Probabilistic earthquake early warning in complex earth models using prior sampling

Andrew Valentine, Paul Käufl & Jeannot Trampert

EGU 2016 – 21st April
www.geo.uu.nl/~andrew
a.p.valentine@uu.nl
A case study: The Whittier and Chino faults
An ideal EEW source determination system…

- Treats all physical effects properly, particularly:
  - Complex wave propagation phenomena
  - Strongly heterogeneous crustal structures

- Operates in a probabilistic framework – includes full treatment of uncertainties

- Has low computational costs during operation

Achieving any two of these is reasonably straightforward – but can we have all three?
Treating physical effects properly...

- Full numerical regional wavefield simulations: SPECFEM 3D (e.g. Peter et al., GJI, 2011).
- 3D structural model for Southern California, CVMH11.9 (Tape, Liu, Maggi & Tromp, Science, 2009).
- Simulation cost: ~100 CPU hours/source
An ideal EEW source determination system...

- Treats all physical effects properly, particularly:
  - Complex wave propagation phenomena
  - Strongly heterogeneous crustal structures

- Operates in a probabilistic framework – includes full treatment of uncertainties

- Has low computational costs during operation

Achieving any two of these is reasonably straightforward – but can we have all three?
MCMC: ‘Posterior’ sampling

Sampling process cannot start until after observations have been made: infeasible for EEW scenarios
An alternative approach: ‘Prior sampling’

1. Draw model parameter(s) at random from the prior distribution
2. Compute synthetic data corresponding to that model
3. Add random noise chosen to represent observational and modeling uncertainties
Joint data-model probability density function
Joint data-model probability density function
Learning a representation of the pdf

How can we represent and store the joint distribution effectively?

- Assume it to be smooth and continuous
- Represent 1D marginal distributions as Gaussian mixture models (GMMs)
- Coefficients of GMM are the outputs of a neural network, called a ‘Mixture Density Network’ (MDN)

\[
p(m|d_0) = \sum_{i=1}^{N} \alpha_i(d_0) \exp \left( -\frac{(m - \mu_i(d_0))^2}{2\sigma_i^2(d_0)} \right)
\]

Prior sampling:

- Separates computationally-intensive sampling stage from time-critical ‘evaluation’ stage

- Is highly inefficient for a single observation – but ideally-suited to a monitoring setting, where the same inverse problem must be solved repeatedly

- Results in ‘conservative’ uncertainty estimates that could be reduced with more targeted sampling

‘Evaluation’ takes a fraction of a second!
Can we make sampling tractable for EEW?

Source parameters of interest:
- 3 location parameters
- 5 moment tensor components
- Source half-duration

9-dimensional model space

Number of samples required to achieve a given sampling density in $D$ dimensions $\sim \exp(D)$
Can we make sampling tractable for EEW?

Source parameters of interest:
- 3 location parameters
- 5 moment tensor components
- Source half-duration

9-dimensional model space

However: Point source seismogram can be expressed as linear combination of 6 independent Green’s functions:

\[ s(M, x, t, \tau, x_c, \oplus) = f(t, \tau) \times \left( \sum_{i=1}^{6} M_i \psi_i(x, t, x_c, \oplus) \right) \]

Expensive sampling only needs to be done in 3 dimensions!

Strategy:
1. Randomly sample locations; compute 6 Green’s functions using SPECFEM3D
2. Randomly sample moment tensors and source half-durations at each location and construct seismograms
150 ($\approx \text{CPU allocation}/(6\times\text{cost per simulation})$) locations distributed randomly in box enclosing Whittier & Chino faults
Record wavefield at:
- 1300 ‘real’ stations in Southern California, plus
- Regular grid of 600 ‘virtual’ stations
Simulate wave propagation for 200s after origin time, complete to 0.5Hz

Waveforms available for download (12Gb!): www.geo.uu.nl/~jeannot

With thanks to: ‘Cartesius’, the Dutch national supercomputer
From 150 sets of Green’s functions we generate ~2 million samples of (location, moment tensor, half duration, seismograms)

Assume all source mechanisms are equally likely – but would be straightforward to prefer alignment with fault

Use only a small subset of stations; filter data at 0.2Hz

Train learning algorithms to provide a smooth representation of the joint data-model distribution
Does it work? – Synthetic test

- Window starts when the first station ‘triggers’
- Wait \( t \) seconds and use recordings from all stations – some may not have recorded signal
- Each window length is implemented as a separate learning algorithm
Does it work? – Synthetic test

- $\gamma$: 0.02 0.06 0.44 0.20
- $\kappa$ (strike): 0.02 0.92 0.72 0.78
- $\sigma$ (rake): 0.10 0.28 0.63 0.56
- $h$ (cos(dip)): 0.06 0.02 0.24 0.03
- $M_w$: 1.74 2.04 2.81 2.55
- depth [km]: 0.24 0.30 0.81 0.86
- $\text{lat}$ [°]: 0.92 1.57 1.80 1.71
- $\text{lon}$ [°]: 0.85 1.82 1.94 1.86
Does it work? – $M_w 5.4$, Chino Hills, 2008
Should we bother with the source at all?

If we have seismic observations at 10 randomly-chosen locations, what can we say about regional peak ground acceleration?

Inputs to system:
- \(N\) receiver locations
- \(N\) seismograms, at those locations
- Point(s) at which PGA estimation is desired

Output: \[ P \left( \log_{10} (PGA) \mid d_1, \ldots, d_N, x_1, \ldots x_N \right) \]
Summary

- By making use of learning algorithms in a ‘prior sampling’ framework, EEW results can be made available within milliseconds.

- It is becoming feasible to take advantage of state-of-the-art numerical wave propagation codes and heterogeneous 3D crustal models.

- The approach can be extended to allow direct inversion for almost any quantity of interest.

Waveform database:
- Download from www.geo.uu.nl/~jeannot

References:
- Käufl, Valentine, de Wit & Trampert, BSSA, 2015. (*Waveform inversion*)
- Käufl, Valentine, de Wit & Trampert, GJI, in press. (*Prior sampling*)
- Käufl, Valentine & Trampert, *in prep.* (*EEW in 3D media*)