

2003 SCEC Annual Report

Project Title: Quantifying Uncertainty in Finite Fault Inversion

Principal Investigators: Peng-Cheng Liu and Ralph J. Archuleta

Institution: Institute for Crustal Studies, University of California, Santa Barbara

Introduction

Inversion of seismic data to determine the kinematic parameters on a finite fault (hereafter called finite fault inversion) is currently the most general method to determine the spatial and temporal rupture process of large earthquakes. Although finite fault inversions have become almost routine in the study of large events, the determined source model contains significant uncertainty. In practice, the uncertainty in the final solution can arise from several factors, such as, the data set being inverted is incomplete.

Seismic records, especially records in the near-source region, are the basic, and generally the most available information, used in finite-fault inversion to resolve detailed aspects of the rupture process. From a statistical point of view, these records are a sample of ground motion generated from an earthquake. Normally the size of the sample is not sufficiently large. In addition to the limited number of observers, their spatial coverage affects the inversion of ground motions (Olson and Anderson, 1988). Obviously the incomplete data set is one of the important sources of uncertainty.

In this report, we apply a bootstrap statistical method to estimate the variation in finite fault parameters induced by an incomplete data set. We reanalyze seismic data from the 1989 M 7.1 Loma Prieta earthquake and quantify the variation in the source model determined for this earthquake. Hartzell and Iida (1990) had quantified the variations in the two slip components of a finite fault. In their study, a rupture velocity is prescribed (in this case, determining two slip components for each subfault is a linear inversion problem), the L2 norm is used as objective function, and the misfit between synthetic and recorded data is implicitly assumed to follow Gauss distribution. Our study is more general. We have estimated the variation in the rupture velocity and risetime, as well as the slip amplitude and rake, of a finite fault source, without any assumption on the probability distribution.

The Implementation of Bootstrap Method

The bootstrap is a re-sampling procedure (Efron, 1982). The practical application of the bootstrap method usually requires generating a large number of bootstrap samples. An estimate $\hat{\theta}^j$ of the parameter θ is derived from each of these samples. In general, a typical bootstrap sample differs from the original sample because some of the observations will be repeated several times, and some will not occur at all. Consequently, the value of $\hat{\theta}^j$ will vary from one bootstrap sample to the next. By statistically analyzing all the resulting $\hat{\theta}^j$ the variance of parameter θ can be determined.

The application of the bootstrap method to quantify the variation in finite fault inversion is straightforward:

1. Randomly generate a bootstrap data set of size N with replacement from the original recorded data set. N is the number of original data.
2. Find a solution of source parameters $\mathbf{P}=\{p_1, p_2, \dots, p_M\}$ by minimizing the difference between bootstrap data and corresponding synthetics where M is the number of free source parameters.
3. Repeat steps 1 and 2 K times.

Having K solutions, $\mathbf{P}^1, \mathbf{P}^2, \dots, \mathbf{P}^K$, we can represent the covariance matrix of the best model $\bar{\mathbf{P}}$ as

$$\mathbf{C}_p = \frac{1}{K} \sum_{k=1}^K (\mathbf{P}^k - \bar{\mathbf{P}})(\mathbf{P}^k - \bar{\mathbf{P}})^T. \quad (1)$$

Where the best model $\bar{\mathbf{P}}$ is defined as the solution determined from inversion of the original data set. The i^{th} diagonal element of the covariance matrix \mathbf{C}_p is the variance of the best source parameter \bar{p}_i . The square root of variance gives the standard deviation, which is a measure of uncertainty in the best solution of the source parameter. It is worth pointing out that the $\bar{\mathbf{P}}$ used in (1) is the best source model instead of the mean of all bootstrap models. This choice is suggested by Chernick (1999) because the resulting estimate should be closer to the true covariance matrix of $\bar{\mathbf{P}}$. For estimating standard deviations, K is recommended to be at least 100 (Chernick, 1999).

Variation in the Source Model of the Loma Prieta Earthquake

To stress the effect of an incomplete data set on the finite fault inversion, we selected only 16 three-component stations in the near-source region of the Loma Prieta earthquake. The Green's functions are calculated from the 1D velocity model used by Wald, *et al.* (1991). Because there is no Q value in this 1D model, we assume $Q_p=2Q_s$ and $Q_s=0.1V_s$, where the units of V_s is m/s. Both the observed data (ground velocities) and Green's functions are bandpassed in the frequency range 0.05–1.0 Hz. Some of the strong motion records do not have trigger time information. For these stations the synthetic shear wave from the hypocenter is aligned with the first impulsive S wave in the data. A two second time shift is also used to account for the triggering of the Loma Prieta earthquake by a small event that did not trigger the strong motion accelerographs (Wald *et al.*, 1991).

We chose the same fault model as Wald, *et al.* (1991)—strike of 128° , dip of 70° to the southwest. The fault measures 41.25 km in length and extends from a depth of 1.5 km to 20.3 km, giving a down-dip width of 20 km. The hypocenter is at 37.04°N , 121.88°W , with a depth of 18 km. For simplicity the fault area is discretized into 15 rectangular elements along strike and 8 elements down-dip for a total of 120 subfaults of equal area (dimensions 2.75 km by 2.5 km). Following the work of Liu and Archuleta (2003), we assign the unknown source parameters to the nodes (or corners) of the subfaults. The source parameters within each subfault are calculated by bilinear interpolating the four nodal quantities of the subfault. In this study, each node has five source parameters: slip amplitude D , rake angle λ , rupture velocity c , as well as the rise and fall times (τ_1 and τ_2) of slip rate function. The times τ_1 and τ_2 control the temporal shape of the slip rate function, which is assumed to have the form

$$\frac{dS(t)}{dt} = \begin{cases} \frac{2D}{4\tau_1 + \tau_1\tau_2} \sin\left(\frac{t}{\tau_1}\right) & 0 < t < \tau_1 \\ \frac{D}{4\tau_1 + \tau_1\tau_2} + \cos\left(\frac{t - \tau_1}{\tau_2}\right) \frac{D}{\tau_2} & \tau_1 < t < \tau_1 + \tau_2 \end{cases} \quad (2)$$

The summation of τ_1 and τ_2 is defined as the rise time of the slip function $S(t)$. The inversion procedure developed by Liu and Archuleta (2003) is employed to determine these source parameters.

We first determine the best finite-fault source model for the Loma Prieta earthquake from the inversion of the original data set (16 three-component records, which give a total of 48 ground-motion velocity time histories). Then we apply a balanced re-sampling technique (Chernick, 1999) to generate the 100 bootstrap data sets from the 48 original seismogram

records. For each of these bootstrap data sets we find a source model. By analyzing all 101 source models in terms of Equation 1, we obtain the estimation of the variation in the best source model. For the best source model, we plot the spatial distributions of slip amplitude, rake, rupture velocity and risetime in Figures 1. The estimated standard deviations of these parameters are shown in Figure 2. (Although large slip is always accompanied with large variation), there are still quite large standard deviations in some regions with small slip amplitude. In fact, the largest standard deviations do not appear in the area with the largest slip amplitudes. A similar relationship can also be found between risetimes and their standard deviations. The variations of both rake angle and rupture velocity are strongly correlated with slip amplitude, the larger the slip amplitude, the less the standard deviation of rake angle or rupture velocity. Moreover rupture velocities have small variation around the hypocenter, albeit where the slip amplitudes are very small. This feature indicates that the start times of seismic waveforms place strict constraints on the initial rupture speed (Hartzell and Langer, 1993).

In addition to displaying the distributions of standard deviations, we also calculate their average values. According to above the discussion, we believe that the coefficient of standard deviation (the ratio of standard deviation and the best value of a parameter) is more suitable for evaluating the variations of slip amplitude and risetime. In general the source parameters are well determined in the region with large slip amplitude. Therefore the standard deviation or the coefficient of standard deviation is weighted with the associated slip amplitude when we average them over the fault plane. The average standard deviation of slip amplitude, rake angle, rupture velocity, and risetime are 0.72 m, 14° , 0.26 km/s, and 0.62 sec, respectively. The coefficients of standard deviation of these parameters are 30%, 10%, 10%, and 47%, respectively. It is clear that the best resolved source parameter is rupture velocity, with only ten percent relative variation. We can infer that seismic waveforms well constrain the determination of rupture velocity. Resolution of the slip amplitude is worse than that of rupture velocity, but better than that of risetime. We believe that the large or small value of rake angle will barely affect its standard deviation. An average standard deviation of 14° reflects the variation of rake angle, which should be similar to the variation of slip amplitude. We are not surprised that risetime has the worst resolution because the frequency band of data is limited to 1 Hz.

Discussion and Conclusion

In the inversion process, the source parameters are determined by minimizing the objective function. The value of the objective function is an integrated factor that measures the average misfit between synthetics and data. Because it is impossible to achieve a perfect fit between the synthetics and data, with a sufficiently large data set in practice, one inversion may result in synthetics fitting the data better at some stations than at others, while another inversion may fit the data equally well but at different stations. In this case, source models derived from the inversion of data are likely different, although they have similar values of the objective function. Bootstrapping the data set provides a way to estimate the possible variation in finite fault parameters induced by an incomplete data set.

In our inversion of the Loma Prieta earthquake, the rupture velocity is resolved well. The resolutions of slip amplitude and rake angle are not as good as rupture velocity, but they might be improved by increasing the volume of the data set, for example, by adding teleseismic data and/or geodetic data. Although the resolution of risetime is the worst, reducing the variation of risetime will be difficult to achieve because the frequency band used in the inversion is limited by our poor knowledge of the Earth structure.

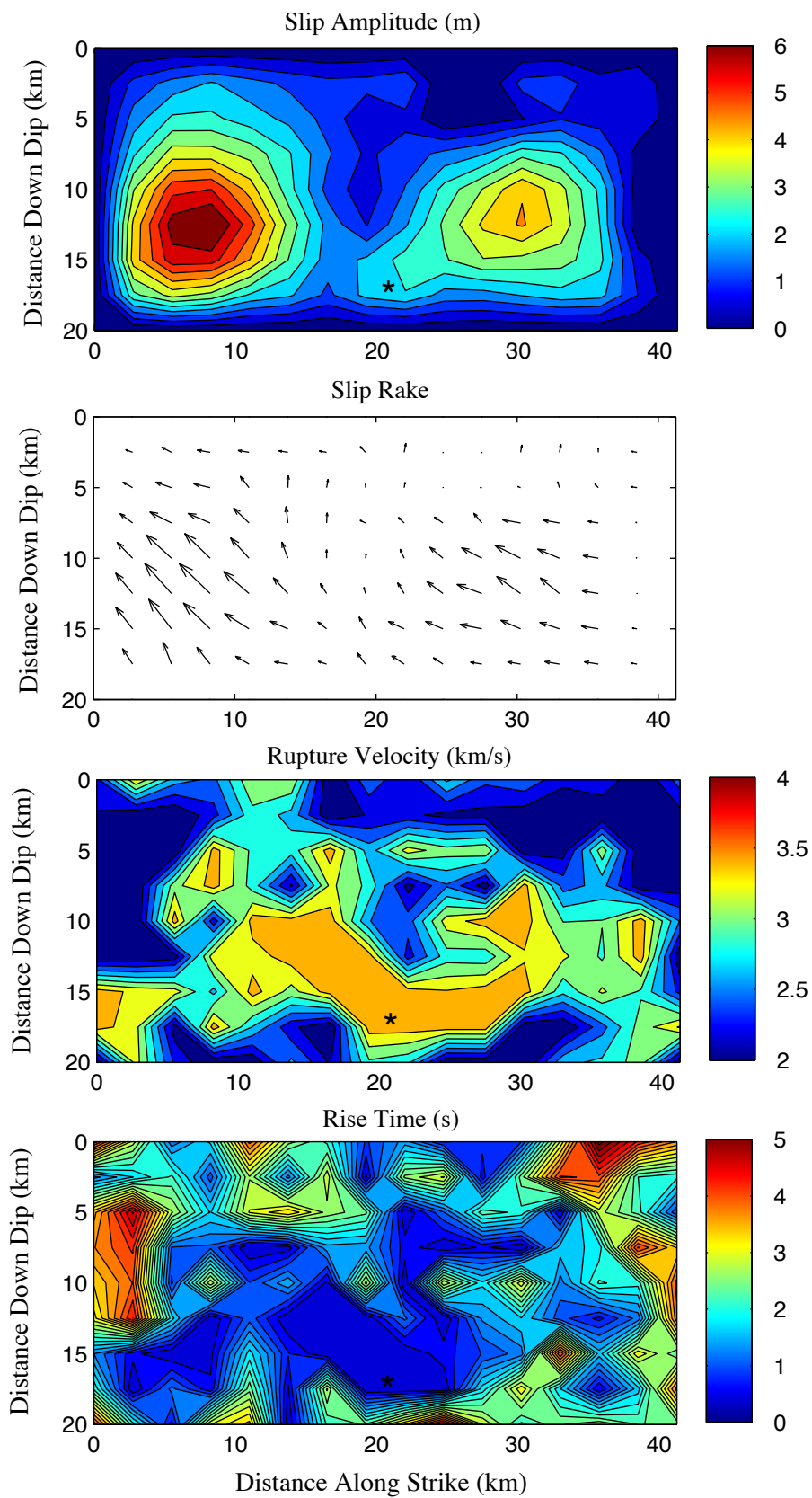


Figure 1. Plots of slip amplitude in meters, rake, rupture velocity in km/s, and rise time in sec.

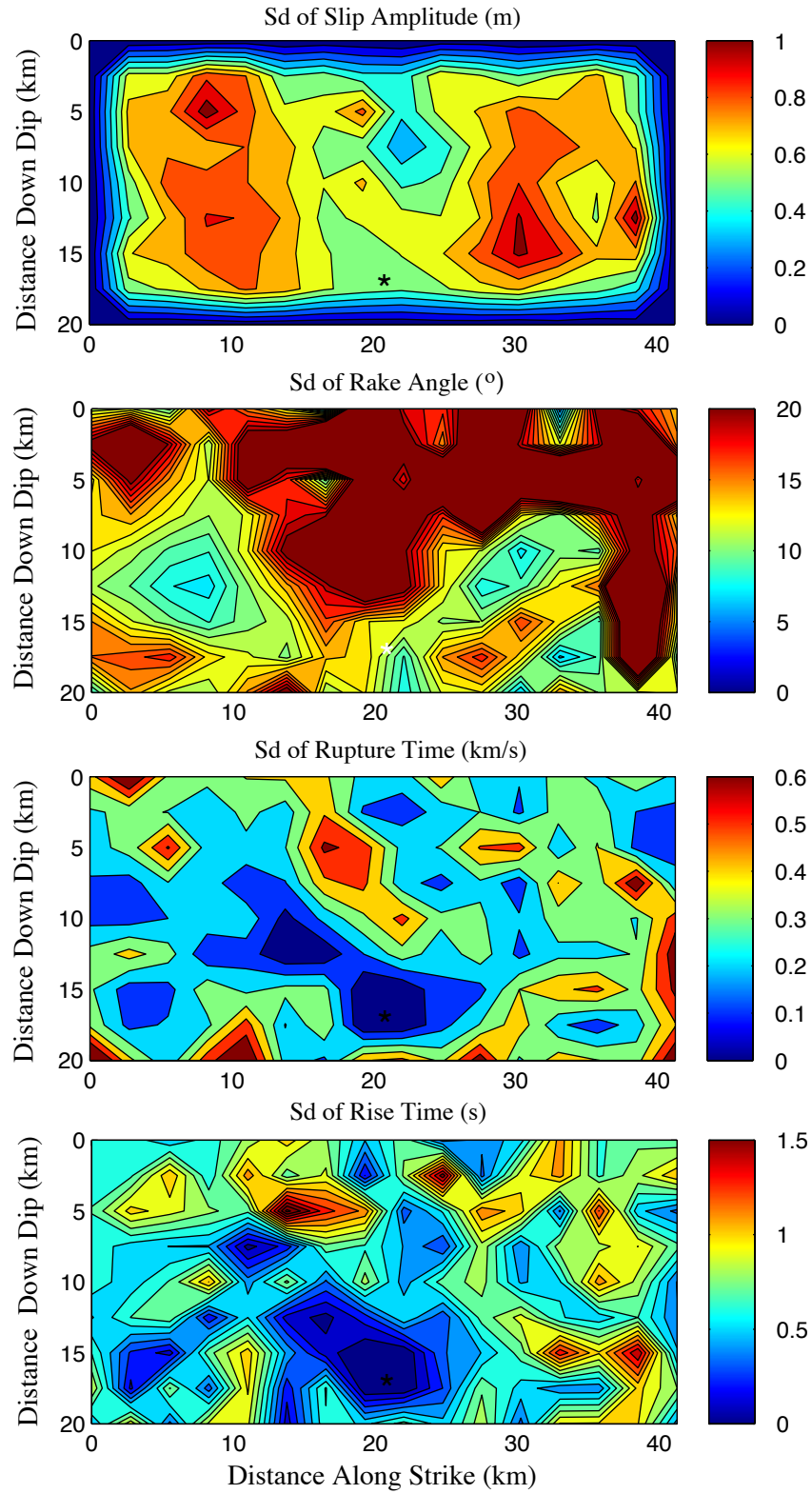


Figure 2. The spatial distributions of standard deviations (Sd) determined for the best solution (Figure 1.) of slip amplitude, rake angle, rupture velocity, and rise time.

References

- Chernick, M.R. (1999). *Bootstrap Method, A Practitioner's Guide*, John Wiley & Sons, Inc.
- Efron, B. (1982). *The Jackknife, the Bootstrap, and Other Resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphia.
- Hartzell and Langer (1993). Importance of model parameterization in finite fault inversion, *J. Geophys. Res.*, 98, 22123-22134 .
- Hartzell and Iida (1990). Source complexity of the 1987 Whittier Narrows, California, earthquake from the inversion of strong motion records, *J. Geophys. Res.*, 95, 12475-12485.
- Liu, P-C and R. Archuleta (2003). A new nonlinear finite fault inversion with 3D Green's functions: Application to 1989 Loma Prieta, California, earthquake. Submitted to *J. Geophys. Res.*.
- Olson, A. H. and Anderson, J. G. (1988). Implications of frequency domain inversion of earthquake ground motions for resolving the space-time dependence of slip on an extended fault, *Geophys. J. R. Astr. Soc.*, **94**, 443-455.
- Wald, D., D. Helmberger, and H. Heaton (1991). Rupture model of the 1989 Loma Prieta earthquake from the inversion of strong-motion and broadband teleseismic data. *Bull. Seism. Soc. Am.* **81**, 1540-1572.