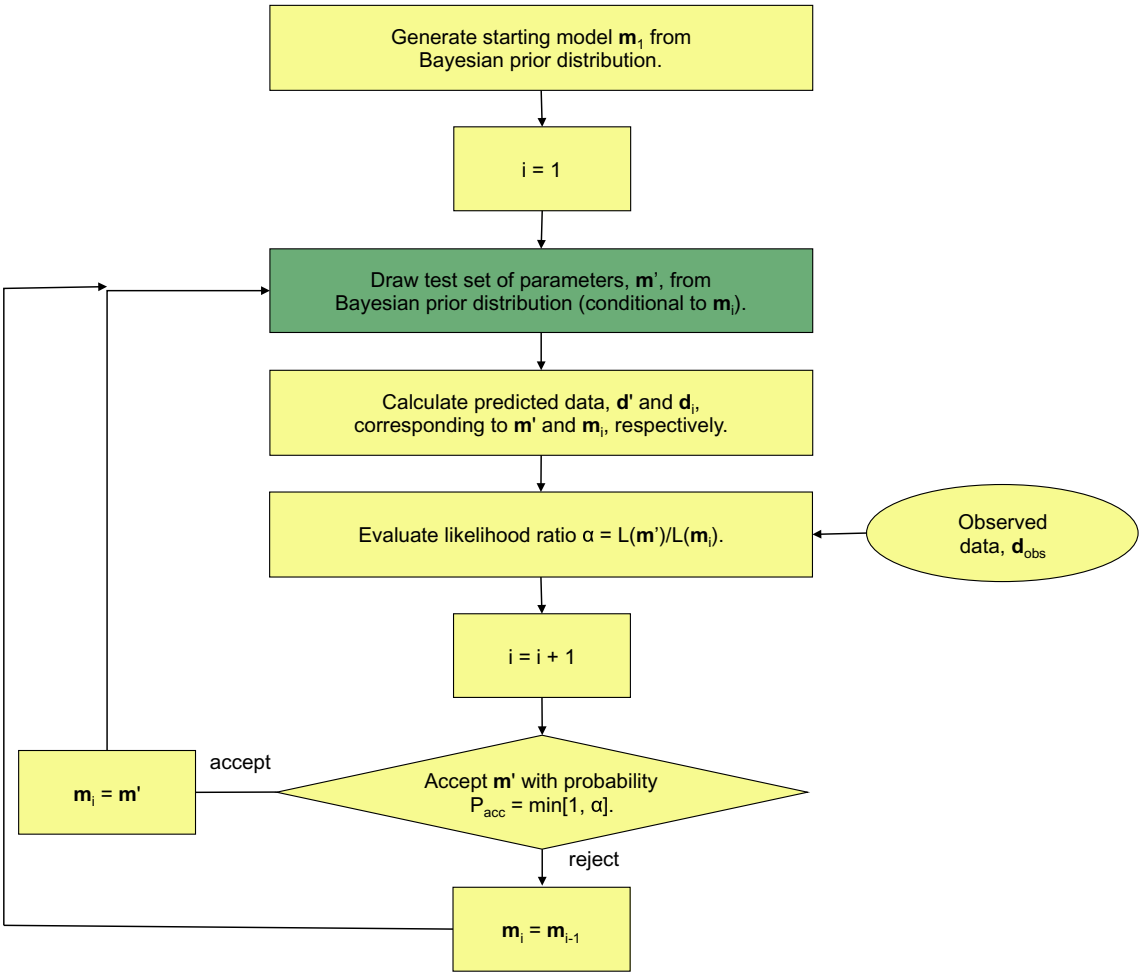

Figure 1. Extended Metropolis algorithm of Mosegaard & Tarantola (1995) for sampling from the Bayesian posterior distribution of model parameters.

selected 'subdomains' of the model grid while treating all of the other values in the grid as known. As demonstrated by Hansen *et al.* (2012), whether sequential simulation occurs in a single iterative process involving all model parameters, or whether it occurs through multiple resimulation steps involving random subsets of the model parameters, stochastic realizations from the prior will be generated as both approaches equilibrate to $\rho(\mathbf{m})$. The advantage of the latter SGR approach in the context of MCMC is that it offers a means of controlling the magnitude of the model perturbations through the size of the subdomain to be resampled. That is, SGR provides a way of making small perturbations along the Markov chain while honouring existing measurements and prior geostatistical information. Any algorithm for performing sequential simulation can be used for SGR within the extended Metropolis algorithm, and various types of hard and soft data can be readily taken into account. Please refer to the paper of Hansen *et al.* (2012) for further details.

In recent studies, two different SGR implementations have been considered for the stochastic solution of near-surface geophysical and hydrological inverse problems. The first involves the resimulation of a contiguous block of model parameters at each MCMC iteration, where the position of the block within the model grid is chosen from a uniform random distribution (Fu & Gómez-Hernández 2009; Cordua *et al.* 2012). In the second implementation, a fixed number of model parameters are also resimulated at each MCMC iteration, but in this case the locations of the parameters are chosen independently and at random (Hansen *et al.* 2008; Irving & Singha 2010; Mariethoz *et al.* 2010; Hansen *et al.* 2012). Many other possibilities for resimulation clearly exist. With respect to the two considered approaches, results of testing on a rather limited number of examples have suggested that the SGR methodology could provide an effective means of generating samples from the Bayesian posterior distribution for a range of near-surface inverse problems. What is critically missing from the existing literature, however, is a detailed and systematic analysis of the method with regard to its performance under different types of subsurface models, numbers of data, and levels of data error, with a view towards eventual practical applications. The latter is essential for understanding under what conditions and algorithmic choices the SGR method will perform best, as well as when the overall approach can be expected to be computationally feasible. Below we explore these questions through a series of numerical examples by evaluating the efficiency of different resampling strategies with regard to their ability to generate statistically independent posterior realizations.

2.4 Output analysis

To determine the number of extended Metropolis iterations required to produce one independent realization from the Bayesian posterior distribution, which serves as our metric for algorithmic efficiency in

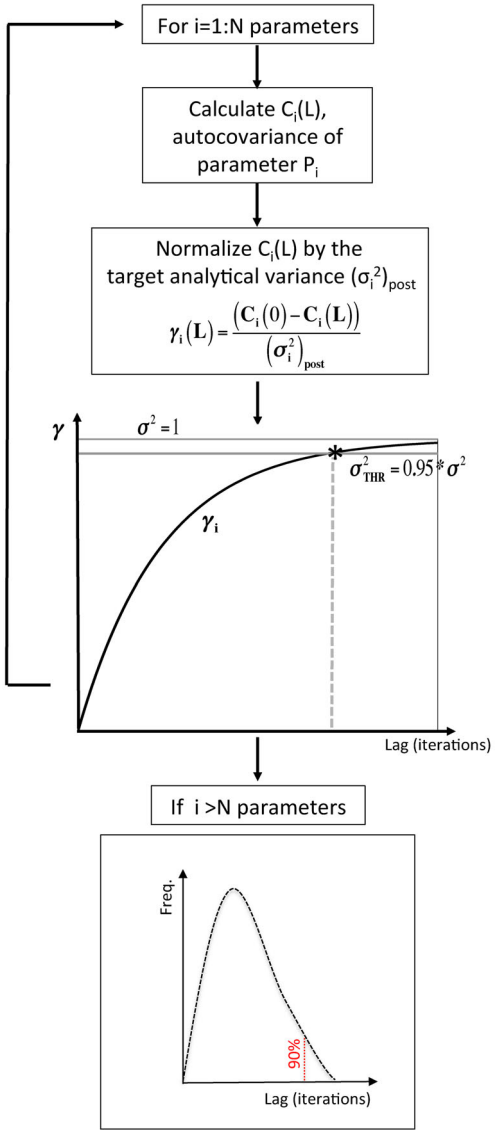

Figure 2. Schematic outline of the autocovariance analysis procedure used in this study to assess the number of MCMC iterations required to generate independent samples from the Bayesian posterior distribution.

this paper, we analyse the statistics of the autocovariance behaviour of each model parameter calculated across the different MCMC samples. Before such an approach can be utilized, we must first discard the pre-burn-in samples from each Markov chain, which are still influenced by the starting configuration of model parameters and thus do not properly represent the posterior (e.g. Gilks *et al.* 1996). To this end, we follow the approach outlined in Cordua *et al.* (2012) which involves examination of the behaviour of the likelihood as a function of iteration. Burn-in is considered to occur when the likelihood begins to fluctuate around an equilibrium level and the residuals appear to describe the expected distribution of the data errors.

Fig. 2 summarizes the overall analysis procedure that we apply to the post-burn-in MCMC samples. For each model parameter i in the simulation grid, the autocovariance $C_i(L)$ is computed across the different realizations for different iteration lag values L. This is then transformed into a normalized variogram function $\gamma_i(L)$ according to the following equation:

$$\gamma_i(L) = \frac{(C_i(0) - C_i(L))}{(\sigma_i^2)_{\text{post}}}, \tag{8}$$

where $(\sigma_i^2)_{\text{post}}$ is the 'true' value of the posterior variance, which we assume known and obtainable analytically in this study in order to validate our simulation results. Further details on this assumption, which is valid only for the case of a linear inverse problem, are provided in Section 3.

The values for $(\sigma_i^2)_{\text{post}}$ are used in eq. (8) to ensure that each variogram normalizes to a sill value of $\sigma^2 = 1$ at a certain iteration lag L. The number of iterations required for a particular model parameter to represent an independent sample can then be computed as the lag at which its normalized variogram intercepts a given threshold value, which is chosen to be close to one. In order to find the number of iterations required to generate one independent realization from the *joint* posterior distribution of *all* model parameters, we use the fact that the vast majority of individual model parameters must have been updated enough to represent independent samples. For this study, we consider the 90th percentile of the iteration lag values computed across all model cells, which corresponds to the number of iterations that are needed with the extended Metropolis algorithm before 90 per cent of model parameters can be considered to be independent.

3 SYNTHETIC STUDY

In the context of a series of numerical examples, we now systematically investigate the performance of SGR as a proposal mechanism within the extended Metropolis algorithm. In particular, we examine with respect to their effect on the number of iterations required to generate statistically independent realizations from the Bayesian posterior distribution: (i) the nature of the MCMC model proposals, as described by their shape (i.e. blocks of cells or randomly selected individual cells) and their size (i.e. number of cells updated in each iteration); (ii) the nature of the data, specifically the number of measurements considered and their associated level of error; and (iii) the degree of spatial correlation between the model parameters. This is done for a number of simple but pertinent straight-ray crosshole georadar tomographic examples exhibiting different model and data characteristics. Although we fully acknowledge that the straight-ray assumption will be violated in cases where significant heterogeneity exists in the distribution of subsurface electromagnetic wave velocity, we use it here because it provides a linear inverse problem with a correspondingly fast forward model for which an analytical solution exists for the Bayesian posterior assuming Gaussian priors (Tarantola 2005). This solution allows us to validate our simulation results and determine the number of MCMC iterations required to produce one independent posterior realization, which again serves as our metric for algorithmic efficiency. In this regard, working with the linear straight-ray problem provides us with a 'best-case' indicator of the efficiency of SGR within the extended Metropolis algorithm, as scenarios involving substantial non-linearity and/or more complex priors could be expected to decrease this efficiency as the posterior distribution deviates from multi-Gaussian.

3.1 Tomographic inverse problem

Crosshole georadar traveltime tomography is a popular geophysical method for imaging the spatial distribution of subsurface electro-

magnetic wave velocity between two boreholes. With this technique, a short electromagnetic pulse is radiated from a transmitter antenna, located in one of the boreholes. The pulse then propagates through the subsurface and is recorded at a receiver antenna, which is located in the adjacent borehole. Assuming that the trajectory of the first-arriving energy travelling between the antennas can be approximated by a straight line or ray, a linear relation can be formulated between the measured traveltime of the first-arriving energy and the distribution of wave 'slowness' (1/velocity) in the subsurface as follows:

$$\mathbf{t} = \mathbf{D}\mathbf{s} + \epsilon, \tag{9}$$

where $\mathbf{t}$ is a vector of traveltime measurements corresponding to different transmitter and receiver configurations, $\mathbf{D}$ is a matrix whose row and column entries contain the distance travelled by ray i in model cell j, respectively, $\mathbf{s}$ is a vector containing the slowness values for each model cell, and ϵ represents the observation or picking errors on the measured traveltimes.

The crosshole georadar tomographic inverse problem consists of using $\mathbf{t}$ to infer $\mathbf{s}$. Under the prior assumptions that (i) the slowness values between the boreholes are distributed according to a multi-Gaussian distribution having mean $\mathbf{s}_0$ and covariance matrix $\mathbf{C}_s$, and (ii) the traveltime errors are also multi-Gaussian with mean zero and covariance matrix $\mathbf{C}_t$, the following analytical expressions can be derived (e.g. Tarantola 2005):

$$\widetilde{\mathbf{s}} = \mathbf{s}_0 + \mathbf{C}_s\mathbf{D}^T\left(\mathbf{D}\mathbf{C}_s\mathbf{D}^T + \mathbf{C}_t\right)^{-1}\left(\mathbf{t}_{\text{obs}} - \mathbf{D}\mathbf{s}_0\right) \tag{10}$$

$$\widetilde{\mathbf{C}}_s = \mathbf{C}_s - \mathbf{C}_s\mathbf{D}^T\left(\mathbf{D}\mathbf{C}_s\mathbf{D}^T + \mathbf{C}_t\right)^{-1}\mathbf{D}\mathbf{C}_s, \tag{11}$$

where $\widetilde{\mathbf{s}}$ and $\widetilde{\mathbf{C}}_s$ are the mean and covariance matrix of the posterior distribution, which in this case is also multi-Gaussian. Note that, based on eqs (10) and (11), Hansen & Mosegaard (2008) developed a means of directly generating independent posterior realizations using sequential simulation. Indeed, for linear geophysical inverse problems with Gaussian model priors and data uncertainties, MCMC methods are not necessary as the latter approach is extremely computationally efficient. Again, however, the analysis presented in this paper for the linear inverse problem allows us to investigate the performance of SGR within MCMC under a best-case scenario where model validation is possible, which in turn allows us to have a sense for its best possible performance under less trivial conditions.

3.2 Considered examples and data

Two heterogeneous radar slowness distributions having different spatial correlation lengths are considered as the 'true' subsurface models to be estimated in this study. The models are specified on a 8.1-m-wide by 14.1-m-deep grid discretized into 0.3-m-square cells, which yields a 47 × 27 domain of 1269 model parameters (Fig. 3). To generate these slowness distributions, we used sequential Gaussian simulation (Deutsch & Journel 1992) assuming a global mean and variance of 10 ns m^{-1} and 0.5 ns^2 m^{-2}, respectively, and an exponential geostatistical model for the spatial covariance. For the first example, shown in Fig. 3(a), the horizontal and vertical correlation lengths were set to 4.8 and 2.7 m, respectively. For the second model, shown in Fig. 3(b), they were set to 2.1 and 0.9 m, respectively.

Having the slowness distributions, we next simulated the acquisition of straight-ray crosshole georadar traveltime data for the two

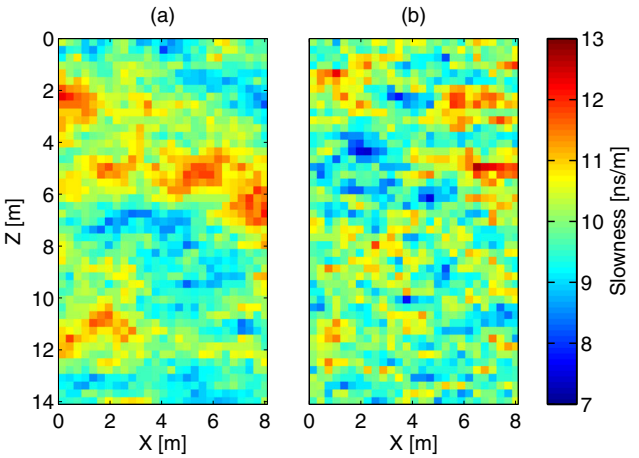

Figure 3. 'True' slowness fields considered for our analysis of the performance of sequential geostatistical resampling within the extended Metropolis algorithm. The fields were generated using sequential Gaussian simulation. (a) Longer-correlation-length example. (b) Shorter-correlation-length example.

Table 1. Six different data scenarios that were considered for each slowness model in Fig. 3.

Scenario	Number of traveltime measurements	Noise standard deviation (ns)
1	2209	0.5
2	2209	0.8
3	2209	1.1
4	576	0.5
5	576	0.8
6	576	1.1

models in Fig. 3. Table 1 summarizes the six different scenarios that were considered in each case. Traveltimes were generated for two different measurement configurations, the first involving a borehole transmitter and receiver increment of 0.3 m from 0.15 to 13.95 m depth, yielding 2209 data, and the second involving a 0.6-m increment, yielding 576 data. Uncorrelated zero-mean Gaussian random noise was then added to each traveltime. In this regard, standard deviation values of 0.5, 0.8 and 1.1 ns were considered for the traveltime errors. Together, these six scenarios allow us to investigate the performance of MCMC under varying degrees of 'data complexity' for the longer- and shorter-correlation-length models. It is important to note that the number of data considered, their levels of error and the size of the model space are all quite typical of relatively simple near-surface geophysical inverse problems. Also note that additional constraints such as borehole logs were not considered in this study to avoid overconditioning of our inversions in an already small-scale model domain.

Assuming a Gaussian prior distribution for the slowness between the boreholes, the Bayesian posterior distribution for the linear traveltime inverse problem is completely described by the mean and covariance given by eqs (10) and (11). Fig. 4 shows the corresponding mean and variance fields for the longer-correlation-length model in Fig. 3(a) for Scenarios 1 and 6. As expected, the mean field for Scenario 6, which considers fewer measurements and higher measurement errors, displays smoother variation than the mean field for Scenario 1, which considers more measurements and smaller measurement errors. The posterior uncertainty of the slowness field for Scenario 6 is also larger than that for Scenario 1. In both cases, the

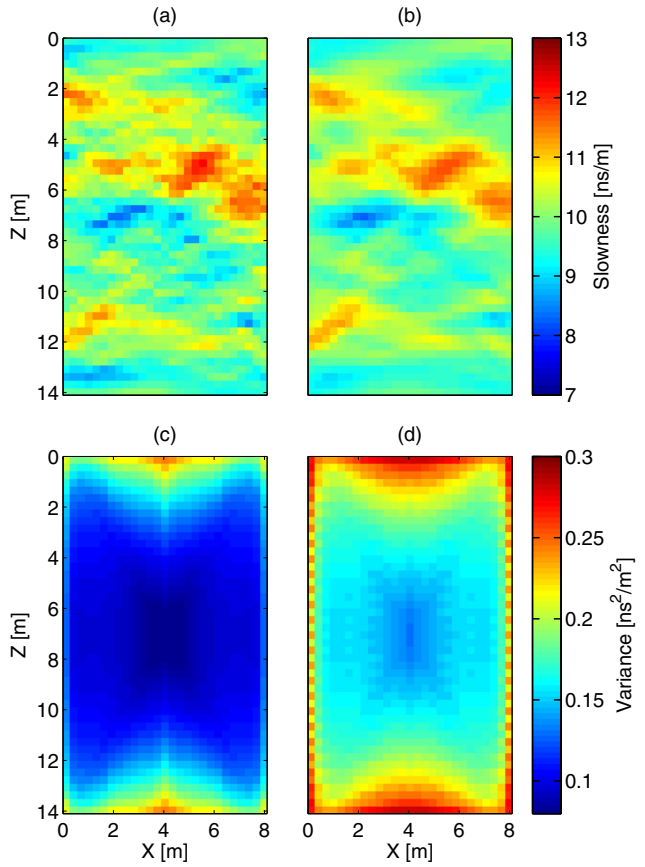

Figure 4. Posterior mean and variance fields computed using eqs (10) and (11) for the slowness distribution in Fig. 3(a) and for inversion Scenarios 1 (a and c) and 6 (b and d), respectively.

estimated uncertainties are highest in the parts of the model space that are characterized by particularly low ray coverage (i.e. at the top and the bottom of the model domain) and lowest in those regions characterized by high ray coverage (i.e. in the central portion of the model domain).

3.3 Results

3.3.1 Effect of the nature of the resampling

We first investigate the effect of the nature of the sequential resampling on the efficiency of the extended Metropolis algorithm to generate independent posterior samples. To this end, we consider both block and randomly selected cells resampling and we vary the total number of cells resampled in each MCMC iteration. For the block resampling, randomly chosen square regions of the model space were perturbed, with the size of the blocks being equal to either 1, 4, 9 or 25 cells. For the randomly selected cells resampling, the same numbers of perturbed cells were considered but their locations were chosen independently. For all of the tests performed, the number of traveltime data and the standard deviation of the data noise were held constant at 2209 and 1.1 ns, respectively, which corresponds to Scenario 3 (Table 1). The number of iterations required for burn-in was in general between 1000 and 3000 and the MCMC algorithm was run for 5 million iterations. To verify the consistency of the results obtained, 10 independent Markov chains were run for each example.

Table 2 and Fig. 5 show the results obtained for the longer- and shorter-correlation-length examples from Fig. 3 in terms of the overall MCMC acceptance rate and the number of iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian posterior distribution (Fig. 2). To calculate the acceptance rate, we divided the total number of accepted proposals by the total number of iterations. As expected, this quantity is strongly dependent upon the number of cells resampled because the latter determines the overall magnitude of the proposed perturbation. That is, the larger the number of cells resampled, the bigger the proposed model change and the greater the risk that it will be rejected because of a larger corresponding change in the likelihood (eq. 7).

Notice in Fig. 5 that, for each example and type of resampling considered, there exists an 'optimal' region for the number of points resampled where the number of iterations required for independent posterior samples is minimized (i.e. maximum efficiency). Increasing the number of cells from this value, the model perturbation in each MCMC iteration increases, thus leading to a decrease in the acceptance rate and a reduction in efficiency. Conversely, decreasing the number of cells from this value means that the acceptance rate will increase but that the posterior samples will show greater correlation, which also leads to a reduction in efficiency. Previous work with MCMC in the current context has suggested that the observed acceptance rate can be used as an approximate indicator of algorithmic performance, in the sense that having an acceptance rate of approximately 25–40 per cent should represent a good balance between exploration and exploitation of the model parameter space (Irving & Singha 2010; Cordua *et al.* 2012; Hansen *et al.* 2012). We see from Fig. 5, however, that this can only be loosely interpreted because the acceptance rate at which the extended Metropolis algorithm will exhibit maximum efficiency is dependent upon the type of resampling considered as well as nature of the subsurface model. For example, for the longer-correlation-length model, block resampling has optimal efficiency at around 45 per cent, whereas randomly selected cells resampling has optimal efficiency at around 20 per cent. For the shorter-correlation-length model, on the other hand, both block and randomly selected cells resampling exhibit optimal efficiency at around 28 per cent. The latter findings are consistent with the work of Gelman *et al.* (2003), who found that the average acceptance rate under optimal efficiency for an uncorrelated multivariate normal distribution is approximately 25 per cent.

Somewhat concerning about the results in Table 2, in the context of practical application of SGR within the extended Metropolis algorithm, is the number of iterations necessary to generate a single independent posterior realization under optimally efficient conditions. In the best case, which for this series of tests involved the shorter-correlation-length example and randomly selected cells resampling of 4 model parameters at each iteration, 54 200 iterations were needed before 90 per cent of the model parameters had been updated enough to be considered independent. Conversely, in the worst case, which involved the longer-correlation-length example and block resampling of nine model parameters, this value increased to 527 550 iterations. Clearly these kinds of numbers combined with a forward model of reasonable complexity could quickly limit the computational tractability of the approach. Previous work on SGR for this problem largely considered examples with relatively small numbers of data (e.g. less than 1000 traveltimes) and relatively large data errors (e.g. traveltime errors with a standard deviation greater than 1 ns). As a result, the efficiency of the algorithm in those cases could be expected to be greater than that seen here. Nonetheless, estimating as we did a grid of 47 × 27 model parameters using 2209

Table 2. Overall MCMC acceptance rate and the number of iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian posterior distribution for different types of sequential geostatistical resampling. Results are shown for both the longer- and shorter-correlation-length examples from Fig. 3, assuming the data from Scenario 3 (Table 1).

Number of cells perturbed	Random cells resampling Acceptance rate (per cent)	Random cells resampling Number of iterations	Block resampling Acceptance rate (per cent)	Block resampling Number of iterations
Longer-correlation-length example				
1	75	848 580	75	848 580
4	46	575 870	44	326 690
9	25	255 580	22	547 550
16	12	259 000	14	1 473 820
25	5	973 640	9	2 000 200
Shorter-correlation-length example				
1	64	128 940	64	128 940
4	28	54 200	28	112 110
9	9	74 700	13	359 760
16	2	183 700	8	2 406 160
25	1	858 380	5	3 002 030

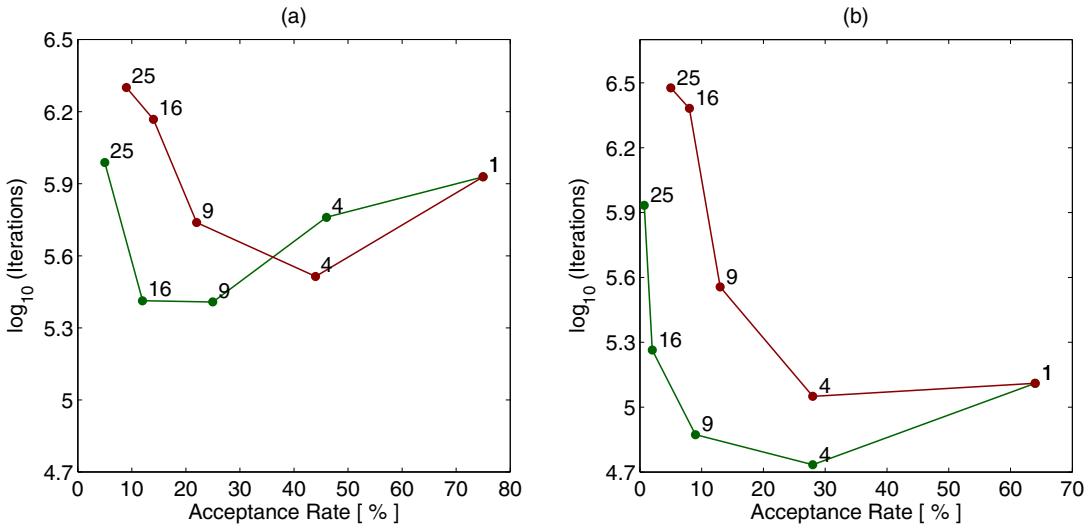

Figure 5. Number of MCMC iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian posterior distribution as a function of the overall acceptance rate. Results are shown for random cells resampling (green) and block resampling (red) and for the (a) longer- and (b) shorter-correlation-length examples in Fig. 3 assuming the data from Scenario 3 (Table 1). The number of cells perturbed is indicated above each data point.

traveltime measurements is a realistic near-surface geophysical inverse problem, and in fact at the low end of complexity in terms of dimension and number of data. We explore this aspect in further detail in the next section when we examine the inversion results for different data characteristics.

Finally, Fig. 5 indicates that, in almost all cases, block resampling is a considerably less efficient means of generating independent posterior realizations than randomly selected cells resampling for the same number of perturbed model parameters. Indeed, with the exception of four-cell resampling in the longer-correlation-length example, the number of iterations required for random cell resampling is significantly less. Moreover, at the point of optimal efficiency, block resampling requires more iterations than randomly selected cells resampling, meaning that it appears to be a less efficient proposal method in general. To better understand this behaviour, we computed the *cell-by-cell* overall acceptance or update rate, calculated by dividing the total number of accepted proposals by the total number of proposed changes for each model cell, for block and randomly selected cell resampling for the longer-correlation-length example with 9 cells perturbed (Fig. 6). Note that both resampling methods show a very similar *global* acceptance rate for this example, which may be misleading (Table 2) given their very different performance characteristics. When randomly selected cells resampling is used, we see that all of the model cells show a very similar update rate of approximately 25 per cent (Fig. 6a), which is a positive characteristic because it means that the MCMC algorithm is not generating independent posterior samples quickly for some model cells, but slowly for others. Indeed, the cells having the slowest update rate control the speed at which independent posterior samples are generated. When block resampling is considered, on the other hand, the cell-by-cell acceptance rate reveals that a high number of model parameter perturbations (∼50 per cent) are accepted in

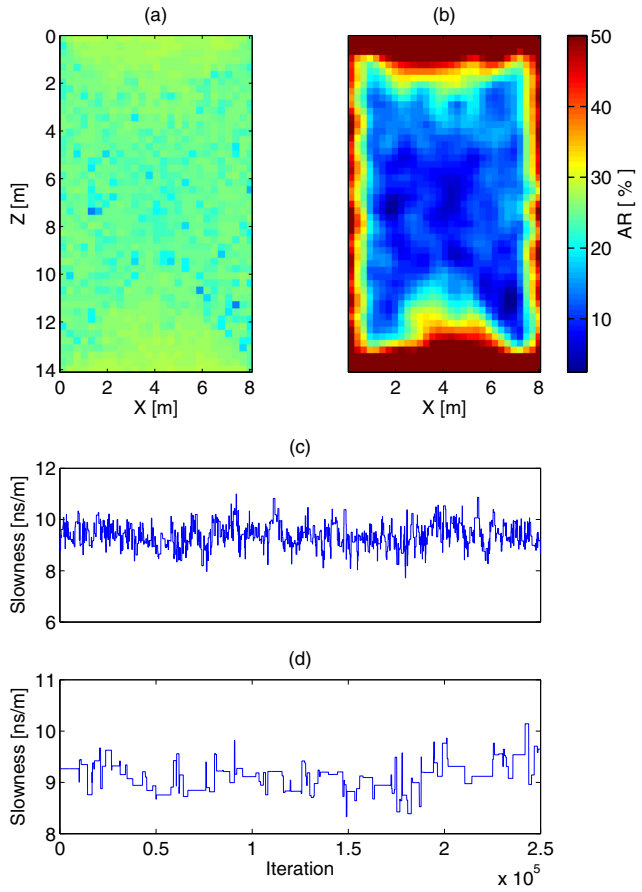

Figure 6. Cell-by-cell acceptance rate (AR) computed for (a) randomly selected cells resampling and (b) block resampling for the longer-correlation-length example shown in Fig. 3(a). Nine model cells were resampled in each MCMC iteration. Also shown are the values of the model parameters located at locations (c) [$x = 4$ m; $z = 0.3$ m] and (d) [$x = 4$ m; $z = 7$ m] as a function of iteration for the block resampling case.

regions characterized by low ray density (e.g. at the top and bottom of the model domain), whereas few accepted transitions (∼5 per cent) occur in regions with high ray density (e.g. in the centre of the model domain) (Fig. 6b). This is illustrated more clearly in the update history plots in Figs 6(c) and (d), where we observe a substantial difference in the update rate between two model parameters located in the upper and central regions of the model, respectively.

The above behaviour is a consequence of the combined effects of the spatially varying sensitivity of the traveltime data along with our inherent inability to effectively control the magnitude of the perturbed parameter values with the standard SGR approach. That is, with block resampling, the model perturbations are localized to a specific region of the parameter grid to which the traveltime data may be more or less sensitive. As a result, the MCMC algorithm is able to accept proposals quickly in regions exhibiting lesser sensitivity, but comparatively slowly in regions exhibiting higher sensitivity. Unfortunately, there is no easy way to change this behaviour (e.g. to tailor the magnitude of the perturbed values in each region of the model in accordance with the sensitivity of the data) because the resampling is done based on the mean and covariance of the prior distribution. Conversely, with randomly selected cells resampling, model parameter perturbations are not localized to one region and thus we avoid having some parts of the model experiencing significantly greater acceptance rates than others, which in turn leads to greater algorithmic efficiency. Note that although one might consider biasing the block resampling to visit highly sensitive regions of the model space more frequently, thereby updating these regions more often, we found that this does not offer a practical solution to the above problem because surrounding regions, which condition the block resampling, become updated less often. This has the effect of reducing the variability of the resampling in the biased regions, even though they are resampled more frequently.

Table 3. Number of MCMC iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian prior and posterior distributions for the longer-correlation-length example shown in Fig. 3(a), assuming the data from Scenario 3 (Table 1). Nine model cells were resampled in each iteration.

Distribution	Random cells resampling	Block resampling
Prior	108 620	27 980
Posterior	215 520	547 550

To gain some final clarity into the above aspects, Table 3 shows the number of MCMC iterations found necessary to generate independent samples from the *prior* distribution (i.e. where no data were considered) versus the posterior distribution (i.e. prior × likelihood) for nine-parameter block and randomly selected cells resampling for the longer-correlation-length example. Here we see that block resampling allows for significantly more efficient sampling of the prior because the overall magnitude of change in a 3×3 resampled block will be greater than that for nine isolated resampled cells whose values are more strongly conditioned by the surrounding data. In other words, the block resampling strategy provides a chain of prior realizations with an overall lower level of correlation between adjacent samples when compared with randomly selected cells resampling. However, as soon as the likelihood function is included into the MCMC algorithm, most of the perturbations performed with block resampling, especially those proposed in locations characterized by a relatively high sensitivity of the traveltime data, are rejected. As a result, efficiency greatly decreases and the number of iterations required to generate independent posterior samples is approximately double that of randomly selected cells resampling. It is important to note that, without a cell-by-cell analysis of the acceptance rate as performed above, it would extremely difficult to identify such problems related to the unequal updating of model parameters within the extended Metropolis algorithm. Indeed, in some cases, we might generate a series of seemingly independent posterior samples with a reasonable overall acceptance rate in which some of the model parameters were never updated.

3.3.2 Effect of data and model characteristics

We next examine the impact of the number of data, level of data error, and degree of model parameter spatial correlation on the efficiency of the SGR approach to generate independent posterior samples. This was again done with a view towards applications to realistic near-surface inverse problems; that is, although it is well understood that increasing the number of data and/or decreasing the level of data error will result in less efficient posterior sampling because of a narrower likelihood function, we wish to understand the magnitude of these effects in the context of what range of behaviour

Table 4. Overall MCMC acceptance rate and the number of iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian posterior distribution for different data and model characteristics as described in Table 1 and Fig. 3, respectively. All values in the table correspond to resampling performed using the number of cells that approximately maximized algorithmic efficiency. Both randomly selected points and block resampling are shown.

	Random cells resampling			Block resampling		
Scenario	Number of cells perturbed	Acceptance rate (per cent)	Number of iterations	Number of cells perturbed	Acceptance rate (per cent)	Number of iterations
			Longer-correlation-length example			
1	3	24	1 119 700	1	35	1 412 500
2	6	23	650 780	4	36	790 530
3	9	25	215 520	4	44	326 690
4	8	23	439 730	4	40	407 810
5	18	24	85 650	9	36	107 230
6	29	25	41 580	16	32	64 510
			Shorter-correlation-length example			
1	2	23	681 150	1	30	1 000 285
2	3	25	123 940	1	34	301 995
3	4	28	54 200	4	29	112 110
4	4	24	76 100	4	27	192 480
5	8	24	19 650	4	41	31 600
6	14	25	8600	9	32	14 300

might be expected in practical situations. Table 4 summarizes the results obtained for the longer- and shorter-correlation-length examples in Fig. 3 when considering each of the different data scenarios described in Table 1 and using both randomly selected cells and block resampling. In each case, we present only the results for the number of resampled cells that was found to give approximately the maximum algorithmic efficiency. In other words, the results in Table 4 are a condensed summary of a wide range of tests performed such that we could determine the number of cells requiring the least number of iterations to generate independent posterior samples. Unfortunately, no metric is available that can provide the latter information before running an MCMC simulation.

We see in Table 4 that, for most of the cases considered, it is not possible to update more than a few model parameters in each MCMC iteration using SGR if our goal is to have the maximum possible algorithmic efficiency. Indeed, in some instances, only one parameter can be perturbed at a time. An unavoidable consequence in these cases is that the extended Metropolis algorithm will require an extremely long time to move through the posterior parameter space, which is reflected in a correspondingly large number of iterations required to generate independent samples. We also see that, as expected, increasing the number of data and/or decreasing the level of data error reduces MCMC efficiency. What is surprising, however, is the enormity of the range in the number of iterations required for 90 per cent of the model parameters to represent independent posterior samples for the different test cases, which again all represent situations that may be encountered in practical applications. In the best case, corresponding to the shorter-correlation-length example with 576 traveltimes, a data error standard deviation of 1.1 ns, and using randomly selected cells resampling, 8600 MCMC iterations were found necessary. In the worse case, which corresponds to the shorter-correlation-length example with 2209 traveltimes, a standard deviation of 0.5 ns, and using block resampling, 1 412 500 iterations were required. Within a single example and using the same type of resampling, the difference between the 'easiest' data case (Scenario 1) and the 'hardest' data case (Scenario 6) was between one and two orders-of-magnitude. Clearly, these results suggest that the number of data and their expected level of error should be carefully taken into account when considering the use of the SGR approach, as in many cases the generation of only a few independent posterior samples could easily overwhelm computational resources. To put this into perspective, consider a relatively fast geophysical forward code that takes one second to compute the data corresponding to a set of subsurface model parameters. When dealing with our worst case where 1 412 500 MCMC iterations were necessary for a single posterior sample, generating 50 such samples would take over two years. On the other hand, in our best performing test case where only 8600 iterations were required, generating the 50 samples would take approximately five days. Again, previous studies involving SGR have generally involved situations closer to the latter case in terms of the number of data and level of data error, and have not included a cell-by-cell analysis of the behaviour of the different model parameters as a function of iteration number. Also note in Table 4 that, as observed earlier, for the same test case under optimally efficient conditions, randomly selected cells resampling consistently requires less iterations to generate independent posterior samples than block resampling.

Let us now consider the effects of model parameter correlation. Table 4 shows that, for longer correlation lengths, it is generally possible to perturb more cells at once because the magnitudes of the proposed perturbations are smaller compared to when shorter correlation lengths are considered. However, we see that relative to each scenario, the MCMC algorithm is considerably more efficient for shorter correlation lengths in that a difference of almost one order-of-magnitude is observed when comparing the number of iterations required to independent posterior samples. To gain further clarity into the latter aspect, Table 5 shows the number of iterations found necessary to generate independent samples from the *prior* distribution for randomly selected cells and block resampling for the longer- and shorter-correlation-length examples. The number of resampled cells in each case was chosen to be consistent with that used to generate posterior samples corresponding to Scenario 3 (Table 4). We clearly see that, for both types of resampling, obtaining independent samples from the prior is computationally more expensive in the longer-correlation-length case. That is, the higher correlation between model parameters makes it easier to become stuck around

Table 5. Number of MCMC iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian prior distribution for the longer- and shorter-correlation-length examples in Fig. 3, assuming the data from Scenario 3 (Table 1). In each case, the number of cells resampled was chosen to be the same as that used for posterior sampling (Table 4).

	Random cells resampling		Block resampling	
Correlation length	Number of cells perturbed	Number of iterations	Number of cells perturbed	Number of iterations
Short	4	21 500	4	13 600
Long	9	108 620	4	80 000

a particular configuration, making changes consistent with the prior but often reverting back to the same configuration. These effects are less pronounced when using block resampling, mainly because the overall magnitude of change in a block is greater than in isolated resampled cells whose values are more strongly conditioned to the surrounding data.

4 GRADUAL-DEFORMATION RESAMPLING

The results presented in the previous section illustrated a number of important challenges associated with the use of SGR as a proposal mechanism within the extended Metropolis algorithm. First, localized model perturbations combined with spatially varying data sensitivities can lead to highly inefficient posterior sampling, in the sense that the update rate of model parameters through the simulation grid can vary significantly. This is especially an issue for block resampling, which was found in almost all cases to be less efficient than randomly selected cells resampling. Secondly, the degree of spatial correlation between the model parameters can have a strong effect on MCMC efficiency, as increasing the correlation tends to reduce the speed at which sequential resampling is able move through the prior space. We also noted that the overall SGR approach, at least as presented in previous work, does not allow for direct control over the magnitude of the resimulated values, thus limiting its flexibility as a proposal mechanism. Here, we investigate the performance of an alternative proposal strategy, based on the principle of gradual deformation (Caers 2007), that aims to address these limitations while at the same time retaining the key advantages of SGR, namely the ability to consider complex geostatistical priors as well as condition proposals to hard and soft data. We focus on updating the entire parameter field simultaneously, as this guarantees an equal acceptance rate for all model cells and has been found in other applications to improve MCMC performance when the model parameters are strongly correlated (e.g. Hassan *et al.* 2009).

Gradual deformation is a procedure by which different stochastic realizations of a parameter field, all having the same overall statistical properties, are iteratively linearly combined in order to obtain a satisfactory match to a set of observed data that depends upon that parameter field. It is in essence an optimization process that allows us to 'warp' one realization towards another until a new realization is obtained that allows for adequate predictions of the data. The key to the method lies in the way that the realizations are combined, which is done so as to preserve the statistical properties of the input fields. Many variations of the approach have been presented (e.g. Roggero & Hu 1998; Hu 2000; Hu *et al.* 2001; Hu 2002; Ravelec-Dupin *et al.* 2002). Of particular interest in the context of our work is the study of Hu *et al.* (2001), who proposed a gradual-deformation methodology based on perturbing the uniform random number vector used to draw from local conditional distributions in sequential geostatistical simulation. With their methodology, this vector (v_1) is first transformed into a standard Gaussian random noise vector (y_1) using the inverse of the cumulative Gaussian distribution G, that is, $y_1 = G^{-1}(v_1)$. Next, y_1 is 'deformed' into a perturbed standard Gaussian random noise vector y using the following 'gradual-deformation rule':

$$y = y_1 \cos(\pi t) + y_2 \sin(\pi t), \tag{12}$$

where parameter $t \in [0, 0.5]$ controls the overall magnitude of the perturbation and y_2 is another independent standard Gaussian random noise vector. The vector y is then back-transformed into a vector of perturbed uniform random numbers v using $v = G(y)$. Vector v is a perturbation of v_1, hence when v is used instead of v_1 to draw from local conditional distributions in sequential simulation, we obtain a perturbation of the original stochastic realization that was generated using v_1. Combining the Gaussian random noise vectors using eq. (12) ensures that the mean and variance are preserved, with the choice of t determining how different vector y is from vector y_1. For example, when $t = 0$, $y = y_1$ (hence $v = v_1$) and no perturbation is produced. As t increases, y (hence v) becomes gradually different from y_1 (hence v_1). In the limit case of $t = 0.5$, $y = y_2$ and a completely independent stochastic realization is generated. Because this methodology works on perturbing the uniform random vector used to draw from local conditional distributions in sequential simulation, it can be used anywhere sequential simulation is applicable, for example even in the context of complex priors defined by multiple-point geostatistics (Caers 2007).

Although gradual deformation has been utilized extensively in the context of iterative optimization to fit a set of subsurface model parameters to measured data, its potential applicability to MCMC methods as a prior proposal generator remains largely unexplored. Table 6 and Fig. 7 show the results we obtained from running the extended Metropolis algorithm with gradual-deformation resampling for different t values on the longer- and shorter-correlation-length examples in Fig. 3. The number of measurements and data error standard deviation correspond to Scenario 3 in Table 1. As expected, larger t values lead to bigger global perturbations and thus more rejections of the proposed models. For both examples, an acceptance rate of approximately 25 per cent is seen to provide the fewest number of iterations required for 90 per cent of the model parameters to represent independent posterior samples. This is similar to the acceptance rate at maximum efficiency seen for randomly selected cells resampling (Table 4). Comparing the performance of gradual deformation with the other resampling methods, however, we see that it offers in many cases a significant improvement in efficiency. For the longer-correlation-length case, for example, 111 610 iterations are required as opposed to 215 520 and 326 690 for randomly selected cells and block resampling, respectively. Indeed, the curves showing the number of iterations versus acceptance rate in Fig. 7 are consistently lower than those for the other sampling methods shown in Fig. 5. Also note that, as observed previously, the resampling is more efficient when shorter correlation lengths are involved. However, this difference is less pronounced than was observed for the randomly selected points and block resampling (Table 4).

We next investigated the best possible performance with gradual-deformation resampling for the different scenarios described in Table 1. Table 7 shows the results obtained. In Fig. 8, we compare these results with those obtained for randomly selected cells and block resampling. Note that, for both the longer- and shorter-correlation-length examples, the gradual-deformation-based proposal strategy

Table 6. Overall MCMC acceptance rate and the number of iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian posterior distribution as a function of the tuning parameter t value used in gradual-deformation resampling. Results are shown for both the longer- and shorter-correlation-length examples from Fig. 3, assuming the data from Scenario 3 (Table 1).

Longer-correlation-length example			Shorter-correlation-length example		
t	Acceptance rate (per cent)	Number of iterations	t	Acceptance rate (per cent)	Number of iterations
0.0022	64	452 900	0.0040	57	88 800
0.0032	50	217 340	0.0060	40	52 180
0.0042	38	189 140	0.0070	32	44 490
0.0052	28	153 840	0.0075	29	45 200
0.0057	25	111 610	0.0080	25	43 580
0.0062	21	113 880	0.0085	23	42 000
0.0082	10	135 540	0.0090	20	42 300
0.0100	5	214 500	0.0100	16	44 360
0.0110	3	267 620	0.0110	12	52 010
0.0130	2	512 620	0.0200	1	534 220

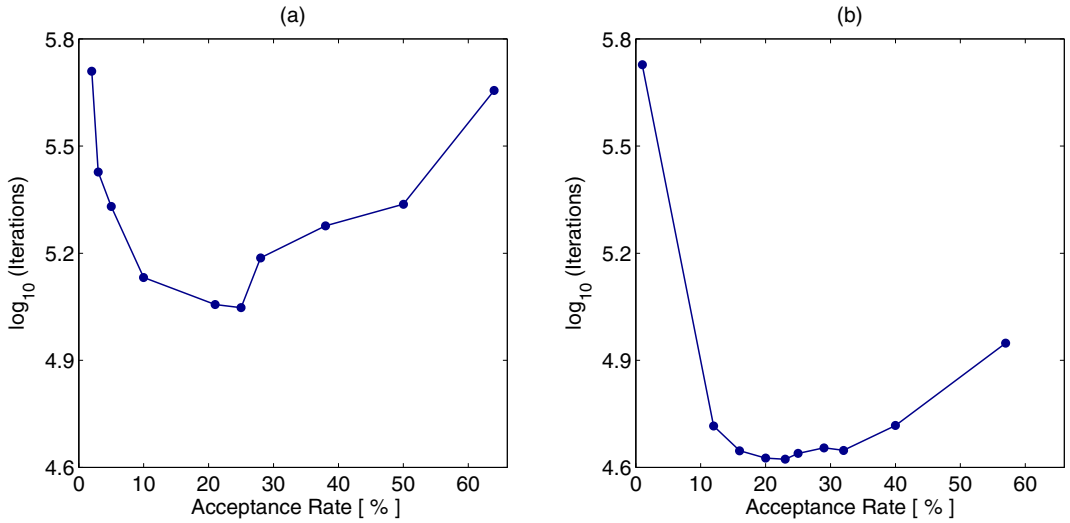

Figure 7. Number of MCMC iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian posterior distribution as a function of the overall acceptance rate for gradual deformation using different values of the tuning parameter t. Results are shown for the (a) longer- and (b) shorter-correlation-length examples in Fig. 3 assuming the data from Scenario 3 (Table 1).

is generally more efficient at drawing independent posterior samples. This is particularly evident for the longer-correlation-length case, but also holds for the shorter correlation length with the exception of Scenarios 5 and 6 where randomly selected points resampling is seen to be slightly more efficient. Although these improvements are promising, perhaps the most important observation from Table 1 is that in most cases the gradual-deformation proposal strategy still leads to a computationally intractable number of iterations required for a single posterior sample. That is, the method, at least when implemented within the extended Metropolis algorithm, does not overcome the computational limitations seen earlier with SGR using randomly selected points and block resampling.

5 DISCUSSION AND CONCLUSIONS

The use of SGR within the extended Metropolis algorithm has received significant recent interest for the Bayesian stochastic inversion of near-surface geophysical data because it naturally allows for the incorporation of complex prior information into the inverse problem, as well as conditioning to a variety of other data such as measurements of physical properties along boreholes. Although previous work has suggested that the approach may be directly applicable to a wide range of problems, the results of our systematic testing indicate that there will exist many practical cases where the methodology, at least as currently formulated, is not computationally feasible. Indeed, we observed that for numbers of data, numbers of model parameters, and levels of data error that are quite typical of small-scale near-surface geophysical inverse problems, SGR required at the very least many thousands of iterations to provide a single independent sample from the Bayesian posterior distribution, with this number being over one million iterations for some examples. This was found to be the case for all three SGR strategies examined, with gradual-deformation resampling providing the most efficient implementation, followed by randomly selected cells resampling and then block resampling. Note that these large numbers of iterations did not pose significant issues for the straight-ray tomographic problem considered in our testing; however they clearly represent a strong limitation in the context of other more realistic forward models requiring more time to compute a set of predicted data, as well as for problems involving greater numbers of parameters (e.g. larger domains, 3-D instead of 2-D).

Table 7. Overall MCMC acceptance rate and the number of iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian posterior distribution using gradual-deformation resampling. Results are shown for different data and model characteristics as described in Table 1 and Fig. 3, respectively. All values in the table correspond to resampling performed using tuning parameter t values that approximately maximized algorithmic efficiency.

Scenario	t	Acceptance rate (per cent)	Number of iterations
Longer-correlation-length example			
1	0.0026	24	916 060
2	0.0041	25	219 330
3	0.0057	25	111 610
4	0.0050	25	145 790
5	0.0080	24	51 700
6	0.0109	25	22 830
Shorter-correlation-length example			
1	0.0037	25	335 170
2	0.0059	24	96 530
3	0.0085	23	42 000
4	0.0069	25	61 710
5	0.0114	24	20 950
6	0.0157	25	10 050

Our results also indicate that simple metrics, for example the overall proposal acceptance rate, are generally not good indicators of MCMC performance in that they can easily mislead the user into thinking that the algorithm is running effectively when it is not. Indeed, examination of the cell-by-cell update rate in our test inversions showed that poor MCMC performance resulting from an unequal updating of the model cells due to spatial variations in data sensitivity may be masked by a reasonable overall proposal acceptance rate. In addition, algorithmic efficiency, as judged here by the number of iterations necessary for 90 per cent of the model cells to represent independent samples, was found to be highly dependent upon the nature of the resampling, the degree of spatial correlation between model parameters, the level of data error, and the number of data considered. Unfortunately, no clear rules appear to exist in order to determine the parameters that will result in an optimal MCMC implementation for a given type of sampling and data/model characteristics, without actually running the algorithm under a variety of different configurations.

In light of the findings in this paper, our general recommendations are the following. With regard to the use of SGR as a proposal mechanism, it appears clear that a 'global' model perturbation approach involving all parameters at once, such as the gradual-deformation resampling method developed in Section 4, will offer improved MCMC performance over localized perturbation approaches such as randomly selected points and block resampling. This will be especially the case in the presence of significant spatial correlation between model parameters, and when strong differences in data sensitivity throughout the model domain require the user to have some degree of control over the magnitude of the perturbed parameter values. In this regard, it is important to note that only a basic gradual-deformation proposal methodology was explored in this paper; the approach indeed offers the flexibility to locally control the strength of the model perturbations through the use of spatially variable t values. Whatever resampling method is considered, however, further research is clearly necessary to improve MCMC performance if the SGR method is to be considered a generally suitable tool for a wide range of near-surface geophysical problems. Although notable computational improvements are obtained with this approach by incorporating as much information as possible into the prior distribution, one still must deal with the 'curse of dimensionality' inherent to pixel-based model parameter representations. To this end, many alternative strategies may be considered over the admittedly simple extended Metropolis approach considered here and in previous efforts. For example, recent advancements in MCMC (e.g. ter Braak & Vrugt 2008; Vrugt *et al.* 2009; Laloy & Vrugt 2012) allow for consideration of much greater numbers of parameters and the use of parallel computing. Another possibility is to consider hybrid-MCMC approaches, such as the one proposed by Sambridge (2013), which involve the use

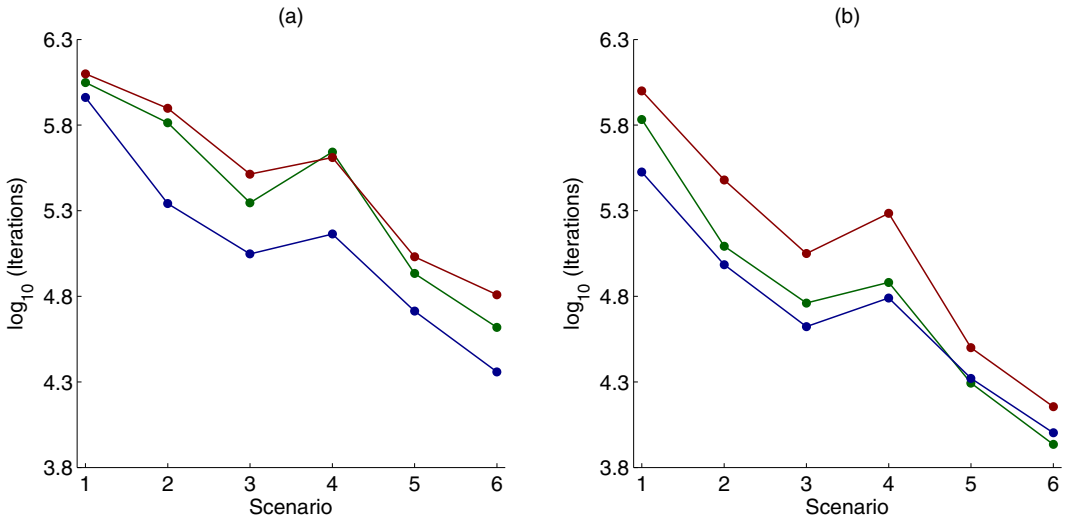

Figure 8. Number of MCMC iterations required for 90 per cent of the model parameters to represent independent samples from the Bayesian posterior distribution for the different data scenarios described in Table 1. All resampling was performed using the number of cells that approximately maximized algorithmic efficiency. Shown are randomly selected points resampling (green), block resampling (red) and gradual-deformation resampling (blue) for the (a) longer- and (b) shorter-correlation-length examples.

of the parallel tempering algorithm together with MCMC sampling techniques. Finally, we recommend that further work be done on how to better incorporate prior information and condition to hard and soft measurements within a reduced model parameterization approach. As noted earlier, such methods have the strong advantage of greatly reducing the dimensionality (and thus complexity) of the inverse problem, and it may be possible to develop approaches for effectively reducing prior bias related to the use of correlated basis functions.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Swiss National Science Foundation.

REFERENCES

Alumbaugh, D.L. & Newman, G.A., 2000. Image appraisal for 2-d and 3-d electromagnetic inversion, *Geophysics,* **65**(5), 1455–1467.

Bosch, M., 1999. Lithologic tomography: from plural geophysical data to lithology estimation, *J. geophys. Res.,* **104**(B1), 749–766.

Bosch, M., Meza, R., Jiménez, R. & Hönig, A., 2006. Joint gravity and magnetic inversion in 3D using Monte Carlo methods, *Geophysics,* **71**(4), G153–G156.

Buland, A. & Kolbjørnsen, O., 2012. Bayesian inversion of CSEM and magnetotelluric data, *Geophysics,* **77**(1), E33–E42.

Caers, J., 2007. Comparing the gradual deformation with probability perturbation method for solving inverse problems, *Math. Geol.,* **39**(1), 27–52.

Caers, J. & Zhang, T., 2004. *Multiple-point Geostatistics: A Quantitative Vehicle for Integrating Geologic Analogs into Multiple Reservoir Models,* AAPG Special Volumes.

Cordua, K.S., Hansen, T.M. & Mosegaard, K., 2012. Monte Carlo full-waveform inversion of crosshole GPR data using multiple-point geostatistical a priori information, *Geophysics,* **77**(2), H19–H31.

Davis, K. & Li, Y., 2011. Fast solution of geophysical inversion using adaptive mesh, space-filling curves and wavelet compression, *Geophys. J. Int.,* **185**(1), 157–166.

Deutsch, C., 2002. *Geostatistical Reservoir Modelling,* Oxford University Press.

Deutsch, C. & Journel, A., 1992. *Gslib: Geostatistical Software Library and User's Guide,* pp. 340, Oxford University Press.

Fu, J. & Gómez-Hernández, J.J., 2009. A blocking Markov chain Monte Carlo method for inverse stochastic hydrogeological modeling, *Math. Geosci.,* **41**(2), 105–128.

Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B., 2003. *Bayesian Data Analysis,* CRC Press.

Gilks, W.R., Richardson, S. & Spiegelhalter, D.J., 1996. *Markov Chain Monte Carlo in Practice,* Vol. 2, CRC Press.

Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation,* Oxford University Press.

Hansen, T.M., Cordua, K.S. & Mosegaard, K., 2012. Inverse problems with non-trivial priors: efficient solution through sequential Gibbs sampling, *Comput. Geosci.,* **16,** 593–611.

Hansen, T.M. & Mosegaard, K., 2008. Visim: sequential simulation for linear inverse problems, *Comput. Geosci.,* **34**(1), 53–76.

Hansen, T.M., Mosegaard, K. & Cordua, K.S., 2008. Using geostatistics to describe complex a priori information for inverse problems, in *8th International Geostatistics Congress,* Santiago, Chile, pp. 329–338.

Hassan, A.E., Bekhit, H.M. & Chapman, J.B., 2009. Using Markov chain Monte Carlo to quantify parameter uncertainty and its effect on predictions of a groundwater flow model, *Environ. Modelling Softw.,* **24**(6), 749–763.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications, *Biometrika,* **57**(1), 97–109.

Hu, L.Y., 2000. Gradual deformation and iterative calibration of Gaussian-related stochastic models, *Math. Geol.,* **32**(1), 87–108.

Hu, L.Y., 2002. Combination of dependent realizations within the gradual deformation method, *Math. Geol.,* **34**(8), 953–963.

Hu, L.Y., Blanc, G. & Noetinger, B., 2001. Gradual deformation and iterative calibration of sequential stochastic simulations, *Math. Geol.,* **33**(4), 475–489.

Huysmans, M. & Dassargues, A., 2009. Application of multiple-point geostatistics on modelling groundwater flow and transport in a cross-bedded aquifer (Belgium), *Hydrogeology Journal,* **17**(8), 1901–1911.

Irving, J. & Singha, K., 2010. Stochastic inversion of tracer test and electrical geophysical data to estimate hydraulic conductivities, *Water Resour. Res.,* **46,** W11514, doi:10.1029/2009WR008340.

Jafarpour, B., 2011. Wavelet reconstruction of geologic facies from nonlinear dynamic flow measurements, *IEEE Trans. Geosci. Remote Sens.,* **49**(5), 1520–1535.

Jafarpour, B., Goyal, V.K., McLaughlin, D.B. & Freeman, W.T., 2009. Transform-domain sparsity regularization for inverse problems in geosciences, *Geophysics,* **74**(5), R69–R83.

Jafarpour, B., Goyal, V.K., McLaughlin, D.B. & Freeman, W.T., 2010. Compressed history matching: exploiting transform-domain sparsity for regularization of nonlinear dynamic data integration problems, *Math. Geosci.,* **42**(1), 1–27.

Kaipio, J. & Somersalo, E., 2005. *Statistical and Computational Inverse Problems,* Vol. 160, Springer.

Laloy, E. & Vrugt, J.A., 2012. High-dimensional posterior exploration of hydrologic models using multiple-try dream (zs) and high-performance computing, *Water Resour. Res.,* **48**(1), doi:10.1029/2011WR010608.

Linde, N. & Vrugt, J.A., 2013. Distributed soil moisture from crosshole ground-penetrating radar travel times using stochastic inversion, *Vadose Zone Journal,* **12**(1), doi:10.2136/vzj2012.0101.

Mariethoz, G., Renard, P. & Caers, J., 2010. Bayesian inverse problem and optimization with iterative spatial resampling, *Water Resour. Res.,* **46**(11), doi:10.1029/2010WR009274.

Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory,* Academic Press.

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.

Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.,* **100**(B7), 12 431–12 447.

Oware, E., Moysey, S. & Khan, T., 2013. Physically based regularization of hydrogeophysical inverse problems for improved imaging of process-driven systems, *Water Resour. Res.,* **49**(10), 6238–6247.

Ramirez, A.L. *et al.*, 2005. Stochastic inversion of electrical resistivity changes using a Markov chain Monte Carlo approach, *J. geophys. Res.,* **110**(B2), B02101, doi:10.1029/2004JB003449.

Ravelec-Dupin, M.L., Noetinger, B. & Hu, L.Y., 2002. The FFT moving average (FFT-MA) generator: an efficient numerical method for generating and conditioning Gaussian simulations, *Math.Geol.,* **32**(6), 701–723.

Roggero, F. & Hu, L., 1998. Gradual deformation of continuous geostatistical models for history matching, in *SPE Annual Technical Conference and Exhibition,* New Orleans, LA.

Sambridge, M., 2013. A parallel tempering algorithm for probabilistic sampling and multimodal optimization, *Geophys. J. Int.,* **196,** 357–374.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.,* **40**(3), 3-1–3-29.

Scholer, M., Irving, J., Looms, M.C., Nielsen, L. & Holliger, K., 2012. Bayesian Markov-chain-Monte-Carlo inversion of time-lapse crosshole GPR data to characterize the vadose zone at the Arrenaes site, Denmark, *Vadose Zone Journal,* **11**(4), doi:10.2136/vzj2011.0153.

Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics, *Math. Geol.,* **34**(1), 1–21.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation,* SIAM.

Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.,* **20**(2), 219–232.

ter Braak, C.J. & Vrugt, J.A., 2008. Differential evolution Markov chain with snooker updater and fewer chains, *Stat. Comput.,* **18**(4), 435–446.

Trainor-Guitton, W. & Hoversten, G.M., 2011. Stochastic inversion for electromagnetic geophysics: practical challenges and improving convergence efficiency, *Geophysics,* **76**(6), F373–F386.

Vrugt, J.A., Ter Braak, C., Diks, C., Robinson, B.A., Hyman, J.M. & Higdon, D., 2009. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. Nonlin. Sci. Numer. Simul.,* **10**(3), 273–290.