Bulletin of the Seismological Society of America, Vol. 97, No. 6, pp. 1935–1949, December 2007, doi: 10.1785/0120060203

Modeling Variable-Stress Distribution with the Stochastic Finite-Fault Technique

by Karen Assatourians and Gail M. Atkinson

Abstract The stochastic finite-fault ground-motion modeling technique is modified to simulate the effects of a variable-stress parameter on the fault. The radiated source spectrum of each subsource that comprises the fault plane is multiplied by a correction spectrum that leaves the low-frequency portion of the spectrum intact and multiplies the high-frequency end of the spectrum by a constant proportional to the stress parameter of each subfault raised to the power of $2/3$; this scaling behavior follows from the Brune source model. The modification causes the response spectra and time series of simulated traces to be sensitive to the stress parameter distribution on the fault surface.

The approach is implemented using an inversion tool that effectively inverts observed response spectral data to derive the stress parameter distribution on the fault surface. It applies the Levenberg–Marquardt nonlinear inversion method to minimize differences of average (log) response spectra ordinates at high frequencies between observations and simulations at all stations.

We perform a number of experiments to study the effects of fault-dip angle, iterations per station, initial guess of the stress distribution, and station distribution on the capabilities of the inversion tool. We also evaluate the ability of the inversion tool to resolve the relative stress parameters of multiple asperities.

Application of the inversion tool to the data of the **M** 6 2004 Parkfield earthquake indicates that an asperity with a high stress parameter is located in the southeast end of the fault, at a depth greater than 4 km; another asperity is located in the center of the fault, but with a lower stress parameter. This distribution is in agreement with results by other researchers.

Introduction

Finite-fault modeling is an important tool for the prediction of ground motion near large earthquakes (Hartzell, 1978; Irikura, 1983; Heaton and Hartzell, 1986; Joyner and Boore, 1986; Somerville *et al.*, 1991; Tumarkin and Archuleta, 1994; Zeng *et al.*, 1994; Beresnev and Atkinson, 1998). The typical approach is to simulate the ground motion for a large earthquake by simulating the rupture of several small earthquakes as subevents that comprise a large fault-rupture event (introduced by Hartzell [1978] and tested by Kanamori [1979]; Irikura [1983]; Heaton and Hartzell [1989]). In the stochastic finite-fault method, the subsources are modeled as stochastic point sources (Boore, 1983, 2003); the method multiplies a theoretical and/or empirical spectrum of source, path, and site effects with the spectrum of band limited windowed Gaussian noise. Transformation to the time domain produces an earthquake time series (accelerogram). By summing properly scaled point source simulations in the time domain, an extended fault-plane source can be modeled. Silva *et al.* (1990) combined an ω^2 point source with an empirical attenuation and duration model to construct a finite-fault radiation model. In similar studies, Chin and Aki (1991) and Schneider *et al.* (1993) followed Silva *et al.* in their choice of source and path models. Beresnev and Atkinson (1999, 2002) also used a similar source and path model but explored the implications of subsource size on the radiated spectrum. Zeng *et al.*, (1994) and Yu *et al.* (1995) used an ω^2 point source with synthetic path effects for modeling finite-fault radiation. Motazedian and Atkinson (2005) introduced a new variation based on a "dynamic corner frequency." In this implementation, the corner frequency is a function of time, and the rupture history controls the frequency content of the simulated time series of each subfault. The rupture begins with a high corner frequency and progresses to lower corner frequencies as the ruptured area grows. Limiting the number of active subfaults (the percentage of pulsing area, which is the maximum possible area of the fault radiating seismic waves divided by the total area of the fault) in the calculation of the dynamic corner frequency influences the spectral shape at intermediate frequencies. The dynamic corner frequency concept allows the simulations to

be independent of subfault size and provides for conservation of seismic moment. Motazedian and Atkinson (2005) developed a FORTRAN code for this approach based on the earlier program FINSIM by Beresnev and Atkinson (1998); the new program is called EXSIM. Like most other stochastic finite-fault approaches, EXSIM does not assign a directivity effect to individual subfaults, but the effects of rupture propagation along the fault (from subfault to subfault) are assumed to mimic the overall directivity effect.

The Motazedian and Atkinson (2005) stochastic finite-fault method uses the Brune (1970, 1971) source model as the theoretical form of the subsource spectrum. This introduces the effect of the stress parameter of the subsource on the high-frequency level of the source spectrum, because the high-frequency spectral level scales as $M_0^{1/3}\Delta\sigma^{2/3}$, where M_0 is the seismic moment and $\Delta\sigma$ is the stress parameter (Boore, 1983, 2003). The Motazedian and Atkinson model assigns a uniform stress parameter to all subsources; thus it does not consider the effects of patches of high stress, or asperities (Aki, 1984). We modified the existing version of the stochastic finite-fault modeling code EXSIM so that it can consider the effect of a variable-stress parameter on the fault surface. To our knowledge, this is the first application of variable-stress modeling with the stochastic finite-fault modeling technique. We also considered modeling the effects of variable rupture velocity and/or variable slip within the code. Forward modeling test results showed that variable rupture velocity has only a slight effect on the predicted response spectra at some stations and negligible effect at others. Thus, we ignore variable rupture velocity. Variable slip acts in a similar way to variable stress in that it enhances predicted ground-motion amplitudes in regions near high-slip patches; in the context of the simple stochastic model, variable slip and variable stress are essentially interchangeable. We chose to parameterize the effect through variable stress rather than slip, due to the direct relationship between stress and high-frequency ground-motion scaling, which is a fundamental feature of the stochastic approach. This is a simplification; it is acknowledged that multiple effects, including slip distribution, may be folded into the inferred stress distribution.

The underpinning of the modification to the stochastic finite-fault method is the development of an analytical form that multiplies the spectrum of subsources in such a way as to leave the low-frequency level of the spectra intact (as the low-frequency level of the spectrum depends on the seismic moment of the subsource and is independent of the stress parameter); the high-frequency level is amplified such that it is proportional to the subfault stress parameter raised to the power of $2/3$ (as per the Brune model). The sensitivity of the response spectra of simulated signals to stress distribution allows development of an inversion tool to derive the stress distribution on the fault surface, based on the response spectra of available records. The inversion problem is a multiparameter nonlinear one, where the stress parameters of subsources are the parameters; it is solved iteratively by the least-squares estimation method of Levenberg–Marquardt (Marquardt, 1963; Lee and Stewart, 1981; Draper and Smith, 1998). The measure of error in the inversion process is the difference between the observed and the simulated logarithm of response spectra, averaged over the high frequency part of the spectrum.

Our approach is to first assign uniform slip and stress parameters to the fault, using the Motazedian and Atkinson (2005) simulation method (EXSIM), and find their optimum average values for a specific event (based on modeling of the recorded ground motions). Then, by keeping the slip uniformly distributed, we allow the inversion program to search for the stress distribution that provides the best possible match of the response spectra of simulations to the response spectra of observations, at high frequencies. The outcome of this approach might be leakage of some slip (moment) effect into stress space, because there is no exact boundary between the low- and high-frequency spectral ranges. Because the high-frequency spectral level is mainly affected by the stress parameter, we consider the slip (or moment) leakage issue to be unimportant.

Variable-Stress Parameter in Stochastic Finite-Fault Modeling

Our modeling approach follows the stochastic finite-fault method of Motazedian and Atkinson (2005), except that we assign weights to subfaults such that the stress parameter on each subfault is the product of the weight and the average stress parameter of the fault. We assume that the following conditions should be satisfied: (1) the low-frequency portion of the spectrum should be independent of stress, because the moment determines the zero-frequency level of the spectrum; (2) as per the Brune source model, the high-frequency amplitude level should be proportional to the subevent stress parameter raised to the power of $2/3$ (Brune, 1970; Boore, 1983, 2003); (3) if the stress parameter of a subfault approaches zero, the contribution of that subfault to the time history should approach zero ($\Delta\sigma_{ij} \rightarrow 0 \Rightarrow w_{ij} \rightarrow 0 \Rightarrow W_{ij} \rightarrow 0 \Rightarrow X_{ij}(f) \rightarrow 0$); and (4) if the stress parameter of a subfault equals the average value for the fault, the subfault stress correction factor (X_{ij}) should be unity.

A functional form that multiplies the spectral components of the ijth subfault, that satisfies the preceding conditions, is

$$X_{ij}(f) = \frac{1 + f^2/f_{C_{ij}}^2}{1 + f^2/(W_{ij} \times f_{C_{ij}}^2)}, \tag{1}$$

where $X_{ij}(f)$ is the multiplicative subfault stress correction factor for a subfault in the ith column and jth row, f is the frequency, f_{Cij} is the corner frequency of the subfault as derived per Motazedian and Atkinson (2005). In equation (1), W_{ij} is the weighting factor:

$$W_{ij} = \frac{N w_{ij}^{2/3}}{\sum_i \sum_j w_{ij}^{2/3}}, \tag{2}$$

where N is the total number of subfaults and w_{ij} is the ratio of a subfault's stress parameter to the average stress parameter of the fault (i and j are the indexes of the column and row of the subfault, respectively):

$$w_{ij} = \frac{\Delta\sigma_{ij}}{\overline{\Delta\sigma}}. \tag{3}$$

Here $\Delta\sigma_{ij}$ is the stress parameter of a subfault and $\overline{\Delta\sigma}$ is the average stress parameter of the fault. Based on the preceding correction factors, the time history of the simulated trace will be in the following form:

$$a(t) = \sum_{i,j} F^{-1}(X_{ij}(f) \times A_{ij}(f)), \tag{4}$$

where $a(t)$ is the accelerogram of the entire fault, $X_{ij}(f)$ is the spectrum stress correction factor calculated in equation (1), $A_{ij}(f)$ is the subfault spectrum, and F^{-1} denotes the inverse Fourier transform.

To demonstrate the performance of this spectral modification to the stochastic finite-fault method, the peak ground accelerations (PGAs) for a grid of points around a hypothetical fault (selected to have the Parkfield fault geometry, for reasons that will become apparent later in the article when we study this event) are simulated, as shown in Figure 1. The stress parameter distribution is assigned to be uniform except for a subfault at the southeast tip of the fault (mid-depth) that is assigned a stress parameter five times larger than the specified average of 50 bars. A uniform soil type with average $V_{S30} = 620$ m/sec (a generic California rock site, Boore and Joyner [1997]) is considered for all sites on the grid. The moment magnitude of the event is assigned as **M** 6.0. The high PGAs are concentrated around the high-stress-parameter end of the fault, as expected.

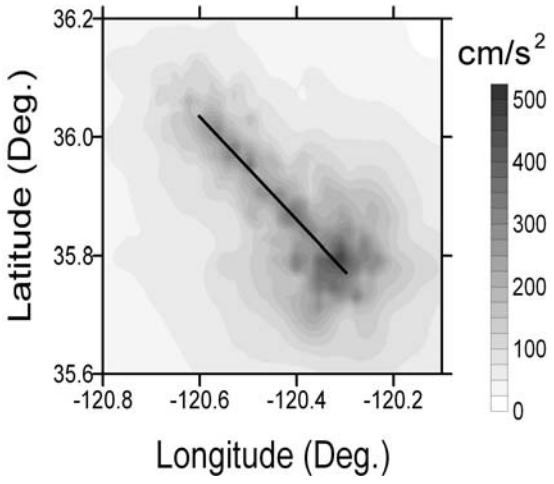

Figure 1. PGA around a fault with the Parkfield fault geometry and M_w 6.0. The stress distribution is assumed to be uniform except for a high-stress subfault at mid row in the southeast tip of the fault.

Because the high-frequency components of the accelerograms recorded close to high-stress-parameter subfaults are amplified, it should be possible to use records of nearby stations to infer the stress parameter distribution on the fault surface during an earthquake based on this model.

Development of an Inversion Tool for Stress Distribution

This section describes the inversion tool used to find the best stress distribution on the fault. First, we introduce the error measures used in the formulation of the inversion process. Then we describe the approach to inverting average log pseudospectral acceleration (PSA) residuals (observed log PSA minus simulated log PSA) to obtain the stress parameter distribution on the fault surface. We demonstrate an inversion shortcut that is based on assuming a functional form for the objective function to be minimized and finding its coefficients empirically. The rationale for the shortcut is to reduce the number of EXSIM simulations (and the processing time) necessary for the construction of the Jacobian matrix (Lee and Stewart, 1981) in each iteration of the inversion.

Error Measurements to Evaluate Model Performance

The inversion process used in this study finds an optimum stress distribution on the fault surface that minimizes the difference between response spectra (PSAs) of observed and simulated accelerograms. These differences are defined by error equations. Two types of error measures are used, both of which utilize the logarithm of PSA ordinates to avoid potential biases arising from pronounced local effects that may occur at some stations. The first measure, Ξ, is the average value of Ξ_j over all stations, where Ξ_j is the average absolute error of the log PSA ordinates of the simulation, with respect to the observation at station j:

$$\Xi_j = \frac{1}{N}\sum_{i=1}^{N} |\log \mathrm{PSA}(O_{ij}) - \log \mathrm{PSA}(S_{ij})|, \tag{5}$$

where $\mathrm{PSA}(S_{ij})$ is the 5%-damped pseudoacceleration ordinate of a simulated signal at the ith frequency for the jth station, $\mathrm{PSA}(O_{ij})$ is the corresponding PSA ordinate of the observed signal, and N is the total number of frequencies. Averaging over K stations results in

$$\Xi = \frac{1}{K}\sum_{j=1}^{K} \Xi_j. \tag{6}$$

The reason for separating the Ξ calculation into two equations is that the number of spectral frequencies (N) is not equal for all stations. This measure is used in the beginning of any inversion iteration to find the appropriate average stress parameter value.

An alternative definition of average error is Ξ_i, where instead of summing the error over frequencies as in equa-

tion (5), it is done over stations. The result is an average error spectrum:

$$\Xi_i = \frac{1}{K}\sum_{j=1}^{K} |\log \mathrm{PSA}(O_{ij}) - \log \mathrm{PSA}(S_{ij})|. \qquad (7)$$

This measure is used to demonstrate how the error evolves as a function of frequency during the inversion (e.g., Fig. 14).

The second error measure, Θ, is an average of square values of ε_j over all stations, where ε_j is the average error of log PSA ordinates of the simulation compared to the observation over a specified frequency range at station j. Calculation of ε_j is performed for frequencies higher than double the earthquake corner frequency, between frequency numbers N_1 to N_2, as

$$\varepsilon_j = \frac{-1}{(N_2 - N_1 + 1)}\sum_{i=N_1}^{N_2} [\log \mathrm{PSA}(O_{ij}) - \log \mathrm{PSA}(S_{ij})], \qquad (8)$$

where N_1 and N_2 are the lower and upper ordinate number of the considered frequency range, respectively. The lower frequency (N_1) is assumed to be about twice the value of the fault's corner frequency, while the upper frequency (N_2) is the maximum frequency available in the simulations and records. Our approach for the selection of N_2 was to choose the smallest of the maximum available frequencies among all records, which makes the programming easier. Our forward modeling experiments show that the average log PSAs obtained are not sensitive to the choice of frequency range; for example, if we restrict consideration to just very high frequencies (say, four times the corner frequency instead of two), we obtain very similar results. The negative sign in equation (8) is for consistency with forthcoming equations. Then,

$$\Theta = \frac{1}{K}\sum_{j=1}^{K} \varepsilon_j^2. \qquad (9)$$

The need for defining two sets of error measures arises due to different optimization approaches for the determination of the average stress parameter on the fault and the stress parameter distribution of subfaults. The determination of the average stress parameter follows an enumerative optimization method. In this method, combinations of possible parameters are searched over a grid of parameter values so as to result in a minimum objective function value. Ξ is a good candidate for the objective function; parameter values providing the minimum Ξ value are the best choice for all stations and all frequencies. The stress parameter distribution determination follows a calculus-based optimization method (Goldberg, 1988). In this method, we need the error at each station as well as its sign. The measure ε_j is a good candidate for this purpose and meets these conditions. As shown in the next section, the inversion tool formulation is based on the minimization of Θ.

Formulation of the Inversion Tool

The formulation of multivariable nonlinear inversion in seismology is a classic in the literature (Lee and Stewart, 1981; Menke, 1989; Tarantola, 1987; Parker, 1994; Tarantola, 2005). We follow the Lee and Stewart (1981) formulation, using the Levenberg–Marquardt inversion algorithm (Marquardt, 1963). This method is selected because it combines beneficial aspects of Gauss–Newton and gradient methods, while avoiding some of their weaknesses. The advantages of the selected approach are as follows: (1) the solution converges quickly, so we will be able to avoid many additional EXSIM simulations; and (2) in many cases an uninformed standard guess works well, so there is a high chance of a converging solution being found for an initial guess based on a uniform-stress distribution. This property of the Levenberg–Marquardt algorithm makes inclusion of additional information for an appropriate initial guess unnecessary.

The iterative optimization equation is

$$\delta(\Delta\sigma) = -[A^T A + \lambda I]^{-1} A^T \varepsilon, \qquad (10)$$

where $\delta(\Delta\sigma)$ is the stress parameter adjustment vector and ε is the residual vector obtained over a range of frequencies for the recording stations; its elements are calculated using equation (8). A is the Jacobian matrix, λ is Levenberg–Marquardt adjustable parameter, and I is the identity matrix. All elements refer to the current iteration. The Jacobian matrix is calculated using the following equation:

$$A = \begin{bmatrix} \frac{\partial \varepsilon_1}{\partial(\Delta\sigma)_1} & \frac{\partial \varepsilon_1}{\partial(\Delta\sigma)_2} & \cdots & \frac{\partial \varepsilon_1}{\partial(\Delta\sigma)_m} \\ \frac{\partial \varepsilon_2}{\partial(\Delta\sigma)_1} & \frac{\partial \varepsilon_2}{\partial(\Delta\sigma)_2} & \cdots & \frac{\partial \varepsilon_2}{\partial(\Delta\sigma)_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varepsilon_n}{\partial(\Delta\sigma)_1} & \frac{\partial \varepsilon_n}{\partial(\Delta\sigma)_2} & \cdots & \frac{\partial \varepsilon_n}{\partial(\Delta\sigma)_m} \end{bmatrix}. \qquad (11)$$

Each element of the Jacobian matrix gives the variation of the residual at a station, per unit variation of one subfault's stress parameter.

Initially, a guessed value of the stress parameter distribution is assumed; the uniform-stress distribution with the best average value for the fault is a good choice. Then the values of the Jacobian matrix elements, the errors at each station, and the overall error based on equation (9) are calculated. Also an initial value is assigned to λ, say 0.001. By substituting these values in equation (10), the correction stress parameter vector, $\delta(\Delta\sigma)$, is derived. By applying the stress correction to the initial guess, a new stress parameter distribution is derived and processed the same way. The procedure repeats until the stress distribution has converged to the best value. If the overall error in a given step Θ reduces, then the value of λ will be multiplied by a number less than unity, to provide for a larger weight of the Gauss–Newton method compared to the gradient method. Conver-

sely, if Θ increases in a step, then λ is multiplied by a number greater than unity, to provide larger weight to the gradient method compared to the Gauss–Newton method. We followed Press *et al.*, (1992) and used 10 and 1/10 for the multipliers of λ, for more than unity and for less than unity, respectively. The iterations repeat for some predefined number of times, or as long as the error reduction is significant. The calculation of the inverse matrix of equation (10) is performed using Gauss–Jordan elimination, and a subroutine by Press *et al.* (1992) is used for this purpose. The results are not sensitive to the selected method for this calculation.

Functional Form of the Objective Function and Jacobian Matrix Elements

The inversion process includes the calculation of elements of the Jacobian matrix, which is the derivative of the error at a site relative to the stress parameter of each individual subfault. Calculation of the matrix elements would require that the number of EXSIM runs be at least two times the number of subfaults, times the number of stations, for each iteration. This is a large number of EXSIM runs, which would make application of this approach time-consuming on a PC. We solve this problem as follows.

Each component of the Jacobian matrix is the derivative of a summation from equation (8) for a station with respect to the stress parameter; that is, $A_{jk} = \partial\varepsilon_j/\partial(\Delta\sigma)_k$, where ε_j is the value in equation (8) and $(\Delta\sigma)_k$ is the stress parameter in the kth subfault. Because the recorded trace's response spectrum is not changing, then

$$A_{jk} = \frac{\partial \varepsilon_j}{\partial(\Delta\sigma)_k} = \frac{\partial}{\partial(\Delta\sigma)_k}\left(\frac{1}{(N_2 - N_1 + 1)}\sum_{i=N_1}^{N_2}[\log_{10}\mathrm{PSA}(S_{ij})]\right)$$

$$\text{or} \quad A_{jk} = \frac{\partial}{\partial(\Delta\sigma)_k}\overline{\log_{10}[\mathrm{PSA}(S_j)]}\Big|_{f_1}^{f_2}. \qquad (12)$$

Equation (12) shows that element A_{jk} of the Jacobian matrix is equal to the derivative with respect to the stress parameter of the average simulated log PSA at station j. We exploit this observation to develop the following inversion shortcut.

Simulations are performed for an earthquake at a range of distances and for varying values of the uniform-stress parameter. The simulation parameters are the same as those used for the entire fault simulation, except that the fault is defined to be the size of a subsource and is assigned a correspondingly small magnitude (based on subsource size). Because these simulations are performed for a subsource, their results are equivalent to point source simulations such as those derived by Boore (2000). The calculated $\overline{\log_{10}[\mathrm{PSA}(S)]}\Big|_{f_1}^{f_2}$ results using the values of Table 1 (except for fault size and magnitude as noted in the preceding discussion) are shown in Figures 2 and 3. Figures 2 and 3 show that there is a linear relationship between $\overline{\log_{10}[\mathrm{PSA}(S)]}\Big|_{f_1}^{f_2}$ and $\log_{10}(\Delta\sigma)$ of a subfault, and between $\overline{\log_{10}[\mathrm{PSA}(S)]}\Big|_{f_1}^{f_2}$ and log distance . The linear relationship between $\overline{\log_{10}[\mathrm{PSA}(S)]}\Big|_{f_1}^{f_2}$, $\log_{10}(\Delta\sigma)$, and $\log_{10}$ distance is expressed as

$$\overline{\log_{10}[\mathrm{PSA}(S)]}\Big|_{f_1}^{f_2} = a\log_{10} R + b\log_{10}\Delta\sigma + c, \qquad (13)$$

where R is the distance from the center of the subfault and a, b, and c are constants that are determined from linear multiple regression. The derivative of equation (13) with respect to the stress parameter will be proportional to A_{jk}, where the proportionality constant is assumed to be a function of distance. The reason for such a proportionality is that the $\overline{\log_{10}[\mathrm{PSA}(S)]}\Big|_{f_1}^{f_2}$ at a station from a fault is not equal to the algebraic sum of $\overline{\log_{10}[\mathrm{PSA}(S)]}\Big|_{f_1}^{f_2}$ of subfaults, in which case A_{jk} would not be a function of distance, R_{jk}. Based on this reasoning the elements of the Jacobian matrix are

$$A_{jk} = \frac{b \times R_{jk}^{\alpha}}{(\Delta\sigma)_k \times \ln 10}, \qquad (14)$$

Table 1
Parkfield Earthquake Simulation Parameters

Parameter	Parameter Value
Fault orientation	Strike 137°, dip 83°
Depth of top	0 km
Fault dimensions	Length 40 km, width 13 km
Number of subfaults	Along length 10, along width 3
Fast Fourier transform points, sample interval	8192, 0.02 sec
Shear wave velocity, density	3.5 km/sec, 2.8 g/cm^3
Rupture velocity	0.8× shear wave velocity
$Q(f)$	$180f^{0.45}$
Moment magnitude	6.0
Kappa	0.035
Geometrical attenuation	If $R < 40$, R^{-1}; else $R^{-0.5}$
Subfault duration	$\propto 0.05 \times$ distance (sec.)
Windowing function	Saragoni–Hart
Amplification function	Boore and Joyner (1997) for $V_{S30} = 620$ m/sec
Damping	5% of critical damping

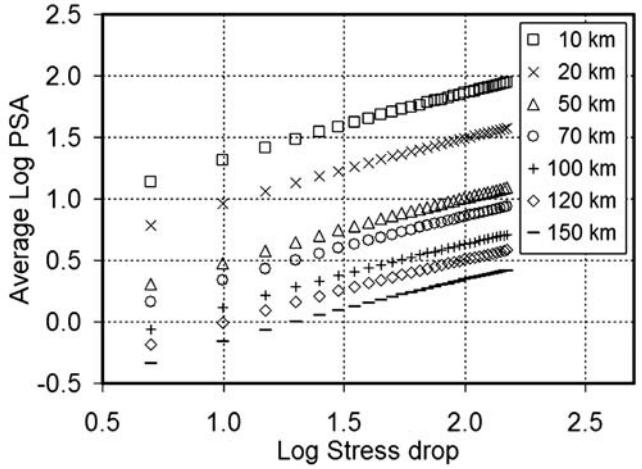

Figure 2. $\overline{\log_{10}[\mathrm{PSA}(S)]}\big|_{f_1=0.45\,\mathrm{Hz}}^{f_2=10\,\mathrm{Hz}}$ of simulations at constant distances from a subfault for varying stress parameters.

where b is the coefficient from equation (13), R_{jk} is the ratio of the distance between station j and subfault k divided by unit distance, ln is a logarithm in natural base, $(\Delta\sigma)_k$ is the stress parameter of the kth subfault, and α has the value of $a/2$ (equation 13). The value of α was determined empirically by checking various values in the range of a, in order to reproduce the Jacobian matrix values of equation (12).

Inversion Procedure

The program for inversion starts by assuming a uniform-stress distribution, and then performs a grid search to find the optimum average stress parameter that minimizes Ξ, based on recorded horizontal-component PSA values. Then it uses ε_j of the stations (equation 8) to construct the Jacobian and data matrices of the inverse problem. It also calculates the Θ error for the current stress parameter distribution. After solving the inverse matrix, the stress parameter correction vector, $\delta(\Delta\sigma)$, is derived and applied to the stress parameter distribution. At this time the first iteration is complete.

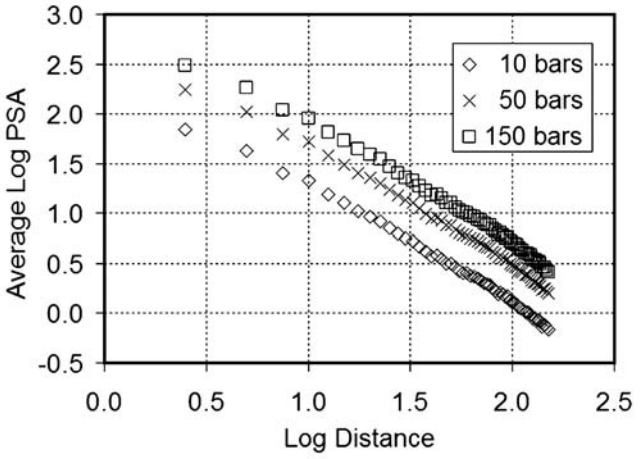

Figure 3. $\overline{\log_{10}[\mathrm{PSA}(S)]}\big|_{f_1=0.45\,\mathrm{Hz}}^{f_2=10\,\mathrm{Hz}}$ of simulations for constant stress parameters at varying distances from a subfault.

In the beginning of the second iteration, the new stress parameter distribution is treated as weights, normalized by the average stress. Then the program does a grid search for the optimum average stress parameter for this distribution. The new optimum average stress parameter is then multiplied by the normalized weights, providing a revised initial guess of the stress distribution, which is used for derivation of the stress parameter correction vector. The error Θ is then calculated again. If the value of Θ is lower than it was in the previous step, then the λ value in equation (10) is multiplied by a number smaller than unity (we used 0.1) for the next step, and the new solution is accepted as a successful iteration. If the value of Θ is higher than it was in the previous stage, then the λ value in equation (10) is multiplied by a number larger than unity (we used 10) for the repeating step; in this case the new solution is not accepted as a successful iteration and is discarded. The same iteration is repeated from the stress parameter distribution of the previous step, but with the new larger λ value. This procedure is repeated for at least 10 iterations, or the number of times that the program can run successfully, whichever is smaller.

Model Performance Studies

It is useful to evaluate the performance of the inversion program under ideal conditions in order to draw conclusions about its applicability and self-consistency. The cases studied here consider the Parkfield fault dimensions, strike, and depth but evaluate the effect of different dip angles and station geometries. The parameters in Table 1 are used for forward modeling. For a given stress distribution and station distribution, we use forward modeling to generate a synthetic PSA dataset, which is then input as data to the inversion routine. The output of the inversion routine is then compared to the initial assigned stress distribution. The similarity of the input and output stress distributions is a measure of the program's performance under the given conditions. The effects of the number of stations, the number of EXSIM iterations per station (to smooth the predicted PSA), the fault dip angle, the station distribution relative to asperity location, and the initial guess in the inversion input are studied. The way that the program identifies multiple asperities with different sizes is also considered.

Forward modeling is performed for a grid of symmetrically distributed stations around the fault as shown in Figure 4. We chose a dense array of 328 stations, to allow the inversion program to have sufficient independent equations to resolve the parameters (the stress parameter of the subfaults).

Forward simulations are performed for the preceding station distribution using EXSIM, given the parameters of Table 1, with an assumed stress parameter distribution and fault dip angle. The correlation coefficient is considered as a measure of similarity between the inversion outputs and

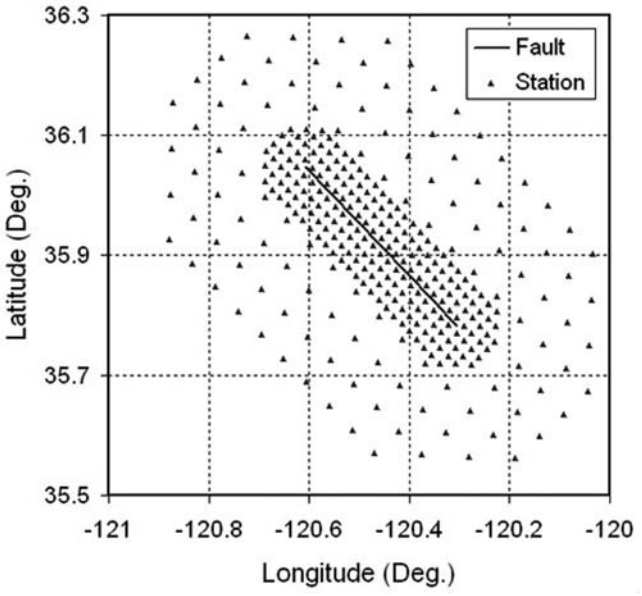

Figure 4. Distribution of hypothetical stations around the fault. Simulation outputs for these stations are used as data for the inversion program.

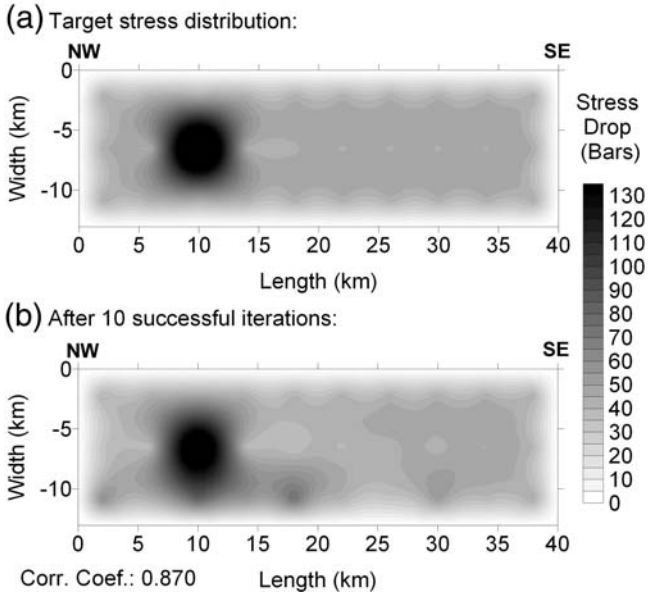

Figure 5. Stress parameter distribution on the fault surface with 25° fault dip angle: (a) target distribution and (b) inversion output after 10 inversion iterations and 10 EXSIM iterations.

target stress distribution (input distribution). A reduction in the value of the error measure Ξ is considered to indicate improvement of the solution with additional iterations.

Effect of EXSIM Iterations and Fault Dip Angle

"EXSIM iterations" is the number of EXSIM simulation trials used for obtaining an average PSA at each site under the considered scenario. In the first experiment, the assumed stress parameter is uniform on the fault except for one subfault, which has a stress parameter five times larger than that of the other subfaults. We consider fault dip angles of 5°, 25°, 45°, and 83° (the actual dip angle of the Parkfield fault).

Figure 5 shows the example performance for a fault with 25°-dip angle, using 10 EXSIM simulation trials to generate smooth PSAs and 10 inversion iterations to find the solution. The inversion process started from a uniform-stress parameter guess. Figure 5a shows the target stress distribution on the fault surface and Figure 5b is the result of inversion after 10 iterations. Error Ξ values for 25° and 83° fault dip angles are plotted in Figure 6. This graph compares the application of 10 or 30 EXSIM simulations per station on the error Ξ in different iterations. Thirty EXSIM iterations can result in a slightly lower error Ξ for steep fault dip angles (in this case 83°), but in general 10 EXSIM trials are sufficient to establish stable average expected values of PSA against which real or simulated data can be compared.

Error Ξ of the inversion solution at each iteration is shown for four fault dip angles, 5°, 25°, 45°, and 83°, in Figure 7. This graph shows that (1) in 10 inversion iterations the error reduces considerably; (2) in the case of 83° and 45° fault dip angles, the Ξ value approaches its asymptotic value in 10 inversion iterations; (3) the larger the fault dip angle, the poorer is the convergence of the solution and the more ambiguous is the ultimate result; and (4) gentle fault dip angles result in a solution that continues to improve with further iterations.

The inversions of synthetic data with varying fault dip angles show that loss of resolution in finding the high-stress-parameter patches on the fault increases gradually with increasing fault dip angle. This can be attributed to the range of angles over which subfaults can be seen as a function of dip angle. The higher the dip angle, the smaller is the angular coverage of stations around subfaults, and the lower is the likelihood of finding the effect of individual subfaults distinctively using the inversion method.

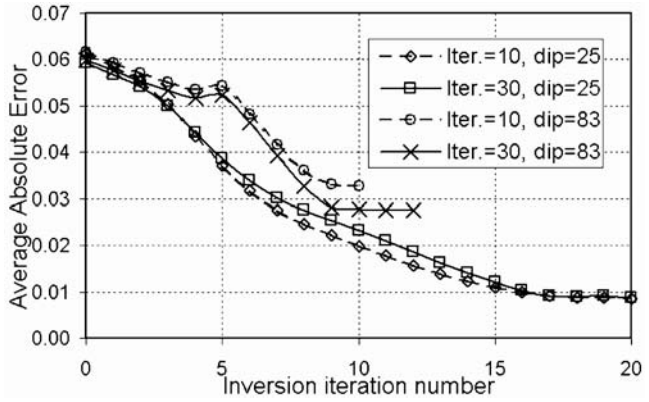

Figure 6. Evolution of error Ξ by inversion iteration for two fault dip angles. 10 and 30 EXSIM iterations were applied to both forward modeling and inversion procedures.

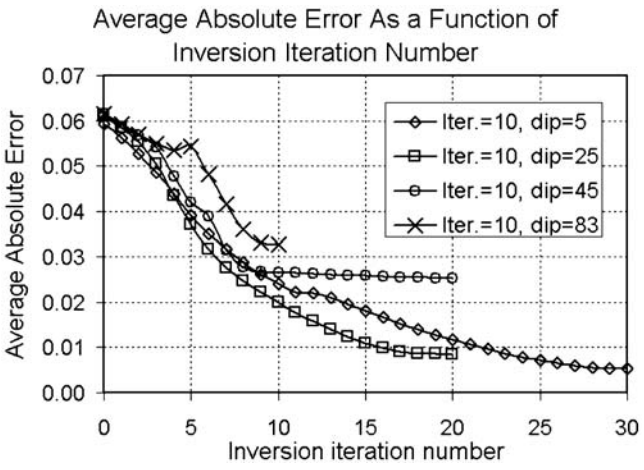

Figure 7. Evolution of error Ξ by inversion iteration number for four fault dip angles. Ten EXSIM iterations were applied to both forward modeling and inversion procedures.

Multiple Asperities

Because the PSA of a record at a site is not the algebraic sum of the PSAs of individual subfaults, it is important to find out if the presence of a high-stress-parameter asperity eliminates the effect of a neighboring low-stress-parameter asperity and to what extent multiple asperities might be detected. We tested the ability of the inversion program to distinguish multiple asperities by considering a stress distribution with three high-stress-parameter patches on a gently dipping fault (Fig. 8a). Improvement of the solution after 2 and 15 inversion iterations is shown in Figure 8b and c. The program successfully found the location and magnitude of the high-stress-parameter subfaults in the initial iterations after starting from a uniform distribution (given a good coverage of stations around the fault and a gentle dip angle). We conclude that the program solution is converging to the answer quickly if the real distribution contains multiple stress parameter patches.

Effect of the Initial Guess

The studies implemented in the preceding discussion consider a uniform-stress parameter distribution as the initial guess for inversion. One of the reasons for choosing the Levenberg–Marquardt inversion method was its good success probability when starting from uniform values of parameters. However, the initial guess plays an important role in nonlinear multiparameter estimation techniques. To check how it impacts our approach, we also considered the effects of three different cases for the initial guess: (1) a very accurate initial guess, (2) a slightly inaccurate initial guess, and (3) a completely wrong initial guess. The findings were as follows.

(1) When the initial guess is accurate, the solution does not change fundamentally from the initial guess, but it reaches the target solution within the first few iterations.

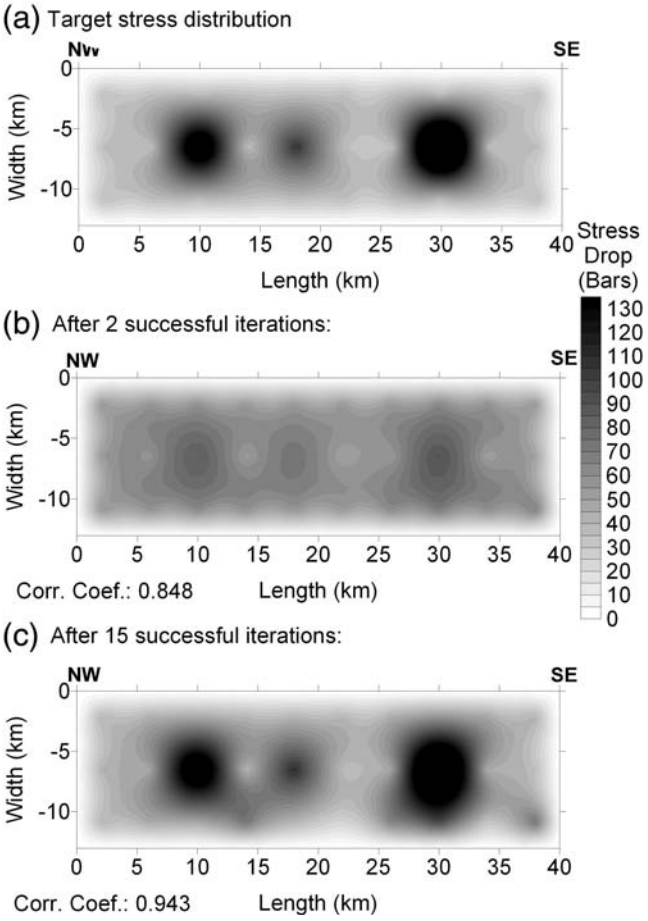

Figure 8. Stress parameter distribution on the fault surface with 5° dip angle: (a) target stress, (b) result of inversion after two iterations, and (c) result of inversion after 15 iterations. Thirty iterations of EXSIM for both forward modeling and inversion were applied.

The correlation coefficient, which is a measure of similarity between solutions and the target distribution, shows that a good initial guess can push the solution towards the target answer very quickly, and such a guess is more effective than assuming a uniform distribution.

(2) When the initial guess is inaccurate but carries some important features of the target distribution, then the inversion program improves the solution in every iteration, but the inaccuracies of the initial guess are not recovered within the first initial iterations. The correlation coefficient changes in every step, suggesting that a uniform-stress distribution is a better choice than an inaccurate distribution as an initial guess.

(3) The experiment for checking the response of the inversion procedure to an entirely wrong initial guess is done in two ways. We considered two wrong initial guesses: (a) a very poor initial guess, where the guessed asperity is in the wrong place and is as large in amplitude as the actual main asperity, and (b) a poor initial guess, where the guessed asperity is in the wrong place but is much smaller than the actual asperity. The experiments show that, if the initial guess is a very poor one, the inversion procedure may fail to

continue iteratively toward the correct solution, but if the initial guess is just poor, then the program can converge toward the right distribution in the long run. Overall, we conclude that the uniform distribution is probably the best initial guess unless we have sufficient information from the intensity distribution or time evolution of amplitudes in records to make a perfect initial guess, which is seldom the case in real applications.

Effect of Station Distribution and Asperity Location

Station distribution plays an important role in locating asperities, because from equation (14) we know that elements of the Jacobian matrix are distance dependent (depending on the distance from the center of the subfault to the recording site). Furthermore, a sparse distribution of stations around a subfault makes the data insufficient for the unique determination of the stress parameter on that subfault. To study the effect of asymmetric station distribution around the fault, we chose the actual station distribution around the 2004 Parkfield earthquake causative fault (Fig. 9). Fault dip angle and EXSIM iteration number in both forward modeling and inversion are 83° and 10, respectively. The other parameter settings of EXSIM are those in Table 1. Figure 9 shows that the station distribution around the southeast end of the fault is much denser than in the northwest end. Forward modeling and inversions are performed for 10 different scenarios, in which we assigned one high-stress-parameter subfault (for each scenario) to lie in a location in the mid row of the fault from northeast to southwest. In case 1 the high-stress subfault is at the northwest end of the fault, and in successive cases the high-stress-parameter subfault moves towards the southeast, in one subfault intervals. The stress of the high-stress-parameter subfault is fixed at five times the value of the other subfaults.

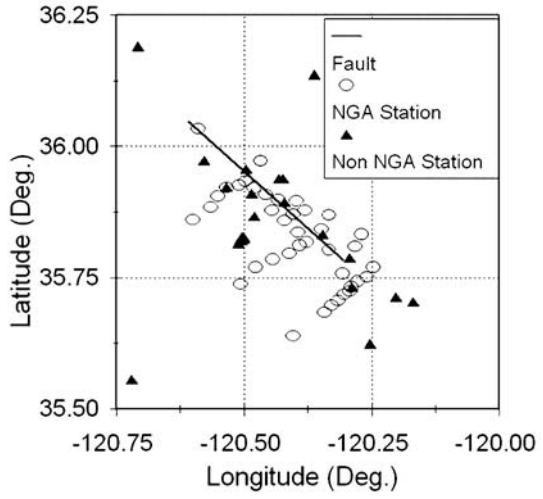

Figure 9. Distribution of stations around the Parkfield fault with known (NGA) and unknown (non-NGA) sites.

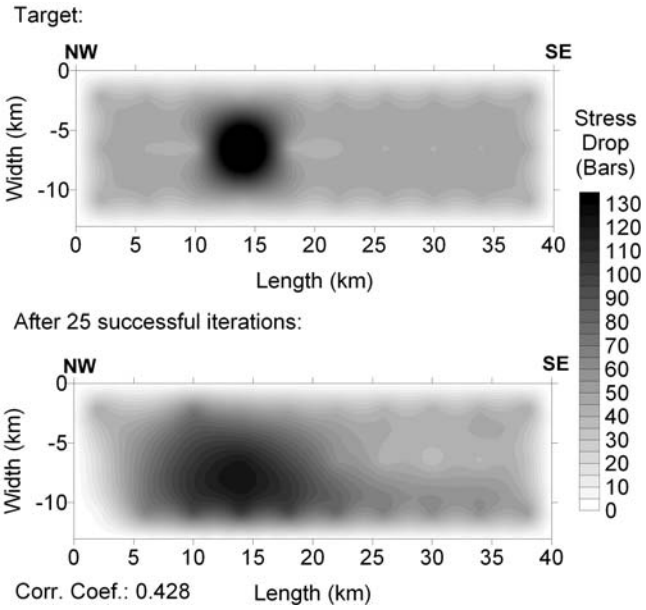

Figure 10. Case 4 stress distribution: (top) target distribution and (bottom) inversion solution after 25 iterations.

The solutions obtained after the third scenario (the third location of a high-stress subfault) start to be meaningful. By the tenth scenario (the high-stress subfault in the southwest), the solution is excellent as measured by a high correlation coefficient and has good Ξ reduction. Because the Parkfield fault is almost vertical, the overall resolution ability of the program is not high. The inversion program fails after the first iteration in the first scenario (the high-stress subfault in the northwest) because of a lack of sufficient number of recording sites around that subfault (there is just one station above the high-stress-parameter subfault and nothing else for some distance). We allowed the inversion program to run for up to 50 iterations for some cases, although the solution improvement in all cases is very small after the tenth iteration. As the asperity is moved towards the southeast, where the station distribution is denser, the ability of the program improves markedly, as shown in Figures 10 and 11. We conclude that if the fault dip angle is steep, then the high-stress-drop asperity may hide itself in an area of very low station density. However, even if there are a few stations around the asperity, the method can still find its most probable location horizontally (but not vertically) for steeply dipping faults.

Example Application of the Method to the Parkfield Earthquake

The **M** 6.0 2004 Parkfield earthquake provided strong-motion data from more than 85 stations, mostly at very close distances to the causative fault. Response spectra of horizontal components from these data are compared with simulations based on our modified version of stochastic finite-fault modeling, for the same stations. A stress distribution on the

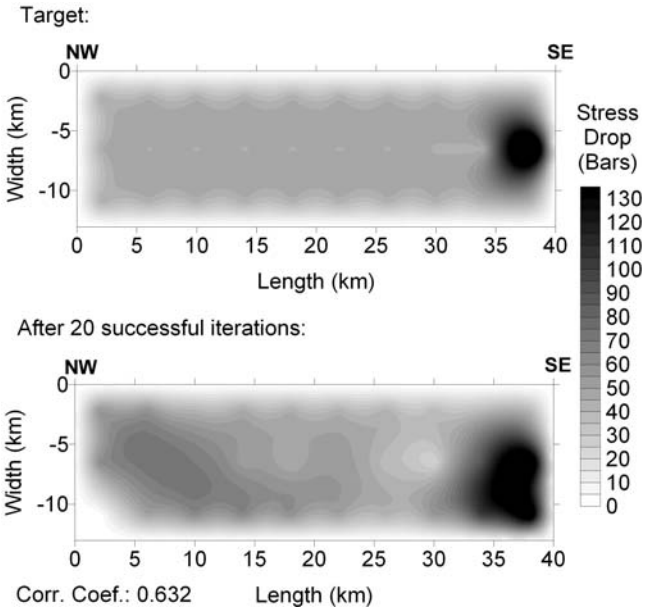

Figure 11. Case 10 stress distribution: (top) target distribution and (bottom) inversion solution after 20 iterations.

fault surface is sought that minimizes the overall differences between the high-frequency pseudoacceleration (PSA) ordinates of the simulated records and observations. The seeking process is performed by means of the Levenberg–Marquardt inversion method (Marquardt, 1963; Lee and Stewart, 1981; Draper and Smith, 1998) as described in the preceding discussion. The PSAs of recorded data are first corrected for site response to the equivalent values for a uniform $V_{S30} =$ 620 m/sec site type, before performing the inversion. The procedure of site correction is explained next. (Note: an alternative approach for site correction would be to modify the synthetic records during EXSIM simulations for every site and then compare site-specific simulated spectra with recorded ones in the inversion process; we opted to correct spectra of the recorded signals to a common reference level for programming reasons).

Data Used and Site Corrections:

The list of stations used in this study is given in Table 2; the distribution of stations around the Parkfield fault is shown in Figure 9. The list is extracted from the COSMOS database (Archuleta *et al.*, 2003; Squibb *et al.*, 2004; Archuleta *et al.*, 2006). The total number of stations is 87, among which 46 are listed in the NGA database (Pacific Earthquake Engineering Research, New Generation Attenuation Relation Project 2005, http://peer.berkeley.edu/nga/index.html). The stations in the NGA database contain V_{S30} (the average shear wave velocity in the upper 30 m) and site type. The V_{S30} values of Table 2 are extracted from the preferred V_{S30} column of the NGA database, as based on the U.S. Geological Survey Northridge assignments (Borcherdt and Fumal, 2002) and on the California Geological Survey assignments (Wills and Silva, 1998; Wills *et al.*, 2000). Other stations with unknown site type were assigned NEHRP class D with the average $V_{S30} = 255$ m/sec; this value is a reasonable average for soil sites in the area. In the error calculations and inversions the data from the non-NGA stations Hog Canyon and Priest Valley were omitted because of poor bandwidth.

The recorded signals' response spectra are first corrected for soil response to produce the corresponding spectra for $V_{S30} = 620$ m/sec (generic rock, Boore and Joyner [1997]) site conditions. This is done in two steps. First, using equation 15 from Boore and Atkinson (2006) the peak horizontal acceleration for $V_{S30} = 620$ m/sec (PHAr) that is expected at each station for **M** 6 is calculated:

$$\ln \text{PHAr} = F_M(\mathbf{M}) + F_D(r_{jb}, \mathbf{M}) + F_S(V_{S30}, \mathbf{M}, r_{jb}) + \varepsilon\sigma_T, \qquad (15)$$

where $F_M(\mathbf{M})$ is a magnitude scaling term, $F_D(r_{jb})$ is a distance scaling term, and $F_S(V_{S30}, \mathbf{M}, r_{jb})$ is the site scaling term. These three terms can be calculated based on values of V_{S30}, **M**, r_{jb} (Joyner–Boore distance), and constants presented in Boore and Atkinson (2006). The constants are different for PHArs larger and smaller than 0.06g. Then the PHAr values are used in equation 16 from Choi and Stewart (2005) to find the nonlinear amplification factors of each site:

$$\ln(F_{ij}) = c \times \ln\left(\frac{V_{S30_{ij}}}{V_{ref}}\right) + b \times \ln\left(\frac{PHA_{r_{ij}}}{0.1}\right), \qquad (16)$$

where F_{ij} is the amplification factor needed to multiply PHAr to predict the peak horizontal acceleration (PHA) (or response spectral ordinates) for a site with a given V_{S30}. b, c, and V_{ref} are coefficients given in tables presented by Choi and Stewart (2005). Note that Boore and Atkinson (2006) adopted a slightly smoothed version of Choi and Stewart's result in their derivation of empirical ground-motion relations in the NGA project. By dividing the ordinates of the acceleration response spectrum at each station by the corresponding amplification factor, the site-corrected PSAs are derived; these are the equivalent PSA values for a site with $V_{S30} = 620$ m/sec.

Inversion of Recorded Parkfield Data

The first step in the modeling of stress distribution on the Parkfield fault is to determine the best uniform-stress model, using the response spectra of all horizontal components (Table 2), site-corrected to $V_{S30} = 620$ m/sec. The basic parameters of the simulation are typical for California and follow previous applications (Motazedian and Atkinson, 2005); they are given in Table 1. The parameters to be determined are the stress parameter and the pulsing percentage. These are allowed to vary on a grid of parameters, with stress from 10–100 bars and pulsing percentage from 10%–

Table 2

List of Stations Used in This Study

Number	Station	Owner*	In NGA list	Site type	V_{S30}†	Latitude	Longitude
1	Big Sur, CA—Pfeiffer State Park	CSMIP	No	?	?	36.252	−121.782
2	Buttonwillow, CA—Highway 58 and Wasco	CSMIP	No	?	?	35.403	−119.449
3	Cambria, CA—Highway 1, Caltrans Bridge grounds	CSMIP	No	?	?	35.593	−121.124
4	Capitola, Ca—fire station	CSMIP	Yes	D	288	36.973	−121.953
5	Coalinga, CA—Slack Canyon; Hidden Valley Ranch	CSMIP	Yes	C	684	36.034	−120.590
6	Coalinga, Ca—fire station	USGS	No	?	?	36.137	−120.363
7	Fresno, CA—NSMP Office	USGS	No	?	?	36.744	−119.728
8	Fresno, CA—VA Medical Center	USGS	No	?	?	36.773	−119.781
9	Greenfield, CA—police station	CSMIP	No	?	?	36.321	−121.243
10	Hog Canyon, CA	NCSN	No	?	?	35.867	−120.480
11	Hollister, CA—South St. and Pine Dr.	CSMIP	Yes	C	370	36.848	−121.397
12	Hollister, Ca—airport building # 3	USGS	Yes	D	271	36.891	−121.404
13	Hollister, Ca—city hall annex	USGS	Yes	D	274	36.851	−121.402
14	King City, CA—two-story hospital	CSMIP	No	?	?	36.206	−121.132
15	King City, CA—Canal and Reich	CSMIP	No	?	?	36.207	−121.131
16	Moss Landing, CA—Highway 1 and Dolan	CSMIP	No	?	?	36.807	−121.778
17	Cholame 1E	CSMIP	Yes	D	338	35.743	−120.277
18	Cholame 2E	CSMIP	Yes	C	376	35.751	−120.259
19	Cholame 2W	CSMIP	Yes	D	184	35.733	−120.290
20	Cholame 3E	CSMIP	Yes	C	376	35.770	−120.247
21	Cholame 3W	CSMIP	Yes	D	338	35.724	−120.294
22	Cholame 4AW	CSMIP	Yes	D	338	35.707	−120.316
23	Cholame 4W	CSMIP	Yes	C	438	35.718	−120.304
24	Cholame 5SW	CSMIP	Yes	D	289	35.697	−120.328
25	Cholame 6W	CSMIP	Yes	D	338	35.684	−120.342
26	Cholame 12W	CSMIP	Yes	C	408	35.639	−120.404
27	Fault zone 1	CSMIP	Yes	D	338	35.758	−120.307
28	Fault zone 3	CSMIP	Yes	C	370	35.803	−120.334
29	Fault zone 4	CSMIP	Yes	D	338	35.836	−120.395
30	Fault zone 6	CSMIP	Yes	C	438	35.859	−120.420
31	Fault zone 7	CSMIP	Yes	C	370	35.871	−120.404
32	Fault zone 8	CSMIP	Yes	C	376	35.878	−120.381
33	Fault zone 9	CSMIP	Yes	C	438	35.879	−120.445
34	Fault zone 11	CSMIP	Yes	C	376	35.896	−120.398
35	Fault zone 12	CSMIP	Yes	D	338	35.899	−120.433
36	Fault zone 14	CSMIP	Yes	D	338	35.908	−120.458
37	Fault zone 15	CSMIP	Yes	C	376	35.921	−120.481
38	Gold Hill	USGS	No	?	?	35.833	−120.346
39	Gold Hill 1W	CSMIP	Yes	D	338	35.818	−120.378
40	Gold Hill 2E	CSMIP	Yes	D	338	35.843	−120.348
41	Gold Hill 2W	CSMIP	Yes	C	376	35.812	−120.391
42	Gold Hill 3E	CSMIP	Yes	C	370	35.870	−120.334
43	Gold Hill 3W	CSMIP	Yes	C	438	35.796	−120.411
44	Gold Hill 4W	CSMIP	Yes	C	438	35.785	−120.444
45	Gold Hill 5W	CSMIP	Yes	C	438	35.770	−120.477
46	Gold Hill 6W	CSMIP	Yes	C	438	35.738	−120.507
47	Highway 46—Cholame Creek Bridge	CSMIP	No	?	?	35.733	−120.289
48	Stone Corral 1E	CSMIP	No	?	?	35.788	−120.294
49	Stone Corral 2E	CSMIP	Yes	C	376	35.810	−120.282
50	Stone Corral 3E	CSMIP	Yes	C	376	35.833	−120.270
51	Temblor	CSMIP	No	?	?	35.705	−120.169
52	Vineyard Canyon 1W	CSMIP	Yes	C	376	35.934	−120.497
53	Vineyard Canyon 2E	CSMIP	Yes	C	712	35.973	−120.467
54	Vineyard Canyon 2W	CSMIP	Yes	D	338	35.927	−120.509
55	Vineyard Canyon 3W	CSMIP	Yes	D	297	35.922	−120.534
56	Vineyard Canyon 4W	CSMIP	Yes	C	376	35.905	−120.551
57	Vineyard Canyon 5W	CSMIP	Yes	C	376	35.885	−120.565
58	Vineyard Canyon 6W	CSMIP	Yes	C	438	35.861	−120.600
59	Donna Lee	USGS	No	?	?	35.939	−120.425
60	Eades	USGS	No	?	?	35.894	−120.421
61	Froelich	USGS	No	?	?	35.911	−120.486
62	Jack Canyon	USGS	No	?	?	35.713	−120.203

Table 2 (*Continued*)

Number	Station	Owner*	In NGA list	Site type	V_{S30}†	Latitude	Longitude
63	Joaquin Canyon	USGS	No	?	?	35.940	−120.432
64	Middle Mountain	USGS	No	?	?	35.958	−120.496
65	Red Hills	USGS	No	?	?	35.624	−120.254
66	Stockdale Mountain	USGS	No	?	?	35.973	−120.579
67	Vineyard Canyon	USGS	No	?	?	35.923	−120.534
68	Work Ranch	USGS	No	?	?	35.814	−120.511
69	Point Buchon, CA—Los Osos	CSMIP	No	?	?	35.274	−120.885
70	Coalinga, CA—Priest Valley	CSMIP	No	?	?	36.191	-120.708
71	Salinas, CA—City Yard—John & Work	CSMIP	Yes	D	271	36.671	−121.642
72	Salinas, Ca—county hospital grounds	CSMIP	Yes	D	271	36.697	−121.634
73	San Luis Obispo, ca—city recreation building 864	USGS	Yes	C	712	35.285	−120.661
74	San Luis Obispo, CA—Lopez Lake Grounds	CSMIP	No	?	?	35.208	−120.457
75	Templeton, Ca—hospital grounds	CSMIP	No	?	?	35.556	−120.720
76	USGS Parkfield Dense Seismograph Array 01	USGS	No	?	?	35.821	−120.507
77	USGS Parkfield Dense Seismograph Array 02	USGS	No	?	?	35.822	−120.506
78	USGS Parkfield Dense Seismograph Array 03	USGS	No	?	?	35.821	−120.505
79	USGS Parkfield Dense Seismograph Array 05	USGS	No	?	?	35.824	−120.503
80	USGS Parkfield Dense Seismograph Array 06	USGS	No	?	?	35.824	−120.503
81	USGS Parkfield Dense Seismograph Array 07	USGS	No	?	?	35.824	−120.503
82	USGS Parkfield Dense Seismograph Array 08	USGS	No	?	?	35.825	−120.501
83	USGS Parkfield Dense Seismograph Array 09	USGS	No	?	?	35.826	−120.501
84	USGS Parkfield Dense Seismograph Array 10	USGS	No	?	?	35.828	−120.500
85	USGS Parkfield Dense Seismograph Array 11	USGS	No	?	?	35.826	−120.502
86	USGS Parkfield Dense Seismograph Array 12	USGS	No	?	?	35.827	−120.504
87	USGS Parkfield Dense Seismograph Array 13	USGS	No	?	?	35.827	−120.505

*Abbreviations used for the owner of the station are as follows: CSMIP—California Strong Motion Instrumentation Program; NCSN—Northern California Seismic Network; and USGS—U.S. Geological Survey.

†Velocity in m/sec.

100%. The calculated error measurement is Ξ (equation 6). Figure 12 plots the model errors; the optimal values are stress ~50 bars and pulsing percentage ~20%. Plotting the errors for each station at a specific frequency shows the level of over/underestimation of the model at different stations around the fault; this is ε_j in equation (8) for a given frequency index ($N1 = N2 = N$). The geographical distribution of ε_j around the fault for 5 Hz is shown in Figure 13. From Figure 13 we infer a high-stress patch near the southeast end of the fault (causing the observed positive residuals).

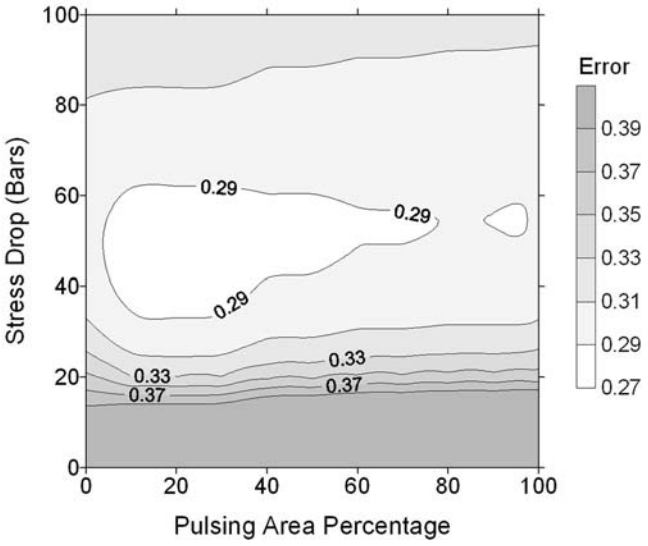

Figure 12. Error Ξ (in $\log_{10}$ units) for different values of the pulsing area and the uniform-stress parameter. Best match with data is for 50 bars, with 20% pulsing area.

We inverted the recorded data (85 stations corrected to $V_{S30} = 620$ m/sec) to obtain the stress distribution of the **M** 6 2004 Parkfield earthquake. The available response spectra in the range of 0.45 to 10 Hz are used as input data to the inversion procedure; this frequency range is considered to be a reasonable choice for the high-frequency range for the **M** 6 Parkfield earthquake based on the overall fault corner frequency.

The inversion program starts from a uniform-stress distribution and stops after just four successful iterations of the inversion program to produce the derived distribution. The error Ξ of the solution is 8.3% lower than that of the initial guess. Average error spectra and the absolute errors at each frequency averaged over all stations are shown in Figure 14. Figure 15 shows the initial and derived stress distributions on the fault.

As was shown in Figure 13, the uniform-stress distribution caused large underestimation of the $\overline{\log_{10}[\mathrm{PSA}(S)]}\Big|_{f_1=0.45\ \mathrm{Hz}}^{f_2=10\ \mathrm{Hz}}$ at the southeast end of the fault, small underestimation at the middle and the northwest end of the fault, and overestimation between these points. Modification of the stress parameter distribution to that of Figure 15b compensates these errors reasonably; this distri-

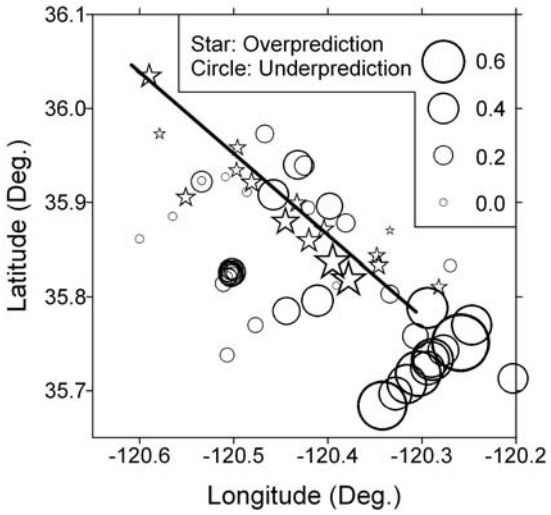

Figure 13. Residuals of simulations at each station ε_j around the fault, for PSA at 5 Hz (in $\log_{10}$ units).

bution is consistent with our visual judgment of errors in Figure 13.

Overall, the inversion clearly indicates that a high-stress-parameter patch is located in the southeast end of the Parkfield fault, at depths greater than 4 km. The relative locations of the hypocenter and the important asperity is in good agreement with the observation by Mai *et al.* (2003, 2005), that the hypocenter is close to high-stress-parameter centers but not in the center of them. Also, the stress parameter distribution of Figure 15, derived by our inversion method, has many similarities to the slip distribution derived by Langbein *et al.* (2005) and Liu *et al.* (2006). Langbein *et al.* (2005) used both strong-motion and Global Positioning System (GPS) data for their solution; Liu *et al.* (2006) inverted strong-motion seismograms to obtain a space–time

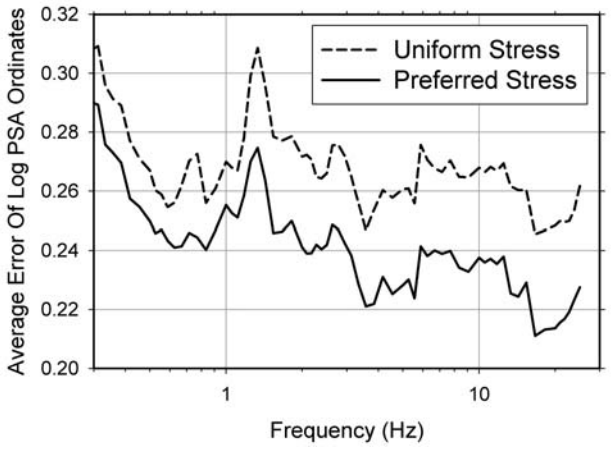

Figure 14. Average error (Ξ_i) at all stations as a function of frequency, for uniform-stress distribution (dashed line) and preferred stress distribution of Figure 15 (solid line).

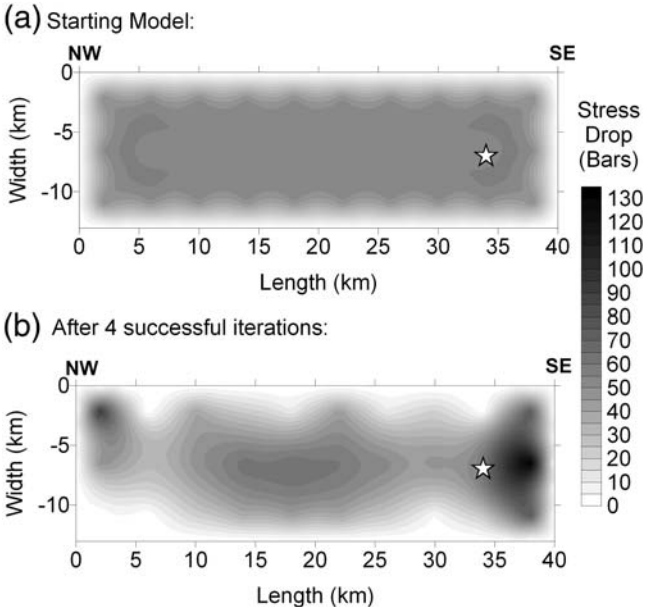

Figure 15. Stress distribution of the observed Parkfield data: (a) initial guess and (b) solution after four successful iterations. The star is the location of the hypocenter.

distribution model of coseismic slip on the fault. Their inversion for the rupture model yielded nonunique solutions; they presented a total of 10 possible solutions, along with an average solution. Figure 16 shows two representative solutions and the average solution (it was redrawn using the information provided in the Liu *et al.* electronic supplement). The implicit assumption made in this comparison is that high-slip areas correspond to high-stress-drop areas.

Discussion and Conclusions

We have modified the stochastic finite-fault model of Motazedian and Atkinson (2005) to allow for variability of the stress parameter across the fault plane. The modification satisfies two conditions: (1) the low-frequency level of the Fourier spectrum of subfaults should not be affected, and (2) the high-frequency level of the spectrum of subfaults should be proportional to the stress parameter raised to the power of $2/3$. This modification leads to enhanced high-frequency motions near high-stress patches on the fault. Distant sites are unaffected.

To make it possible to calculate the stress parameter distribution on the fault surface from observed response spectra of an earthquake, an inversion tool is developed based on the Levenberg–Marquardt method. The inversion generally converges quickly from an initial assumed uniform-stress distribution.

The horizontal components of the **M** 6 2004 Parkfield earthquake recorded at 85 locations are used in a test application of the inversion method. First, the response spectra of the records are corrected to equivalent values for a site condition of $V_{S30} = 620$ m/sec. Then, the corrected PSAs are

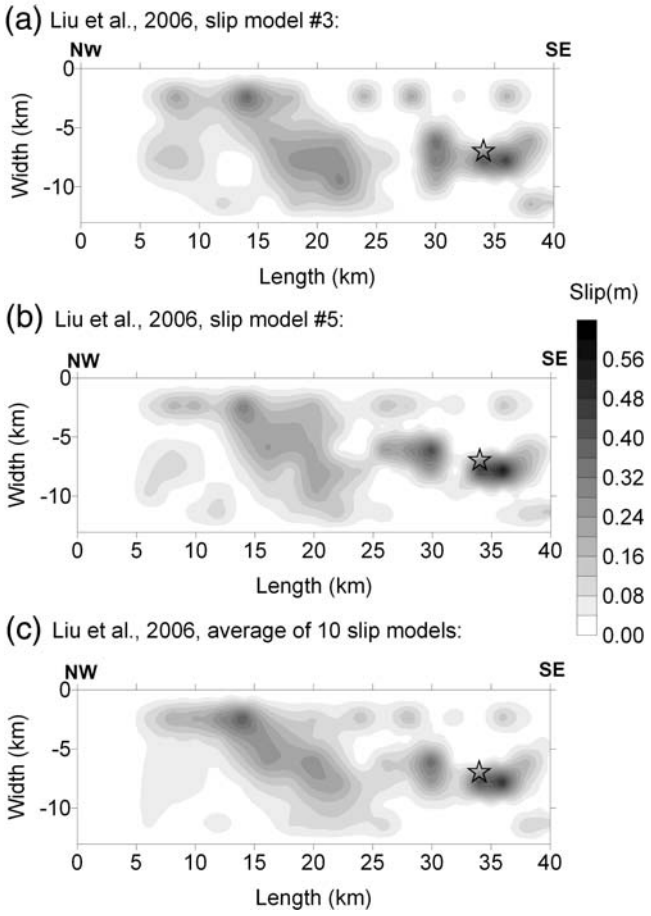

Figure 16. Slip distribution on the Parkfield fault surface derived from strong-motion seismographic data (from Liu *et al.* [2006]): (a) solution number 3, (b) solution number 5, and (c) average of 10 solutions.

inverted for the stress parameter distribution on the fault surface. The inversion of the Parkfield data clearly demonstrates the presence of an asperity in the southeast end of the fault at depths between 4 and 8 km. This is in agreement with the conclusion of Mai *et al.* (2003, 2005) that the hypocenter and the main asperity large slip area are very close to each other.

We examined the ability of the inversion tool to recover an input stress distribution under various conditions and conclude the following:

1. Gentle fault dip angles allow an accurate determination of the stress distribution, while steep fault dip angles limit its resolution. With gentle fault dip angles, the stations see the subfaults across a fuller range of angles and distances, and this makes the equations for the inversion distinct. The method performs very well if the fault dip angle is less than 45°, although it performs reasonably well with dip angles more than 45°, at least in terms of locating the horizontal position of asperities.
2. The method has the capability to resolve multiple asperities, with as little as one subfault distance between them. It can also find the relative stress parameter levels of the asperities. With multiple asperities the solution converges faster than is the case for a single asperity.
3. If the initial guess of the stress distribution is an accurate one, the solution will converge with few inversion iterations. If an accurate initial guess is not obvious, then the uniform distribution provides a reasonable initial guess. If the initial guess is a strong asperity in the incorrect location, then the program may fail to converge; otherwise, it will be able to push the answer toward the real distribution eventually.
4. As is the case in other inversion techniques, a dense station distribution around the fault will result in high accuracy and precision of the high-stress asperities. If the station distribution is sparse, or there are no stations near the asperity, then the program may fail to find an asperity or provide it with low resolution.

The variable-stress version of stochastic finite-fault modeling provides more modeling resolution than the previous stochastic approaches and has the ability to reduce the error between simulated and recorded response spectra by a significant amount. The sensitivity of the new model to the variable-stress parameter on the fault surface allows its application in an inversion tool. We may thus deduce the stress parameter distribution on the fault surface based on strong-motion observations from stations around the fault. With a reasonable number and distribution of stations around a fault, the method is able to find the asperities, although a steep fault dip angle limits the resolution of the depth of the asperities. The ability to interpret the strong-motion observations in terms of source property variability along the fault should lead to improved understanding of the driving mechanisms behind ground-motion observations and their spatial variability. It does not lead to improved predictive capabilities for future earthquakes in the short term, but may in the long term as information on stress distributions is compiled for a large number of events.

Acknowledgments

We thank Dariush Motazedian for stimulating discussions of finite-fault modeling and helpful suggestions. Also we thank David Boore, Arthur McGarr, and an anonymous reviewer for their comments. This research was funded by the Natural Sciences and Engineering Research Council of Canada.

References

Aki, K (1984). Asperities, barriers, characteristic earthquakes and strong motion prediction. Special section. Fault behavior and the earthquake generation process, *J. Geophys. Res.* **89,** 5867–5872.

Archuleta, R. J., J. H. Steidl, and M. Squibb (2003). COSMOS Virtual Data Center: a Web-based portal to the world's strong motion data. Abstracts of the 73rd annual meeting of the Seismological Society of America, *Seism. Res. Lett.* **74,** 218.

Archuleta, R., J. Steidl, and M. Squibb (2006). The COSMOS VDC: a search engine for worldwide strong-motion data, *Annual Meeting of the SSA: 100th Anniversary Earthquake Conference*, http://db.cosmos-eq.org/ (last accessed March 2006).

Beresnev, I. A., and G. M. Atkinson (1998). FINSIM—a FORTRAN program for simulating stochastic acceleration time histories from finite faults, *Seism. Res. Lett.* **69,** 27–32.

Beresnev, I. A., and G. M. Atkinson (1999). Generic finite-fault model for ground-motion prediction in eastern North America, *Bull. Seismol.l Soc. Am.* **89,** 608–625.

Beresnev, I. A., and G. M. Atkinson (2002). Source parameters of earthquakes in eastern and western North America based on finite-fault modeling, *Bull. Seismol. Soc. Am.* **92,** 695–710.

Boore, D. M. (1983). Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra, *Bull. Seismol. Soc. Am.* **73,** 1865–1894.

Boore, D. M. (2000). SMSIM version 2.0; FORTRAN programs for simulating ground motions from earthquakes, *U.S. Geol. Surv. Open-File Rept. 00-0509*, 55 pp.

Boore, D. M. (2003). Simulation of ground motion using the stochastic method: Seismic motion, lithospheric structures, earthquake and volcanic sources. the Keiiti Aki volume, *Pure Appl. Geophys.* **160,** 635–676.

Boore, D. M., and G. M. Atkinson (2006). Boore–Atkinson provisional NGA empirical ground-motion model for the average horizontal component of PGA, PGV and SA at spectral periods of 0.05, 0.1, 0.2, 0.3, 0.5, 1, 2, 3, 4, and 5 seconds, a report to the PEER–Lifelines Next Generation Project, http://quake.usgs.gov/~boore, (revised 27 October 2006; last accessed December 2006).

Boore, D. M., and W. B. Joyner (1997). Site amplifications for generic rock sites, *Bull. Seismol. Soc. Am.* **87,** 327–341.

Borcherdt, R. D., and T. E. Fumal (2002). Shear-wave velocity compilation for Northridge strong-motion recording sites, *U.S. Geol. Surv. Open-File Rept. 02-0107*, 16 pp.

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.* **75,** 4997–5009.

Brune, J. N. (1971). Tectonic stress and the spectra of seismic shear waves from earthquakes: correction, *J. Geophys. Res.* **76,** 5002.

Chin, B., and K. Aki (1991). Simultaneous study of the source, path, and site effects on strong ground motion during the 1989 Loma Prieta earthquake: a preliminary result on pervasive nonlinear site effects. The 1989 Loma Prieta, California, earthquake and its effects, *Bull. Seismol. Soc. Am.* **81,** 1859–1884.

Choi, Y., and J. P. Stewart (2005). Nonlinear site amplification as function of 30 m shear wave velocity, *Earthq. Spectra* **21,** 1–30.

Draper, N. R., and H. Smith (1998). *Applied Regression Analysis*, Wiley, New York.

Goldberg, D. E. (1988). *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, Massachusetts.

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

Heaton, T. H., and S. H. Hartzell (1986). Source characteristics of hypothetical subduction earthquakes in the northwestern United States, *Bull. Seismol. Soc. Am.* **76,** 675–708.

Heaton, T. H., and S. H. Hartzell (1989). Estimation of strong ground motions from hypothetical earthquakes on the Cascadia subduction zone, Pacific Northwest, *Pure Appl. Geophys.* **129,** 131–201.

Irikura, K. (1983). Semi-empirical estimation of strong ground motions during large earthquakes, *Bull. Disaster Prev. Res. Inst.* **33,** 63–104.

Joyner, W. B., and D. M. Boore (1986). On simulating large earthquakes by Green's-function addition of smaller earthquakes, in *Earthquake Source Mechanics*, American Geophysical Monograph **37,** 269–274.

Kanamori, H. (1979). A semi-empirical approach to prediction of long-period ground motions from great earthquakes, *Bull. Seismol. Soc. Am.* **69,** 1645–1670.

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

Lee, W. H. K., and S. W. Stewart (1981). *Principles and Applications of Microearthquake Networks*, Academic Press, New York.

Liu, P., S. Custódio, 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.

Mai, P. M., P. Spudich, and J. Boatwright (2003). Hypocenter locations in finite-source rupture models: Abstracts of the 73rd annual meeting of the Seismological Society of America, *Seism. Res. Lett.* **74,** 208.

Mai, P. M., P. Spudich, and J. Boatwright (2005). Hypocenter locations in finite-source rupture models, *Bull. Seismol. Soc. Am.* **95,** 965–980.

Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Ind. Appl. Math.* **11,** 431.

Menke, W. (1989). *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press, San Diego, California.

Motazedian, D., and G. M. Atkinson (2005). Stochastic finite-fault modeling based on a dynamic corner frequency, *Bull. Seismol. Soc. Am.* **95,** 995–1010.

Parker, R. L. (1994). *Geophysical Inverse Theory*, Princeton Univ. Press, Princeton, New Jersey.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical recipes in FORTRAN: the art of scientific computing*, Cambridge U. Press, New York.

Schneider, J. F., W. J. Silva, and C. Stark (1993). Ground motion model for the 1989 m 6.9 Loma Prieta earthquake including effects of source, path, and site, *Earthq. Spectra* **9,** 251–287.

Silva, W., R. Darragh, C. Stark, I. Wong, J. C. Stepp, J. F. Schneider, and S. J. Chiou (1990). Methodology to estimate design response spectra in the near-source region of large earthquakes using the band-limited-white-noise ground motion model, in *Proceedings of the U.S. National Conference on Earthquake Engineering 487*.

Somerville, P. G., M. K. Sen, and B. P. Cohee (1991). Simulation of strong ground motions recorded during the 1985 Michoacan, Mexico, and Valparaiso, Chile, earthquakes, *Bull. Seismol. Soc. Am.* **81,** 1–27.

Squibb, M. B., R. J. Archuleta, J. Steidl, T. L. Henyey, T. H. Jordan, and J. K. McRaney (2004). COSMOS Virtual Data Center; SSA 2004: abstracts of the annual meeting, *Seism. Res. Lett.* **75,** 247.

Tarantola, A. (1987). *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier, Amsterdam.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Paramenter Estimation*, Soc. Ind. Appl. Math., Philadelphia, Pennsylvania.

Tumarkin, A. G., and R. J. Archuleta (1994). Empirical ground motion prediction: Earthquake source mechanics, *Ann. Geophys.* **37,** 1691–1720.

Wills, C. J., and W. Silva (1998). Shear-wave velocity characteristics of geologic units in California, *Earthq. Spectra* **14,** 533–556.

Wills, C. J., M. D. Petersen, W. A. Bryant, M. S. Reichle, G. J. Saucedo, S. S. Tan, G. C. Taylor, and J. A. Treiman (2000). A site conditions map for California based on geology and shear wave velocity: Seismological Society of America 2000. Abstracts of the 95th annual meeting, *Seism. Res. Lett.* **71,** 248.

Yu, G., K. N. Khattri, J. G. Anderson, J. N. Brune, and Y. Zeng (1995). Strong ground motion from the Uttarkashi, Himalaya, India earthquake: comparison of observations with synthetics using the composite source model, *Bull. Seismol. Soc. Am.* **85,** 31–50.

Zeng, Y., J. G. Anderson, and G. Yu (1994). A composite source model for computing realistic synthetic strong ground motions, *Geophys. Res. Lett.* **21,** 725–728.

Department of Earth Sciences
Carleton University Ottawa
Ontario K1S 5B6 Canada
kassato2@connect.carleton.ca
gma@earthsci.carleton.ca

Manuscript received 28 September 2006