PROBABILITY DISTRIBUTIONS FOR GEOACOUSTIC INVERSION

Stan Dosso
School of Earth and Ocean Sciences, Univ. of Victoria, P.O. Box 3055, Victoria, B.C., CANADA V8W 3Y2

BACKGROUND

Determining seabed geoacoustic properties from ocean acoustic fields represents a strongly nonlinear inverse problem with no direct solution. Global inversion methods, such as simulated annealing (SA) and genetic algorithms (GA), provide a practical approach based on searching the multi-dimensional parameter space for the geoacoustic model that minimizes the mismatch between measured and modelled fields [1–3]. However, these approaches provide only the best-fit solution, with no indication of the range of acceptable model parameters. Recently, GA have been used as an importance sampling technique to estimate properties of the *posteriori* probability distribution (PPD) for the geoacoustic inverse problem [4]. However, the sampling distribution of GA is unknown, and hence the PPDs constructed in this manner can suffer from both unknown errors and biases.

An alternative approach, based on SA at a fixed temperature (i.e., sampling rather than minimizing), samples directly from the PPD [5]. This procedure, known as the Metropolis Algorithm (MA) in statistical mechanics, can be used to construct an accurate and unbiased PPD, which can then be displayed in terms of marginal distributions for individual parameters. This procedure is described and illustrated here for geoacoustic inversion.

THEORY

Both SA inversion and the procedure described here for constructing PPDs for inverse problems are based on an analogy with statistical mechanics. In statistical mechanics, the probability P of a system $\mathbf{m}$ being in an energy state $E(\mathbf{m})$ is given by the Boltzmann distribution

$$P(\mathbf{m}) = \frac{\exp\left[-E(\mathbf{m})/T\right]}{\sum_{\mathbf{m}} \exp\left[-E(\mathbf{m})/T\right]}, \tag{1}$$

where T is the absolute temperature. In the early days of scientific computing, Metropolis *et al.* devised a simple numerical procedure to simulate $P(\mathbf{m})$. The MA consists of applying random perturbations to the system (or model) $\mathbf{m}$, and accepting these perturbations if

$$\xi \leq \exp[-\Delta E/T], \tag{2}$$

where ξ is a random number drawn from a uniform probability distribution on [0, 1]. It can be proved that this procedure converges asymptotically, i.e., in the limit of a large number of perturbations, the MA samples accurately and without bias from the Boltzmann distribution [5]. The evolution of the system to its global-minimum energy configuration can be simulated by applying the MA while slowly reducing the temperature T to collapse $P(\mathbf{m})$ about its groundstate.

To apply these concepts to data inversion and appraisal, consider a data set $\mathbf{d}$ with the error on each datum consisting of an independent, zero-mean, Gaussian-distributed random variable with standard deviation σ, and assume that the *a priori* information regarding the model $P(\mathbf{m})$ consists of a uniform distribution between known upper and lower limits. The model PPD is given by Bayes theorem which, for the above assumptions, can be written

$$P(\mathbf{m}|\mathbf{d}) = P(\mathbf{d}|\mathbf{m})\,P(\mathbf{m})/P(\mathbf{d}) = \frac{\exp\left[-E(\mathbf{m})\right]}{\int_{\mathbf{m}} \exp\left[-E(\mathbf{m})\right]\,d\mathbf{m}}, \tag{3}$$

where

$$E(\mathbf{m}) = [\mathbf{d} - \mathbf{d}(\mathbf{m})]^T\,\mathbf{C}_{\mathbf{D}}^{-1}\,[\mathbf{d} - \mathbf{d}(\mathbf{m})]. \tag{4}$$

In (4), $\mathbf{C}_{\mathbf{D}} = \langle \mathbf{n}\mathbf{n}^T \rangle$ represents the data covariance matrix containing all sources of uncertainty, and $\mathbf{d}(\mathbf{m})$ represents the replica data computed for model $\mathbf{m}$. Noting the similarity between eqs (3) and (1), two approaches to the inverse problem are available: (i) Maximize $P(\mathbf{m}|\mathbf{d})$ by minimizing $E(\mathbf{m})$ using the MA while slowly reducing T. This defines the method of SA, and yields the maximum-likelihood solution. (ii) Construct $P(\mathbf{m}|\mathbf{d})$ by applying the MA at a fixed temperature $T = 1$. This approach yields the full PPD for the model.

Typically, for acoustic matched-field inversion, the amplitude and absolute phase of the acoustic source are not known, i.e., we must treat

$$\mathbf{d}(\mathbf{m}) = Ae^{i\theta}\,\mathbf{w}(\mathbf{m}) \tag{5}$$

where $\mathbf{w}(\mathbf{m})$ represents the modelled acoustic field and A and θ represent the unknown amplitude and phase. For $\mathbf{C}_{\mathbf{D}} = \sigma\mathbf{I}$, the source characteristics can be removed by optimizing $E(\mathbf{m})$ over $Ae^{i\theta}$ to obtain [4]

$$E(\mathbf{m}) = [1 - B(\mathbf{m})]\,|\mathbf{d}|^2/\sigma, \tag{6}$$

where B is the normalized Bartlett processor

$$B(\mathbf{m}) = \frac{|\mathbf{d}^* \cdot \mathbf{w}(\mathbf{m})|^2}{|\mathbf{d}|^2\,|\mathbf{w}(\mathbf{m})|^2}. \tag{7}$$

Ideally, an independent estimate of the data error σ is available; however, if σ is unknown, optimizing over σ leads to the following maximum likelihood solutions [4]

$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}}\,[1 - B(\mathbf{m})], \tag{8}$$

$$\hat{\sigma} = |\mathbf{d}|^2[1 - B(\hat{\mathbf{m}})]/N, \tag{9}$$

and the model PPD is given by

$$P(\mathbf{m}|\mathbf{d}) \propto \exp\{[1 - B(\mathbf{m})]\,|\mathbf{d}|^2/\hat{\sigma}\}. \tag{10}$$

To apply the above procedure, the maximum-likelihood solution $\hat{\mathbf{m}}$ is computed according to (8) using an optimization scheme such as SA, $\hat{\mathbf{m}}$ is then used to define $\hat{\sigma}$ according to (9), and finally the MA is applied to construct the model PPD by sampling (10), i.e., by sampling an energy function

$$E(\mathbf{m}) = [1 - B(\mathbf{m})]\,|\mathbf{d}|^2/\hat{\sigma} \tag{11}$$

at temperature $T = 1$. An efficient method to achieve equilibrium for the MA is to start at high T, cool rapidly to $T = 1$, and then accumulate results for a large number of iterations. The efficiency can be further improved by reducing the perturbation sizes during cooling in a manner that reflects the various parameter sensitivities (this reduces the number of perturbations rejected during sampling). We have developed an efficient scheme based on using a perturbation size of 10× the running average of the last 30 accepted perturbations to scale Cauchy perturbation distributions for each parameter.

EXAMPLE

To illustrate the calculation of PPD for geoacoustic inversion, this section considers data collected by the SACLANT Undersea Research Centre in the Mediterranean Sea off the west coast of Italy [6]. Acoustic fields were recorded on a 48-sensor vertical array due to a swept-frequency source (300–850 Hz) towed over a track with nearly constant bathymetry (average water depth approximately 135 m). Based on known geology of the region, the geoacoustic model is taken to consist of the sediment thickness h, sediment and basement sound speeds c_s and c_b, source range and depth r and z, water depth at the source and array D_1 and D_2, and array tilt T (measured as a horizontal displacement of the top hydrophone). A hybrid inversion algorithm that combines SA with the local downhill simplex method [7] was applied to determine the parameter values that minimize the Bartlett mismatch with the measured acoustic fields for a frequency range of 300–400 Hz. Independent inversions were carried out for six source ranges from 2 to 7 km to examine the variability of the results. Given that the environmental parameters are expected to remain relatively constant with range, this variation should provide a rough indication of the relative parameter uncertainties. The model PPD was computed using the MA with unknown σ, as described above, for the acoustic data recorded for a source range of 4 km.

The results of the inversion and appraisal are shown in Fig. 1. The marginal PPDs in Fig. 1 show considerable variation in the relative uncertainty of the various parameters. For instance, the basement sound speed c_b is determined with much less uncertainty (narrower marginal distribution) that the sediment sound speed c_s. The results of the six independent inversions (crosses)

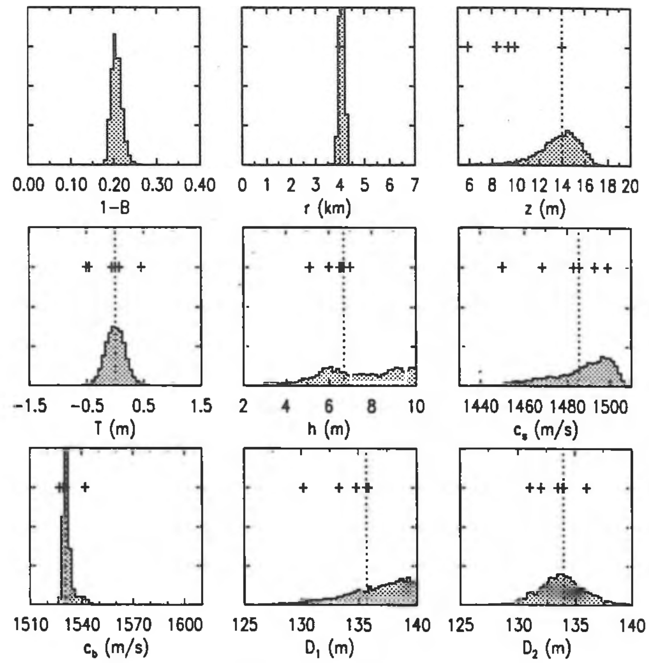

Fig. 1 Marginal PPDs for mismatch and geoacoustic model parameters for the source at 4-km range. Crosses indicate the parameter values determined for data collected at all six ranges; dotted lines indicate values determined for the 4-km source. For the geoacoustic parameters, the range of ordinate values indicates the parameter search interval.

support this, with a much greater variation between the inversion results for c_s than for c_p.

REFERENCES

[1] M. D. Collins, W. A. Kuperman & H. Schmidt, 1992. Nonlinear inversion for ocean-bottom properties, J. Acoust. Soc. Am. **92**, 2770–2783.

[2] S. E. Dosso, M. L. Yeremey, J. M. Ozard & N. R. Chapman, 1993. Estimation of ocean-bottom properties by matched-field inversion of acoustic field data, IEEE J. Oceanic Eng. **18**, 232–239.

[3] P. Gerstoft, 1995. Inversion of acoustic data using a combination of genetic algorithms and the Gauss-Newton approach, J. Acoust. Soc. Am. **97**, 2181-2190.

[4] P. Gerstoft & C. F. Mecklenbrauker 1998. Ocean acoustic inversion with estimation of *a posteriori* probability distribution, J. Acoust. Soc. Am. **104**, 808–819.

[5] M. K. Sen & P. L. Stoffa, 1996. Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion, Geophys. Prosp. **44**, 313–350.

[6] M. R. Fallat, P. L. Nielson & S. E. Dosso, 1999. Hybrid geoacoustic inversion of broadband Mediterranean Sea data, Submitted: J. Acoust. Soc. Am.

[7] M. R. Fallat & S. E. Dosso, 1999. Geoacoustic inversion via local, global and hybrid algorithms, J. Acoust. Soc. Am. **105**, 3219–3230.