COMPARISON OF MONTE CARLO METHODS FOR MODEL PROBABILITY DISTRIBUTION DETERMINATION IN SAR INTERFEROMETRY

Andrew Hooper\textsuperscript{1} and Tim J. Wright\textsuperscript{2}

\textsuperscript{1}Delft Institute of Earth Observation and Space Systems, Delft University of Technology, Delft, Netherlands
\textsuperscript{2}School of Earth and Environment, Leeds University, Leeds, U.K.

ABSTRACT

A commonly-applied Monte Carlo approach to derive the probability distribution for the model parameters when inverting InSAR data, involves adding simulated noise to the unwrapped interferometric phase. A non-linear inversion technique, such as simulated annealing is then used to find the least-squares best-fit solution for the model parameters, given the data plus the simulated noise. This process is repeated some number of times, with the resulting distribution of solutions representing the posterior probability distribution of the model parameters. An alternative approach uses a Markov chain Monte Carlo algorithm to generate the posterior probability distribution directly, in a Bayesian manner. Here we confirm that the two strategies give similar results, for both simulated and real interferometric SAR data, in the case where there are no prior constraints on the model parameters. We find also that the Markov chain algorithm is two orders of magnitude more efficient in terms of the number of forward models that need to be run. The Markov chain algorithm has a second advantage in that prior constraints on the model parameters can be incorporated.

Key words: InSAR; Inversion; Monte Carlo; Markov chain.

1. INTRODUCTION

Interferometric SAR can provide invaluable data on the surface deformation related to various geophysical phenomena. In order to increase understanding of the underlying processes, the data are usually inverted to provide constraints on model parameters of interest. Typically the solution sought is a joint probability distribution for these model parameters.

A commonly-applied Monte Carlo approach [1] to derive the probability distribution is to first characterize the noise in terms of its covariance function, usually in one dimension. This covariance function is then used to simulate correlated noise, which is added to the unwrapped interferometric phase. A non-linear inversion technique, such as simulated annealing is then used to find the least-squares best-fit solution for the model parameters, given the data plus the simulated noise. This process is repeated some number of times, with the resulting distribution of solutions representing the posterior probability distribution of the model parameters.

An alternative approach [2] again relies on characterising the noise in terms of a covariance function, but then uses a Markov chain Monte Carlo algorithm to generate the posterior probability distribution directly, in a Bayesian manner. One implementation of this utilises the Metropolis-Hastings algorithm: in each iteration, a trial value for all model parameters is selected by taking a random step from the current model values. The observations are then predicted using the trial values (i.e. the forward model is applied), from which the joint probability density of the predicted observations is calculated. The trial model then either becomes the new current model or is rejected, with the chance of being retained being the ratio of the probability densities of the trial and the current models (if \( \geq 1 \) then the trial model is always retained). A new trial set of model parameters is then calculated, and so on.

One advantage of the Markov chain approach is that prior constraints on the probability distribution of the model parameters can be incorporated into the inversion process. However, in the case where there is no information on the prior probability of the model parameters, the two methods should give the same posterior probability distribution.

The first method requires many forward models to be run.
Figure 2. Marginal plots of the posterior probability distributions for the simulated data set generated using (a) the simulated noise method and (b) the Markov chain method. Red lines and black dots represent the true values. Less points are available to generate the histograms for (a) leading to a coarser resolution.
2. SIMULATED DATA

We simulated surface displacements for a steeply-dipping rectangular fault in an elastic halfspace with both right-lateral and reverse motion. We converted these into line-of-sight phase, assuming a wavelength of 5.67 cm and an incidence angle of 23°, and added realistic atmospheric and decorrelation noise (Fig. 1).

We solved for the model parameters of the fault using both simulated noise and Markov chain algorithms. In both cases we used the true variance-covariance matrix in the inversion. The results are broadly similar, with the shapes of the individual histograms and the correlations between pairs of parameters characterised similarly (Fig. 2). In all cases the true values lie within the posterior probability distributions. Another way to compare the results is to look at the 95% confidence bounds for each parameter from both methods, and the agreement is again good (Fig. 3). In order to fit the data onto one plot, the values were first normalised by the true values.

To estimate how many iterations were necessary for each method to successfully characterise the posterior probability distribution, we looked at how the 95% bounds evolved with each iteration (Fig 4). For the simulated noise method, convergence was achieved after about 1000 iterations, whereas with the Markov chain method, approximately 400,000 iterations were necessary. However on average 37,000 forward model runs were necessary for every simulated noise iteration, requiring $3 \times 10^7$ model runs in total. For the Markov chain method, only one forward model run was required for every iteration, meaning that $4 \times 10^5$ runs were needed altogether. Thus, the Markov chain is some two orders of magnitude more efficient, for this example at least.

3. NENANA MOUNTAIN EARTHQUAKE

We applied the Markov chain algorithm to find the probability distribution for a rectangular fault dislocation that slipped during the 2002 Nenana Mountain earthquake, using one of the descending interferograms (ifm3)
formed by Wright et al., [3]. We made the same assumptions as [3], namely, that the slip was 1 m and that the noise was fully characterised by an exponential covariance function with standard deviation 7.3 nm and an e-folding length scale of 11.5 km. The results are shown in Fig. 6, together with the results from [3] using the simulated noise method. As only 100 iterations were performed in the latter case, the histogram resolution is much coarser. Nevertheless, the characteristic shape of the univariate and bivariate histograms are broadly similar. This is further demonstrated by the degree of correlation between both methods of the normalised 95% confidence bounds for each parameter (Fig 5).

4. CONCLUSIONS

Two Monte Carlo algorithms, one involving simulated noise and the other a Markov chain algorithm, give similar results in the case where there are no prior constraints on the model parameters. However, two orders of magnitude fewer model runs are required for the Markov chain method. The Markov chain algorithm has a second advantage in that prior constraints on the model parameters may also be incorporated.

ACKNOWLEDGMENTS

ASAR data were provided by ESA under a Cat-1 proposal.

REFERENCES


Figure 6. Marginal plots of the posterior probability distributions for the Nenana Mountain earthquake, generated using (a) the simulated noise method and (b) the Markov chain method. Red lines and black dots represent the true values. Less points are available to generate the histograms for (a) leading to a coarser resolution.