Bulletin of the Seismological Society of America, Vol. 97, No. 6, pp. 1911–1934, December 2007, doi: 10.1785/0120070080

Stability and Uncertainty of Finite-Fault Slip Inversions: Application to the 2004 Parkfield, California, Earthquake

by Stephen Hartzell, Pengcheng Liu, Carlos Mendoza, Chen Ji, and Kristine M. Larson

Abstract The 2004 Parkfield, California, earthquake is used to investigate stability and uncertainty aspects of the finite-fault slip inversion problem with different *a priori* model assumptions. We utilize records from 54 strong ground motion stations and 13 continuous, 1-Hz sampled, geodetic instruments. Two inversion procedures are compared: a linear least-squares subfault-based methodology and a nonlinear global search algorithm. These two methods encompass a wide range of the different approaches that have been used to solve the finite-fault slip inversion problem. For the Parkfield earthquake and the inversion of velocity or displacement waveforms, near-surface related site response (top 100 m, frequencies above 1 Hz) is shown to not significantly affect the solution. Results are also insensitive to selection of slip rate functions with similar duration and to subfault size if proper stabilizing constraints are used. The linear and nonlinear formulations yield consistent results when the same limitations in model parameters are in place and the same inversion norm is used. However, the solution is sensitive to the choice of inversion norm, the bounds on model parameters, such as rake and rupture velocity, and the size of the model fault plane. The geodetic data set for Parkfield gives a slip distribution different from that of the strong-motion data, which may be due to the spatial limitation of the geodetic stations and the bandlimited nature of the strong-motion data. Cross validation and the bootstrap method are used to set limits on the upper bound for rupture velocity and to derive mean slip models and standard deviations in model parameters. This analysis shows that slip on the northwestern half of the Parkfield rupture plane from the inversion of strong-motion data is model dependent and has a greater uncertainty than slip near the hypocenter.

Introduction

Inversion for the kinematic slip history on a finite-fault plane can be traced back to Trifunac (1974). In this study, Haskell (1969) dislocation theory was used to find the slip on a relatively small number of subfaults by a least-squares fit to selected strong-motion records from the 1971 San Fernando earthquake. Although the formulation of the model was simplistic, the conceptualization of the problem was insightful. The next significant advances in finite-fault inversion methodology were made by Olson and Apsel (1982) and Hartzell and Heaton (1983). Their contributions lie in the addition of stabilizing constraints to the inversion, the use of layered half-space Green's functions, and the inclusion of teleseismic waveform data. The desire for finer resolution of faulting details led these investigators to use smaller spatial parameterization of fault slip in the least-squares problem, $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is the matrix of synthetic subfault ground-motion records, $\mathbf{x}$ is the solution vector of slip on the subfaults, and $\mathbf{b}$ is the observed data. However, because $\mathbf{A}$ is usually an ill-conditioned matrix, such that a small change in the data leads to a large change in the solution, stabilizing constraints are required. Hartzell and Heaton (1983) investigated three types of constraints: (1) moment minimization, (2) smoothing, and (3) filtering of singular values. By using moment minimization, elements of $\mathbf{x}$ are forced to zero that do not significantly contribute to matching the data. Smoothing requires the difference between spatially adjacent values of $\mathbf{x}$ to also be zero. By so doing, the starting and stopping phases of individual subfaults are minimized and the best-fitting smooth solution is obtained. Filtering of singular values obtains a solution based on the larger singular values of the $\mathbf{A}$ matrix. Hartzell and Heaton (1983) favored moment minimization and smoothing as being more physical constraints, which have since been used by many other investigators. In addition, Hartzell and Heaton (1983) applied a nonnegative constraint on the elements of the solution vector to prevent nonphysical backslip and destructive interference between subfaults. Both Olsen and Apsel (1982) and Hartzell and Heaton (1983) stressed the nonuniqueness

of the solution and showed the variability in results as a function of different rupture velocities, data sets, and stabilization constraints.

The investigation of nonuniqueness and uncertainty issues in finite-fault slip inversions has been further developed by a number of authors. Olsen and Anderson (1988) solved the problem in the frequency domain with a least-squares (L_2) norm and dramatically illustrated the variability in slip for different station configurations using synthetically generated data for a known slip distribution. Not surprisingly, the station distribution that performed well consisted of a line of stations paralleling the fault strike and offset by one-half the maximum fault depth. In a series of papers Das and Kostrov (1990, 1994) and Das and Suhadolc (1996) inverted forward-generated test problems using primarily an L_1 norm fit to the data. They evaluated the value of different stabilizing constraints and concluded that a positivity constraint on the slip is essential for recovering the correct approximate solution. They further concluded that they were unable to fit the data adequately or reproduce the correct forward model if the inversion was constrained to use a different rupture velocity, the fault was smaller than the actual dimensions, or the wrong velocity structure was used. They suggested relying only on those aspects of the fault model that remain unchanged with different assumptions about the model. Hartzell *et al.* (1991) considered the effect of different inversion norms by comparing L_1 and L_2 derived slip distributions for the 1989 Loma Prieta earthquake. They concluded that significant differences existed between the two norms but that robust features could be identified that were common to both.

Global search algorithms have been implemented to solve the nonlinear problem where the rupture velocity, defined as the average rupture velocity from the hypocenter, is allowed to vary as a function of position on the fault (Zeng and Anderson, 1996; Hartzell *et al.*, 1996; Ji *et al.*, 2002; Liu and Archuleta, 2004). Few studies have compared slip distributions from linear and nonlinear formulations, but Hartzell *et al.* (1996) provide a good example. In their study of the 1994 Northridge earthquake, they demonstrated that a linear least-squares inversion, in which each subfault is allowed to rupture multiple times, gave a similar solution to a global search method. By allowing each subfault to rupture repeatedly with a different amount of slip, the rise time or slip rate function and, to a lesser extent, the rupture timing are made a function of position on the fault. The global search method allows for variable rupture timing and thus variable average rupture velocity and a parameterized slip rate function. The good agreement in the case of the Northridge earthquake may be attributed in part to the apparently small variations in rupture velocity. Cohee and Beroza (1994) compared the multiple-time-window linear inversion method with a linearized iterative method that allowed slip during one time interval only but with some variation in rupture velocity. From their tests based on the 1992 Landers earthquake, they concluded that both methods yielded considerable similarities in slip amplitudes and in the gross rupture timing. They further noted that seismic moment based on strong-motion data alone can be poorly resolved and dependent on the model parameterization and inversion method, with the multiple-time-window method giving a 20% larger estimate than the one-window result.

Other studies have considered the variability of the solution with different velocity models. Sekiguchi *et al.* (2000) applied a linear least-squares method to the 1995 Hyogo-ken Nanbu earthquake and showed that significant distortion resulted in the match to a forward-generated slip model when the individual Green's functions were shifted in time equivalent to a misestimation of the velocity structure by a standard deviation of 3.5%. However, the gross features of the model were still discernible. Graves and Wald (2001) and Wald and Graves (2001) investigated the resolution of 1D and 3D velocity models with a linear least-squares inversion method for the geometry of the 1994 Northridge earthquake. They found that synthetically generated data for a 3D velocity model could only partially resolve the slip distribution when the Green's functions used in the inversion were based on an alternative 3D velocity model or a 1D velocity model. Liu and Archuleta (2004), using a global search method for slip amplitude, rake, rise time, and rupture velocity, compared results from 1D and 3D velocity models for the 1989 Loma Prieta earthquake. They found that many of the features of the rupture models were similar but noted some clear differences in the spatial distribution of slip. They advised using a range in Earth models to sample the variability in faulting models.

Recently, studies by Custodio *et al.* (2005) and Liu *et al.* (2006) have investigated uncertainty in finite-fault source inversions associated with the data and the global search process, respectively. Custodio *et al.* (2005) looked at the effect of subdividing the 2004 Parkfield strong-motion data set into 12 regions and inverting different subsets of the data. Liu *et al.* (2006) found several slip models for the 2004 Parkfield earthquake based on different initial random seeds in their global search method. Piatanesi *et al.* (2007) used a similar methodology to find an ensemble of models for the 2000 western Tottori, Japan, earthquake that fit the data approximately equally well.

The studies mentioned in the preceding discussion represent only a small fraction of the numerous works on finite-fault slip inversions. Although the general trends noted by the aforementioned authors regarding the use of stabilizing constraints, alternative parameterizations, and velocity models are certainly correct, it is difficult to draw overarching quantitative conclusions because each study is highly dependent on the amount and distribution of data. In this article, we use the 2004 Parkfield earthquake as a test case for stability and uncertainty issues because of its large strong-motion data set that is relatively uniformly distributed along the fault strike. Thus, conclusions based on this earthquake should be representative of source studies with good strong-motion coverage. Studies using synthetic data sets are useful for gaining insight into the behavior of finite-fault in-

versions. However, synthetic data sets are difficult to construct that fairly represent all the poorly defined complexities that exist in real data.

The 2004 Parkfield Earthquake

The magnitude 6.0 Parkfield earthquake of 28 September 2004 represented the culmination of a long-anticipated event between the creeping section of the San Andreas fault to the north of the epicenter and the locked section to the south (Bakun *et al.*, 2005; Langbein *et al.*, 2005). Because of the fairly regular occurrence of moderate-sized earthquakes near Parkfield in 1881, 1901, 1922, 1934, and 1966 (Bakun and Lindh, 1985), this section of the San Andreas fault has been heavily instrumented.

Data

We use the strong-motion records from 54 near-field accelerometers deployed by the California Geological Survey (Shakal *et al.*, 2006) and the U.S. Geological Survey (Borcherdt *et al.*, 2006) (Fig. 1 and Table 1). These stations give dense coverage along the entire length of the fault rupture under investigation. The records are integrated to either velocity or displacement and are band-pass filtered from 0.2 to 1.0 Hz or 0.2 to 2.0 Hz, depending on our application, with a time step of 0.05 sec. A four-pole, zero-phase shift,

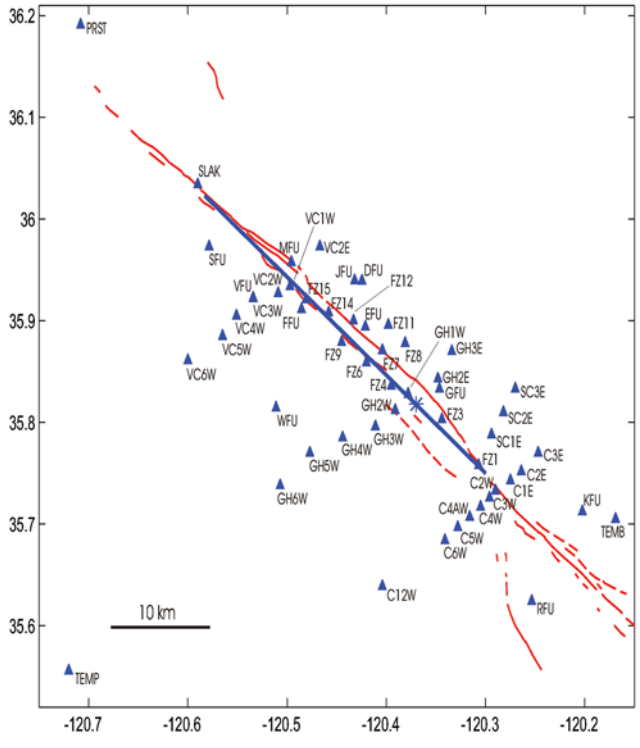

Figure 1. Distribution of the 54 strong-motion stations used in this study. Fault traces are shown in red. The surface trace of the model rupture plane is given by the solid blue line, and the location of the epicenter is given by the asterisk.

Table 1
Strong Motion Stations Used in Inversions

Station	Latitude (° N)	Longitude (° W)	Quaternary
FZ1	35.758	120.307	Q
FZ3	35.803	120.344	Q
FZ4	35.836	120.395	Q
FZ6	35.859	120.420	
FZ7	35.871	120.404	Q
FZ8	35.878	120.381	
FZ9	35.879	120.445	
FZ11	35.896	120.398	
FZ12	35.900	120.433	Q
FZ14	35.908	120.458	Q
FZ15	35.921	120.481	
VC6W	35.861	120.600	Q
VC5W	35.885	120.565	Q
VC4W	35.905	120.551	
VC3W	35.922	120.534	Q
VC2W	35.927	120.509	
VC1W	35.934	120.497	
VC2E	35.973	120.467	
C12W	35.639	120.404	Q
C6W	35.684	120.341	Q
C5W	35.697	120.328	Q
C4W	35.717	120.305	Q
C4AW	35.707	120.316	Q
C3W	35.726	120.296	Q
C2W	35.733	120.290	Q
C1E	35.743	120.275	Q
C2E	35.752	120.264	
C3E	35.770	120.247	Q
GH6W	35.738	120.507	
GH5W	35.770	120.477	
GH4W	35.785	120.444	
GH3W	35.796	120.411	Q
GH2W	35.812	120.391	Q
GH1W	35.828	120.378	Q
GH2E	35.843	120.348	
GH3E	35.870	120.334	
SC1E	35.788	120.294	Q
SC2E	35.810	120.282	
SC3E	35.833	120.270	
TEMB	35.705	120.169	
TEMP	35.556	120.720	Q
PRST	36.191	120.708	
SLAK	36.034	120.590	
DFU	35.939	120.425	
EFU	35.894	120.421	
FFU	35.911	120.486	
GFU	35.833	120.346	
JFU	35.940	120.432	
KFU	35.713	120.203	
MFU	35.958	120.496	
RFU	35.624	120.254	
SFU	35.973	120.579	
VFU	35.923	120.534	Q
WFU	35.815	120.511	

butterworth filter is used. The low-frequency limit is determined by instrument capabilities and the amplitude of the long-period signal in the data records. We use the two horizontal components of motion at each station to avoid the complexity of modeling lower-amplitude higher-frequency *P* waves on the vertical component. We also consider Global Positioning System (GPS)-derived horizontal displacements from 13 continuously recorded stations of the SCIGN network with a sampling rate of 1 sample/sec (Fig. 2). Unlike the strong-motion stations, the GPS locations are more spatially concentrated and do not give as wide a coverage of the fault plane. Offsets were estimated every second for 45 min before and after the earthquake using the techniques of Larson *et al.* (2003), augmented with modified sidereal filtering (Choi *et al.*, 2004). Common-mode errors were minimized by stacking positions from sites outside the Parkfield region (Wdowinski *et al.*, 1997; Bilich, 2006). We define the GPS static offsets as the difference between the position averaged over the 100-sec interval before the earthquake and the position averaged over the 30–130-sec interval after the earthquake. The continuous GPS data reveal that a significant afterslip occurred within the first half hour following the earthquake. Therefore, because we are interested in just the seismic portion of the displacements, use of campaign GPS data is problematic and is not considered in this study.

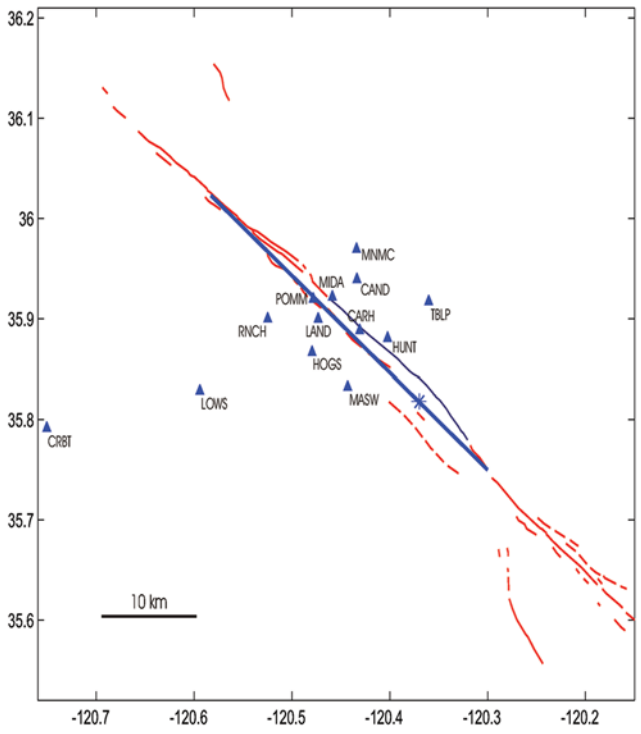

Figure 2. Distribution of the 13 continuous, 1-Hz sampled, geodetic stations. Details are the same as Figure 1.

Velocity Structure

Because of the long-standing interest in the region of the Parkfield earthquakes, a considerable amount is known about the local velocity structure. Seismic velocities are known to be higher west of the San Andrea fault near Parkfield (Eberhart-Phillips and Michael, 1993; Thurber *et al.*, 2004, 2006). We therefore use a different 1D velocity structure for the southwest and northeast sides of the fault (Tables 2 and 3). The *P*-wave velocities are interpreted from the 3D velocity structure of Thurber *et al.* (2004, 2006). The *S*-wave velocities are derived from the *P*-waves assuming a Poisson solid, except for the slower near-surface sedimentary layers. Two different velocity profiles are considered for both sides of the fault: one with stiffer surface material and a higher shear-wave velocity of 1.1 km/sec more appropriate for Cenozoic and Mesozoic rocks, and a second with added surface layers having shear-wave velocities of 0.3 and 0.5 km/sec and thicknesses of 0.05 km appropriate for Quaternary deposits. The mapped surface deposits are used to assign the Quaternary layers to selected stations indicated by a Q in Table 1. The velocity profiles minus the Quaternary surface layers are similar to those used by Liu *et al.* (2006). In addition, low and high attenuation models are considered (types I and II in Tables 2 and 3). The thicknesses of Quaternary deposits and Q-values in the Parkfield area are not known in sufficient detail; therefore, the values used here are arbitrary but useful for testing the sensitivity of the finite-fault inversion to shallow variations in velocity and attenuation. A detailed 3D velocity and Q structure would be a more accurate model but is beyond the scope of this study.

Fault-Plane Model

The slip during the earthquake is assumed to lie on a plane with a strike and dip of 140º SE and 89º SW, respectively, based on the location of aftershocks (Thurber *et al.*, 2006). The fault extends over the depth range from 0.5 to 14.5 km, although deeper faulting is allowed in certain tests. The hypocenter is located at 35.818º N, 120.370º W at a depth of 8.2 km (Liu *et al.*, 2006). These coordinates put the rupture on the southwest fracture zone subparallel to the San Andreas fault, consistent with aftershock locations. The model fault length is 40 km, 30 km to the northwest of the epicenter and 10 km to the southeast.

Inversion Methodologies

Two main inversion procedures are compared in this study: the linear least-squares method of Hartzell and Heaton (1983) and the global search algorithm of Liu and Archuleta (2004). These two methods have been selected because they are fundamentally different and together encompass a wide range of the general methodologies that have been applied to kinematic fault slip inversion. Both methods use the same Green's functions, calculated using the frequency–wavenumber method of Zhu and Rivera (2002), and the same fault plane parameters mentioned previously.

Table 2
1D Velocity Models for Parkfield

Thickness (km)	V_p (km/sec)	V_s (km/sec)	Density (gm/cc)	Q_p Type I	Q_s Type I	Q_p Type II	Q_s Type II
Southwest Side of Fault							
1.0	2.0	1.1	2.0	200	100	70	35
1.0	3.5	2.0	2.3	600	300	140	80
1.0	4.5	2.6	2.3	780	390	300	200
1.0	5.4	3.1	2.7	930	465	400	240
3.0	5.8	3.4	2.7	1020	510	500	280
2.0	6.2	3.6	2.8	1080	540	500	280
3.0	6.8	3.9	2.8	1170	585	510	300
14.0	7.0	4.0	2.8	1200	600	700	400
—	8.0	4.5	3.4	1350	675	1200	650
Northeast Side of Fault							
1.0	2.0	1.1	2.0	200	100	70	35
1.0	3.5	2.0	2.3	600	300	140	80
3.0	4.5	2.6	2.3	780	390	300	200
5.0	5.4	3.1	2.7	930	465	400	240
3.0	6.5	3.8	2.8	1140	575	510	300
13.0	7.0	4.0	2.8	1200	600	700	400
—	8.0	4.5	3.4	1350	675	1200	650

The approach of Hartzell and Heaton (1983) divides the fault plane into equal-area subfaults. The ground displacement for a unit amount of slip on each subfault is calculated by summing the contributions of point sources uniformly distributed over the area of the subfault. Two different subfault sizes are considered: 2 by 2 km and 1 by 1 km. The point source spacing is 125 m. Slip is allowed to occur on each subfault in a series of discrete time intervals. The functional form of the slip rate function is assumed to be constant for each time interval. However, a complicated or extended rise time function can be constructed on each subfault, if required by the data, by the linear combination of slips over the individual time intervals. The rupture velocity is constant, with some flexibility in the timing of phases offered by the different time intervals of slip. Different rupture velocities are obtained by applying the appropriate rupture delay time to each point source on the fault. The rake can be made to vary over a range of 90º by taking linear combinations of Green's functions for two different rakes at right angles on each of the subfaults. The inversion is stabilized by appending smoothing, **S**, and moment minimization, **M**, constraints to the least-squares problem, $\mathbf{Ax} = \mathbf{b}$, resulting in

$$\begin{pmatrix} \mathbf{C}_d^{-1}\mathbf{A} \\ \lambda_1\mathbf{S} \\ \lambda_2 \end{pmatrix}\mathbf{x} \cong \begin{pmatrix} \mathbf{C}_d^{-1}\mathbf{b} \\ 0 \\ 0 \end{pmatrix},$$

where λ_1 and λ_2 are linear weights, whose magnitudes control the trade-off between satisfying the constraints and fitting the data. $\mathbf{C}_d$ is an *a priori* data covariance matrix that is used as a data scaling matrix. The data covariance matrix is diagonal and normalizes each data record to have a peak amplitude of 1.0. Thus, each record has nearly equal weight in the inversion. The solution vector **x** is obtained using the Householder reduction method that invokes a positivity constraint on the solution (Lawson and Hanson, 1974), that is, each of the elements of the vector **x** is greater than or equal to zero. In this study the constraint weights, λ_1 and λ_2, are determined by a trial-and-error process that balances fitting the data with smoothing and minimizing the slip. Other methods can be applied such as Bayesian information criterion (ABIC) (Sekiguchi *et al.*, 2000) or a cross validation procedure (Hansen, 1998). These methods of determining constraint weights deserve further investigation, but we leave them to subsequent studies.

With the subfault formalism of Hartzell and Heaton (1983) we also consider the L_1 norm solution using the linear programming routine of Bartels and Conn (1980). The choice of a particular norm deals with the distribution of errors in the data, the distribution of data around the source, and the characteristics of the source model one considers desirable. Least-squares analysis (L_2) assumes errors in the data have a Gaussian distribution, while the L_1 norm assumes that errors have an exponential distribution (Menke, 1984). An exponential distribution with the same mean and variance as a Gaussian distribution has a longer tail. Therefore, the probability of having a few outlying points with an exponential distribution is higher, and an L_1 minimization should be able to accept a few bad data points by attaching less weight to them. Another reason for considering

Table 3
1D Velocity Models for Parkfield, Quaternary Surface Layers

Thickness (km)	V_p (km/sec)	V_s (km/sec)	Density (gm/cc)	Q_p	Q_s
0.05	1.2	0.3	1.9	100	50
0.05	1.8	0.5	1.9	100	50

alternative norms to L_2 was mentioned by Olson and Anderson (1988). The L_2 norm minimizes the sum of squares of the differences between the data and the predicted values and the sum of the squares of the solution vector. Therefore, Olson and Anderson (1988) argue that fault slip amplitude will tend to be less than the true amplitude, and to compensate, the spatial and temporal distribution of slip may be altered to focus energy toward stations. Thus, station geometry plays an important role in the ability to resolve the trade-offs between slip amplitude and its spatial and temporal distribution. The L_1 norm minimizes the sum of the absolute values of the differences between and data and the predictions and the sum of the absolute values of the solution vector. Therefore, the same trade-offs exist but with a different weighting.

The method of Liu and Archuleta (2004) uses a simulated annealing algorithm to search for the fault parameters, slip, rake, average rupture velocity from the hypocenter, and rise time, for a grid work of points on the fault. For the Parkfield earthquake, fault parameters are calculated at a 2 by 2 km spacing, but they are interpolated to a spacing of 167 by 167 m to produce smoothly varying values. Synthetics based on all the interpolated point sources are used at each step of the search process. With this approach searches are performed over a range in rupture velocities, which are converted to rupture delay times and applied to the individual point sources by dividing the distance to the hypocenter by the rupture velocity. One of the slip rate functions used in this study has the two-parameter formulation defined by

$$\dot{s}(t) = \begin{cases} A[\sin(t\pi/2T_1)] & \text{if } 0 \le t \le T_1 \\ (A/2)\{1+\cos[\pi(t-T_1)/T_2]\} & \text{if } T_1 \le t \le T_1+T_2 \end{cases},$$

where $A = 2\pi/(4T_1 + \pi T_2)$. The total rise time T is the sum of the rise, T_1, and fall, T_2, of the slip rate function. A linear least-squares problem minimizes the L_2 norm of the data and synthetics; however, with a global search algorithm, we are free to specify the form of the objective function to be minimized. We consider the L_1 and L_2 norms as well as the following objective function used by Sen and Stoffa (1991) and Hartzell *et al.* (1996):

$$E(\mathbf{M}) = \sum_{i=1}^{N_d} W_i \left(1 - \frac{2\sum_{t_b}^{t_e}[u_o(t)u_s(t)]}{\sum_{t_b}^{t_e} u_o^2(t) + \sum_{t_b}^{t_e} u_s^2(t)}\right) + W_c(\text{constraints}),$$

where $\mathbf{M}$ is the vector of model parameters; N_d is the number of data records; W_i is a weight reflecting the quality of the data; (t_b, t_e) is the time window used in the inversion for the observed, u_o, and synthetic, u_s, records; and W_c is a weight used to adjust the trade-off between fitting the data and satisfying the constraints. The function $E(\mathbf{M})$ is a trade-off between a cross correlation and the L_2 norm, considering signal shape as well as amplitude information. For Parkfield we use only the horizontal ground-motion records and set all $W_i = 1.0$. $E(\mathbf{M})$ is scale invariant with each record contributing equally. Similar to the method of Hartzell and Heaton (1983) two constraints are applied: smoothness and target seismic moment. The target seismic moment constraint functions similarly to a moment minimization constraint and is only strongly enforced when the difference between model and target seismic moment is more than 15%. The target seismic moment used for Parkfield is 9.0×10^{24} dyne cm (Liu *et al.*, 2006).

Comparison of Inversion Results

In this section we investigate the sensitivity of the Parkfield earthquake slip inversion to several different data and model parameters for which choices or assumptions must be made in any kinematic slip inversion. Through this process we determine which parameters are important to the solution and which have relative insensitivity. As we have noted previously, our conclusions are dependent on the distribution of data, which is one of the best for the moderate-sized Parkfield earthquake. Our initial comparisons are made using the subfault-based linear least-squares method of Hartzell and Heaton (1983) because of its low CPU requirements and ease of application. Unless otherwise stated, the following inversion results are based on the strong-motion velocity records band-pass filtered from 0.2 to 2.0 Hz, with five rupture intervals, each having a triangle slip rate function with a duration of 0.2 sec, a fixed rake of 180°, and the high Q model in Table 1 with Quaternary surface layers at designated sites. Subsequently, we will show that very similar results are obtained with the nonlinear global search method of Liu and Archuleta (2004) when an L_2 norm is used with the same bounds on the model parameters. The global search method will then be used to investigate other norms, inversions with nonlinear model parameters, and the standard deviations of model parameters.

Constant Rupture Velocity

Not every inversion methodology requires a constant rupture velocity, and we will explore the effects of removing this restriction later. However, insight can be gained from examining the changes to the solution with different constant rupture velocities. Figure 3 (frames a1, b1, and c1) shows slip distributions for three different rupture velocities that fall within the range reported for crustal earthquake slip models (Somerville *et al.*, 1999), and Table 4 gives the associated moments and L_2 norms for these models.

As the rupture velocity is increased, the source at the hypocenter stays nearly constant because the difference in timing of the subfaults near the hypocenter is small. However, the deeper slip to the northwest of the hypocenter changes significantly due to the larger accumulated timing difference at this distance from the hypocenter. Interestingly, there is not a simple relationship between source distances

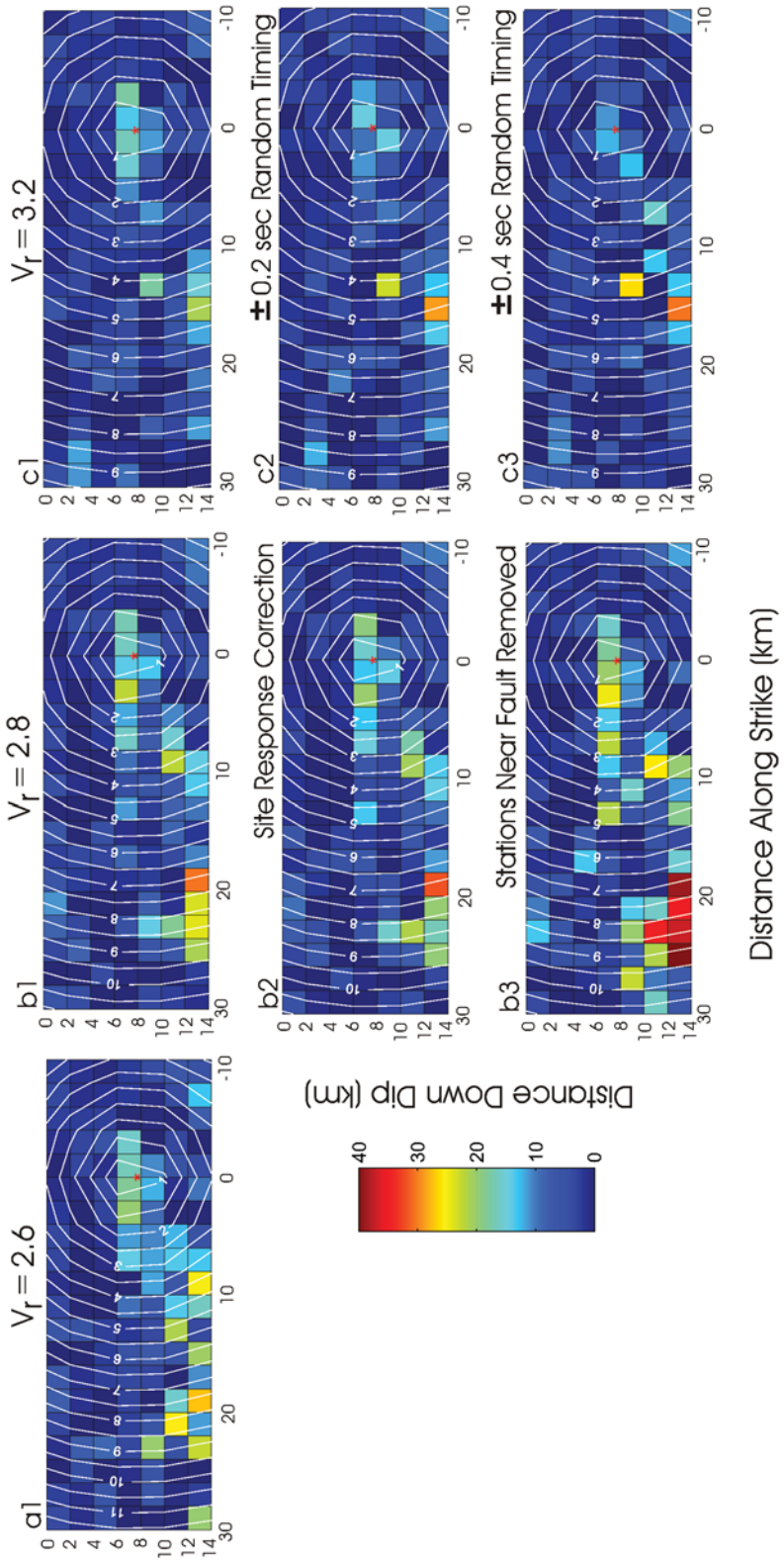

Figure 3. Comparison of different inversions for slip (in centimeters) using the linear least-squares method. Frames a1, b1, and c1 compare results for constant rupture velocities of 2.6, 2.8, and 3.2 km/sec, respectively. Frames b2 and b3 assume a rupture velocity of 2.8 km/sec and show the effects of correcting the Parkfield strong-motion records for site amplification and removing stations near the trace of the fault, respectively. Frames c2 and c3 assume a rupture velocity of 3.2 km/sec and show the effect of adding uniformly distributed random timing errors to the alignment of the data and synthetic waveforms. In this figure and following ones, the hypocenter is indicated by a red star, and rupture front contours are shown in white with units of seconds.

Table 4
Results for Linear Least-Squares Inversions, Different Constant Rupture Velocities

V_R (km/sec)	$\mathbf{M}_o$ (dyne cm)	$\|L_2\|$
2.6	1.06×10^{25}	42.03
2.8	0.94×10^{25}	41.49
3.2	0.69×10^{25}	41.38

down the fault and rupture velocity. One may expect sources to move outward, away from the hypocenter, as the rupture velocity increases to match fixed onset times of phases in the records. These shifts occur when the station distribution is poor, but for the dense station coverage at Parkfield, simple shifting of the sources is not allowed. The moment is a poorly defined quantity, decreasing with increasing rupture velocity, reflecting larger synthetic amplitudes from larger directivity effects. As the rupture velocity increases, the L_2 norm decreases. We conclude that rupture velocity has an important effect on the solution and for most cases should be considered a variable quantity (Spudich and Cranswick, 1984; Fletcher *et al.* 2006).

Alignment of Data and Synthetics

Part of the inversion process of waveform data, whether it is linear or nonlinear, involves the alignment of recorded data with synthetic waveforms. This step is usually accomplished by lining up the first break of the first-arriving *P* wave (for vertical records) or *S* wave (for horizontal records) with the first break of the *P* or *S* synthetic waveform from the hypocenter. Because the velocity structure is not known accurately enough, this alignment cannot generally be accomplished using theoretical travel times and can be subjective. We test the effect of random errors in timing by adding uniformly distributed random time shifts. These added errors in timing can also serve as a partial proxy for uncertainties in the velocity structure, recognizing that changes to the velocity structure would also affect the waveforms as well as the timing of phases. Figure 3 (frames c1, c2, and c3) and Table 5 show the results for different timing assumptions for the *S* phase. Three cases are considered: our best timing estimate, a uniformly distributed random error of ± 0.2 sec about our best estimate, and a uniformly distributed random error of ± 0.4 sec about our best estimate. As the random time shift increases, the definition of the source at the hypocenter decreases to a point that we consider unacceptable, and the L_2 norm increases significantly. The *S*-wave first breaks are clear enough for Parkfield that a misalignment of 0.2 sec between the data and the synthetic would be detectable at the majority of the Parkfield stations, and a misalignment of 0.4 sec would be detectable at nearly every station. Therefore, we do not consider misalignment of the data and synthetic records to be a significant issue for the Parkfield earthquake, but proper alignment may be a major source of error for other events with emergent or complex initiations.

Table 5
Results for Linear Least-Squares Inversions, Alignment of Data and Synthetics

Timing (sec)	$\mathbf{M}_o$ (dyne cm)	$\|L_2\|$
Best	0.69×10^{25}	41.38
± 0.2	0.70×10^{25}	42.52
± 0.4	0.66×10^{25}	43.31

Site Response

Finite-fault rupture inversions have typically been pursued without detailed consideration of site effects. If empirical Green's functions are used (Hartzell, 1978, 1989), they can incorporate site effects; however, the appropriate size, number, and distribution of empirical Green's functions are not always available. More commonly, site effects are approximated by using different 1D velocity structures for hard and soft soil sites. Liu *et al.* (2006) weighted the strong-motion records in the inversion by the inverse of a measure of the local site amplification. The local amplification was obtained by taking a spectral ratio between the Coalinga earthquake recording at the site with an average of the low-amplitude Coalinga earthquake recordings at Parkfield sites. In this way, sites with larger site response are given lower weight in the inversion.

In our study, we try an alternative approach that does not discount any data but attempts to correct it to a level that is more consistent with the calculated Green's functions. We also estimate site response using the Coalinga earthquake records. The digital U.S. Geological Survey (USGS) stations (the last 11 stations listed in Table 1) were not installed at the time of the Coalinga earthquake and are used in the inversion unaltered. For the remaining stations, we select station FZ6 (Fig. 1) on Cenozoic rock as our reference site. This station has one of the lowest-amplitude and simpler waveforms for the Coalinga earthquake. Site response spectra are then estimated by the ratio of the root mean square (rms) average of the two horizontal spectra (after correction for distance) at each site with respect to the corresponding spectrum at FZ6. The resulting spectra are close to a level of 1.0 below 0.7 Hz and rise at higher frequencies. To smooth over peaks and troughs, we calculate average levels in the frequency bands 0.7–1.0, 1.0–1.3, 1.3–1.6, and 1.6–2.0 Hz. A few of these average values reach a level of five, but most lie between 1 and 2. The data records are corrected by these average levels before inversion. No correction is made below 0.7 Hz. The results are shown in Figure 3 (frame b2). There is little difference with the corresponding inversion that makes no site response correction, Figure 3 (frame b1). This result was not expected but is probably due to the averaging effect of many strong-motion records with different site characteristics. The few stations with high site re-

sponse (Shakal *et al.*, 2006) are outweighed by numerous nearby stations with low site response. For an earthquake that is less well recorded, we would expect a few stations with anomalous amplitudes to be more influential. The moment, being a long-period parameter, is also relatively invariant, 0.91×10^{25} dyne cm with site response correction compared to 0.94×10^{25} dyne cm without. Similar conclusions are reached in our following sensitivity tests to near-surface velocity structure.

Winnowing of the Data

The sensitivity of the solution to the data set is initially tested by discarding 13 stations that lie near the trace of the modeled fault plane (FZ1, FZ3, FZ4, FZ6, FZ7, FZ12, FZ14, FZ15, C2W, GH1W, MFU, PRST, and SLAK) (Fig. 1). These stations represent 23% of the total data set. Later this question is treated in a statistically meaningful way using the bootstrap method. We select stations near the fault trace reasoning that these are the most ambiguous with respect to the appropriate velocity structure to be applied, northeast or southwest (Table 2). Figure 3 (frame b3) shows the results. The source at the hypocenter remains nearly unchanged, but the sources further down the fault to the northwest are affected more strongly. We conclude from this test that the solution is sensitive to the number of stations and their distribution, but given the large number of strong-motion records for the Parkfield earthquake, the removal of these particular stations still yields a similar spatial distribution of slip. This question is treated more rigorously in the following. However, the moment is significantly affected, having a value of 1.31×10^{25} dyne cm when the stations are removed.

Slip Rate Function

Another parameter that must be specified in kinematic fault inversions is the slip rate function. To this point our tests have used a 0.2-sec duration triangle with five possible slip windows each lagged from the previous one by 0.2 sec. To test the sensitivity of this choice we try an alternative slip rate function defined by a 0.4-sec triangle with five possible slip windows each lagged from the previous one by 0.2 sec. With this parameterization the total possible rise time is 1.2 sec, close to the previous maximum of 1.0 sec. In addition, the data is low passed at 1.0 Hz instead of 2.0 Hz. The results with this new slip rate function and filtering are shown in Figure 4 (frame a2). The slip distribution is nearly the same as the inversion using a 0.2 sec triangle (frame a1), and the moment is within 3%. The first node in the spectrum of a triangle occurs at $2/T$, where T is the duration of the triangle, which for our two choices is well above the frequency limit of the inversion. The two-parameter slip rate function of Liu and Archuleta (2004) also gives nearly identical results as a simple triangle. Therefore, our filtered data is insensitive to the particular form of the slip rate function with little impact on the results as long as the spectrum of the slip rate function is relatively flat below the corner frequency of the earthquake or the upper frequency limit of the inversion.

Shallow Velocity Structure

Uncertainties in the velocity structure have been shown to be a significant contributor to slip inversion error (Graves and Wald, 2001; Wald and Graves, 2001; Liu and Archuleta, 2004). A full investigation of velocity structure related errors is beyond the scope of this study; however, we performed limited testing of the near-surface velocity structure to assess its impact. Figure 4 (frame a3) shows the results of using the same high velocity surface structure at all the strong-motion sites (Tables 2 and 3), that is, no Quaternary layers are used. Figure 4 (frame a4) shows the effect of excluding the Quaternary surface layers as well as switching to the low Q attenuation model (type 2 in Table 2). Neither slip distribution deviates significantly from the case that includes Quaternary layers and the high Q model (Fig. 4, frame a1). Further, the moments of these models are within 13% of one another. These findings are consistent with our conclusions from the site response corrections. The natural frequencies of the Quaternary layers ($V_s/4h$, where h is layer thickness) are above 1 Hz, in a range of low power for the velocity waveforms. The insensitivity to Q is not surprising given the close proximity of the strong-motion stations to the source.

Subfault Size

To investigate the effect of the subfault size on the slip inversion, we cut the length and width of each subfault in half, from 2×2 km to 1×1 km (Fig. 4, frame b1). Comparing the result with the solution using larger subfaults (Fig. 4, frame a1), we see that the areas of significant slip have not moved and that the moments are nearly identical (0.96×10^{25} dyne cm for 1×1 km subfaults, and 0.94×10^{25} dyne cm for 2×2 km subfaults). However, there are two differences that are apparent with the smaller subfaults. First, the L_2 norm is significantly lower (38.80 for 1×1 km subfaults, and 41.49 for 2×2 km subfaults) because of the larger number of free parameters. Second, there is now slip on more randomly distributed subfaults, many of which are relatively isolated patches. We will give evidence in the following discussion using the bootstrap method that slip on these isolated subfaults is model dependent with a large uncertainty.

Inversion of Displacement Waveforms

Figure 4 (frame b2) shows the result of inverting displacement waveforms rather than velocity records. Filtering remains the same, band pass from 0.2 to 2.0 Hz, and the slip rate function also remains the same, 0.2-sec triangle over five time slices. The smaller subfaults are used with dimensions 1×1 km. Because of later long-period arrivals that are partly surface waves and partly double-integration

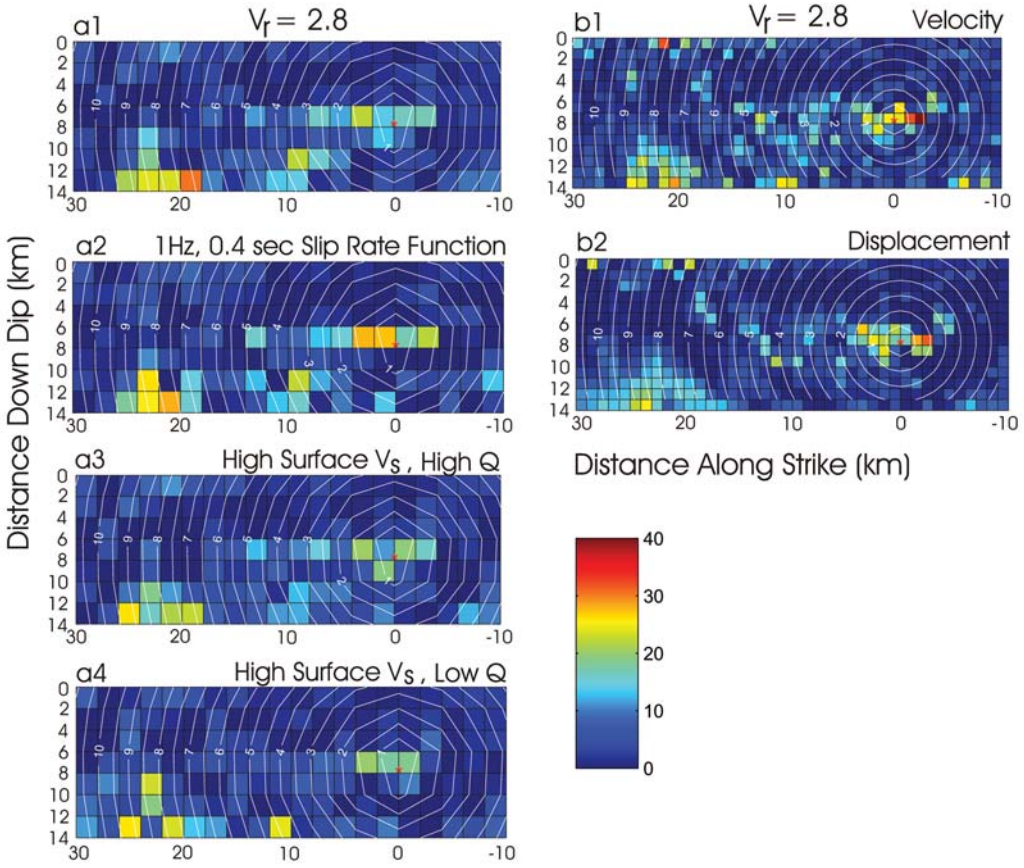

Figure 4. Comparison of different inversions for slip (in centimeters) using the linear least-squares method. All frames in this figure assume a rupture velocity of 2.8 km/sec. Frame a1 employs a 0.2-sec duration triangle slip rate function and strong-motion records low-pass filtered at 2 Hz. Frame a2 uses a 0.4-sec duration triangle slip rate function and strong-motion records low-pass filtered at 1 Hz. Frames a3 and a4 show the effects of using alternative velocity and Q models. Frames b1 and b2 compare results for the inversion of velocity and displacement waveforms and subfault sizes of 1 by 1 km.

processing noise not represented in our 1D synthetics, the displacement records are tapered after about 8 sec of *S*-wave arrival. The results are very similar to the corresponding velocity inversion in Figure 4 (frame b1) but with a smaller moment of 0.79×10^{25} dyne cm. The moment is sensitive to the duration of the untapered record length, which makes inversion of the displacement waveforms more problematic.

Comparison of Different Inversion Methodologies and Data

Figure 5a (frames a1 and a2) compares results of the linear least-squares subfault-based inversion with the nonlinear global search algorithm for an L_2 norm. In order to compare with the linear least-squares method, the global search algorithm is constrained to use a constant rupture velocity of 2.8 km/sec and a fixed rake of 180°. Plotting of the results for the two methods is different because the linear least-squares method assumes constant slip over subfaults, and the global search algorithm defines the model parameters at points on a grid with linear interpolation between these values. Furthermore, the global search method assumes zero slip for the grids on the bottom and sides of the fault. Given these distinct differences between the two methods, the results in Figure 5a are quite similar, in that both have large slip at the hypocenter, both have slip distributed along the deeper section of the fault to the northwest with maxima near 10 and 20 to 26 km. Both models also show the same shallow slip at 22 km. Comparisons were also run where the rake is allowed to vary between 90° and 180°, resulting in similar agreement between the linear and nonlinear methods in both rake angle as well as slip (not shown). The linear programming method of Bartels and Conn (1980) with a subfault parameterization also gives results close to the nonlinear global search method for an L_1 norm and a fixed rupture velocity with fixed or variable rake (not shown).

Our slip inversion results can be compared with the independent identification of source locations done by Fletcher *et al.* (2006). In their study, Fletcher *et al.* (2006) use the USGS Parkfield Dense Seismograph Array (UPSAR) to trace the movement of the rupture front. Plane-wave analysis is used to identify the sources of high-frequency arrivals at the array. Figure 5a (frame a3) shows their source locations

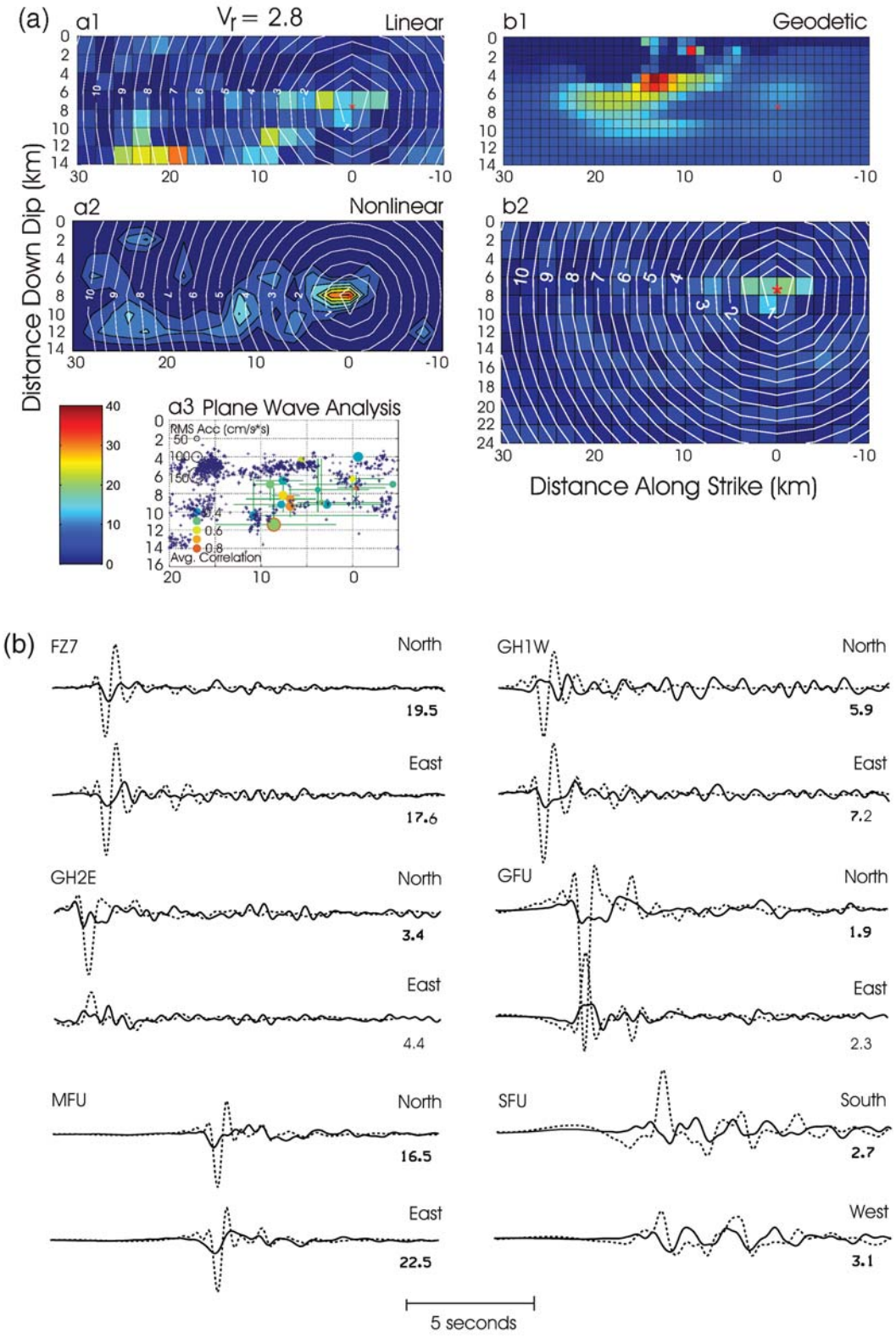

Figure 5. (a) Frames a1 and a2 compare linear least-squares and nonlinear global search slip inversion results for a constant rupture velocity of 2.8 km/sec and an L_2 norm. Frame a3 shows the primary sources identified by Fletcher *et al.* (2006) using plane-wave array analysis. Frame b1 is the slip obtained from the linear least-squares inversion of the geodetic data in Figure 2. Frame b2 shows the effect of extending the fault plane to a depth of 25 km for the linear least-squares inversion and a rupture velocity of 2.8 km/sec. (b) Prediction of strong-motion velocities (dashed curves) by the geodetic slip model in (a, frame b1) compared with the data records (solid curves) to a frequency of 2 Hz. The forward prediction of the strong motion assumes a constant rupture velocity of 2.8 km/sec. Data and synthetics are plotted on the same scale. Peak observed motion is given in cm/sec.

as circles, where the size of the circle is proportional to the source strength and the color indicates the degree of correlation of the arrivals. The location of their largest identified sources (three with the same location), indicated by the large green circle approximately 10 km northwest of the hypocenter at a depth of about 12 km, agrees with the location of the first deep source in the finite-fault slip inversion. Their analysis does not extend northwest to the distance of the other deep source, 20 to 25 km from the hypocenter; however, we have seen in Figure 3 that the positioning of these sources is dependent on the rupture velocity. Note also that the estimated error in the location of these sources is considerable (indicated by horizontal and vertical green bars). Fletcher *et al.* (2006) inferred that the rupture initiated at a speed of approximately $0.9V_s$ but slowed to a speed of $0.6V_s$ after about 2 sec. Such variations in rupture velocity cannot be encompassed by the linear inversion as parameterized here.

A further comparison can be made with the inversion of the independent geodetic data set (Fig. 2). A subfault-based linear least-squares inversion is employed with smoothing and minimization constraints. Static offsets are calculated using the layered half-space method of Xie and Yao (1989). The east–west components of displacement are downweighted by a factor of 0.5 relative to the north–south components because their smaller amplitudes are considered to be less reliably determined; however, results with and without this weighting are similar. The resulting slip distribution is shown in Figure 5a (frame b1). Similar results have been obtained by Langbein *et al.* (2006), Murray and Langbein (2006), and Johanson *et al.* (2006); however, the corresponding moment estimates have considerable scatter. These studies obtained moments of 1.1×10^{25}, 1.3×10^{25}, and 2.16×10^{25} dyne cm, respectively. Our preferred model yields an estimated moment of 1.38×10^{25} dyne cm, but solutions with moments twice as large have been obtained with no significant improvement in the fit to the data. Although large afterslip has been documented by the aforementioned authors, one key to obtaining the correct coseismic moment is the use of properly weighted stabilization constraints. The large range in moment that we observed for the Parkfield earthquake from geodetic data can be attributed to different weighting of these constraints.

Besides the difficulty of obtaining the correct moment from the geodetic data, the slip model has little resemblance to the ones based on strong-motion data. The one persistent source in the inversion of the seismic waveform data is the large slip at the hypocenter, located 10 to 15 km southeast of the slip maximum in the geodetic inversions. The geodetic inversions show little slip at the hypocenter. As we demonstrate in the following discussion, the large source at the hypocenter from the inversion of strong-motion data has the lowest uncertainty on the entire fault plane. The difference between the strong-motion and geodetic solutions may be due to the distribution of geodetic stations, which are concentrated near the center of the model fault plane (Fig. 2). Because there is no timing information, the geodetic offsets are most economically modeled by a fairly compact source directly under the station distribution. Figure 5b compares the prediction of the geodetic slip model at some of the poorer modeled strong motion stations located near the *SH* radiation pattern maxima of the large source in the geodetic solution. To make this comparison we have assumed a constant rupture velocity of 2.8 km/sec. Figure 5b shows that the geodetic solution does not fit the amplitudes, polarities, or pulse widths of the strong-motion records. Slip on the northern half of the fault plane may also have a significantly longer rise time that is not captured by the high-pass-filtered strong-motion data. As we develop further in the following discussion, the standard deviation of the slip is high on this part of the fault plane from the inversion of strong-motion data alone. Joint inversions of different data sets will result in weighted averages of the slip models based on the separate data sets. Each data set has its unique errors and uncertainties, which are usually poorly known, making it difficult to select an unbiased weighting scheme for all the data. We do not address the issue of joint inversions in the present study.

Model Fault Size

We observe that much of the slip to the northwest of the hypocenter in the models presented to this point lies near the bottom of the model fault plane between a depth of 10 and 14 km. To test the effect of the size of the fault plane, we extend the depth of the fault to 25 km, using the same subfault-based linear inversion as in Figure 5a (frame a1). The result is shown in Figure 5a (frame b2). The prominent source at the hypocenter remains unchanged, but slip on the rest of the fault is not represented by distinct sources but by lower-amplitude diffused slip. Although there is still no significant slip below 14.5 km (the maximum depth of the previous models), there are also no other strong sources. This dramatic result implies that slip on the northwestern half of the fault is highly model dependent. As we show in the following discussion, the choice of the inversion norm significantly affects the solution on this section of the fault.

Comparison of Different Inversion Norms and Model Parameterizations

Figures 6, 7, and 8 compare results for the three inversion norms, L_2, L_1, and the hybrid cross-correlation/L_2 norm (referred to hereafter as the cross-correlation norm), respectively. All examples use the nonlinear global search method to facilitate comparisons between norms and to allow us to lift restrictions on rupture parameters. For each norm the effects of relaxing the constraints on the model parameters of rake, rupture velocity (V_R), and rise time (T_R) are illustrated. The moments and corresponding norms are given in Tables 6, 7, and 8. The final residual errors can be compared only within a given norm because of differences in norm calculation.

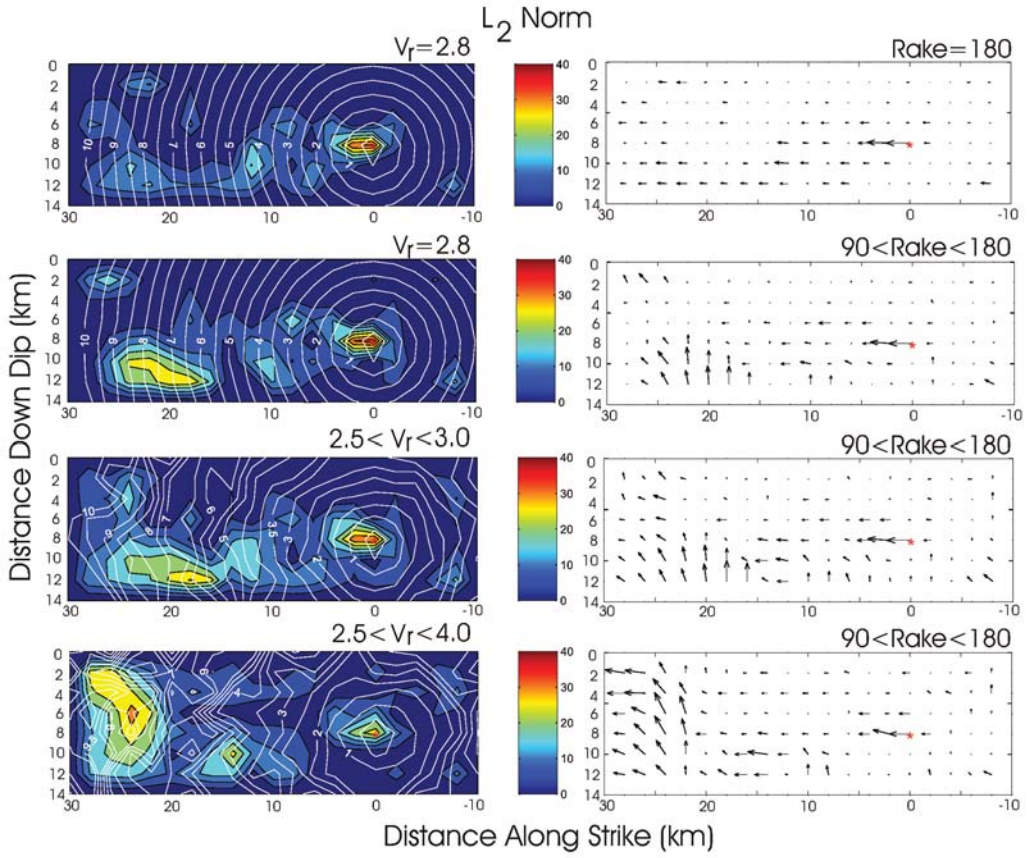

Figure 6. Comparison of different inversions for slip (left, in centimeters) and rake (right) using the nonlinear global search method and an L_2 norm. Each inversion has different limits on the model parameters: rupture velocity, rake, and rise time. In this figure and following figures the length of the rake vector is proportional to the slip.

The L_1 and L_2 results track one another closely in terms of slip, rake, and rupture velocity as the constraints on the rupture parameters are relaxed. The cross-correlation norm shows similar rupture timing but noteworthy differences in slip and rake angle from the L_1 and L_2 solutions. The source at the hypocenter is common to all three norms, but sources further to the northwest for the cross-correlation norm lie at shallower depths, near the depth of the hypocenter. The deep sources for the L_1 and L_2 norms are primarily dip-slip faulting (for the models that have rake a variable), which is difficult to justify in the Parkfield strike-slip regime. Including the vertical components of ground motion may add more control to the dip-slip component of motion, but we note that Liu *et al.* (2006) also obtained significant amounts of non-strike-slip faulting using the vertical components. In addition, the *SV* phase is down by a factor of 5 on average for the *SH* phase, with considerable complexity. The cross-correlation norm shows some dip-slip faulting, but it is lower in amplitude and occurs near the northwestern end of the fault, occurring as perhaps a fault-end effect. For this reason we favor the results of the cross-correlation norm. Besides the differences in slip distribution introduced by the relaxation of a fixed rake, which are most prominent for the L_1 and L_2 norms, bounds on the rupture velocity also have an important effect on the solution. These results point out the weakness of putting forth one particular slip model as the best, which is often done. Earthquake sources are better characterized in terms of general features that are robust under reasonable model permutations. However, to indicate the degree to which the data can be fit, Figure 9 compares data and synthetic velocity waveforms to a frequency of 2 Hz for the cross-correlation norm solution in Figure 8 with $2.5 \leq V_R \leq 4.0$ km/sec and $90° \leq$ rake $\leq 180°$. The waveforms are very similar to those of Liu *et al.* (2006).

Variable Rupture Velocity and Cross Validation

Tables 6, 7, and 8 show that residual errors decrease as the bounds on the model rupture parameters are expanded and, in particular, as the upper limit on rupture velocity is raised. However, this fact alone is not enough to say that the rupture velocity is high or that it is supershear. In our models (Figs. 6, 7, and 8) there is a region of high rupture velocity initiating at the hypocenter and traversing the fault plane up and to the southeast. The velocities in this region

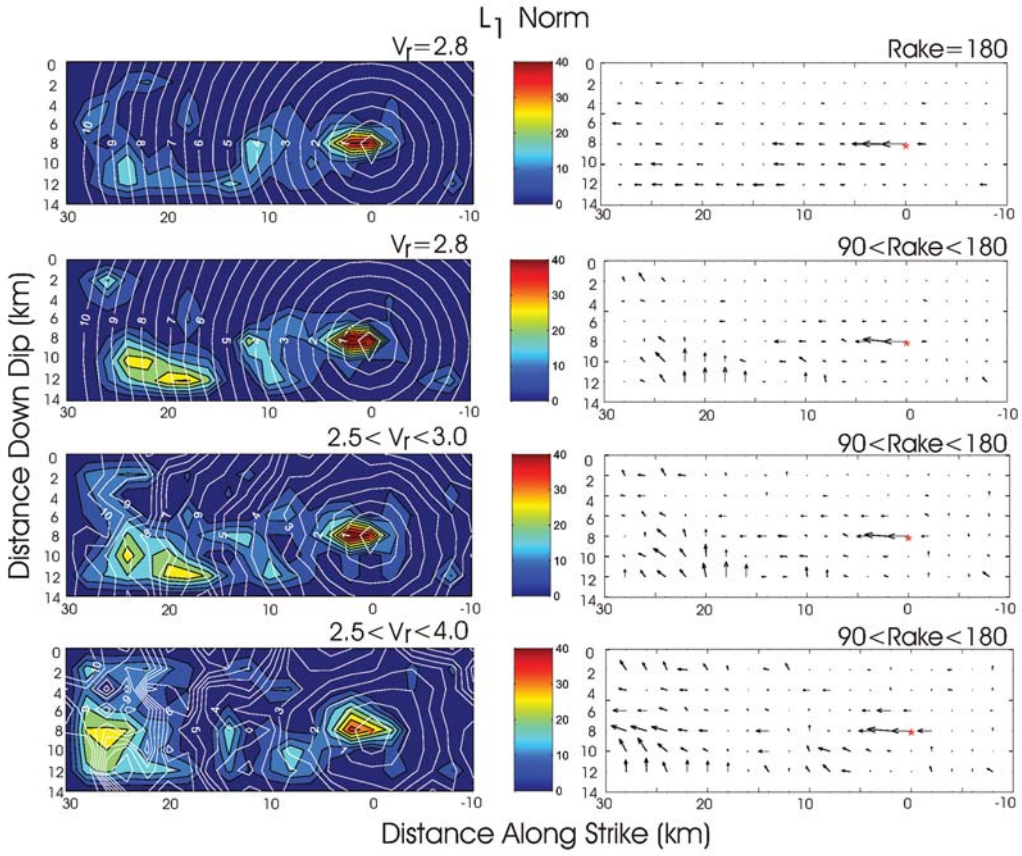

Figure 7. Comparison of different inversions for slip (left, in centimeters) and rake (right) using the nonlinear global search method and an L_1 norm. Each inversion has different limits on the model parameters: rupture velocity, rake, and rise time.

increase past the local shear-wave velocity as the upper bound on rupture velocity is raised. The inversion favors this solution in order to fit the large amplitudes along the Cholame Limb of the Parkfield strong-motion array (Fig. 1, stations C12W, C6W, C5W, C4W, C3W, C1E, C2E, and C3E). However, the improvement in the fit is marginal, and the large amplitude on the Cholame Limb may be due to structural effects that are not in our Green's functions. Cross validation is an unbiased method of assessing a change in model parameterization (Hansen, 1998; Chernick, 1999). We use this technique to investigate the upper limit on the average rupture velocity from the hypocenter, the rupture parameter in the code of Liu and Archuleta (2004). The cross-correlation norm is used with $90° \leq \text{rake} \leq 180°$ and $0.2 \leq T_R \leq 2.0$ sec.

In cross validation, data are omitted from the inversion, and a solution is obtained using the remaining data. This solution is then used to predict the omitted data, and the resulting error at the predicted stations is calculated. For evaluation of the upper bound on rupture velocity, the technique assumes that although high rupture velocities may be advantageous to fitting some records, they will simultaneously be damaging to the fit at other stations. We use three different subdivisions of the stations: two using spatial groupings of the stations, illustrated in Figure 10, and one random selection process. The spatial groupings in Figure 10a and 10b divide the available strong-motion stations into 6 and 10 regions, respectively. These subdivisions are motivated by the fact that rupture directivity effects have strong directional dependence. A similar station subdivision to Figure 10b was used by Custodio *et al.* (2005). The third case selects 12 stations at random to be excluded from the inversion. For the two spatial groupings, 6 and 10 independent inversions are run, respectively, leaving out a different set of stations each time and summing the prediction errors. For the randomly selected stations, 25 independent inversions are run. These sets of inversions are repeated using different upper bounds on the rupture velocity. Figure 11 plots the resulting sum of the errors at the predicted stations versus the upper bound on rupture velocity. All three station selection processes yield consistent results. The prediction errors oscillate below an upper bound on rupture velocity of about 3.6 km/sec. Above 3.6 km/sec, the errors increase. These results suggest that the rupture did not exceed 3.6 km/sec over a significant portion of the fault surface.

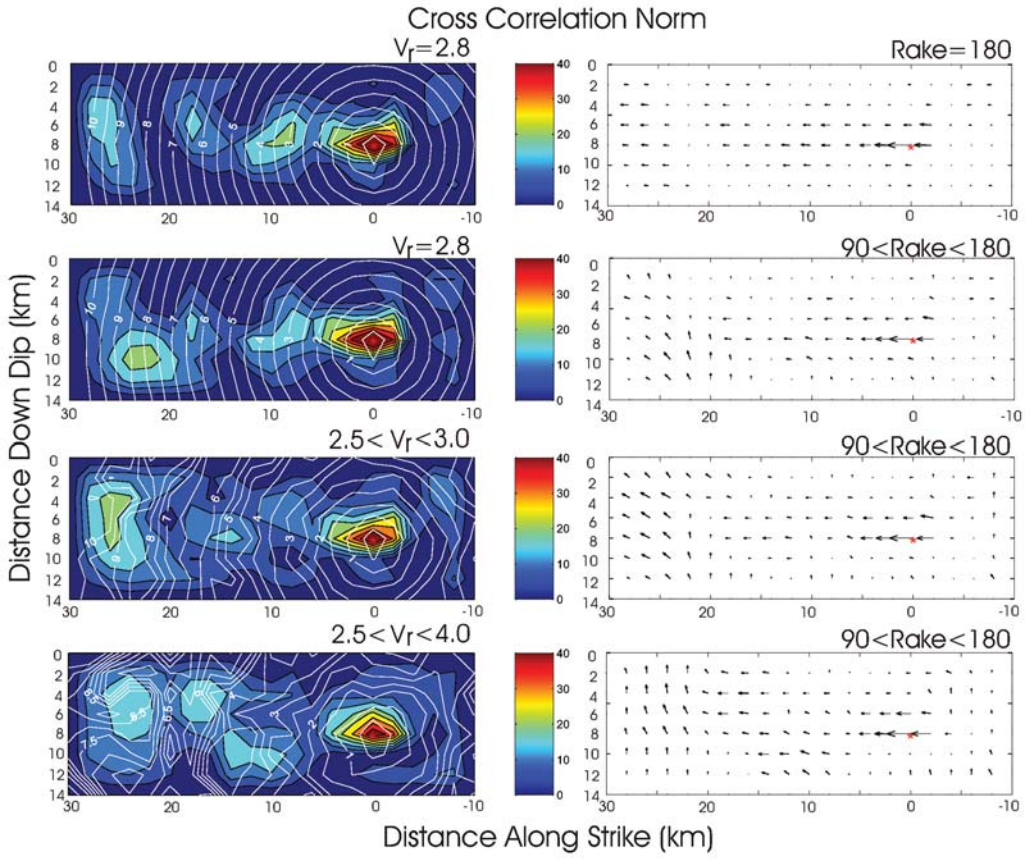

Figure 8. Comparison of different inversions for slip (left, in centimeters) and rake (right) using the nonlinear global search method and the cross-correlation norm. Each inversion has different limits on the model parameters: rupture velocity, rake, and rise time.

Mean and Standard Deviation of Model Parameters and the Bootstrap Method

Two large sources of error in finite-fault slip inversions are inaccurate or inappropriate Green's functions and noisy and inadequate data. This article does not address velocity structure issues except for near-surface effects. However, uncertainty issues associated with the data can be estimated using the bootstrap method (Chernick, 1999). The bootstrap method involves a resampling from the original data set. In our implementation, we use a balanced resampling with replacement, in which each datum or record occurs the same number of times in the bootstrap procedure but may be repeated within any one inversion. These requirements are easily satisfied by generating a string of random integers with values from one to n of length nm, where n is the number of stations with data and m is the number of inversions to be performed. The minimum number of inversions required to obtain a statistically meaningful standard deviation is 100 (Chernick, 1999). Therefore, $n = 54$ and $m = 100$. The 100 inversions then use successive 54-station blocks designated by the random string.

The mean slip model and standard deviations of the rupture parameters from the bootstrap are shown in Figures 12 and 13 for the cross-correlation and L_2 norms, respectively, using the nonlinear global search method. These two bootstraps have the following model parameter search bounds: $90^{\circ} \leq \text{rake} \leq 180^{\circ}$, $2.5 \leq V_R \leq 4.0$ km/ sec, and $0.2 \leq T_R \leq 2.0$ sec. The important features to note include the stability of the source at the hypocenter with its relatively low standard deviations in rupture parameters and the broad distributed slip to the northeast of the hypocenter with its relatively large standard deviations in rupture parameters. The large uncertainty in slip on the northwestern part of the fault plane may be unexpected given the large amount of strong-motion data for the Parkfield earthquake and the many finite-fault inversions that have been presented to date for Parkfield and other earthquakes with relative certainty. However, most finite-fault studies do not consider the large statistical sampling of inversions in this article. Custodio *et al.* (2005) in their investigation of the Parkfield earthquake obtain a second source approximately 10 km northeast of the hypocenter; however, the standard deviation in slip based on 12 inversions is nearly as large as the slip itself. In addition, we note that the rupture velocity for the mean models stays below about 3.5 km/sec except for a small region around the hypocenter. As we pointed out previously from the cross

Figure 9. Comparison of synthetic (red) and data (blue) velocity waveforms (cm/sec) to a frequency of 2 Hz for the cross-correlation norm solution in Figure 8 with model parameter bounds 2.5 ≤ rupture velocity ≤ 4.0 km/sec and 90° ≤ rake ≤ 180°. Data and synthetics are plotted on the same scale. (*Continued*)

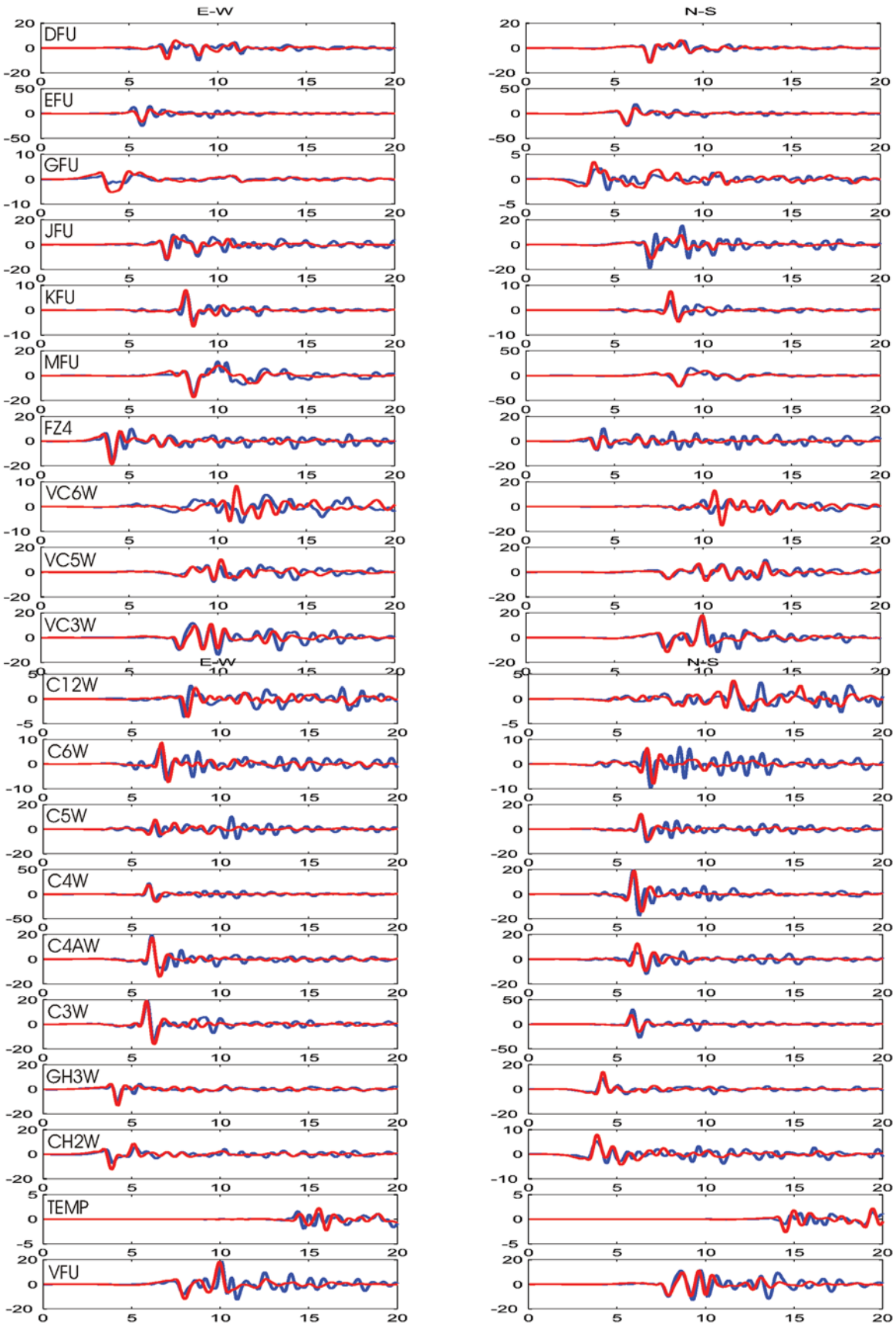

Figure 9. Continued.

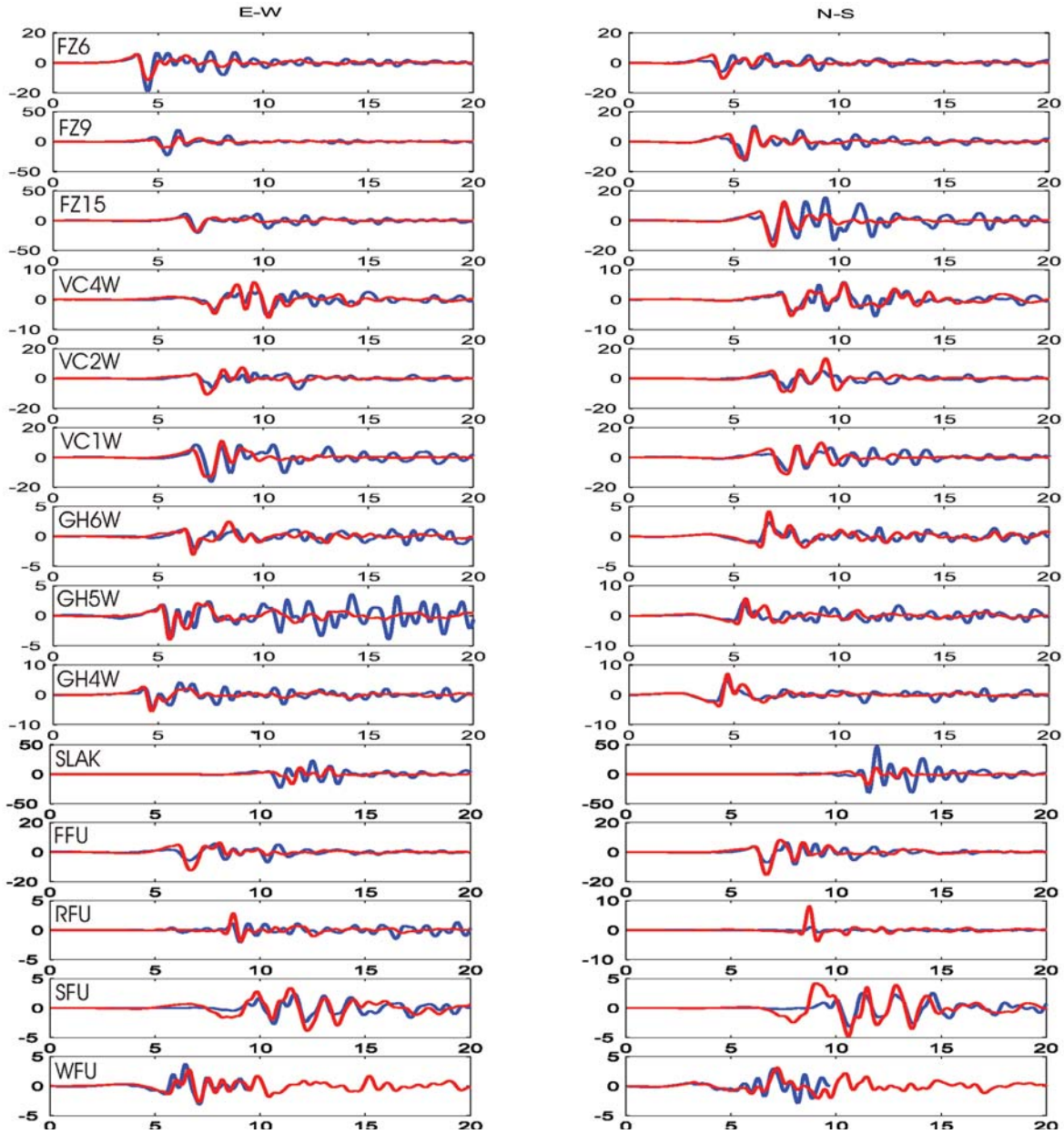

Figure 9. Continued.

validation tests, the high velocities near the hypocenter may be related to inadequacies in the Green's functions.

To test the hypothesis of lower uncertainty in slip with the addition of *a priori* constraints on the model parameters, we set the upper limit on the rupture velocity to 3.6 km/sec based on the cross validation results and restrict the rake to 180° based on tectonic considerations. Figure 14 shows the bootstrap results for these more limiting constraints and the

Table 6

Results for Nonlinear Global Search Inversions, L_2 Norm

Model Parameterization	Moment (dyne cm)	$\|L_2\|$
$V_R = 2.8$, rake $= 180$, $0.2 \le T_R \le 1.0$	0.78×10^{25}	1.64
$V_R = 2.8$, $90 \le$ rake ≤ 180, $0.2 \le T_R \le 1.0$	0.81×10^{25}	1.58
$2.5 \le V_R \le 3.0$, $90 \le$ rake ≤ 180, $0.2 \le T_R \le 1.0$	1.03×10^{25}	1.50
$2.5 \le V_R \le 4.0$, $90 \le$ rake ≤ 180, $0.2 \le T_R \le 2.0$	1.03×10^{25}	1.45

Table 7
Results for Nonlinear Global Search Inversions, L_1 Norm

Model Parameterization	Moment (dyne cm)	$\|L_1\|$
$V_R = 2.8$, rake $= 180$, $0.2 \le T_R \le 1.0$	0.78×10^{25}	3.66
$V_R = 2.8$, $90 \le$ rake ≤ 180, $0.2 \le T_R \le 1.0$	0.81×10^{25}	3.59
$2.5 \le V_R \le 3.0$, $90 \le$ rake ≤ 180, $0.2 \le T_R \le 1.0$	0.95×10^{25}	3.50
$2.5 \le V_R \le 4.0$, $90 \le$ rake ≤ 180, $0.2 \le T_R \le 2.0$	0.95×10^{25}	3.48

cross-correlation norm. Comparing the results with those in Figure 12, we see that, other than the change in rake angle, the mean slip distribution and standard deviation of the slip are nearly the same. The mean distributions and standard deviations in rise time are also very similar. However, the standard deviation in rupture velocity is now approximately four to five times lower with the upper bound on the rupture velocity of 3.6 km/sec compared with an upper bound of 4.0 km/sec. Thus, although the uncertainty in slip is not reduced, an *a priori* small change in the upper bound on rupture velocity has resulted in a significant change in its uncertainty. This result points out the importance of *a priori* model assumptions.

Discussion and Conclusions

Care must be taken in finite-fault slip inversions to avoid making overreaching conclusions because of the strong dependence of the inversions on the distribution, number, and type of data. For the 2004 Parkfield earthquake, there is an extensive near-field strong-motion data set; indeed, it is perhaps the best coverage to date for a moderate-sized earthquake. Therefore, in this regard, our conclusions about uncertainty of faulting parameters for the Parkfield earthquake may represent the most optimistic case. However, even this statement needs to be tempered by the fact that the M 6.0 Parkfield earthquake lies at the lower magnitude boundary of events for which finite-fault slip inversions are typically performed. An M 6.5 or larger earthquake may have yielded a more robust solution with the same number of stations. The earthquake's small size puts particular limitations on the inversion process. No useable teleseismic waveforms are available, except for a handful of short-period records, which do not provide useful constraints on the details of the source process. In addition, the ground-motion records are a higher frequency than are encountered in most finite-fault inversions and, as such, push the limits of the modeling process. With these caveats in place, we summarize the modeling experience for the 2004 Parkfield earthquake and the ramifications for other finite-fault inversions.

First, the slip inversion is strongly dependent on the rupture velocity. For methodologies that use a constant rupture velocity, slip is redistributed for different values of rupture velocity. Higher rupture velocities can push slip further out from the hypocenter to preserve the arrival times of phases. However, for Parkfield the changing pattern of slip is more complicated because the high density of stations does not allow a simple shift in the slip. Methodologies that allow for a variable rupture velocity are preferable to ones that use a constant value. However, modelers must be aware that rupture velocity can trade off with deficiencies in other areas of the model. Olson and Anderson (1988) showed that strong trade-offs can exist between the slip amplitude and the directivity effects of the propagating rupture. These effects can be overcome only by an appropriately positioned data set. For Parkfield, large rupture velocities, even supershear values, are favored by large amplitude records on the Cholame Limb. These large values were discounted, however, by a cross validation process that judges the degradation to other stations. Large amplitude ground motion can also be the result of the focusing of rays by the velocity structure. If these structural components are not in the model, then rupture velocity may act as a proxy.

Second, the inversion norm has a significant effect. This conclusion may be surprising only because most finite-fault inversion studies do not investigate multiple norms. Hartzell *et al.* (1991) demonstrated that important differences existed between L_1 and L_2 norm inversions for the slip during the 1989 Loma Prieta earthquake. Das and Suhadolc (1996) compared results from least-squares singular value decomposition and an L_1 inversion with different constraints and preferred the L_1 solutions. For Parkfield some clear differences

Table 8
Results for Nonlinear Global Search Inversions, Cross-Correlation Norm

Model Parameterization	Moment (dyne cm)	Cross-Correlation
$V_R = 2.8$, rake $= 180$, $0.2 \le T_R \le 1.0$	1.03×10^{25}	0.555
$V_R = 2.8$, $90 \le$ rake ≤ 180, $0.2 \le T_R \le 1.0$	1.03×10^{25}	0.547
$2.5 \le V_R \le 3.0$, $90 \le$ rake ≤ 180, $0.2 \le T_R \le 1.0$	1.03×10^{25}	0.511
$2.5 \le V_R \le 4.0$, $90 \le$ rake ≤ 180, $0.2 \le T_R \le 2.0$	1.03×10^{25}	0.483

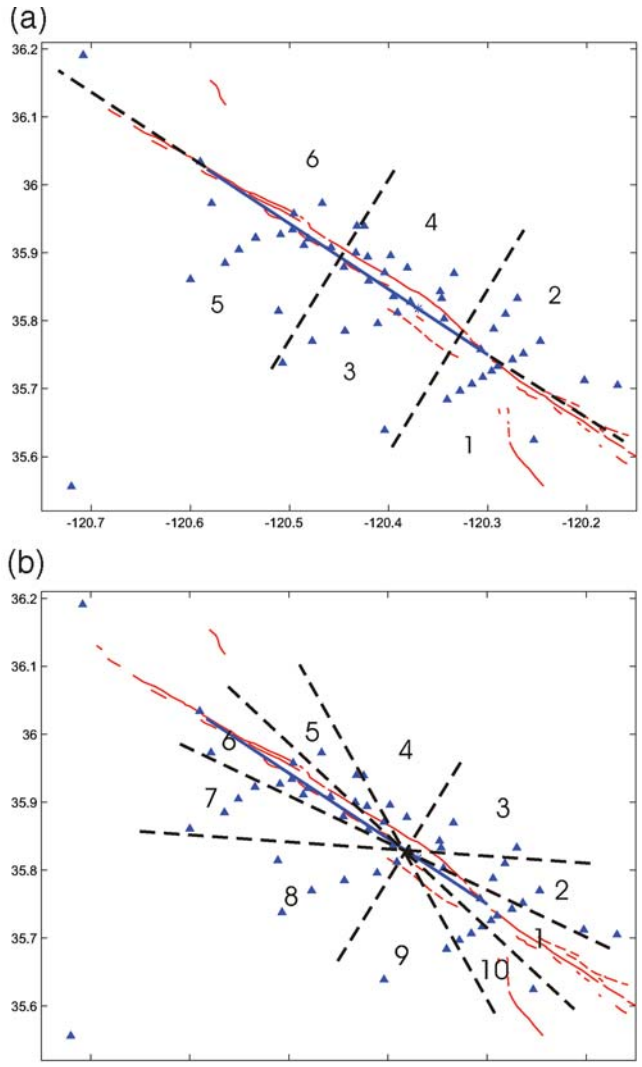

Figure 10. Two subdivisions of the strong-motion stations used with the cross validation method to estimate the maximum limit on average rupture velocity from the hypocenter.

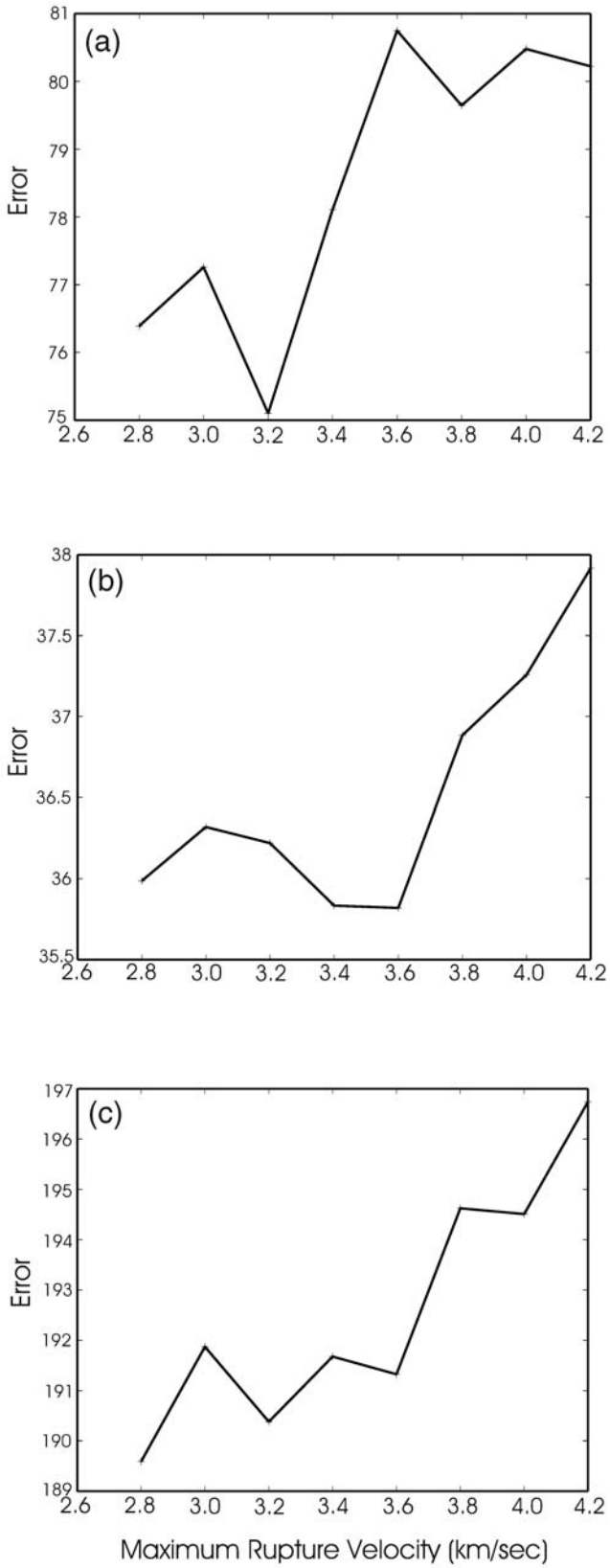

Figure 11. Prediction cross validation error as a function of the upper limit on average rupture velocity from the hypocenter for the subdivision of the strong-motion stations into the (a) 6 and (b) 10 sets, shown in Figure 10, and (c) randomly selected stations.

were noted between the three norms considered. The L_1 and L_2 solutions tracked each other more closely than the cross-correlation norm. Slip northwest of the hypocenter tends to be deeper for the L_1 and L_2 norms and is composed of more dip-slip motion. With the cross-correlation norm, slip on the northwest part of the fault plane tends to remain close to the depth of the large source at the hypocenter, and dip-slip faulting is generally limited to the northern end of the fault with smaller amplitudes. For these reasons, we favor the cross-correlation norm results. However, caution must be exercised in extending this conclusion to other finite-fault inversion problems. Hartzell *et al.* (1996) obtained similar slip histories for the 1994 Northridge earthquake using the linear, subfault-based, least-squares technique used in this study and a nonlinear global search method using the cross-correlation norm also in this study. Clearly, more experience is

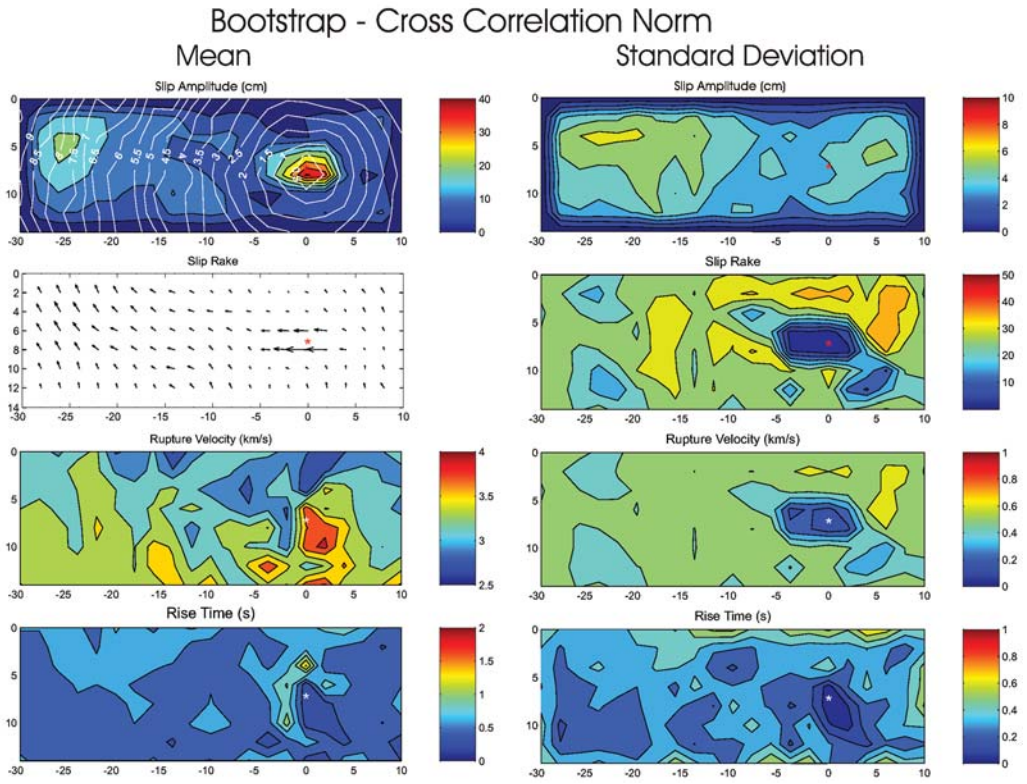

Figure 12. Mean slip model and standard deviation of model parameters from 100 bootstraps of the strong-motion data using the cross-correlation norm.

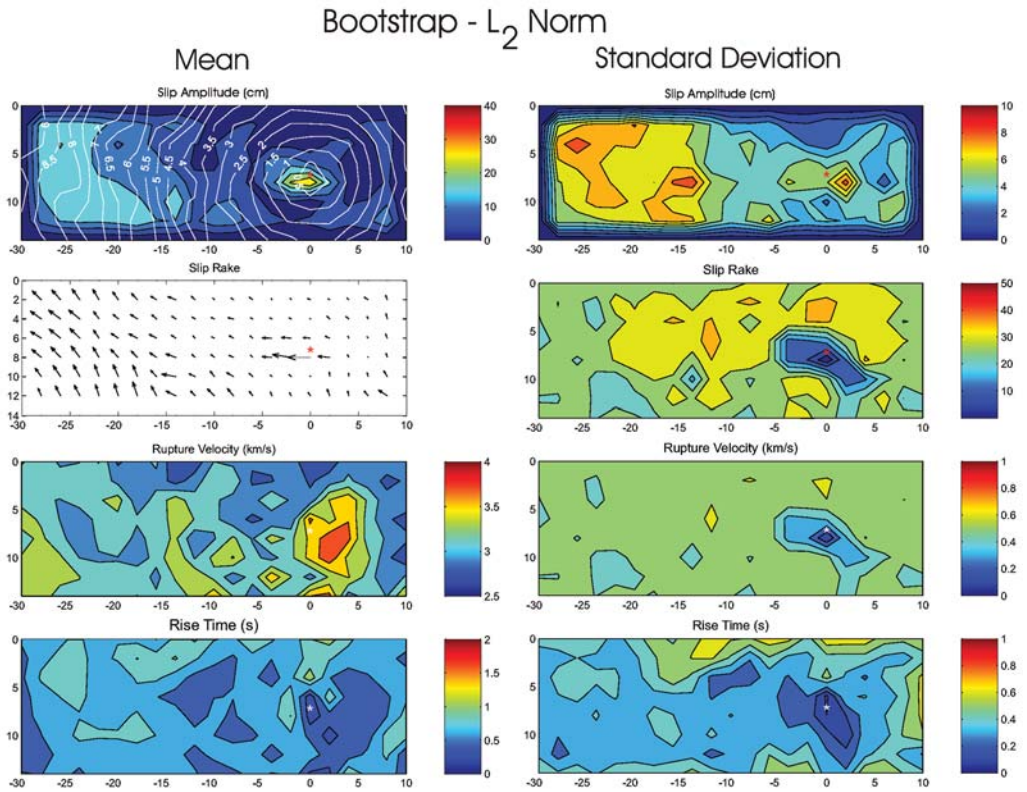

Figure 13. Mean slip model and standard deviation of model parameters from 100 bootstraps of the strong-motion data using an L_2 norm.

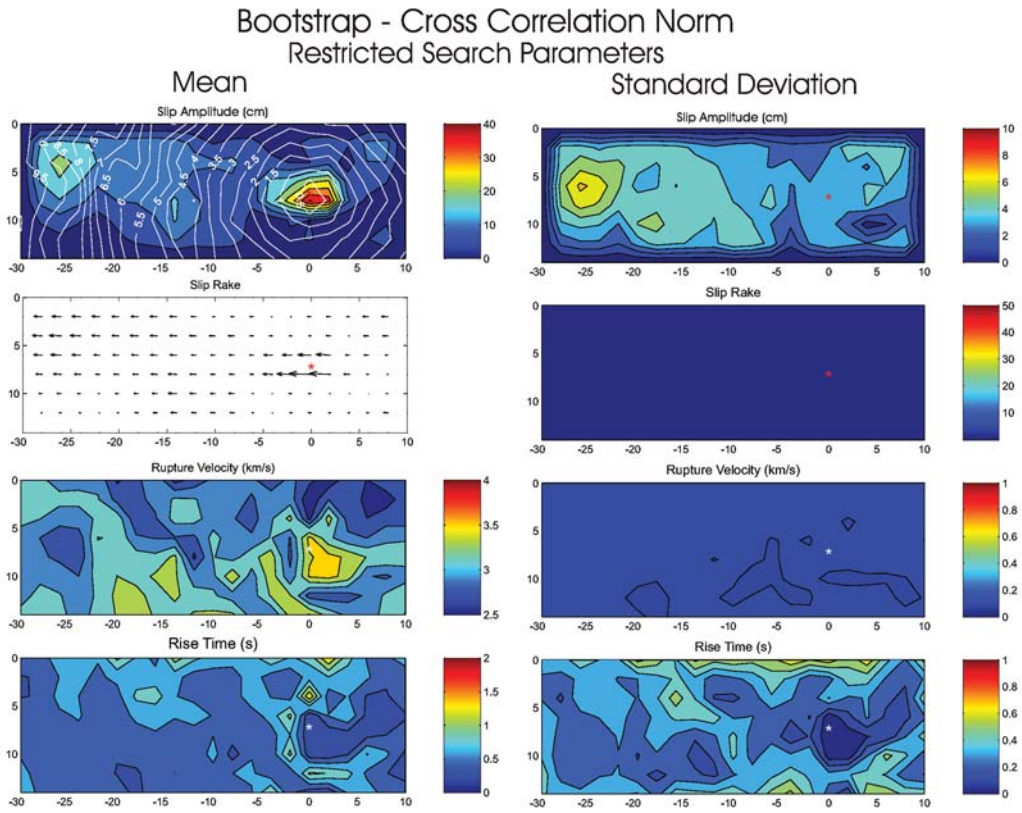

Figure 14. Mean slip model and standard deviation of model parameters from 100 bootstraps of the strong-motion data using the cross-correlation norm and an upper limit of 3.6 km/sec for the rupture velocity and a fixed rake of 180º.

needed in comparing inversion norms for the earthquake problem.

Third, the limits on the model parameters have a significant impact on the outcome. We have shown that very similar results are obtained for different model parameterizations and methodologies if the same limits are imposed on the model parameters and the same inversion norm is used. However, independent of the methodology adopted, as the limits on the model parameters are expanded, the solution changes significantly. We have discussed these changes for the Parkfield earthquake with the upper limit on rupture velocity and suggested how cross validation might be used to put estimates on this parameter. The slip history was also shown to be sensitive to the limits put on the rake angle as well as the area of the fault plane. These conclusions are not surprising and emphasize the importance that constraints from other data sources, such as teleseismic waveforms or tectonic considerations, can have on the finite-fault inversion process.

Fourth, errors associated with the data set, resulting from recording or processing noise or incompatibilities between the Green's functions and the data, should not be understated. We saw through application of the bootstrap method that slip on the northwestern half of the Parkfield rupture has the largest standard deviation. Varying the data set randomly over 100 inversions migrates the slip from one location to another, without well-defined source locations. This result does not reflect a deficiency in the inversion methodology but rather a basic limitation of the data/Green's functions and would be different for each application. Also, alignment of the first arrival of the phase to be modeled in the synthetics with the data records is an important step regardless of the methodology used but was accomplished with confidence in the case of the Parkfield earthquake.

Finally, we can summarize features of the slip inversion problem for the Parkfield earthquake that have a much lower impact on the solution. In the inversion of velocity or displacement records, low-pass filtered below 2.0 Hz, reasonable variations in the velocity structure in the top 100 m have little effect on the result. The low-pass-filtered data we use cannot distinguish between slip rate functions of similar durations, such as a triangle or the two-parameter function of Liu and Archuleta (2004). With the use of stabilizing constraints, including moment minimization and smoothing, the major slip features of the model are shown to be insensitive to a factor of 4 change in the area of the subfaults.

Acknowledgments

The first author benefited from discussions with Rowena Lohman and David Perkins on the topics of cross validation and the bootstrap method.

The manuscript was significantly improved by comments from Art Frankel, Chuck Mueller, Joe Fletcher, and an anonymous reviewer.

References

Bakun, W. H., and A. G. Lindh (1985). The Parkfield, California, earthquake prediction experiment, *Science* **229,** 619–624.

Bakun, W. H., B. Aagaard, B. Dost, W. L. Ellsworth, J. L. Hardebeck, R. A. Harris, C. Ji, M. J. S. Johnston, J. Langbein, J. J. Lienkaemper, A. J. Michael, J. R. Murray, R. M. Nadeau, P. A. Reasenberg, M. S. Reichle, E. A. Roeoffs, A. Shakal, R. W. Simpson, and F. Waldhauser (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake, *Nature* **437,** 969–974.

Bartels, R. H., and A. R. Conn (1980). Linearly constrained discrete L_1 problems, *ACM Trans. Math. Softw.* **6,** 594–608.

Bilich, A. (2006). Improving the precision and accuracy of geodetic GPS: applications to multipath and seismology, *Ph.D. Thesis*, University of Colorado.

Borcherdt, R. D., M. J. S. Johnston, G. Glassmoyer, and C. Dietel (2006). Recordings of the 2004 Parkfield earthquake on the general earthquake observation system array: implications for earthquake precursors, fault rupture, and coseismic strain changes, *Bull. Seismol. Soc. Am.* **96,** S73–S89.

Chernick, M. R. (1999). *Bootstrap Method, A Practitioner's Guide*, Wiley, New York.

Choi, K. H., A. Bilich, K. M. Larson, and P. Axelrad (2004). Modified sidereal filtering: implications for high-rate GPS positioning, *Geophys. Res. Lett.* **31,** no. 22, L22608.

Cohee, B. P., and G. C. Beroza (1994). A comparison of two methods for earthquake source inversion using strong motion seismograms, *Ann. Geofis.* **37,** 1515–1538.

Custodio, S., P.-C. Liu, and R. J. Archuleta (2005). The 2004 $\mathbf{M_w}$6.0 Parkfield, California, earthquake: inversion of near-source ground motion using multiple data sets, *Geophys. Res. Lett.* **32,** L23312.

Das, S., and B. V. Kostrov (1990). Inversion for seismic slip rate history and distribution with stabilizing constraints: application to the 1986 Andreanof Islands earthquake, *J. Geophys. Res.* **95,** 6899–6913.

Das, S., and B. V. Kostrov (1994). Diversity of solutions of the problem of earthquake faulting inversion: application to SH waves for the great 1989 Macquarie Ridge earthquake, *Phys. Earth Planet. Interiors* **85,** 293–318.

Das, S., and P. Suhadolc (1996). On the inverse problem for earthquake rupture: the Haskell-type source model, *J. Geophys. Res.* **101,** 5725–5738.

Eberhart-Phillips, D., and A. J. Michael (1993). Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield region, central California, *J. Geophys. Res.* **98,** 15,737–15,758.

Fletcher, J. B., P. Spudich, and L. M. Baker (2006). Rupture propagation of the 2004 Parkfield, California, earthquake from observations at the UPSAR, *Bull. Seismol. Soc. Am.* **96,** S129–S142.

Graves, R. W., and D. J. Wald (2001). Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions, I. Strong motions, *J. Geophys. Res.* **106,** 8745–8766.

Hansen, Per C. (1998). *Rank-Deficient and Discrete Ill-Posed Problems, Numerical Aspects of Linear Inversion*, Society for Industrial and Applied Mathematics, Philadelphia.

Hartzell, S. (1978). Earthquake aftershocks as Green's functions, *Geophys. Res. Lett.* **5,** 1–4.

Hartzell, S. (1989). Comparison of seismic waveform inversion results for the rupture history of a finite fault: application to the 1986 North Palm Springs, California, earthquake, *J. Geophys. Res.* **94,** 7515–7534.

Hartzell, S. H., and T. H. Heaton (1983). Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake, *Bull. Seismol. Soc. Am.* **73,** 1553–1583.

Hartzell, S., P.-C. Liu, and C. Mendoza (1996). The 1994 Northridge, California, earthquake: investigation of rupture velocity, risetime, and high-frequency radiation, *J. Geophys. Res.* **101,** 20,091–20,108.

Hartzell, S. H., G. S. Stewart, and C. Mendoza (1991). Comparison of L_1 and L_2 norms in a teleseismic waveform inversion for the slip history of the Loma Prieta, California, earthquake, *Bull. Seismol. Soc. Am.* **81,** 1518–1539.

Haskell, N. A. (1969). Elastic displacements in the near-field of a propagating fault, *Bull. Seismol. Soc. Am.* **59,** 865–908.

Ji, C., D. J. Wald, and D. V. Helmberger (2002). Source description of the 1999 Hector Mine, California, earthquake, part I: Wavelet domain inversion theory and resolution analysis, *Bull. Seismol. Soc. Am.* **92,** 1192–1207.

Johanson, I. A., E. J. Fielding, F. Rolandone, and R. Burgmann (2006). Coseismic and portseismic slip of the 2004 Parkfield earthquake from space-geodetic data, *Bull. Seismol. Soc. Am.* **96,** S269–S282.

Langbein, J., J. R. Murray, and H. A. Snyder (2006). Coseismic and initial postseismic deformation from the 2004 Parkfield, California, earthquake, observed by Global Positioning System, electronic distance meter, creepmeters, and borehole strainmeters, *Bull. Seismol. Soc. Am.* **96,** S304–S320.

Langbein, J., R. Borcherdt, D. Dreger, J. Fletcher, J. L. Hardebeck, M. Hellweg, C. Ji, M. Johnston, J. R. Murray, and R. Nadeau (2005). Preliminary report on the 28 September 2004, M 6.0 Parkfield, California earthquake, *Seism. Res. Lett.* **76,** 10–26.

Larson, K. M., P. Bodin, and J. Gomberg (2003). Using 1-Hz GPS data to measure deformations caused by the Denali fault earthquake, *Science* **300,** 1421–1424.

Lawson, C., and R. Hanson (1974). *Solving Least Squares Problems*, Prentice Hall, New Jersey.

Liu, P.-C., and R. J. Archuleta (2004). A new nonlinear finite fault inversion with three-dimensional Green's functions: application to the 1989 Loma Prieta, California, earthquake, *J. Geophys. Res.* **109,** B02318.

Liu, P.-C., S. Custodio, and R. J. Archuleta (2006). Kinematic inversion of the 2004 **M** 6.0 Parkfield earthquake including an approximation to site effects, *Bull. Seismol. Soc. Am.* **96,** S143–S158.

Menke, W. (1984). *Geophysical Data Analysis: Discrete Inverse Theory*, Academic, Orlando, Florida.

Murray, J., and J. Langbein (2006). Slip on the San Andreas fault at Parkfield, California, over two earthquake cycles and the implications for seismic hazard, *Bull. Seismol. Soc. Am.* **96,** S283–S303.

Olson, A. H., and J. G. Anderson (1988). Implications of frequency-domain inversion of earthquake ground motions for resolving the space-time dependence of slip on an extended fault, *Geophys. J.* **94,** 443–455.

Olson, A. H., and R. J. Apsel (1982). Finite faults and inverse theory with applications to the 1979 Imperial Valley earthquake, *Bull. Sesimol. Soc. Am.* **72,** 1969–2001.

Piatanesi, A., A. Cirella, P. Spudich, and M. Cocco (2007). A global search inversion for earthquake kinematic rupture history: application to the 2000 western Tottori, Japan earthquake, *J. Geophys. Res.* **112,** B07314.

Sekiguchi, H., K. Irikura, and T. Iwata (2000). Fault geometry at the rupture termination of the 1995 Hyogo-ken Nanbu earthquake, *Bull. Seismol. Soc. Am.* **90,** 117–133.

Sen, M. K., and P. L. Stoffa (1991). Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics* **56,** 1624–1638.

Shakal, A., H. Haddadi, V. Graizer, K. Lin, and M. Huang (2006). Some key features of the strong-motion data from the **M** 6.0 Parkfield, California, earthquake of 28 September 2004, *Bull. Seismol. Soc. Am.* **96,** S90–S118.

Somerville, P., K. Irikura, R. Graves, S. Sawada, D. Wald, N. Abrahamson, Y. Iwasaki, T. Kagawa, N. Smith, and A. Kowada (1999). Characterizing crustal earthquake slip models for the prediction of strong ground motion, *Seism. Res. Lett.* **70,** 59–80.

Spudich, P., and E. Cranswick (1984). Direct observation of rupture propagation during the 1979 Imperial Valley earthquake using a short baseline accelerometer array, *Bull. Seismol. Soc. Am.* **74,** 2083–2114.

Thurber, C., S. Roecher, H. Zhang, S. Baher, and W. Ellsworth (2004). Fine-scale structure of the San Andreas fault and location of the SAFOD target earthquakes, *Geophys. Res. Lett.* **31,** L12S02.

Thurber, C., H. Zhang, F. Waldhauser, J. Hardebeck, A. Michael, and D. Eberhart-Phillips (2006). Three-dimensional compressional wave-speed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region, *Bull. Seismol. Soc. Am.* **96,** S38–S49.

Trifunac, M. D. (1974). A three-dimensional dislocation model for the San Fernando, California, earthquake of February 9, 1971, *Bull. Sesimol. Soc. Am.* **64,** 149–172.

Wald, D. J., and R. W. Graves (2001). Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions, II. Combining seismic and geodetic data, *J. Geophys. Res.* **106,** 8767–8788.

Wdowinski, S., Y. Bock, J. Zhang, P. Fang, and J. Genrich (1997). Southern California permanent GPS geodetic array: spatial filtering of daily positions for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake, *J. Geophys. Res.* **102,** 18,057–18,070.

Xie, X., and Z. X. Yao (1989). A generalized reflection-transmission coefficient matrix method to calculate static displacement field of a dislocation source in a stratified half space, *Chin. J. Geophys.* **32,** 191–205.

Zeng, Y., and J. G. Anderson (1996). A composite source modeling of the 1994 Northridge earthquake using genetic algorithm, *Bull. Seismol. Soc. Am.* **86,** S71–S83.

Zhu, L., and L. A. Rivera (2002). A note on the dynamic and static displacements from a point source in multilayered media, *Geophys. J. Int.* **148,** 619–627.

U.S. Geological Survey
Denver Federal Center
Box 25046, MS 966
Denver, Colorado 80225
(S.H.)

U.S. Bureau of Reclamation
Denver Federal Center
Box 25007, MS 86-68330
Denver, Colorado 80225-0007
(P.L.)

Centro de Geociencias
UNAM Campus Juriquilla
Apartado Postal 1-742
76001 Queretaro, Oro. Mexico
(C.M.)

Department of Earth Science
University of California
Santa Barbara, California, 93117
(C.J.)

Department of Aerospace Engineering Sciences
University of Colorado
Boulder, Colorado, 80309-0429
(K.M.L.)

Manuscript received 23 March 2007