View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Online Research Database In Technology

Technical University of Denmark

Reducing complexity of inverse problems using geostatistical priors

Hansen, Thomas Mejer; Mosegaard, Klaus; Cordua, Knud Skou

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Hansen, T. M., Mosegaard, K., & Cordua, K. S. (2009). Reducing complexity of inverse problems using geostatistical priors. Paper presented at International association of Mathematical geoscience (IAMG 09), Stanford, CA, United States.

DTU Library
Technical Information Center of Denmark

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Reducing complexity of inverse problems using geostatistical priors *

Thomas Mejer Hansen Klaus Mosegaard Knud Skov Cordua

October 2, 2009

Abstract

In a probabilistic formulation of inverse problems the solution can be given as a sample of the posterior probability distribution. All realizations retained in the posterior sample are consistent with both an assumed prior model and observed data. Some inverse problems are unsolvable, in that one can practically never hope to generate a posterior sample, others are just 'difficult' and require special methods to become tractable, while others again are easily solved. We discuss how difficult nonlinear inverse problems can be handled such that their complexity, i.e. the time taken to obtain a posterior sample, can be reduced significantly using informed priors based on geostatistical models. We discuss two approaches to include such geostatistically based prior information. One is based on a parametric description of the prior likelihood that applies to 2-point based statistical models, and another approach makes use of conditional re-simulation to sample the prior that works for both 2-point and multiple point random models. The latter approach is shown to be superior in terms of computational efficiency. We quantify the information content given by a specific choice of prior model. This enables us to obtain a lower limit of, for example, the size of a grid cell in a grid-parametrized parameter space. The resulting decrease in effective dimension of the parameter space provides a much more efficient sampling of the posterior with orders of magnitude increase in computational efficiency.

1 Introduction

Mosegaard and Tarantola (1995) present a generalized Metropolis algorithm that allows analysis of non-linear inverse problems with complex prior information. Yet, to date most applications of the generalized Metropolis algorithm, or solutions of non-linear inverse problems in general, rely on relatively simple a priori assumptions. The generalized Metropolis algorithm does not require one to explicitly know the formula describing the prior information. A black box that can sample according to the prior pdf is the only requirement (Mosegaard and Tarantola 1995). Geostatistical simulation algorithms provide just such tools for generating random realizations from very complex stochastic prior models. They can rely on both traditional 2-point covariance-based statistics and multiple point statistics where the prior model is inferred from a training image

*Proceeding from IAMG 09, August 23-28, 2009, Stanford, CA

(Guardiano and Srivastava 1993; Strebelle 2002). Recent developments of multiple-point-based simulation allow one to quantify geologically realistic prior information. Here we shall discuss how the choice of a priori models, based on geostatistical random function models, affect the complexity of an inverse problem. In particular, we shall see how the computational efficiency is affected by a specific choice of prior model and the way it is implemented for use with the generalized Metropolis algorithm.

2 The generalized Metropolis algorithm

Consider a typical forward problem, where data d is a function g of some model m (typically the subsurface)

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

An inverse problem is the problem of inferring properties of m, based on observations d and some knowledge about the mapping function g, typically related to physical theory. Tarantola (2005) and Mosegaard (2006) formulate a probabilistic approach to solving inverse problems where prior information is described by the prior probability density function (pdf) $\rho_M(m)$. A probabilistic measure of the data fit associated to a given model is given by the likelihood $L_m(m)$. The solution to such an inverse problem is a probability density function, denoted the a posteriori pdf, and is proportional to the product of the prior pdf and the likelihood (where k is a normalization factor):

$$\sigma_M(m) = k\, \rho_M(m)\, L_M(m) \tag{2}$$

In case g is a linear function, and both $\rho_M(m)$ and $L_M(m)$ can be described by Gaussian statistics, Hansen et al. (2006) and Hansen and Mosegaard (2008) propose a non-iterative, efficient approach using sequential simulation to generate samples of the a posteriori pdf. It is, however, more common that g is a nonlinear operator, and the Gaussian prior assumptions about $L(m)$ and $\rho_M(m)$ is rather restrictive. Mosegaard and Tarantola (1995) suggest a generalized Metropolis Monte Carlo algorithm for sampling the a posterior distribution $\sigma_M(m)$ in the general case where g is nonlinear. The method allows inclusion of complex prior information and arbitrary, complex noise model.

Consider a Markov chain where m_n is a realization of the prior pdf $\rho_M(m)$, and m_{n+1} is in the neighborhood of m_n, but still a realization of $\rho_M(m)$. Further, assume that the likelihood with respect to observed data can be calculated as L(m_n) and L(m_{n+1}), respectively. Then m_{n+1} is accepted as a realization of the a posteriori pdf with probability

$$P_{accept} = \begin{cases} 1 & \text{if L}(m_{n+1})>\text{L}(m_n) \\ \text{L}(m_{n+1}) \,/\, \text{L}(m_n) & \text{otherwise} \end{cases} \tag{3}$$

If m_{n+1} is rejected, m_{n+1} becomes m_n (that is, m_n is repeated). Performed iteratively, this algorithm will sample the a posteriori pdf. In each iteration, one needs to perturb the current model consistently with prior information, compute the likelihood of the perturbed model, and finally generate a random number between 0 and 1 to decide whether the perturbed model is to be accepted. Here we shall focus on, how geologically realistic prior information can be considered for use the generalized Metropolis algorithm, and the effect it has on the computational complexity of the inverse problem.

3 Quantifying prior information using geostatistics

Generally speaking geostatistics is an application of random functions to describe spatial phenomena, typically in form of spatial variability in earth models. Geostatistical simulation algorithms have been developed to efficiently generate realizations of a number of random function models. Geostatistical simulation algorithms can be divided into two groups where the underlying random function model is based on 2-point or multiple-point statistics (Guardiano and Srivastava 1993; Strebelle 2002). 2-point-based geostatistical algorithms take into account spatial variability between pairs of data locations. In case the distribution of the model parameters are Gaussian, one can completely define the underlying random function model using a Gaussian pdf

$$\rho_M(m) \ = c \ exp\left[-\frac{1}{2}(m-m_{prior})^t C_{M_{prior}}^{-1}(m-m_{prior})\right] \tag{4}$$

where m_{prior} is the prior mean, and $C_{M_{prior}}$ is the prior covariance matrix. Multiple-point-based geostatistical models have no parametric description. Instead the multiple point statistics are inferred from a training image. The methodology was initially proposed by Guardiano and Srivastava (1993), and Strebelle (2002) developed the first computationally feasible algorithm for categorical training images. Zhang et. al (2006) suggested another multiple-point-based algorithm where patterns from a continuous or categorical training image are used to generate stochastic realizations with features from the training image. Using these techniques one can generate realizations of random function models that reproduce geologically realistic spatial variability. State of the art implementation of these algorithms are available through for example SGeMS (Remy et al. 2008).

Sampling a prior using the Metroplis algorithm When a parametric description of the random function model is available, as is the case for a Gaussian covariance-based priors (eqn. 4), a separate Metropolis sampler can be used to sample the prior information, (Bosch et al. 2005; Bosch et al. 2006; Jiménez and Bosch 2008).

Sampling a prior using sequential re-simulation An alternative approach to sample geostatistics-based prior information is proposed by Hansen et al. (2008). They suggest to sequentially re-simulate part of the model parameters, conditioned to the remaining, fixed model parameters:

1. In the current model m_i, select a region in the model space, and denote all model parameters in this area as unknown, m_u. The rest of the model parameters are considered known m_k.

2. Perform sequential simulation of m_u, conditioned to m_k. This generates m_{i+1}, which is also a realization of the prior model.

3. Set m_i=m_{i+1} and go to 1.

When this algorithm is run iteratively it will perform a random walk in the model parameter space, guided by the underlying, prior random function given by the geostatistical sampling algorithm.

Sequential re-simulaton can be used with any geostatistical algorithm (both 2-point and multiple-point based) capable of performing conditional simulation. Such algorithms include SGSIM, DSSIM (Soares 2001), SNESIM (Strebelle 2002), FILTERSIM (Wu et al. 2008) but to name a few.

3.1 Comparing sequential resimulation and the Metroplis algorithm for sampling geostatistical-based priors.

At each step in the generalized Metropolis algorithm, one must choose to randomly visit a new model in the vicinity of the current model. The size of the area of 'vicinity' reflects the exploratory nature of the algorithm. We shall refer to the distance or size of the 'vicinity' as the step length. Choosing the correct step length is essential for the computational efficiency of the sampling algorithm. If the step length is small, the acceptance rate will be high, but only model parameters relatively close to the initial model will be considered. If the step length is high the algorithm will be more exploratory but the acceptance probability will be small. Gelman et al. (1996) found that an acceptance rate between 20-40% leads to an algorithm with both relatively high acceptance rate and relative high exploration. Thus, any method used to sample prior information should allow tuning of the exploratory step length in such a way that the exploration of the posterior can be optimized.

Using sequential re-simulation to sample the prior, the step length can be chosen between two extremes. It is least exploratory when only one model parameter is considered for re-sampling per iteration, and most exploratory when all model parameters are resampled in each iteration. The latter approach is equivalent of generating a new uncorrelated sample of the prior. Thus we can directly control the exploratory nature of our prior sampling algorithm.

When using the Metropolis sampler to sample prior information, we need to select an appropriate degree of exploration for both the prior Metropolis sampler and the a posteriori Metropolis sampler. If the step length needed for the prior Metropolis sampler is smaller than the step length leading to a optimal posterior sampler, we suggest to run the prior Metropolis sampler in cascade to the posterior Metropolis sampler, as suggested by Mosegaard and Tarantola (1995). This will reduce the number of evaluations of the likelihood, which is typically an expensive calculation.

To evaluate the computational efficiency when using the Metropolis algorithm for sampling the prior, we consider the computational requirements needed to achieve a step of a certain length / amount of exploration. We quantify the step length as the correlation coefficient between and initial model and the corresponding perturbed model. As a reference model we consider a 2D model of 100x100 grid points, with a point distance of 5 meters. We consider two types of covariance models as priors, a) an exponential covariance model and b) a Gaussian covariance. We consider ranges from 0.5m to 20m. For any given set of covariance model and range, we start a Metropolis sampler from an initial unconditional realization, using a uniform proposal distribution, and tune the algorithm to ensure an acceptance rate of about 30%. The Metropolis algorithm is run for 300000 iteration and we locate the iteration step at which a correlation coefficient between the initial model and the current model is 0.0, 0.1, 0.5 and 0.9 respectively, reflecting different exploration levels. We do this 100 times, for 100 different starting models, to obtain the average iteration number needed to obtain a certain correlation coefficient between the initial model and the perturbed model. The results are summarized in tables 1 and 2, where '∞' means that no model with the given correlation coefficient was found in 300000 iterations.

The main result of tables 1-2 is that as the range increases with respect to the spatial sampling distance, so does the number of iterations needed. The more spatial correlation in the prior, the harder it is to sample it using a Metropolis sample.

A Gaussian covariance model imposes harder constraints on small scale variability than the exponential covariance model, since the first order derivative of a Gaussian covariance model is zero

Range	0.5m	1m	2m	5m	10m	20m
CC=0.9	$8.5 10^{1}$	$8.5 10^{1}$	$8.8.2 10^{1}$	$1.1 10^{2}$	$2.1 10^{2}$	$5.2 10^{2}$
CC=0.5	$6.0 10^{2}$	$5.9 10^{2}$	$5.7 10^{2}$	$8.6 10^{2}$	$3.1 10^{3}$	$1.4 10^{4}$
CC=0.1	$1.6 10^{3}$	$1.7 10^{3}$	$1.8 10^{3}$	$3.9 10^{3}$	$2.2 10^{4}$	$8.6 10^{4}$
CC=0.0	$2.6 10^{3}$	$2.8 10^{3}$	$2.4 10^{3}$	$7.7 10^{3}$	$3.1 10^{4}$	$1.3 10^{5}$

Table 1: Number of iterations to reach a correlation coefficient of 0.9, 0.5, 0.1 and 0.0 between the initial model and the perturbed model for an exponential covariance model with different ranges. 100x100 model parameters describe a 500m x 500m physical space.

Range	0.5m	1m	2m	5m	10m	20m
CC=0.9	$8.5 10^{1}$	$8.5 10^{1}$	$8.8 10^{1}$	$1.4 10^{2}$	$3.5 10^{5}$	∞
CC=0.5	$6.0 10^{2}$	$6.0 10^{2}$	$5.8 10^{2}$	$1.2 10^{3}$	∞	∞
CC=0.1	$1.6 10^{3}$	$1.6 10^{3}$	$1.8 10^{3}$	$5.4 10^{3}$	∞	∞
CC=0.0	$2.6 10^{3}$	$2.8 10^{3}$	$2.5 10^{3}$	$8.0 10^{3}$	∞	∞

Table 2: Number of iterations to reach a correlation coefficient of 0.9, 0.5, 0.1 and 0.0 between the initial model and the perturbed model for a Gaussian covariance model with range. 100x100 model parameters spanning 500x500m.

at zero offset, and non-zero for the exponential covariance model. This means that a realization of a Gaussian covariance model will be very smooth. Such smoothing is in direct contradiction to the models that are proposed by the uniform proposal distribution of the Metropolis algorithm: using a uniform prior to propose models, almost none of the models proposed will be smooth. In fact, the probability of suggesting a smooth model, using a uniform prior, tends to zero as the spatial sampling distance tends to zero, or the range becomes long with respect to this distance. Increasing the range therefore implies that the step length must be reduced in order to maintain an acceptance ratio of about 30%. The longer the range (or the smaller the sampling distance) the smaller the step length and hence the more iterations are needed to obtain a certain level of exploration.

This is the reason that the Metropolis sampler, using a uniform prior, performs increasingly worse as the sampling distance is reduced. Also, due to the smoothness assumption inherent in sampling from a Gaussian covariance model, it is more difficult to sample from a Gaussian covariance than from an exponential covariance model.

This simple example illustrates that using the Metropolis sampler with a uniform prior to sample Gaussian-based prior information leads to a computationally inefficient algorithm for anything but very low dimensional problems. In addition, the stronger the spatial correlation the more CPU expensive it becomes to use the Metropolis algorithm to sample Gaussian models.

In contrast, using sequential re-simulation, the exploratory nature of the prior sampler can easily be adjusted to the sampling problem at hand. The 'step length' is simply chosen as the number of model parameters needed to obtain a given acceptance ratio for the generalized Metropolis sampler. No models will be suggested that are in conflict with prior information. In addition, the sequential re-simulation sampler work with any conditional sequential simulation program, both 2-p and multiple point based, where as the Metropolis approach to sample the prior is limited to priors with a parametric description of the random function, such as covariance based 2-p pri-

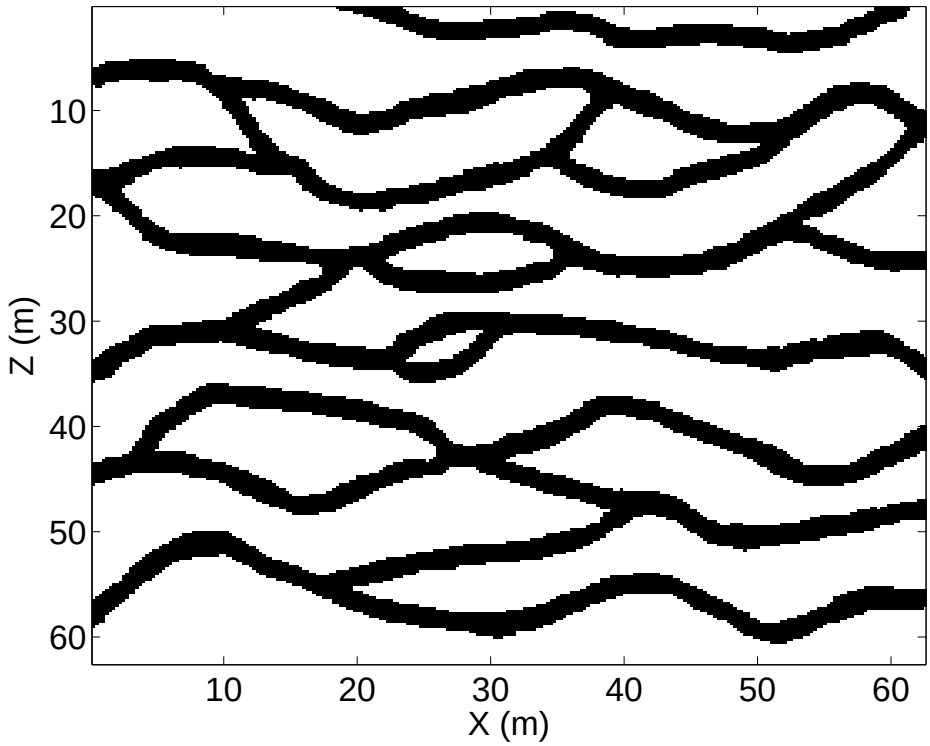


Figure 1: Training image used to generate the reference model. Black channel structures have a velocity of 0.09 m/ns. The background velocity (white) has a velocity of 0.13 m/ns.

ors. The sequential re-simulation sampler provides an efficient and non-biased approach to sample complex prior information, that is well designed for use with the generalized Metropolis algorithm.

4 Synthetic case study: Application of non-linear inverse problems with complex prior information

Figure 1 is a channel-based training image from which we generate an unconditional realization, Figure 2, using the single normal equation simulation algorithm, SNESIM (Strebelle 2002). This will be our reference velocity model for a synthetic cross borehole inversion problem. Travel times are computed, traveling from the 20 sources located to the left in Figure 2, to 40 receivers located to the right. 3% Gaussian noise is added to the synthetic travel-time delay data, Figure 3, and used as observed data. We now consider solving this inverse problem, using the generalized Metropolis algorithm to generate samples of the a posteriori probability distribution, given the observed data, the assumed noise model, and an assumed prior model.

We consider a number of prior models based on both 2-point and multiple-point based random models. All prior models are assumed to have the correct mean and variance, as obtained from the training image in Figure 1. The first 6 prior models are based on 2-point random models. The pure nugget model assumes no spatial correlation, and thus all model parameters are a-priori considered uncorrelated. The 'Gau(1)', 'Gau(3)' and 'Gau(8)' prior models are based on a Gaussian covariance model with an isotropic range of 1m, 3m and 8m respectively. The SGSIM prior is

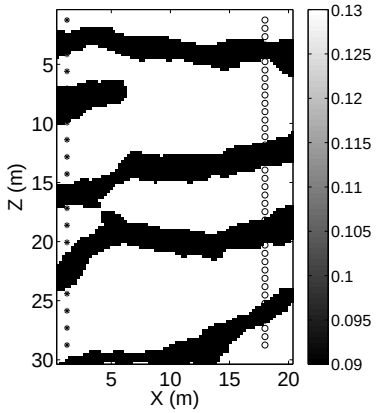


Figure 2: Reference velocity model and location of sources (*) and receivers (o)

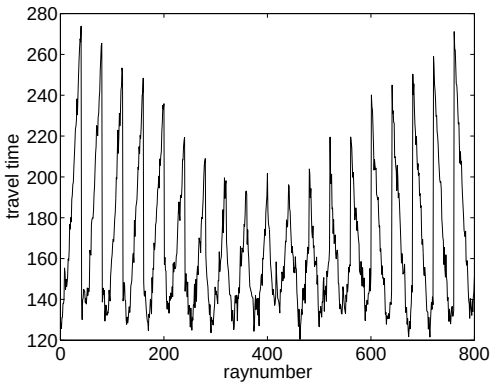


Figure 3: Calculated first arrival travel time using the recording geometry shown in Figure 2. 3% normally distributed noise was added to the travel times.

based on a covariance model, inferred from the training image in Figure 1, as an exponential covariance model with a horizontal range of 6.6m a vertical range of 2.2m. The DSSIM prior is as the SGSIM prior, except that the correct distribution from the training image is used. Finally, the true TI prior is the training image in Figure 1, and the last prior use the training image rotated 90 degrees clockwise. The 4 prior models, 'Pure Nugget', 'SGSIM', 'DSSIM' and 'True TI' reflects 4 prior models with increasing order of statistics consistent with the training image from which the reference model was generated.

Figure 5 shows the initial models for the 8 considered prior random models. For each of these prior models a generalized Metropolis algorithm was run for 35000 iterations. Figure 4 shows the negative log-likelihood of all models accepted by the Metropolis algorithm for 35000 iterations. Figure 6-7 shows the current model at iteration 20000, 25000, 30000, and 35000.

The initial phase of running the generalized Metropolis algorithm is a search for a location in the model parameter space where the forward responses of the models fit data within their uncertainty. This is the 'burn-in' phase. One can locate the end of the burn-in process from Figure 4 as the iteration number where the log-likelihood curve flattens out around a level of approximately $-N/2 = -400$, where N is the number of data. When the burn-in phase is over the algorithm starts sampling the a posteriori probability distribution.

As can be seen from Figure 4 it is only when using the pure Nugget prior, and possible the 'Gau(1)' prior, that we do not get past the burn-in phase.

As consistent prior information is increased we complete the burn-in phase faster. Thus, for the 'Pure Nugget', 'SGSIM', 'DSSIM' and 'True TI' prior, the burn-in is completed at ∞, 4000, 3800, 1000 iterations respectively, indicating that the computational complexity is reduced as consistent information content is increased. Also, even if the Gaussian priors 'Gau(1)','Gau(3)', and 'Gau(8)', do not reflect true subsurface variability, they perform order of magnitudes better than using the uninformed 'Pure Nugget' prior.

This simple example suggests that an otherwise relatively easy inverse problem, such as the inversion of first arrival time data, becomes virtually unsolvable using an uninformed 'Pure Nugget' prior.

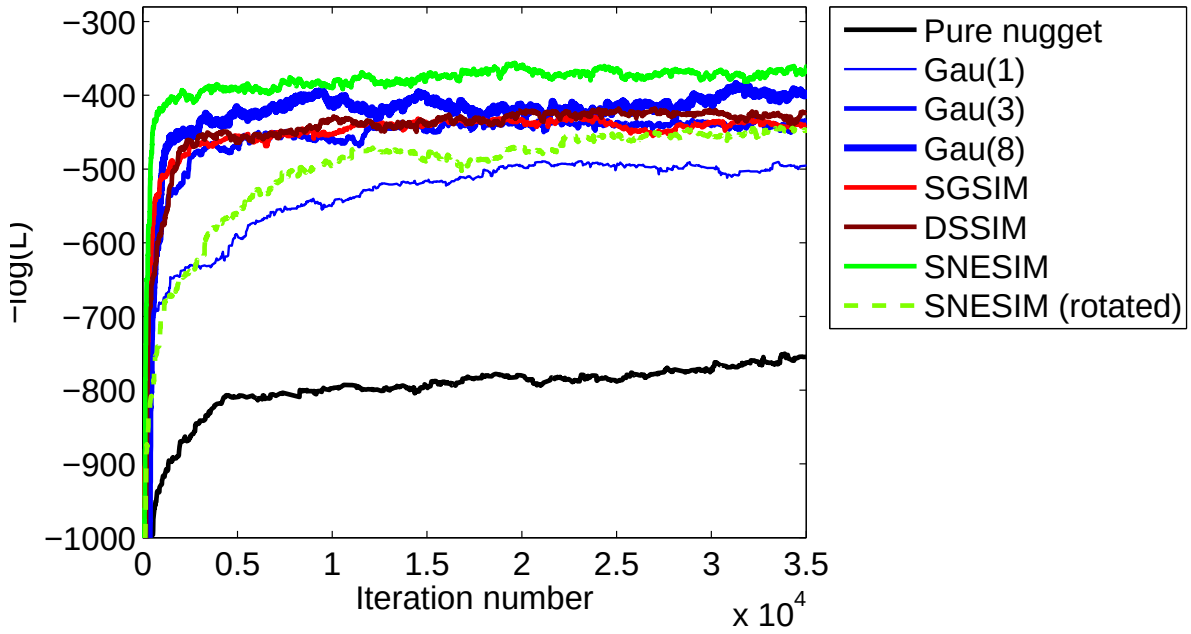


Figure 4: Negative log-likelihood as a function of iteration number for different choices of prior model

5 A measure of information content

Realizations of a Gaussian random function can be generated from an eigenvalue decomposition of the covariance function, C, using the Karhunen-Loeve expansion

$$y = E\Lambda^{1/2} z \tag{5}$$

where E is the matrix of eigenvectors of the covariance matrix C, Λ is a diagonal matrix of corresponding eigenvalues, and z is a series of random Gaussian numbers with mean 0 and variance 1 (Sarma et al. 2008). One can create approximate realizations of the random function with covariance model C using only a limited number of the eigenvectors with the highest eigenvalues. When most eigenvalues of Λ are close to zero the approximation tends to be close to perfect. The accuracy of using a limited set of eigenvalues and corresponding eigenvectors can be quantified using an energy spectrum. An energy spectrum of the eigenvalues plots the cumulative sum of the eigenvalues (normalized to 1 as maximum value) as a function of the number of eigenvalues considered, sorted in decreasing order. From the energy spectrum one can thus find how many of the eigenvalues need to be considered in eqn. 5 to retain a certain level of energy (Sarma et al. 2008), which indicates that the cumulative energy spectrum can be used as an indicator of the information content in any given choice of Gaussian based prior models.

As an example we consider different spatial sample sizes for a model of size 10x10 meter. A grid spacing of 1m results in 10*10=100 model parameters. A grid spacing of 0.25*0.25, results in 40*40=1600 model parameters. We consider prior models with a range from 0 meters to 20 meters, ranging from a prior model assuming no spatial connectivity to a prior model assuming very strong spatial connectivity. Figure 8a shows the percentage of all available eigenvectors needed to represent 95% of the energy content for each given prior. Figure 8b shows the number of eigenvectors

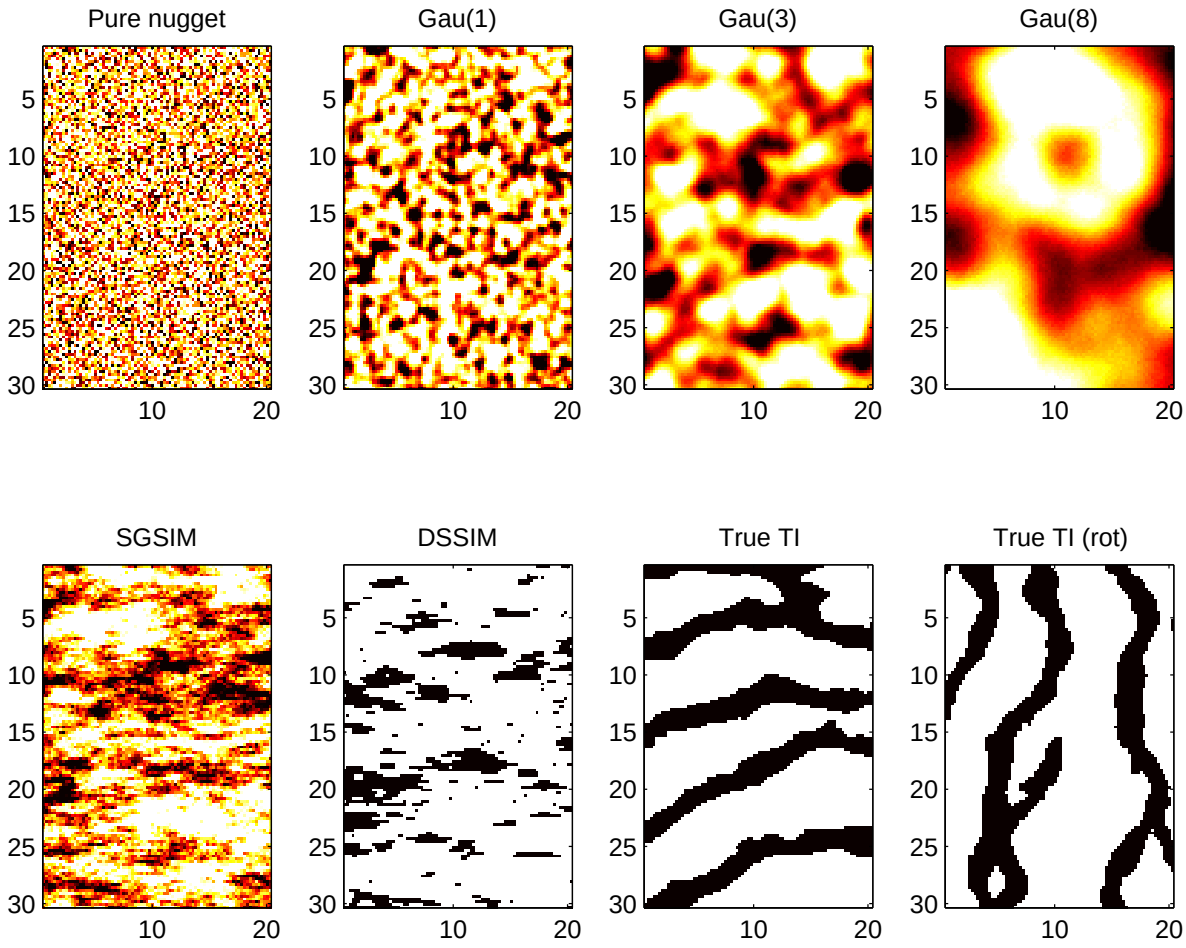


Figure 5: Initial (prior) model

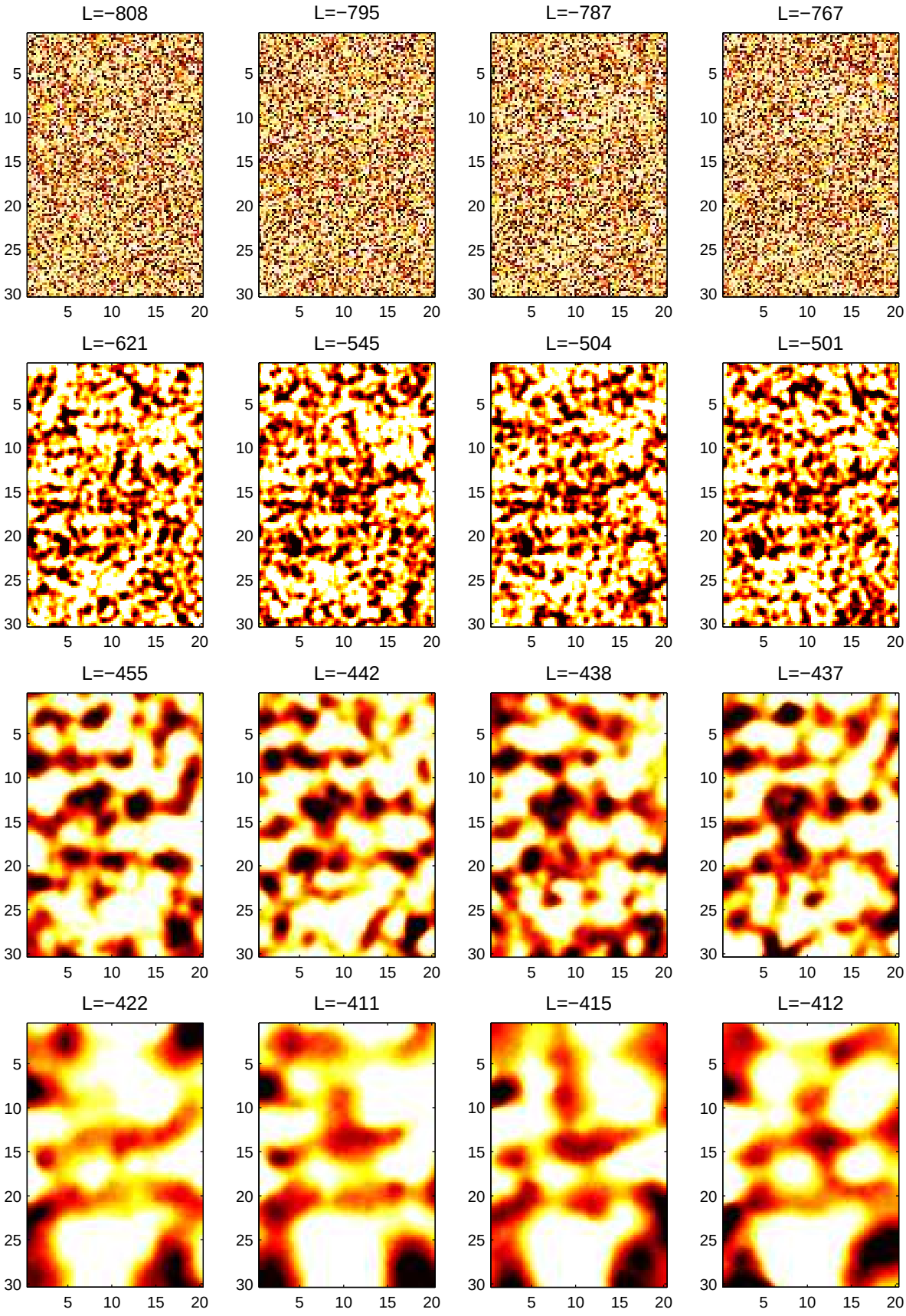


Figure 6: Current model at iteration number 20000, 25000, 30000 and 35000 using the 'Pure Nugget', 'Gau(1)', 'Gau(3)', and 'Gau(8)' prior models (listed from to to bottom). The title is the negative log likelihood for the given model.

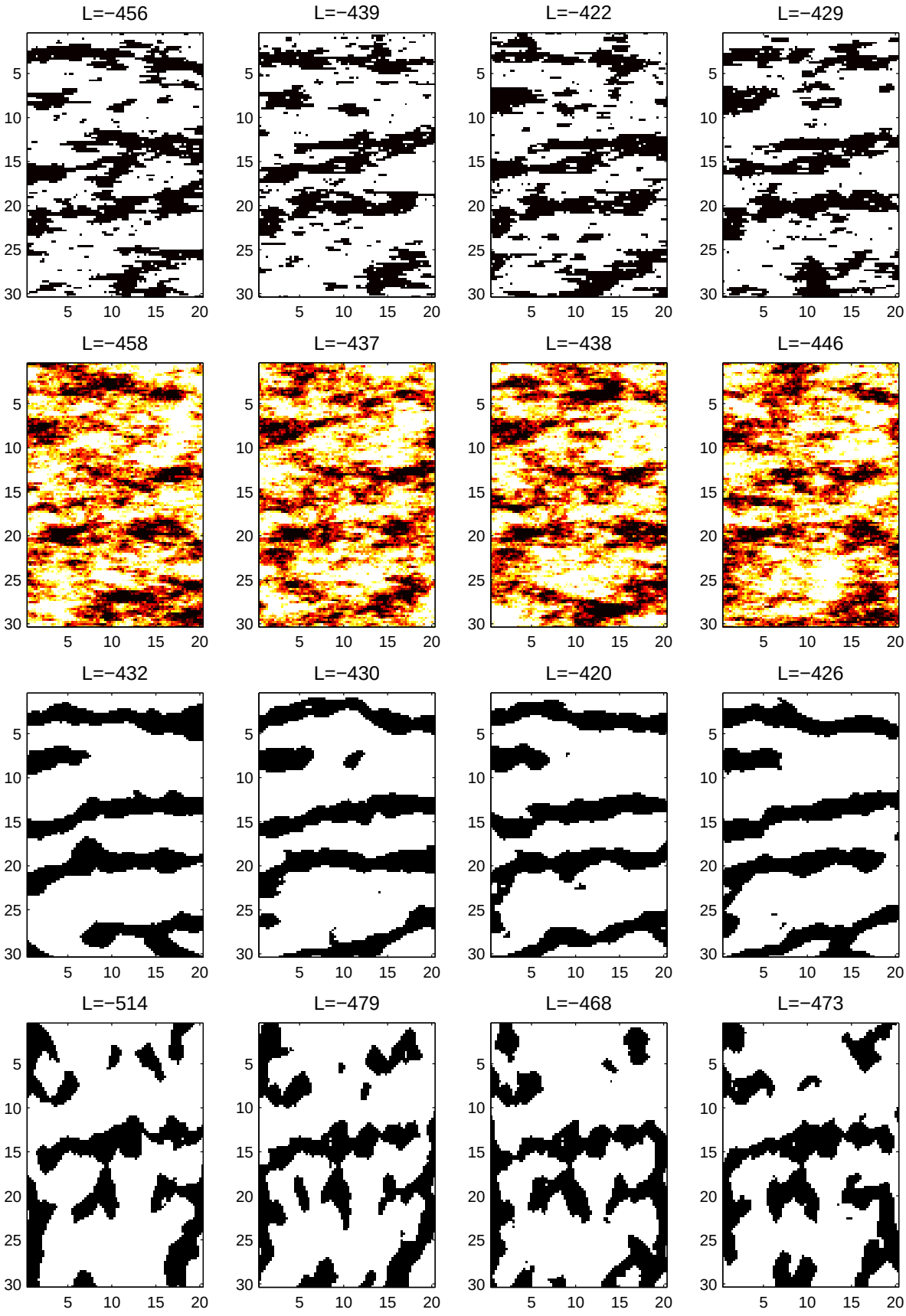


Figure 7: Current model at iteration number 20000, 25000, 30000 and 35000 using the ’SGSIM’, ’DSSIM’, ’True Ti’, and ’True TI (rot)’ prior models (listed from to to bottom). The title is the negative log likelihood for the given model.

needed to represent 95% of the energy content for each given prior.

The number of eigenvectors needed to describe 95% percent of the cumulative eigenvalue content using an uncorrelated prior, is 95% of the total number of eigenvectors (i.e. the total number of model parameters). Thus, increasing the number of model parameters by a factor of N increases the number of effectively free parameters with a factor N. The number of eigenvectors needed to describe 95% of the cumulative eigenvalue content using a highly correlated prior is about $N = 8$, independently of the considered range. Thus, increasing the number of model parameters by a factor of N does NOT increase the number of effective free parameters, which means that for a prior model with spatial correlation there exists a upper limit to the effective number of free parameters inherent in the choice of prior. The longer the range with respect to the sampling distance, the higher the apparent dimension reduction.

There results are important for several reasons. Considering a uncorrelated prior model, sometimes referred to as an uninformed prior, leads to an inverse problem that will be unsolvable except for very small problems with very few model parameters. On the other hand, if one prior assumes some spatial correlation, one can increase the number of considered model parameters dramatically without increasing the number of free parameters of the prior significantly.

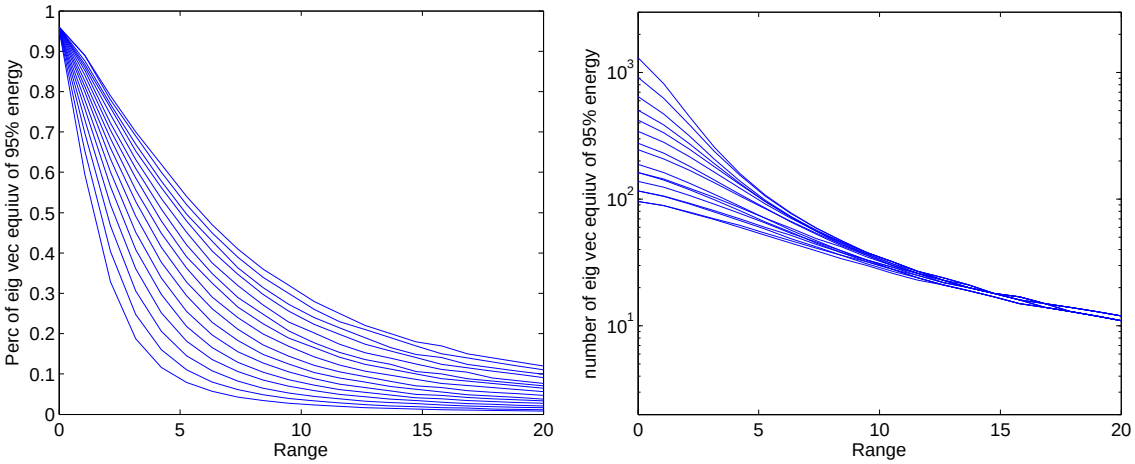
This indicate that, for a specific choice of prior model, there is an upper limit to the complexity of the inverse problem: A threshold for the spatial sampling distance dx exists, below which the complexity/hardness of the inverse problem will not increase. It may be computationally more expensive to sample the prior and compute the likelihood function for a model that is more densely sampled, but the problem of sampling the posterior probability density function will be equally hard. In other words, for a given choice of prior there is a lower limit for the spatial sampling distance below which the complexity of the inverse problem becomes constant, and independent of the number of model parameters.

The analysis presented here is only valid for Gaussian-based random models (2-point statistics). Kernel PCA might be a viable approach to quantifying the information content of multiple-point based priors (Sarma et al. 2008).

6 Discussion

Choosing a prior model with (too) many free parameters will make the sampling problem harder. Choosing a prior model with as many free parameters as model parameters, will result in an unsolvable problem, as the number of model parameters tends to infinity. Choosing a prior model with too few free parameters will restrict the solution space, increase the computation time and perhaps make it impossible to match the data within uncertainties. Our findings suggest that a well-informed prior, i.e. a prior consistent with the true subsurface, will perform computationally more efficiently than a prior in conflict with the true subsurface. For underdetermined inverse problems, one cannot avoid choosing a prior model.

If one tries to choose a 'neutral' prior model and relies on for example a uniform prior (which in reality is a very specific prior assumption) we have shown that the inverse problem becomes unsolvable unless the number of model parameters are very few. Not only does an 'informed' prior increase the computational efficiency of sampling the prior (in that it will allow longer steps to be taken by the sampling algorithm, reducing the time between independent samples) it also has direct effect on the shape of the posterior model. If the choice of prior is 'sound' it will reduce



(a) Percentage of eigenvalues needed to reproduce 95% of the total eigen value energy. Top curve reflects dx=0.25, bottom curve reflects dx=1m

(b) Number of eigenvalues needed to reproduce 95% of the total eigen value energy. Top curve reflects dx=0.25, bottom curve reflects dx=1m

Figure 8: Information content.

the complexity of sampling the posterior, since it will be much easier to locate the area of high probability. These are two major reasons why the choice of prior model has significant impact on the type of nonlinear inverse problems that can be solved using the generalized metropolis algorithm.

7 Conclusions

The complexity of an inverse problem is hugely dependent on the prior, (not only the physics / likelihood) and the way it is introduced into the inverse problem. We considered two alternatives for quantifying 2-point-based Gaussian random models into nonlinear inverse problems where the a posteriori distribution is sampled by the generalized Metropolis algorithm. The first method is a simple utilization of the Metropolis sampler using a uniform proposal distribution. We find this approach to be applicable only to very low-dimensional inverse problems and for prior assumption with little-to-no spatial correlation. We also consider a method based on sequential re-simulation that is both easy to implement and computationally eficient. In addition, it works with any sequential simulation algorithm, both 2-point and multiple point based.

A synthetic case study, utilizing both 2-point and multiple-point-based prior models, shows the application of the sequential re-simulation for sampling the prior as part of an application of the generalized Metropolis sampler. It illustrates how quite complex priori information can be quantified by, for example, training images and used to provide realizations of the a posteriori pdf that honor both data and the complex a priori information. We find that the more consistent information is added to the inverse problem in this manner, the workload of sampling the a posteriori probability function is reduced. The specific choice of using a non-spatially correlated prior leads to a sampling problem that is impossible to solve. Even for the relatively easy inverse problem considered here (first arrival travel time inversion), a spatially uncorrelated prior results in a sampling problem that cannot be solved. On the other hand, a prior with just a small amount of spatial

correlation (much smaller than the apparent correlation length of the actual subsurface) makes the problem relatively easy.

The synthetic case study indicates that the complexity of the inverse problem is reduced as the spatial correlations become stronger. Using PCA analysis we compute the energy spectrum of an eigenvalue analysis performed for a number of different covariance models. We find that for given a specific spatial correlation length and spatial sampling distance we compute the number of eigenvectors needed to reproduce 95% of the energy in the energy spectrum. We find that using any prior model with spatial correlation, the number of eigenvectors needed to reproduce 95 % of the energy spectrum is smaller than 95% of the number of model parameters. Thus, a prior with spatial correlation results in an apparent reduction of the number of free parameters of the prior. In addition, for a given choice of correlation length there is a lower limit to the spatial sampling distance, below which the apparent number of free parameters of the prior is constant. Thus, decreasing the spatial sample size will not increase the number of free parameters, even though the number of model parameters increase.

We find that using sequential re-sampling for sampling complex prior information (as quantified by geostatistical simulation algorithms) provides an effective approach for adding complex prior information to non-linear inverse problems when solved by the generalized Metropolis algorithm. Not only does it provide geologically realistic solutions to inverse problems, it also reduces the effective dimension of the inverse problem to be solved, and hence reduces the computational requirements. The stronger the spatial prior information, the larger the dimension reduction will be.

8 Acknowledgements

We used VISIM (part of mGstat,http://mgstat.sourceforge.net) for 2-point based simulation and SNESIM, by Sebastien Strebelle, for single normal equation simulation.

References

Bosch, M., P. Barton, S. Singh, and I. Trinks (2005). Inversion of traveltime data under a statistical model for seismic velocities and layer interfaces. *Geophysics 70*, R33.

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

Gelman, A., G. Roberts, and W. Gilks (1996). Efficient metropolis jumping hules. In J. Bernardo, K. Berger, A. Dawid, and A. Smith (Eds.), *Bayesian Statistics 5,*, pp. 599–608. Clarendon press, Oxford.

Guardiano, F. and R. Srivastava (1993). Multivariate geostatistics: beyond bivariate moments. *Geostatistics-Troia 1*, 133–144.

Hansen, T. and K. Mosegaard (2008). VISIM: Sequential simulation for linear inverse problems. *Computers and Geosciences 34*(1), 53–76.

Hansen, T. M., A. G. Journel, A. Tarantola, and K. Mosegaard (2006). Linear inverse Gaussian theory and geostatistics. *Geophysics 71*(6), R101–R111.

Hansen, T. M., K. Mosegaard, and K. C. Cordua (2008). Using geostatistics to describe complex a priori information for inverse problems. In J. M. Ortiz and X. Emery (Eds.), *VIII International Geostatistics Congress*, Volume 1, Mining Engineering Department, University of Chile, pp. 329–338.

Jiménez, R. and M. Bosch (2008). Geostatistical inversion of gravity and well-log data. *Revista de la Facultad de Ingeniera de la U.C.V 23*(2), 71–81.

Mosegaard, K. (2006). *Monte Carlo Analysis of Inverse Problems*. University of Copenhagen.

Mosegaard, K. and A. Tarantola (1995). Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research 100*(B7), 12431–12447.

Remy, N., A. Boucher, and J. Wu (2008). *Applied Geostatistics with SGeMS: A User's Guide*. Cambridge Univ Pr.

Sarma, P., L. Durlofsky, and K. Aziz (2008). Kernel Principal Component Analysis for Efficient, Differentiable Parameterization of Multipoint Geostatistics. *Mathematical Geology 40*(1), 3–32.

Soares, A. (2001, November). Direct sequential simulation and cosimulation. *Math. Geol 33*(8), 911–926.

Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Math. Geol 34*(1), 1–20.

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

Wu, J., A. Boucher, and T. Zhang (2008). A SGeMS code for pattern simulation of continuous and categorical variables: FILTERSIM. *Computers and Geosciences 34*(12), 1863–1876.

Zhang, T., P. Switzer, and A. Journel (2006). Filter-based classification of training image patterns for spatial simulation. *Mathematical Geology 38*(1), 63–80.