USING GEOSTATISTICS TO DESCRIBE COMPLEX A PRIORI INFORMATION FOR INVERSE PROBLEMS

THOMAS M. HANSEN[1,2], KLAUS MOSEGAARD[2] and KNUD S. CORDUA[1]
[1]Institute of Geography & Geology, University of Copenhagen, Øster Voldgade 10, DK-1350 Copenhagen K, Denmark
[2]Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 28, DK-2100 Copenhagen East, Denmark.

in
GEOSTATS 2008 , Proceedings of the Eighth International Geostatistics Congress, Editors, Julián M. Ortiz amd Xavier Emery, pages 329-338.

USING GEOSTATISTICS TO DESCRIBE COMPLEX A PRIORI INFORMATION FOR INVERSE PROBLEMS

THOMAS M. HANSEN[1,2], KLAUS MOSEGAARD[2] and KNUD S. CORDUA[1]

[1]Institute of Geography & Geology, University of Copenhagen, Øster Voldgade 10, DK-1350 Copenhagen K, Denmark

[2]Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 28, DK-2100 Copenhagen East, Denmark.

ABSTRACT

Inverse problem theory deals with the problem of inferring properties of the subsurface based on indirect physical measurements. Although inverse problem theory can deal with complex a priori information in cases when samples from the a priori pdf can be generated, it is rarely done in practice. Instead rather simple prior constraints are often assumed, that may have little geological and petro-physical justification. Hence the solutions to inverse problems are rarely consistent with geological information. Geostatistics is a discipline where methods have been developed for simulating realistic geological subsurface structures. The first breakthrough was the application of classical covariance based geostatistics and sequential simulation. In the recent years the application of multiple point statistics have led to impressive simulation algorithms which can be used to simulate realistic geological structures. We will make use of geostatistical simulation algorithms to describe prior information for Bayesian formulated inverse problems. Using a classical georadar cross borehole tomography inverse problem, we will show how the a posteriori probability density function can be sampled, to provide solutions to inverse problems that not only match observed geophysical data within their uncertainties, but also the subsurface geology as described by the geostatistical model of the prior information. From such a sample complex questions can be answered probabilistically. The combination of geostatistics and inverse problem theory can also significantly reduce the complexity of the inverse problem. We demonstrate that when selecting an appropriate prior model based on geostatistics the computational needs for solving the inverse problem can be drastically reduced (orders of magnitude in computational gain).

INTRODUCTION

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

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

Inverse problem deals with the problem of inferring properties of m, based on observations of d and some knowledge about the mapping function g, typically related to physical theory. Tarantola (2005) formulate a probabilistic approach to solving inverse problems where a prior information is described the a priori probability density function (pdf) $\rho_M(m)$. A pdf describing the likelihood with respect to the geophysical data observations is given by $L_M(m)$. The solutions to such an inverse problem is a probability density function, denoted the a posteriori pdf, and is given as a product of the a priori pdf and the likelihood:

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

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 is rather restrictive.

Mosegaard and Tarantola (1995) suggest a modified Metropolis-Hasting Monte Carlo algorithm for sampling the a posteriori distribution, $\sigma_M(m)$, in case g is nonlinear, that allows for the inclusion of complex a priori information.

Consider a Monte Carlo series, where m_n is a realization of the prior pdf $\rho_M(m)$, and m_{n+1} is perturbed with respect to 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 the probability P_{accept}:

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

If m_{n+1} is accepted, m_n becomes m_{n+1}, otherwise the model m_{n+1} is rejected. After this rejection step eqn 3 is applied again. When this is performed iteratively, the algorithm will sample the a posteriori pdf. In each iteration one thus needs to perturb the current model, consistently with a priori 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.

This methodology has been applied to many non-linear inverse problems, such as Dahl-Jensen et al. (1998), Khan et al. (2000), Voss et al. (2006) but to name a few. In all these studies the choice of prior model have been very simple.

Here we will consider using geostatistics to describe complex, and hopefylly realistic, a priori information, and use this to solve non-linear inverse problems

using the modified Metropolis-Hastings algorithm as suggested by Mosegaard and Tarantola (1995).

Geostatistical optimization methods

Hu (2000,2002) and Le Ravalec-Dupin and Noetinger (2002) describe how the gradual deformation method (GDM) can be used for optimization, in order to generate a model consistent with for example geophysical observations and a prior model based on 2-point statistics. Caers (2006) propose to use the probability perturbation method (PPM) to generate model consistent with multiple point (and 2p) based statistics. Even though each generated model using these methods may be consistent with both data observations and prior information, the variability of such models is not ensured to match the variability given by the a posteriori pdf. Thus, such a set of generated model (using different initial random seed) does not provide a sample of the a posteriori pdf.

SAMPLING THE PRIOR PROBABILITY

Consider the a priori model described by $\rho_M(m)$. The only requirement for using $\rho_M(m)$ as a priori information to solve inverse problems, using the methodology of Mosegaard and Tarantola (1995), is that one must be able to generate realizations of $\rho_M(m)$. In addition one must be able to gradually perturb a current model to another model consistent with a priori information. The actual shape of $\rho_M(m)$ need not be known.

Geostatistical simulation algorithms are well suited for this task (Journel and Zhang, 2006). 2-point geostatistical methods, based on Gaussian stochastic models, $\rho_M(m)$ can be completely described by a Gaussian model (using a mean and a covariance). Using mp-geostatistics, $\rho_M(m)$ is described by a training image from which higher order statistics can be extracted. Realizations of 2p and mp-geostatistical models can be efficiently generated using sequential simulation (Gomez-Hernandez and Journel, 1993; Strebelle, 2002).

Perturbed simulation

We make use of a very simple form of gradually deforming a realization of the prior pdf using 'perturbed' simulation. We randomly remove a set of model parameters from a realization, and re-simulate these data conditional to the rest of the data:

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}. m_{i+1} is also a realization of the prior model.
3. Set m_i=m_{i+1} and go to 1.

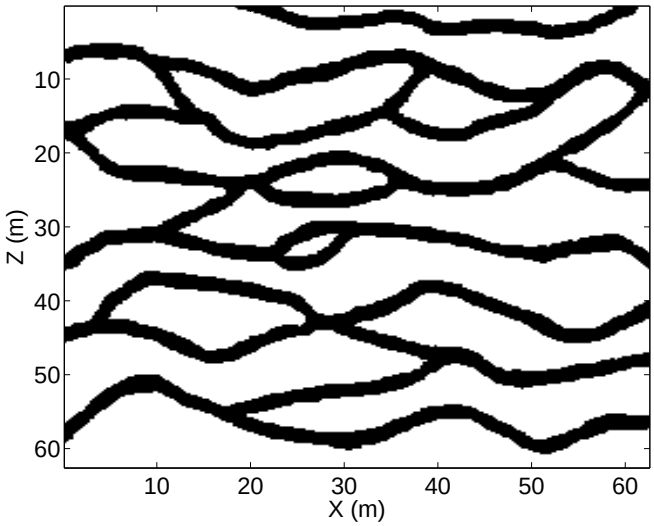

Figure 1: Training image used to generate the reference model. black channel structures has a velocity of 0.09 m/ns. The background velocity (white) has a velocity of 0.13 m/ns.

The previously mentioned GDM and PPM methods could also be applied in order to sample the prior pdf. PPM is slightly more complex to use but will allow more flexibility in the control of the degree of perturbation. GDM only applies to 2 point based simulation. The simple perturbed simulation approach outlined above offers great flexibility in terms of selecting a region to re-simulate, and it is extremely easy to use, and therefore we use this method to perform the gradual deformation in the examples below.

BAYESIAN INVERSION USING A GEOSTATISTICAL PRIOR

Figure 1 shows a training image (TI), reflecting channel structures with a velocity of 0.13 m/ns embedded in background material of velocity 0.09 m/ns. Figure 2 shows a realization based on the TI in Figure 1. We consider this as a reference model. The reference model has a mean of 0.1189 m/ns, a variance of 3.2e-4, and an exponential variogram model with a horizontal range of 6.6m, a vertical range of 2.2 m and a sill of 3.2e-4 match a experimental semi-variogram of the TI. The velocity field is a bimodal velocity field with P(m=0.09)=0.3 and P(m=0.13)=0.7.

Mimicking a cross borehole geo-radar inversion problem, 20 sources and 40 receivers are located evenly spaced vertically in two boreholes, located at x=1.25 and x=18m, Figure 2. Using the eikonal solution to the wave equation (Zelt and Barton, 1998), first arrival times are computed for all 800 sets of source and receiver and 3% uncorrelated Gaussian noise added. This data set of travel time data is the observed data, d_{obs}, in the following and shown in Figure 3

We consider four different models of prior information with an increasing level of

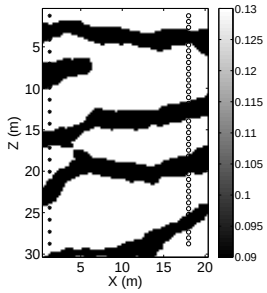

Figure 2: Reference velocity model and location of sources (*) and receivers (o)

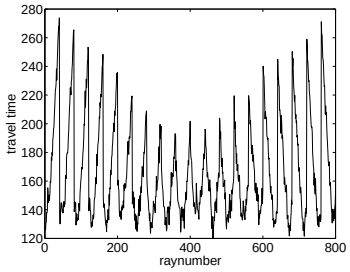

Figure 3: Calculated first arrival travel time using the recording geometry shown in Figure 2. 3% normally distributed noise was added to the travel times.

prior statistical information:

PURE NUGGET Sequentual Gaussian simulation, with a pure nugget model (no spatial correlation), and the correct mean and variance as obtained from the training image.

SGSIM Sequentual Gaussian simulation, using the correct mean, variance, and covariance as obtained from the training image.

DSSIM Direct sequential simulation, using the correct mean, variance, covariance, and target distribution as obtained from the training image.

TI Single normal equation simulation (SNESIM) using the training image in Figure 1 (Strebelle, 2002).

4 independent realizations of the a priori pdf using the four methods described above are shown in Figure 4.

As we assume Gaussian uncertainties, Cd, of the travel time data, the likelihood of a model proposition, m_i, can be calculated as:

$$L_{m_i}(i) = exp(-0.5(g(m_i) - t_{obs})' Cd^{-1}(g(m_i) - t_{obs})) \quad (4)$$

The modified Metropolis-Hastings algorithm is run for 35000 iterations, and thus 35000 different models are considered for each choice of prior model. A randomly selected area of 6x6 pixels is re-simulated at each iteration. Figure 5 show the likelihood function as a function of iteration number for the four different choices of prior model.

In the beginning the algorithm will be in the 'burn in' phase, seeking a suitable place in the model parameter space where a posteriori acceptable models can be found. Once the likelihood stabilizes and varies around approximately L=N_d/2=-400 (where N_d=800, is the number of data) the a posteriori pdf will be sampled. The faster the burn in time, the more CPU efficient the inversion procedure.

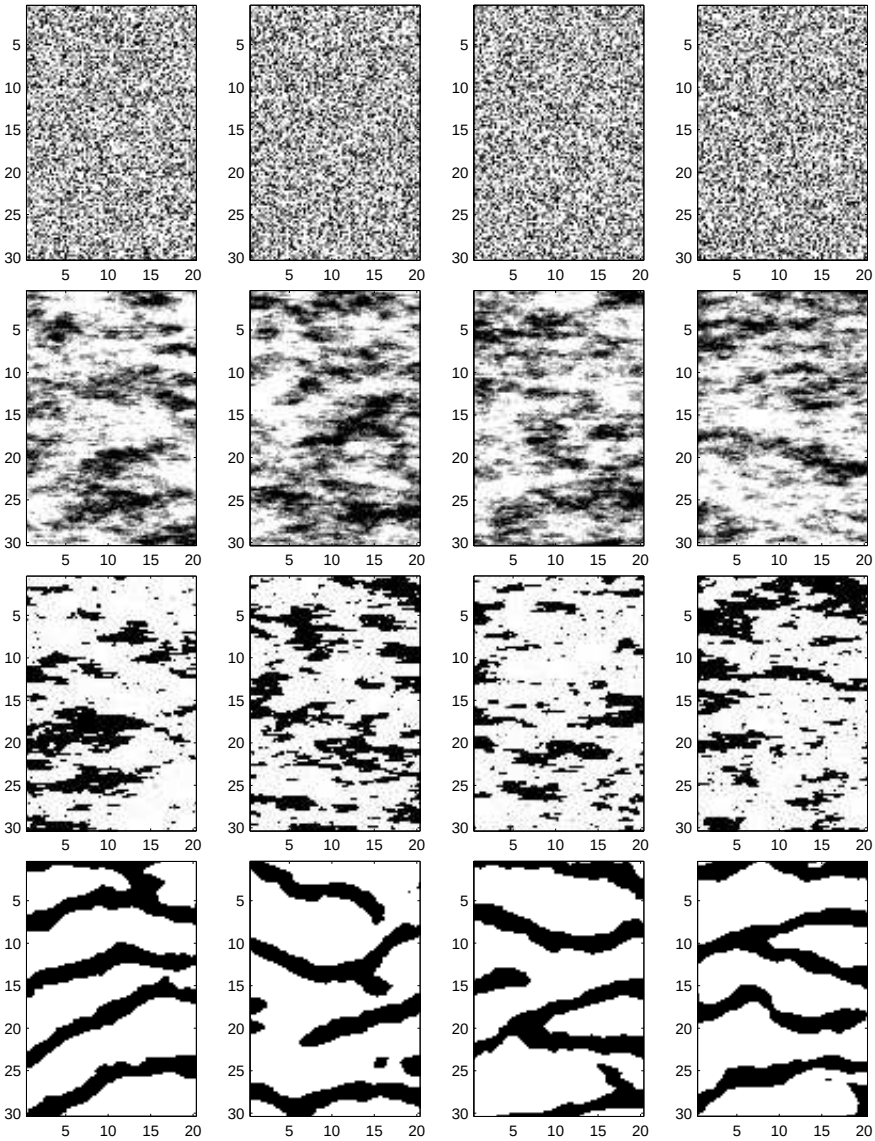

Figure 4: 4 unconditional realizations of the a priori pdf in case of top row) a pure nugget model, 2nd row) SGSIM prior, 3rd row) DSSIM prior, and 4th row) training image prior (from Figure 1).

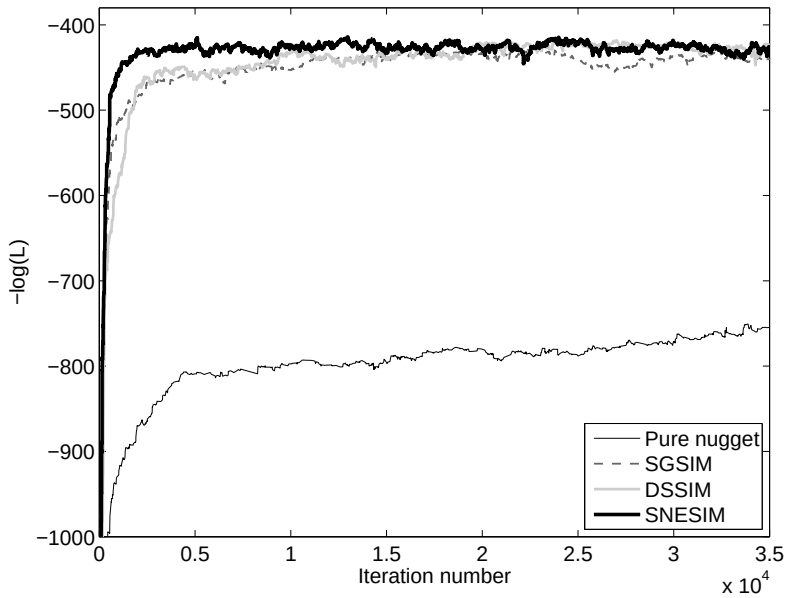

Figure 5: The likelihood of the currently accepted model as function of iteration number in the modified Metropolis-Hastings algorithm for four types of a priori model.

Burn in has been accomplished at around 3000 iterations using the true TI and around 15000 iterations for both the SGSIM and DSSIM prior. For the pure nugget prior the algorithm is still in burn in stage a iteration number 35000. If the linear trend in -log(L) observed from iteration number 5000 is extrapolated, burn in will we accomplished around iteration number 330000. This simple example thus suggest that assuming a pure nugget model results in a 330000/5000=66 times computational overhead as compared to using the true TI. Relying on 2p-geostatistical algorithms, using the true semi-variogram model is many times more efficient than using a pure nugget model. For both the SGSIM and DSSIM prior the burn in has been accomplished around iteration number 10000-15000. In other words the results indicate that the more consistent the prior model is with the actual subsurface the smaller are the computational requirements needed to generate a posteriori samples.

Note that at the iteration number of burn in, only one realization of the posterior is obtained. To obtain a new independent realization, enough iterations must be performed such that two realizations becomes independent.

Figure 6 shows the currently accepted models at iterations number 8000, 16000, 24000 and 32000 for the four types of prior models. Comparing to Figure 4 it is clear that high and low velocity structures tend to be focused in corresponding areas of the reference model, Figure 2, except for the case using a pure nugget prior.

In the 35000 iterations no independent realizations were generated using the

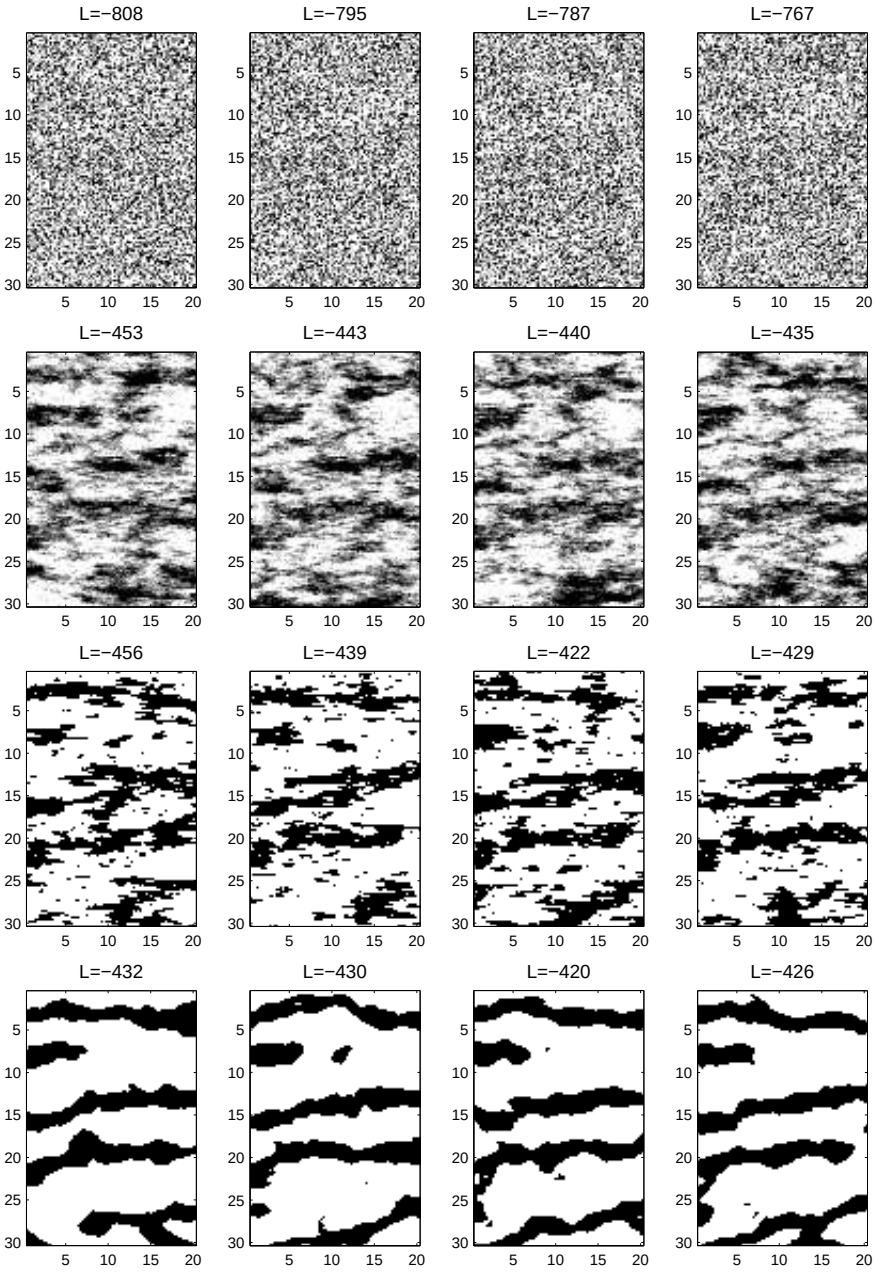

Figure 6: 4 models from the a posteriori pdf generated by the modified Metropolis-Hastings algorithm at iteration number 20000,25000,30000 and 35000 in case considering an a priori model based on: top row) a pure nugget model, 2nd row) SGSIM prior, 3rd row) DSSIM prior, and 4th row) training image prior. Compare to the unconditional realizations of the prior pdf of Figure 4

pure nugget prior. Using the SGSIM, DSSIM and TRUE TI prior, independent realizations were obtained for every 15000, 6000, and 2500 time steps after burn in respectively

All in all 1, 4, and 14 independent realizations of the a posteriori pdf was obtained using the SGSIM, DSSIM and TRUE TI prior during the 35000 time steps, and none using the pure nugget prior.

Even though burn in is accomplished around the same iteration number using SGSIM and DSSIM priors, the use of DSSIM prior is more than 3 times as efficient as using an SGSIM prior in terms of generating independent realizations. Using the true TI is 14 times as efficient as using SGSIM and about 3 times as efficient as using DSSIM.

Thus using the true TI prior, all four models in Figure 6 (bottom row) are independent realizations of the a posteriori pdf. This is enough to illustrate that such realization can provide the base of a statistical analysis of complex statements such as S1:"A high velocity layer at location (1,23) is connected to a high velocity layer at location (20,19)". Whether this statement is correct can be easily probabilistic answered simply by counting the number of occurrences of S1.

$P(S1|\rho_{TrueTI})= 12/14 = 0.85$

It is thus highly probable that statement S1 is correct considering the true TI as prior model. From the reference model, Figure 2, we can see that statement S1 is indeed correct. Using the DSSIM prior, we only found 4 independent realizations and therefore the statistical analysis is quite weak, but we find:

$P(S1|\rho_{DSSIM})= 0/4 = 0$

Thus even though data observations can be fitted within their uncertainties, using both the true TI and DSSIM as prior information, it is not desirable to try to extract higher order statistics, such as P(S1), using a covariance based prior model. This is because covariance based simulation assume maximum entropy, while the subsurface is in this case really better described by a low entropy TI (Journel and Zhang, 2006). One should only trust higher order statistics from covariance based simulations if there is reason to expect high entropy.

CONCLUSIONS

We have shown how both 2-point covariance based, and multiple point training image based geostatistical algorithms can be used as a priori information to sample the a posteriori distribution of non-linear inverse problems.

This enables generation of actual samples of the a posteriori pdf, from which rather complex information can be extracted simply by counting the occurrences of a considered event, from a sample of the a posteriori pdf. We have also illustrated that the high entropy assumption associated with covariance based simulation, makes it unusable for inferring higher order statistics from the a posteriori pdf for a case

when the subsurface is really lower entropy.

It might be tempting to assume as little as possible when solving an inverse problem. However, we have shown that an a priori choice of no spatial correlation, leads to an extremely inefficient sampling. The use of reasonable model of spatial correlation, even based on 2 point geostatistics, reduce the computational complexity by an order of a magnitude. We also found that the choice of prior model most consistent with the subsurface (when using the actual training image used to generate the reference model) provided both the most computational efficient burn in and sampling.

ACKNOWLEDGEMENTS

We used VISIM (http://imgp.gfy.ku.dk/visim) for 2-point based simulation and SNESIM, by Sebastien Strebelle, for single normal equation simulation.

REFERENCES

Caers, J and Hoffman, T (2006). *The probability perturbation method: A new look at bayesian inverse modeling*. In Mathematical Geology, vol. 38, no. 1, pp. 81 – 100.

Dahl-Jensen, D, Mosegaard, K, Gundestrup, N, Clow, G, Johnsen, S, Hansen, A and Balling, N (1998). *Past Temperatures Directly from the Greenland Ice Sheet*. In Science, vol. 282, no. 5387, p. 268.

Gomez-Hernandez, J and Journel, A (1993). *Joint sequential simulation of multi-Gaussian fields*. In Geostatistics Troia, vol. 92, pp. 85–94.

Hansen, T and Mosegaard, K (2008). *VISIM: Sequential simulation for linear inverse problems*. In Computers and Geosciences, vol. 34, no. 1, pp. 53–76.

Hansen, TM, Journel, AG, Tarantola, A and Mosegaard, K (2006). *Linear inverse Gaussian theory and geostatistics*. In Geophysics, vol. 71, no. 6, pp. R101–R111.

Hu, LY (2000). *Gradual deformation and iterative calibration of gaussian-related stochastic models*. In Mathematical Geology, vol. 32, no. 1, pp. 87–108.

Hu, LY (2002). *Combination of dependent realizations within the gradual deformation method*. In Mathematical Geology, vol. 34, no. 8, pp. 953 – 963.

Journel, A and Zhang, T (2006). *The Necessity of a Multiple-Point Prior Model*. In Mathematical Geology, vol. 38, no. 5, pp. 591–610.

Khan, A and Mosegaard, K (2002). *An inquiry into the lunar interior: A nonlinear inversion of the Apollo lunar seismic data*. In J. Geophys. Res, vol. 107, no. E6, pp. 19–44.

Le Ravalec-Dupin, M and Noetinger, B (2002). *Optimization with the gradual deformation method*. In Mathematical Geology, vol. 34, no. 2, pp. 125 – 142.

Mosegaard, K and Tarantola, A (1995). *Monte Carlo sampling of solutions to inverse problems*. In Journal of Geophysical Research, vol. 100, no. B7, pp. 12431–12447.

Strebelle, S (2002). *Conditional simulation of complex geological structures using multiple-point statistics*. In Math. Geol, vol. 34, no. 1, pp. 1–20.

Tarantola, A (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM. ISBN 0-89871-572-5.

Voss, P, Mosegaard, K and Gregersen, S (2006). *The Tornquist Zone, a north east inclining lithospheric transition at the south western margin of the Baltic Shield: Revealed through a nonlinear teleseismic tomographic inversion*. In Tectonophysics, vol. 416, no. 1-4, pp. 151–166.

Zelt, C and Barton, P (1998). *Three-dimensional seismic refraction tomography- A comparison of two methods applied to data from the Faeroe Basin*. In Journal of Geophysical Research, vol. 103, no. B4, pp. 7187–7210.