Geophysical Research Abstracts
Vol. 13, EGU2011-11227-2, 2011
EGU General Assembly 2011

"No Data Left Behind" - Efficient waveform processing for global finite-frequency tomography

Seyed Kasra Hosseini zad (1), Karin Sigloch (1), Simon Stähler (1), and Tarje Nissen-Meyer (2)

(1) Earth Sciences, Ludwig-Maximilians-Universität, Munich, Germany ({hosseini,sigloch,staehler}@geophysik.uni-muenchen.de), (2) Institute of Geophysics, ETH Zurich, Zurich, Switzerland (tarjen@ethz.ch)

The field of seismic tomography is undergoing a rapid shift towards waveform-based methods that explicitly account for scattered wave energy. Our work addresses the problem of the whole-mantle geometry as sampled by body waves, which has three key characteristics:

1. Largest possible amount of relevant data: ideally includes all broadband stations worldwide. Rapid data increase due to new station installations.

2. Body waves have shortest possible wavelengths: they yield maximum image resolution but are very expensive to model computationally.

3. Large-scale mantle structure is well represented by weak perturbations to spherical symmetry. These lateral variations, i.e. our tomographic target, are well represented by first-order scattering.

This fortunate symmetry allows for large computational efficiency if we solve by matrix inversion (linear, single-iteration). At the heart of this approach will be a full-waveform kernel library. The expensive sensitivity kernels are computed only once, through a spherically symmetric reference model, but to the highest relevant frequencies (~0.5 Hz dominant, computation time several months). They get stored as full spatio-temporal wavefields. Kernels of arbitrary, waveform-derived observables (e.g., crosscorrelation travel times, amplitudes), for arbitrary temporal and spectral windows, get synthesized from these wavefield kernels almost on the fly. The computational cost of the inverse step is limited to the cost of the matrix inversion itself. This allows for large flexibility at inversion time: data can be added or rejected on the fly, arbitrary time frequency windows may be tried (rejecting noisy passbands in broadband seismograms), inversion cost does not scale with frequency, and the number of sources or receivers poses no limitation in practice. This flexibility and transparency is ideally suited to dealing with the vast amounts of mixed-quality data in global tomography.

Finite-frequency measurements compare predicted and observed waveforms in a cross-correlation sense. Hence the data processing chain must run symmetrically over both observed and synthetic seismograms. To the maximum extent possible, we want to replace human quality check and intervention by robust automated quality diagnosis, based on the temporal, spectral, and spatial similarity of predictions and observations. A big technical challenge is the robust, automated deconvolution of the source time function and source parameters, a crucial capability at the high frequencies that we want to exploit.

We present a flow chart for a processing chain that executes these steps as automatically and efficiently as possible. The goal is an algorithm that collects and organizes the waveform dataset for one event, identifies good time-frequency windows, calculates the source time function, retrieves corresponding kernels from the kernel library, and assembles them for matrix inversion.

Selecting a flexible and powerful programming language is an important consideration. Python with its efficient high level data structures, seems to be an ideal choice. Moreover, Python and many of its modules are open source, and it is possible to wrap external shared C/C++, Java and FORTRAN libraries. With ObsPy, a powerful library for seismological data processing already exists.