Large uncertainties, limited information — Is there still hope for seismic tomography?

Andrew Valentine & Jeannot Trampert • Department of Earth Sciences, Universiteit Utrecht
andrew@geo.uu.nl • www.geo.uu.nl/∼andrew

Universiteit Utrecht

Overview

Seismic tomography attempts to improve an earth model by matching synthetic (forward-modelled) seismograms to global recordings of ground motion following earthquakes. However, these calculations depend upon knowledge of the spatio-temporal location of the event, and the manner in which energy was released. Typically, this information is obtained via a second inverse problem, using a similar—if not identical—dataset.

Some questions:

- How do we quantify the uncertainty on a set of seismic source parameters? *The earth model used during determination contains errors, and forward modelling may not be exact—and, of course, the data contains noise and measurement errors.*
- How do these uncertainties propagate into the updated earth model? Can we do anything to mitigate such effects?
- Do we have independent information on sources and structure? Can we know both accurately? *Valentine & Woodhouse (2010) suggest that a joint inversion for sources and structure leads to more accurate earth models, but less accurate sources — can this be explained?*
- Can we quantify the information content of seismograms? How do we best access that information?

Uncertainty in source parameters

Earthquakes are typically described by a 'centroid–moment-tensor' (CMT) source: the spatio-temporal location of the centroid of rupture, and a symmetric tensor describing the energy release. There are ten independent parameters: six independent moment tensor components, three spatial co-ordinates, and time. We represent these by the ten-element 'source vector' $\mathbf{f}$. CMTs are determined (Dziewonski *et al.*, 1981) by minimising the least-squares misfit between data, $\mathbf{d}$, and synthetic, $\mathbf{s}$, defined

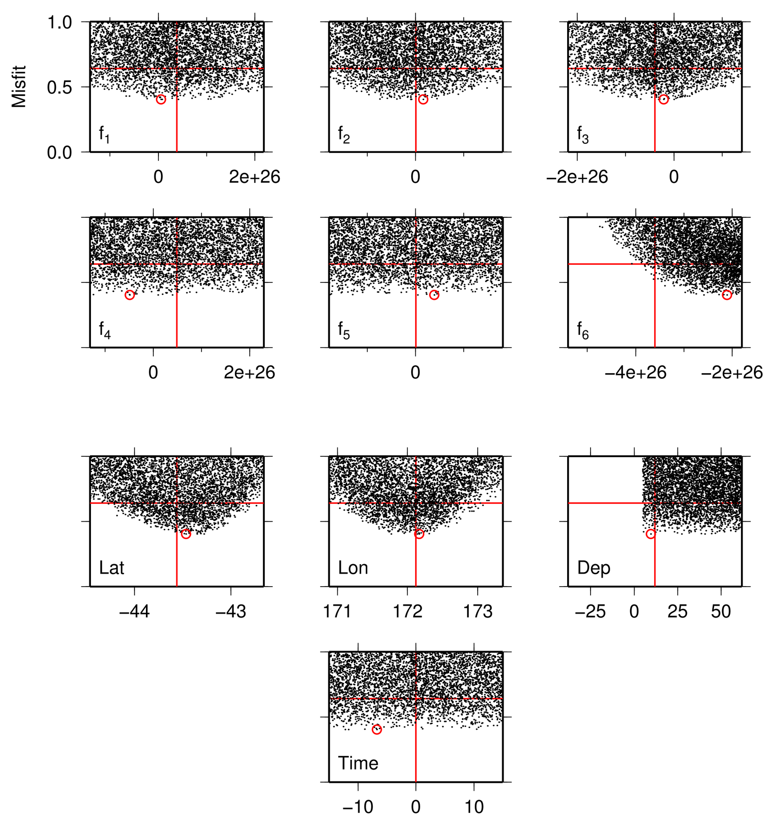
$$m^2 = \frac{(\mathbf{d}-\mathbf{s})^{\mathbf{T}}(\mathbf{d}-\mathbf{s})}{\mathbf{d}^{\mathbf{T}}\mathbf{d}}\,, \qquad (1)$$

leading to the standard iterative solution

$$\mathbf{f_{i+1}} = \mathbf{f_i} + \left(\mathbf{A^T A}\right)^{-1}\mathbf{A^T}\left(\mathbf{d}-\mathbf{s}\right)\,. \qquad (2)$$

Here, $\mathbf{A}$ is a matrix of partial derivatives; the starting point $\mathbf{f_0}$ is chosen by assuming zero energy release, and an spatio-temporal location obtained by analysis of body wave arrival times.

However, $\mathbf{A}$ and $\mathbf{s}$ depend on assumptions about the structure of the Earth, and the manner in which waves propagate through it. The assumptions used are certainly incorrect; how does this affect the validity of the solution to eq.(2)? How do we make realistic uncertainty estimates for these parameters?



Above: 'data' was calculated for a known source using a high-resolution model and numerical wave propagation (S40RTS/SPECFEM; 'body wave' portion of seismogram). We then compute synthetics in a low-resolution model with an approximate technique (M84C/mode summation & PAVA; similar to that used for source determination) for 10,000 sources 'close' to the correct source, and compute their misfit (eq.(1)) against the 'data'. Red lines denote 'correct' source; minimum-misfit solution is circled. Currently quoted uncertainties are tiny, and do not encompass the true solution.

Inversion for Earth structure

Methods for structure inversion are less standardised than those for source inversion. However, one common method (based on Woodhouse & Dziewonski, 1984; Tarantola & Valette, 1982) is also based on minimising the least-squares misfit, eq.(1). We assume that sources are known correctly, and attempt to bring data and synthetic into agreement by adjusting the earth model. Typically, this involves solving an inverse problem of the form

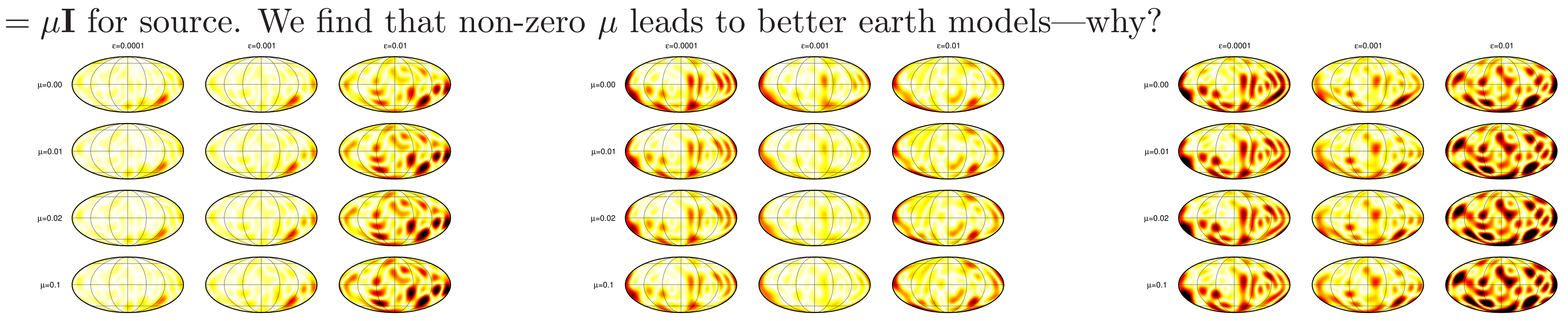
$$\mathbf{p_{i+1}} = \mathbf{p_i} + \left(\mathbf{A^T A} + \mathbf{D}\right)^{-1}\left[\mathbf{A^T}\left(\mathbf{d}-\mathbf{s}\right) - \mathbf{D}\left(\mathbf{p_i}-\mathbf{p_0}\right)\right]\,, \qquad (3)$$

where $\mathbf{p_i}$ represents the earth model parameters, $\mathbf{p_0}$ is the prior model and $\mathbf{D}$ is a regularization matrix.

As with source inversion, there are assumptions inherent to this process that may affect accuracy. More obviously, any errors in the seismic sources—which, by definition, have been determined in a sub-optimal earth model—may propagate into the updated earth model. How do we understand and minimise the effects of this?

Joint inversion for sources and structure

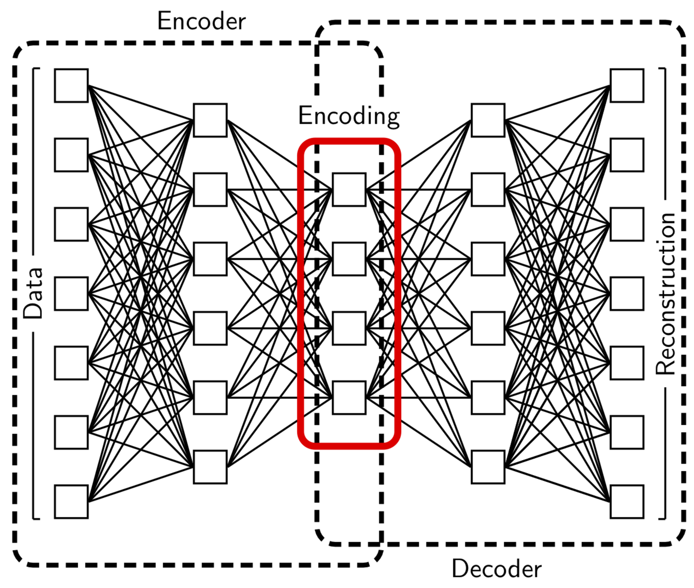
It is possible to form a single inverse problem for earth structure and all sources simultaneously, and solve this by an efficient two-step algorithm in which source corrections are determined implicitly within structure inversion (see Valentine & Woodhouse, 2010). In doing so, it is natural to extend regularization to the source terms; we use $\mathbf{D} = \epsilon\mathbf{I}$ for structure, and $\mathbf{D} = \mu\mathbf{I}$ for source. We find that non-zero μ leads to better earth models—why?



From left to right: errors in model recovery at 50km, 200km and 350km depth, synthetic test.

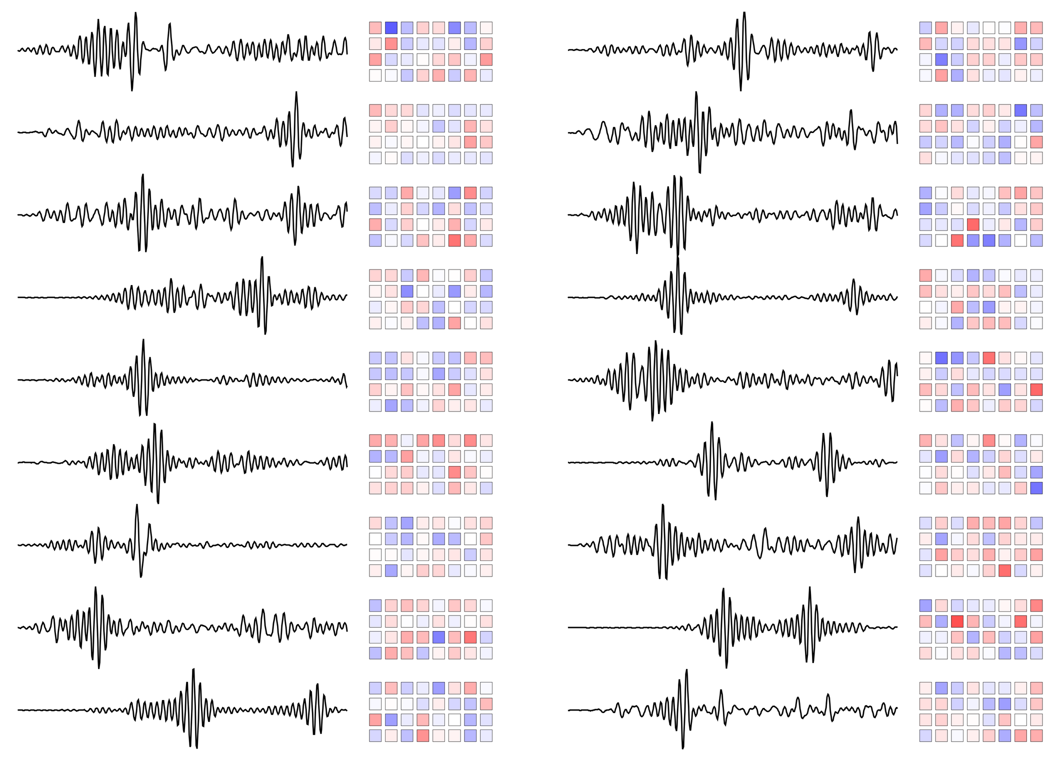
Information and waveforms?

Hinton & Salakhutdinov (2006) introduced the concept of 'autoencoder networks'—neural networks designed to find lower-dimensional representations of complex datasets.

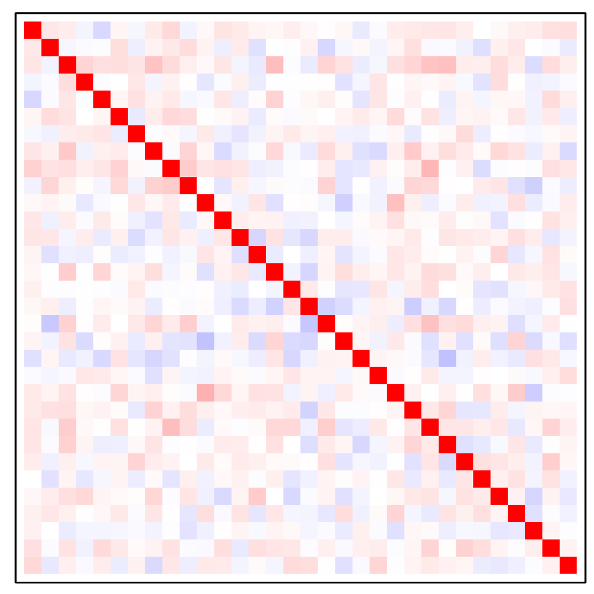


Left: Architecture of an autoencoder network. The network may be regarded as a connected encoder/decoder pair.

The network is 'trained' to find the optimal representation for a dataset using a given number of free parameters. For example, we can take a dataset of 512-point seismograms, and generate an encoding system that allows these to be represented using 32 numbers. Some examples of waveforms and their encoded representation follow:



We find that the waveforms, $\mathbf{s}$, corresponding to the 32 'unit encodings' $(1,0,\ldots,0)$, $(0,1,\ldots,0)$ etc. are close to forming an orthonormal set: we plot $M_{ij} = \frac{\mathbf{s}_i\cdot\mathbf{s}_j}{\mathbf{s}_i\cdot\mathbf{s}_i}$



Does this allow us to assess the information content of particular waveforms, or classes of data? The average amount of information per example x in a dataset, X, is given by the 'information entropy' of the set, e.g.

$$H(X) = \sum_i p(x_i)\log_2 p(x_i)\,. \qquad (4)$$

This can naturally be related to the 'compressibility' of the system; in particular, the relative compressibilities of two datasets is related to their relative information entropies. Can we make use of this? Should we be worried that seismic datasets might not be as information-rich as one would expect?

Acknowledgements

QUEST: Quantitative Estimation of Earth's Seismic Sources and Structure
Marie Curie Actions

References

Dziewonski, A.M., Chou, T.-A., & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res*, 86, pp.2825–2852.

Hinton, G.E. & Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks, *Science*, 313, pp.504–507.

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

Valentine, A.P. & Woodhouse, J.H., 2010. Reducing errors in seismic tomography: Combined inversion for sources and structure, *Geophys. J. Int.*, 180, pp.847–857.

Valentine, A.P. & Trampert, J., *in review*. Compression, quality assessment and searching of seismograms: Data-space reduction via autoencoder networks.

Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle: Three-dimensional modelling of Earth structure by inversion of seismic waveforms, *J. Geophys. Res.*, 89, pp.5953–5986.