Solid Earth Discuss.,
doi:10.5194/se-2016-87-RC1, 2016

Interactive comment on "Fully probabilistic seismic source inversion – Part 2: Modelling errors and station covariances" *by* Simon C. Stähler and Karin Sigloch

T. Bodin (Referee)

thomas.bodin@ens-lyon.fr

Received and published: 12 July 2016

This paper proposes a new likelihood function for Bayesian seismic source inversion. First, a decorrelation misfit function is presented to quantify the distance between estimated and observed waveforms. The decorrelation is defined as D=1-CC, where CC is the waveform cross-correlation coefficient. It is demonstrated that this misfit function performs better than more classical L1 or L2 "point to point" norms, when trying to infer the depth of an earthquake.

In a second time, it is observed that the ensemble of waveform decorrelation coefficients for a large set of high-quality deterministic source solutions, follow a log normal distribution. The parameters of this log-normal distribution (mean and covariance) are

then used to construct an empirical likelihood function for Bayesian inference.

This is a very well written paper. It addresses an important problem that is often overlooked in global seismology : data noise (both observational and theoretical errors) in seismic waveforms may be strongly correlated, and it is important to use a proper model for data noise to avoid biases in waveform inversion.

Major comment

I have a major comment about how the likelihood function is presented. The authors define a convenient misfit measure \phi, and then use the distribution p(\phi) as a likelihood function. However, this distribution does not reflect the distribution of data errors, but instead the distribution of misfit values. A likelihood function must be derived from an assumption about the distribution of data errors and residuals must be defined as a difference between observed and predicted data vectors. In this way, the distribution of residuals follows the statistics of data noise. This is not the case for decorrelation residuals.

Bayesian inversion is based on having observed data d and model parameters m such that p(d|m), the conditional distribution of the observed data given parameters m, follows the statistics of data errors. This function can be interpreted as a function of m for fixed d to produce the likelihood. In this, the observed data d are fixed, measured quantities, independent of the parameters m. Then Baye's theorem can be used to combine p(d|m) with prior information to produce the posterior p(m|d). But here p(\phi) can't be interpreted this way as it does not strictly represent the probability of observing the data.

I think the authors should acknowledge this issue, and clarify certain points:

Page 4 Line 1: a probability distribution on the misfit => a probability distribution on the data. Page 1 line 15: the phrase "the likelihood function of misfit D" does not make sense. A likelihood function depends on a data noise model, not on a choice of misfit.

Page 7 line 12: This is not correct. Taking the distribution of any functional of observed and predicted waveforms does not give a likelihood function. Only the distribution of the difference between observed and predicted data gives a likelihood function.

Minor comments

1. Maybe the authors should refer to the recent work of Zacharie Duputel about characterizing uncertainties in source inversions.

Z. Duputel, P. S. Agram, M. Simons, S. E. Minson and J.L. Beck, 2014. Accounting for prediction uncertainty when inferring subsurface fault slip. Geophys. J. Int., v. 197, p. 464-482

Z. Duputel, L. Rivera, Y. Fukahata, H. Kanamori, 2012. Uncertainty estimations for seismic source inversions, Geophysical Journal International, v. 190, iss. 2, p. 1243-1256.

2. See also this recent paper:

Point source moment tensor inversion through a Bayesian hierarchical model M Mustać, H Tkalčić Geophysical Journal International 204 (1), 311-323

3. The log-normal likelihood goes to zero when the similarity is maximized (when D=0), right? Can this be seen as a way to penalize overfitting solutions?

4. To verify the validity of the synthetic noise added to waveforms in (16) and (17), it could be possible to check whether the distribution of decorrelation values between different realizations of u_pert and u_iˆc is log-normal, right?

5. The cumulative histogram for observed misfit values is difficult to see on Figure 5b.

6. The empirical likelihood function is constructed from a set of pre-computed deterministic source solutions. Does the Bayesian solutions differ a lot from the deterministic solutions? It would be interesting to compare the distribution of residuals obtained from both methods. Are the Bayesian residuals also log-normally distributed ?

Thomas Bodin

Interactive comment on Solid Earth Discuss., doi:10.5194/se-2016-87, 2016.