



جامعة الملك عبد الله  
للعلوم والتقنية  
King Abdullah University of  
Science and Technology

## Uncertainty in Earthquake Source Imaging Due to Variations in Source Time Function and Earth Structure

Item Type	Article
Authors	Razafindrakoto, Hoby; Mai, Paul Martin
Citation	Uncertainty in Earthquake Source Imaging Due to Variations in Source Time Function and Earth Structure 2014, 104 (2):855 Bulletin of the Seismological Society of America
Eprint version	Publisher's Version/PDF
DOI	<a href="https://doi.org/10.1785/0120130195">10.1785/0120130195</a>
Publisher	Seismological Society of America (SSA)
Journal	Bulletin of the Seismological Society of America
Rights	Archived with thanks to Bulletin of the Seismological Society of America Copyright © 2015 by the Seismological Society of America
Download date	04/08/2022 17:20:07
Link to Item	<a href="http://hdl.handle.net/10754/555762">http://hdl.handle.net/10754/555762</a>

# Uncertainty in Earthquake Source Imaging Due to Variations in Source Time Function and Earth Structure

by Hoby N. T. Razafindrakoto and P. Martin Mai

**Abstract** One way to improve the accuracy and reliability of kinematic earthquake source imaging is to investigate the origin of uncertainty and to minimize their effects. The difficulties in kinematic source inversion arise from the nonlinearity of the problem, nonunique choices in the parameterization, and observational errors. We analyze particularly the uncertainty related to the choice of the source time function (STF) and the variability in Earth structure. We consider a synthetic data set generated from a spontaneous dynamic rupture calculation. Using Bayesian inference, we map the solution space of peak slip rate, rupture time, and rise time to characterize the kinematic rupture in terms of posterior density functions. Our test to investigate the effect of the choice of STF reveals that all three tested STFs (isosceles triangle, regularized Yoffe with acceleration time of 0.1 and 0.3 s) retrieve the patch of high slip and slip rate around the hypocenter. However, the use of an isosceles triangle as STF artificially accelerates the rupture to propagate faster than the target solution. It additionally generates an artificial linear correlation between rupture onset time and rise time. These appear to compensate for the dynamic source effects that are not included in the symmetric triangular STF. The exact rise time for the tested STFs is difficult to resolve due to the small amount of radiated seismic moment in the tail of STF. To highlight the effect of Earth structure variability, we perform inversions including the uncertainty in the wavespeed only, and variability in both wavespeed and layer depth. We find that little difference is noticeable between the resulting rupture model uncertainties from these two parameterizations. Both significantly broaden the posterior densities and cause faster rupture propagation particularly near the hypocenter due to the major velocity change at the depth where the fault is located.

*Online Material:* Figures of inverted rupture models, trade-off curve, and posterior probability density function.

## Introduction

An improved understanding of earthquake physics and modern seismic-hazard assessment is a key requirement for reducing earthquake risks. In this context, robust kinematic source models are needed that are realistic representations of the slip history of an earthquake on a fault and provide information on the space–time evolution of ruptures. In the 1980s, [Olson and Apsel \(1982\)](#) and [Hartzell and Heaton \(1983\)](#) developed the linear multitime window inversion method to estimate rupture speed and slip amplitude on the fault. In this approach, the slip history at each point of the fault is presented as a superposition of several elementary slip functions, which allow the point to slip more than once. Applying this multitime window approach, [Sekiguchi and Itawa \(2002\)](#), for example, analyzed the rupture process of the 1999 Kocaeli earthquake, and [Zhou \*et al.\* \(2004\)](#) examined the 1999 Chi-

Chi earthquake; this latter study also included resolution analysis of the source history. Related studies revealed that this approach is capable of capturing details of source complexity ([Wald and Heaton, 1994](#); [Graves and Wald, 2001](#); [Wald and Graves, 2001](#)). More recently, [Lee \*et al.\* \(2011\)](#) applied this technique and detected an episode of slip reactivation during the rupture propagation of the 2011 Tohoku-Oki earthquake.

As computational resources increased, seismologists started to consider the full nonlinear formulation of source inversion without linearization. In this case, a predefined analytical source time function (STF) is required, and one estimates the space–time rupture evolution in terms of slip velocity, rupture time, slip duration, and rake angle at each point of the fault plane. This nonlinear formulation may help

to better understand the underlying dynamic rupture process (Tinti, Spudich, and Cocco, 2005). Optimization strategies such as simulated annealing (e.g., Hartzell *et al.*, 1996; Bouchon *et al.*, 2000; Delouis *et al.*, 2002; Liu and Archuleta, 2004), neighborhood searches (e.g., Vallée and Bouchon, 2004), and genetic algorithms (e.g., Emolo and Zollo, 2005) have been applied to estimate the slip rate, rise time, rupture time, and rake angle that characterize the kinematic rupture process. These techniques are based on an iterative approach in which random or new generations of rupture models are created for each iteration. The reliability of the model is then assessed based on the fits of the corresponding synthetics with the observed data. Once convergence is achieved, all models that fit the data within some chosen limit can be accepted. These studies either rely only on one model that fits the data best or consider a limited number of models to compute a corresponding mean model.

However, a single model is never fully representative of the model space in a nonlinear inverse problem. Therefore, even though macroscopic parameters such as seismic moment and average displacement may be almost identical in all models for the same earthquake, the variability of slip on the fault may be significant. For example, kinematic models for the 1999 Izmit earthquake (Bouchon *et al.*, 2000; Delouis *et al.*, 2002; Sekiguchi and Itawa, 2002; Clévéde *et al.*, 2004) obtained using various data sets, making alternative assumption on the source parameterization and deploying different inversion algorithms, clearly illustrate this nonuniqueness and the resulting variations in the final rupture models. Hence, without appropriate uncertainty assessment of each of these models, it is difficult to examine their reliability (i.e., their common and stable features), as well as the limits of resolution.

Custodio *et al.* (2005) and Hartzell *et al.* (2007) used the 2004 Parkfield earthquake to show how the inverted source model depends on chosen data and assumptions, such as the misfit function and size of fault plane. Their findings call for rigorous and quantitative analysis of uncertainty to better understand the main factors that lead to variability in the models. Uncertainty analyses were also carried out for the 1989 Loma Prieta earthquakes by Hartzell *et al.* (1991) and Emolo and Zollo (2005), who defined a Gaussian probability density function (PDF) around the best-fitting model obtained through a genetic algorithm to estimate kinematic source models. Piatanesi *et al.* (2007) performed an uncertainty estimation from statistical analysis of the ensemble of models generated by an optimization algorithm, whereas Monelli and Mai (2008) proposed the use of a Bayesian inference technique to estimate the model uncertainty by mapping posterior PDFs of source parameters. Producing PDFs of the model parameters defines the admissible solution space more comprehensively.

In this study, we extend the approach of Monelli *et al.* (2009) by incorporating the spatial correlation between neighboring nodes and using the regularized Yoffe function (Tinti, Fukuyama, *et al.*, 2005) as STF, which is compatible with dynamic rupture simulations. Additionally, we account

for epistemic uncertainty (e.g., Abrahamson and Bommer, 2005) associated with inadequate knowledge about physical assumptions regarding a specific model under investigation. Among the different sources of uncertainty related to source inversion, we highlight the epistemic uncertainty associated with Earth's crustal structure. We present results using synthetic tests developed in the context of the source inversion validation exercise (SIV; Mai *et al.*, 2007; Page *et al.*, 2011). The reference model stems from a spontaneous dynamic rupture calculation that assumes random initial stress on an 80°-dipping strike-slip fault, with 90° strike, producing a right-lateral strike-slip event with seismic moment  $M_0 = 1.06 \times 10^{19}$  N·m. We use the corresponding three-component velocity waveforms at 40 well-distributed local sites (Fig. 1) to conduct the inversion; seismograms are filtered using a Butterworth bandpass in the range of 0.01–1 Hz. With this test, we assess the ability of kinematic source inversion to retrieve the physical aspects of the rupture process from a dynamic rupture simulation.

### Formulation of Kinematic Source Imaging

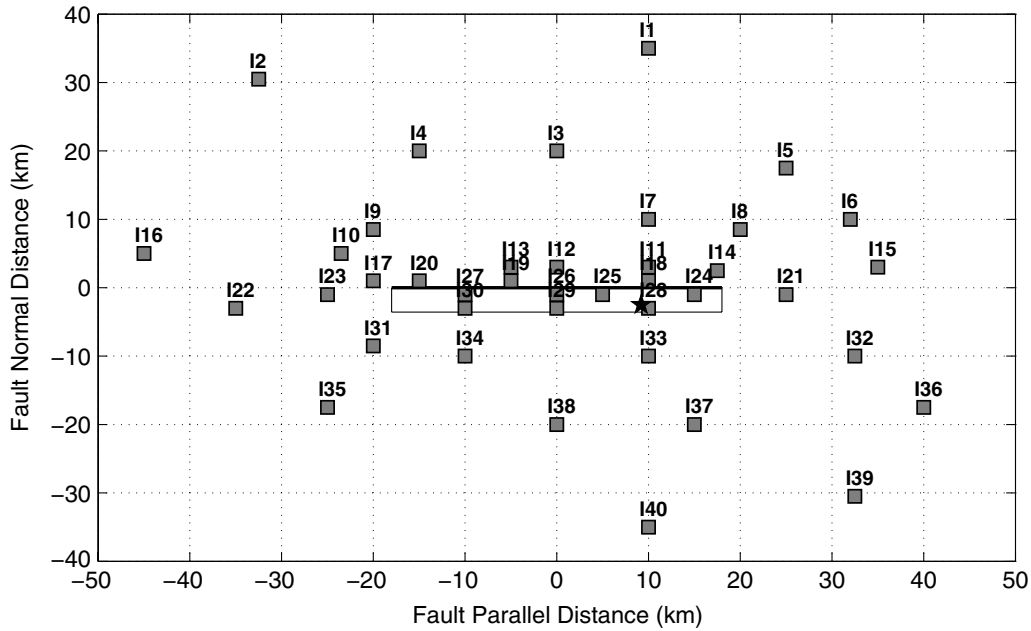
We use the representation theorem in the formulation of Spudich and Archuleta (1987) to infer kinematic rupture parameters. The ground velocity recorded at the Earth's surface is expressed as a convolution of a local STF with the corresponding tractions on the fault, written in the frequency domain as

$$\dot{u}_k(y, w) = \iint_{\Sigma} \dot{s}(x, w) T_k(x, w; y, 0) d\Sigma, \quad (1)$$

in which  $\dot{u}_k(y, w)$  is the Fourier transform of  $\dot{u}_k(y, t)$ , the  $k$ th component of ground velocity at an observation point  $y$ ;  $\dot{s}(x, w)$  is the Fourier transform of  $\dot{s}(x, t)$ , the slip-rate function at a point  $x$  on the fault;  $T_k(x, w; y, 0)$  is the Fourier transform of the traction vector at a point  $x$  on the fault caused by a point impulse force in the  $k$  direction at the observer location  $y$  and angular frequency  $w$ . This formulation obeys the reciprocity theorem, meaning that the location of source and observation points can be switched, and the exact same response will be obtained. Assuming an appropriate Earth structure in the source region, we compute the traction,  $T_k$ , applying a frequency–wavenumber integration method (Spudich and Xu, 2002).

In the framework of inverse theory, the representation theorem in equation (1) can be cast as  $d = g(\theta)$ , in which  $d$  is the data composed of the observed (velocity) waveforms,  $\theta$  contains the parameters that define the earthquake source properties (e.g., peak slip rate, rise time, rupture onset time, and rake), and  $g$  is an operator that describes the Earth structure response. The purpose in this inverse problem formulation is to infer the source parameters  $\theta$ , given synthetic data  $d$  and the calculated Earth response  $g$ .

One may use any optimization technique to obtain one or more models with minimum misfit. However, several models with very different source-parameter distributions



**Figure 1.** Source–station geometry. The data consist of waveforms at 40 stations (gray squares), the black star denotes the epicenter, and the black line marks the surface projection of the 80°-dipping fault.

may fit the data equally well (Mai *et al.*, 2007). This limitation of the optimization can be overcome using a Bayesian technique (Monelli and Mai 2008; Monelli *et al.*, 2009) that infers posterior PDFs to characterize the ensemble of all possible parameters (peak slip rate, rise time, and rupture time) describing the STFs that are consistent with the data and the available prior information. This technique comprises a stochastic process in which the major sources of uncertainty include data noise, improper knowledge about Earth structure, nonunique fault parameterization, and intrinsic randomness of physical processes. To understand the effects of these factors and to eventually minimize them, our study introduces variability in the Earth structure and the assumed STF in the uncertainty analysis.

### Kinematic Source Parameterization

We assume rupture occurs on a planar fault with known length and width; for real earthquakes, these parameters can be estimated based on geodetic information and aftershock distributions, albeit with considerable uncertainty. However, the determination of fault geometry and its associated uncertainty is beyond the scope of this study and will be left for a future extension of this work. The fault plane is discretized into different nodes. The smaller the cell size, the more detailed information we may obtain on the rupture process. However, the data cannot resolve a cell size finer than the shortest wavelength corresponding to the highest frequency considered. In this study, we use a frequency range up to 1 Hz corresponding to a wavelength of 3–3.5 km in the upper crustal layers ( $z < 18$  km) of the given Earth model. Because the shortest wavelength determines the finest useful grid spacing on the rupture plane, we choose a grid spacing of 3 km

along strike and 4 km along dip. Inside each cell, source parameters (peak slip rate, rise time, rupture onset time, and rake) are assumed to be constant. The final slip on each node is not inferred, but calculated from the chosen STF. In our parameterization, the total number of parameters is then 195.

In this study, we test the effects of two different STFs, an isosceles triangle and the regularized Yoffe function. The peak slip rate corresponds to the height of this function, rise time is defined as its duration, and the rupture onset time is the time when a particular point on the fault starts to slip. For the regularized Yoffe STF, an additional parameter needs to be considered: the duration of slip acceleration (acceleration time  $T_{acc}$ ). We choose  $T_{acc}$  deterministically and invert for peak slip velocity, rise time, and rupture onset time.

### Bayesian Inference

In a Bayesian approach, the purpose is to quantify the posterior distribution that maps the solution space. It is defined as the probability density of model parameters  $\theta$  given data  $d$ . Following Bayes theorem, the posterior PDF  $\pi(\theta|d)$  is obtained by combining the prior probability distribution  $\pi(\theta)$  on model parameters with the likelihood  $f(d|\theta)$  of observing these data given the model parameters  $\theta$ :

$$\underbrace{\pi(\theta|d)}_{\text{posterior probability}} = \frac{\overbrace{f(d|\theta)}^{\text{likelihood}} \overbrace{\pi(\theta)}^{\text{prior probability}}}{\pi(d)} = \frac{f(d|\theta)\pi(\theta)}{\int f(d|\theta)\pi(\theta)d\theta}. \quad (2)$$

In equation (2),  $\pi(d)$  is the evidence, or marginal likelihood, which is independent of the model  $\theta$ . Therefore, we can

Table 1  
Search Ranges for Peak Slip Rate, Rise Time, and Rupture Speed

Case	Peak Slip Rate (cm/s)	Rise Time (s)			Rupture Speed (km/s)
		Triangle	Yoffe-0.1*	Yoffe-0.3†	
Inner nodes	0–400	0–6	0.5–10	1–10	1.5–7.1
Nodes at boundary	0–200	0–3	0.5–10	1–10	1.5–7.1

\*Yoffe source time function with acceleration time of 0.1 s.

†Yoffe source time function with acceleration time of 0.3 s.

consider it as a normalization constant (Mosegaard and Tarantola, 1995),  $k^{-1}$ , and equation (2) can be written as

$$\pi(\theta|d) = kf(d|\theta)\pi(\theta). \quad (3)$$

We use an iterative stochastic Markov chain Monte Carlo approach to generate samples of  $\theta$  from a proposed distribution. To search the vast space of potential kinematic rupture models, we apply the Metropolis algorithm (Metropolis *et al.*, 1953) that produces a chain of random models in which each candidate only depends on the previously generated model. Each candidate model is accepted as a new model based on the acceptance probability. The procedure of the Markov chain Monte Carlo, based on the Metropolis algorithm, works as follows:

1. Draw the initial state  $\theta$  in the model parameter space.
2. Generate a candidate sample of  $\theta$ , denoted as  $\theta'$ , from a proposed distribution  $q(\theta'|\theta)$ , a Gaussian centered at the current value of  $\theta$  in our case. The step size of the Markov chain is proportional to the variance of  $q(\theta'|\theta)$  and tuned according to the rejection rate. Too large a step size, for example, leads to many rejections. We choose the step size for the peak slip rate, rupture time, and rise time as 2.5 cm/s, 0.05 s, and 0.05 s, respectively, leading to acceptance rates of about 30%, in agreement with suggested acceptance rates of 30%–50% (Tarantola, 2005).
3. Accept or reject the candidate  $\theta'$  based on the acceptance probability, defined as

$$\alpha(\theta, \theta') = \min\left[1, \frac{f(d|\theta')}{f(d|\theta)}\right]. \quad (4)$$

4. Draw a uniform random number  $u \in [0, 1]$ ; if  $\alpha(\theta, \theta') > u$ , accept the candidate sample  $\theta = \theta'$ ; if  $\alpha(\theta, \theta') < u$ , reject the candidate sample (keep the current state).
5. Repeat the process to generate the next samples.

The samples obtained through this algorithm approximate the posterior distribution  $\pi(\theta|d)$  (Hoff, 2009). To minimize the effect of initial values and autocorrelation among samples, we discard an initial portion of the Markov chain and consider only every  $k$ th sample (thinning the chain). The choice of  $k$  is arbitrary, and different numerical experiments are needed to choose  $k$  appropriately. In our case, we select  $k = 100$ , as larger value of  $k$  leads to a smaller number of samples to infer the posterior, and the marginal posterior becomes nonsmooth. On the other hand, smaller  $k$  returns

nearly the same posterior, showing that the chain length has a larger effect on the posterior than does the thinning interval  $k$  (Arakawa *et al.*, 2009).

### Prior Distribution

The prior distribution incorporates our *a priori* knowledge about the model parameters. We use a uniform distribution for each parameter, assuming no further information is available aside from the range of possible values:

$$\pi(\theta) = \frac{1}{\theta_{\max} - \theta_{\min}} \quad \text{for } \theta_{\min} \leq \theta \leq \theta_{\max}, \quad (5)$$

in which  $\pi(\theta)$  is the prior density, and  $\theta_{\min}$  and  $\theta_{\max}$  are the limits of the sample space for the model parameter  $\theta$  (see Table 1). They are chosen based on physical considerations for peak slip rate and the temporal parameters (rupture time and rise time). Furthermore, for rupture time we define the range based on the rupture speed such that we allow the rupture to propagate at subshear and at intersonic speed (Archuleta, 1984; Dunham and Archuleta, 2004). However, we do not permit rerupturing of the fault. We choose the rupture speed between 0.5 and 2.0 times the shear wavespeed at the shallowest and the deepest layers of the fault zone, respectively, which correspond to a rupture speed between 1.5 and 7.1 km/s.

Along the edges of the assumed rupture plane, the range for slip is chosen such that slip occurs with a maximum possible peak slip rate of 200 cm/s, which is about half of the value at the inner nodes. This choice is more realistic than in Monelli and Mai (2008) and Monelli *et al.* (2009), who did not allow this area to slip at all. Monelli *et al.* (2009) also showed that there is a trade-off for the peak slip rate at neighboring points, characterized by a strong anticorrelation. Hence, considering them as totally independent parameters may be problematic. Therefore, we take into account the spatial correlation between neighboring patches (Jónsson *et al.*, 2002) using a 2D spatial Laplacian filter, respecting the distance between patches and the boundary of the fault plane (see Appendix).

### Likelihood Function

The likelihood function is the output of the forward modeling and is used to assess how well the model explains the data. It can be written as

$$f(d|\theta) = c \exp[-S(\theta)], \quad (6)$$

in which  $c$  is a normalization constant and  $S(\theta)$  quantifies the misfit. Assuming the data uncertainty is Gaussian, we can write

$$S(\theta) = \frac{1}{2} \{[g(\theta) - d]^T C_D^{-1} [g(\theta) - d]\}, \quad (7)$$

in which  $g(\theta)$  is the forward modeling operator,  $d$  is the data, and  $C_D$  is the covariance matrix of data errors. The expressions in equations (3) and (7) show that the covariance matrix of data errors directly determines the form of the posterior and therefore is crucial in a Bayesian approach. However, in earthquake seismology, quantifying data errors is difficult as different sources (instrumental, processing noise, and unmodeled Earth structure) need to be included. To overcome this problem, Monelli *et al.* (2009) propose an empirical formulation for the likelihood function. In contrast, Yagi and Fukahata (2008) and Bodin *et al.* (2012) follow a hierarchical Bayesian approach in which all parameters that define the data error are considered as unknown hyperparameters (Gelman *et al.*, 1995). Here, we adopt the empirical formulation of the likelihood function (Monelli *et al.*, 2009) as

$$f(d|\theta) = c \exp[-\phi(\theta)] \quad (8)$$

$$\phi(\theta) = \frac{S(\theta) - S(\theta_{\text{best}})}{S(\theta_{\text{best}})} \times 100, \quad (9)$$

in which  $S(\theta_{\text{best}})$  represents the fit of the best model  $\theta_{\text{best}}$ , obtained using an optimization based on the evolutionary algorithm (Beyer, 2001), and  $S(\theta)$  is the misfit function in equation (7).

Incorporating the spatial constraint of parameters at neighboring nodes, the likelihood function is expressed using the penalized  $L_2$  norm on  $S(\theta)$ :

$$S(\theta) = \|d - g(\theta)\|_2 + \alpha^2 \|D\theta\|_2, \quad (10)$$

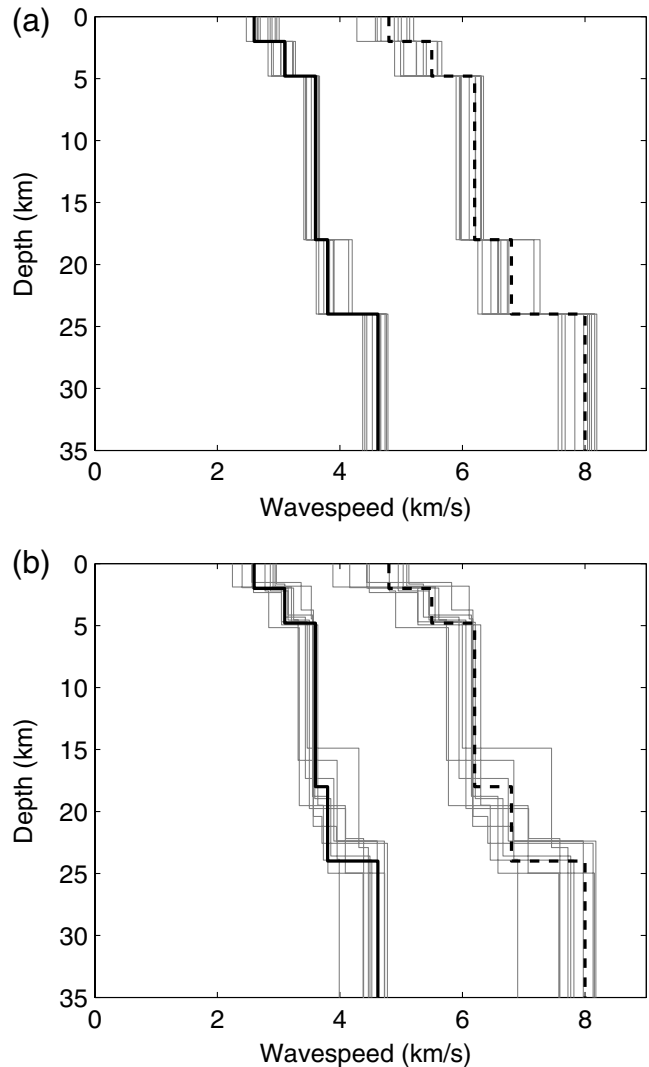
in which  $D$  is the Laplacian operator and  $\alpha$  is the parameter that controls the smoothness of the model, which can be estimated deterministically from a trade-off curve.

#### Parameterization of the Data Covariance Matrix

Following Gouveia and Scales (1998) and Tarantola (2005), the data covariance matrix  $C_D$  (equation 7) represents the combination of observational errors ( $C_{\text{obs}}$ ) and forward modeling uncertainty ( $C_{\text{theory}}$ ):

$$C_D = C_{\text{obs}} + C_{\text{theory}}. \quad (11)$$

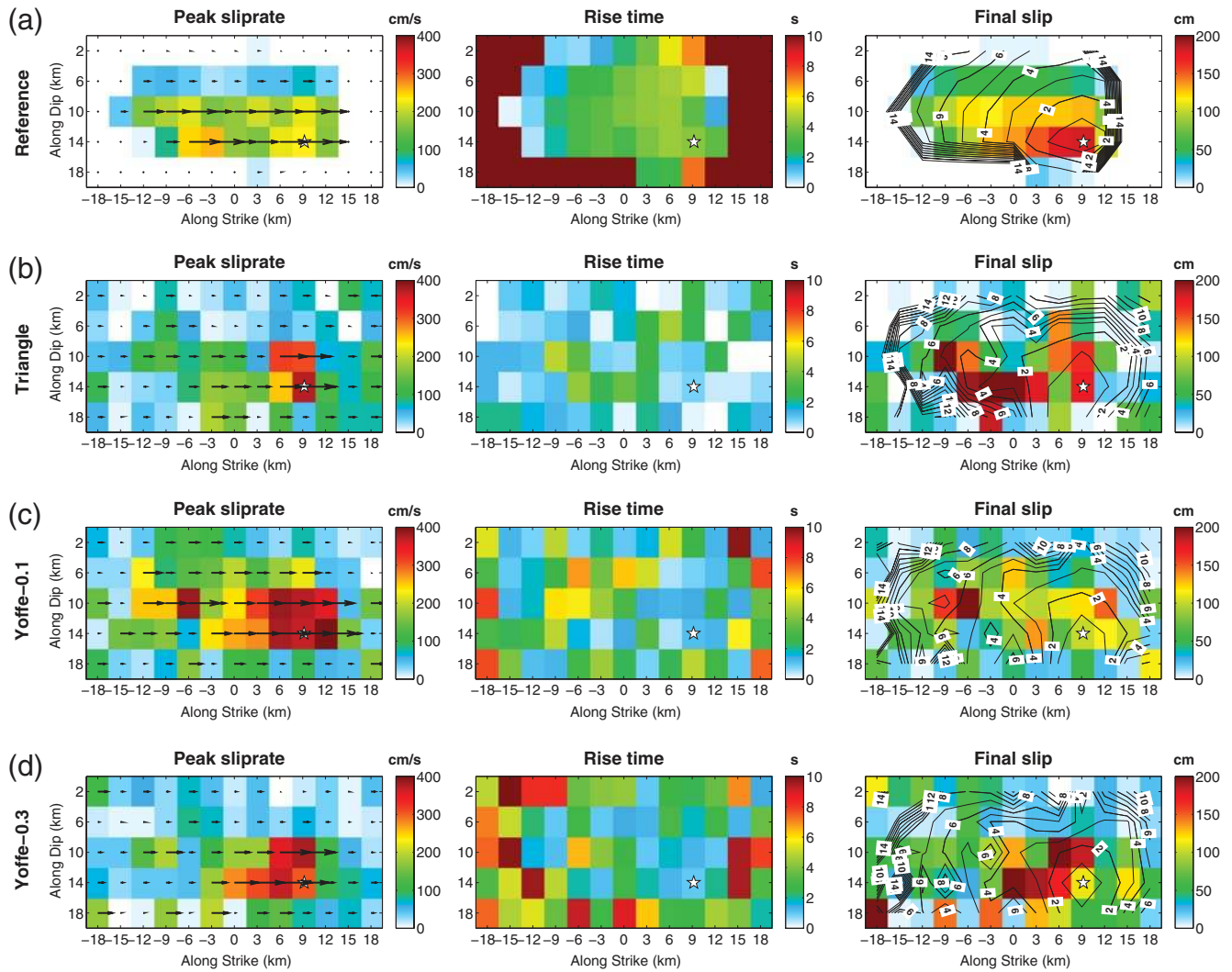
Application of this formulation in seismic source inversion is described in Duputel *et al.* (2012). Following the representation theorem (equation 1), the kinematic source model de-



**Figure 2.** Reference 1D crustal models with two different types of uncertainties. The solid black line and dashed black line show the reference  $S$  and  $P$  wavespeeds, respectively; the gray lines around them correspond to random samples of  $V_P$  and  $V_S$ , drawn from a normal distribution with 10% standard deviation. (a) Wavespeed variability only; (b) wavespeed and layer depth variabilities.

pends on the choice of an Earth model. However, generally we do not have complete knowledge of the true velocity structure. For this source of modeling error, we define the covariance matrix  $C_{\text{theory}}$  as the covariance describing the uncertainty in Earth structure. By computing the variability in the synthetics obtained from the combination of the best model ( $m_{\text{best}}$ ) and variations on crustal models (Fig. 2), we deduce the data covariance that we insert in equations (7) and (9).

Two cases are studied to include uncertainty of crustal structure: (1) accounting for wavespeed uncertainty only and (2) accounting for simultaneous variations in wavespeed and layer depth. Mooney (1989) proposes an uncertainty of 3%–4% for the seismic velocities in global crustal models. However, to be conservative on a local scale when consid-



**Figure 3.** (a) Reference and inverted rupture models using an evolutionary algorithm assuming different STFs, (b) isosceles triangle, (c) Yoffe with acceleration time of 0.1 s, and (d) Yoffe with acceleration time of 0.3 s. Arrows in the left column indicate slip direction and amplitude. Contour lines in the right column mark the rupture time distribution (contour interval is 1 s).

ering heterogeneity near the surface and the trade-off between Moho depth and the crustal velocity, we assume a standard deviation of 10% for the wavespeeds near the surface and around the Moho. In all cases, we maintain the ratio  $V_P/V_S$  around 1.73, the standard value for rocks without significant fluid content. For layer depths, we consider a standard deviation of 10%.

### Modeling Results

We present here the result of the inversion using the data in the SIV exercise (Mai *et al.*, 2007; Page *et al.*, 2011). The reference model (Fig. 3a) stems from a spontaneous dynamic rupture simulation, with heterogeneous initial stress conditions and a constant slip weakening distance in the inner part of the fault plane. Rupture occurs over a rectangular fault of about  $36 \text{ km} \times 16 \text{ km}$  with a slip patch around the hypocenter, between 10 and 14 km depth, extending about 20 km

along strike, and with a maximum of 180 cm. The peak slip-rate distribution comprises two patches, with maximum values of about 250 cm/s near the hypocenter and at 12 km distance from the hypocenter in the strike direction.

We assess the quality of the synthetics based on the variance reduction (Cohee and Beroza, 1994):

$$\text{VR} = \left( 1 - \frac{[d - g(\theta)]^T C_D^{-1} [d - g(\theta)]}{d^T C_D^{-1} d} \right) \times 100\%, \quad (12)$$

in which  $g(\theta)$  is the modeling result,  $d$  is the data,  $C_D$  is the covariance matrix of data errors, and superscript  $T$  denotes transposition.

Using the Metropolis algorithm, we produce 600,000 samples of rupture models, each of which consists of 65 node points at which we estimate the kinematic source parameters. Discarding an initial portion of the Markov chain and considering every 100th sample, this leads to a total sample size

of 5500 to build the PDF. As we incorporate the variability in crustal structure and the STF, including two different types of Earth structure variabilities (composed of 11 Earth structures each) and three different STFs (an isosceles triangle and two regularized Yoffe functions), respectively, we produce about  $5500 \times 22$  (Earth structures)  $\times 3$  (STF) rupture model samples to understand these variabilities. We then examine the posterior PDFs and the corresponding statistical estimates to assess the accuracy of the resulting source parameters at different points on the fault. We also show the 2D marginal posterior PDFs, to visualize the correlation between parameters at the same location and between neighboring points for the same parameter.

### Source Time Function Variability

In this section, we present inversion results using the input 1D crustal model (Fig. 2) without crustal structure uncertainty. The posterior PDFs in this part are generated using the empirical likelihood function (equation 9), which requires the level of fit corresponding to the model that best fits the data.

In Figure 3, we present the optimal models obtained from an evolutionary algorithm (Beyer, 2001; Monelli *et al.*, 2009), using as STF an isosceles triangle and a regularized Yoffe function. For the Yoffe function, we tested different values of acceleration time, 0.1 and 0.3 s, respectively, following Tinti, Fukuyama, *et al.* (2005), who suggested for  $T_{\text{acc}}$  a value between 0.1 and 0.38 s. All models show high slip and slip-rate patches around the hypocenter with maximum values between 300 and 400 cm/s for the peak slip rate, and 100–200 cm for the calculated slip, respectively. For the Yoffe function with  $T_{\text{acc}} = 0.1$  s (Yoffe-0.1), we find the peak slip-rate patch extends about 12 km along strike. However, for both the isosceles triangle STF and the Yoffe function with  $T_{\text{acc}} = 0.3$  s (Yoffe-0.3), a more compact region of high slip rate is inferred, with an extension of about 6 km along strike. In terms of rupture time, there appears to be a discrepancy between source models, showing that the rupture propagates faster for the isosceles triangle than for the regularized Yoffe (see contours on final slip panels, Fig. 3).

To examine the resolution of each inverted slip model and how it compares to the reference model, we calculate the spatial cross correlation between reference (Fig. 3a) and inverted models, defined as

$$\rho = \frac{\sum_{j=1}^n R_j I_j}{\left[ \sum_{j=1}^n R_j^2 \sum_{j=1}^n I_j^2 \right]^{1/2}}, \quad (13)$$

in which  $n$  is the number of points on the fault,  $R_j$  and  $I_j$  represent the reference and inverted rupture models, respectively, at one point  $j$  on the fault (e.g., Graves and Wald, 2001; Shao and Ji, 2012). The value of  $\rho$  could vary from 0 (no correlation) to 1 (best correlation). We apply this formulation to the maps of peak slip rate, slip, and rise time (see

Table 2

Spatial Cross-Correlation Coefficient between the Reference and Each Inverted Model

Model	Peak Slip Rate	Slip	Rise Time
Triangle	0.71	0.74	0.60
Yoffe-0.1*	0.83	0.84	0.76
Yoffe-0.3†	0.72	0.73	0.80

\*Yoffe source time function with acceleration time of 0.1 s.

†Yoffe source time function with acceleration time of 0.3 s.

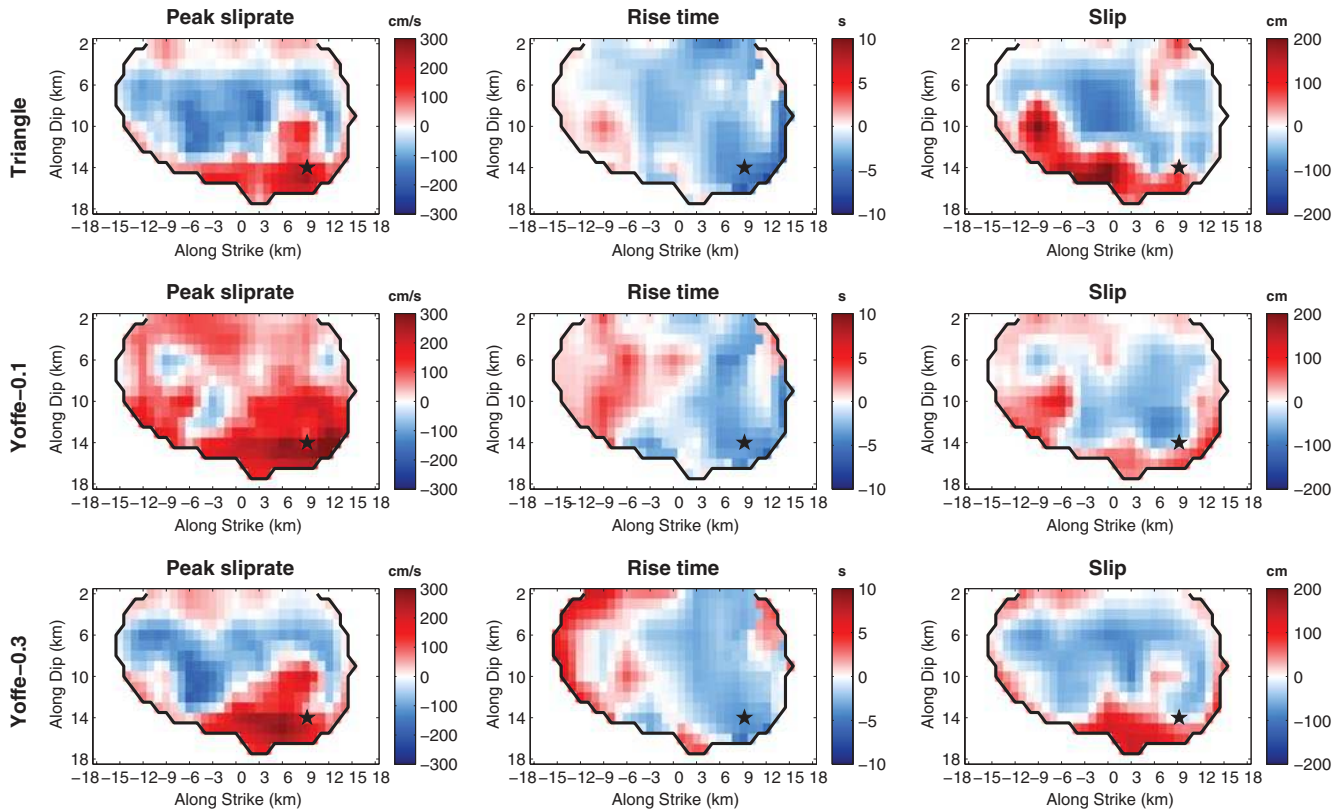
Table 2), and find that the Yoffe-0.1 rupture model is most similar to the reference model, with slip and peak slip-rate correlations of 0.83 and 0.84, respectively. The isosceles triangle and Yoffe-0.3 models have similar ability to reproduce the feature of the reference model, however, with lower spatial cross correlation, 0.74 and 0.73 for the slip correlations, and 0.71 and 0.72 for the peak slip-rate correlations. However, in terms of rise time, the Yoffe-0.3 model is most similar to the reference model with cross correlation of 0.80 compared to 0.60 and 0.76 for the isosceles triangle and Yoffe-0.1 models, respectively.

In Figure 4, we investigate the discrepancies between reference and inverted rupture models in terms of amplitude differences of source parameters at individual point on the fault. For this error analysis, we do not consider the area of zero slip in the reference model (Fig. 3a), the region into which the dynamic rupture did not propagate. In Figure 4, this area is marked with a black contour line. Figure 4 shows that the residuals are not entirely random, but follow certain trends. The peak slip-rate residuals for the isosceles triangle and Yoffe-0.3 show similar pattern, with overpredicting around the hypocenter and underpredicting away from the hypocenter. However, the peak slip-rate residual for Yoffe-0.1 shows a different pattern, with higher slip rates over almost the entire rupture area. The discrepancies of peak slip rate near the hypocenter are related to the spontaneous nucleation of the dynamic rupture simulation, which is difficult to resolve by inversion. The medians of the ratio of reference to inverted peak slip rate over the entire rupture area are 1.23, 1.03, and 0.66 for the isosceles triangle, Yoffe-0.3, and Yoffe-01, respectively.

To unambiguously define rise time for the three different STFs used, we use the time to accumulate 5%–95% of the total slip in this error analysis. The rise time residual appears to be overestimated only near the edge of the rupture and underestimated over most of the fault. This originates from the almost negligible seismic radiation in the tail of the dynamic STF, which is difficult to constrain using a simple analytical STF. Finally, for the slip amplitude, the residuals for the three STFs show similar shape, with large residuals along the edge of the area of zero slip. The medians of the ratio of reference to inverted slip are 1.21, 0.98, and 1.30 for the isosceles triangle, Yoffe-0.1, and Yoffe-0.3, respectively.

Next, we compare the STF of the dynamic reference model and the inverted solutions using the three different





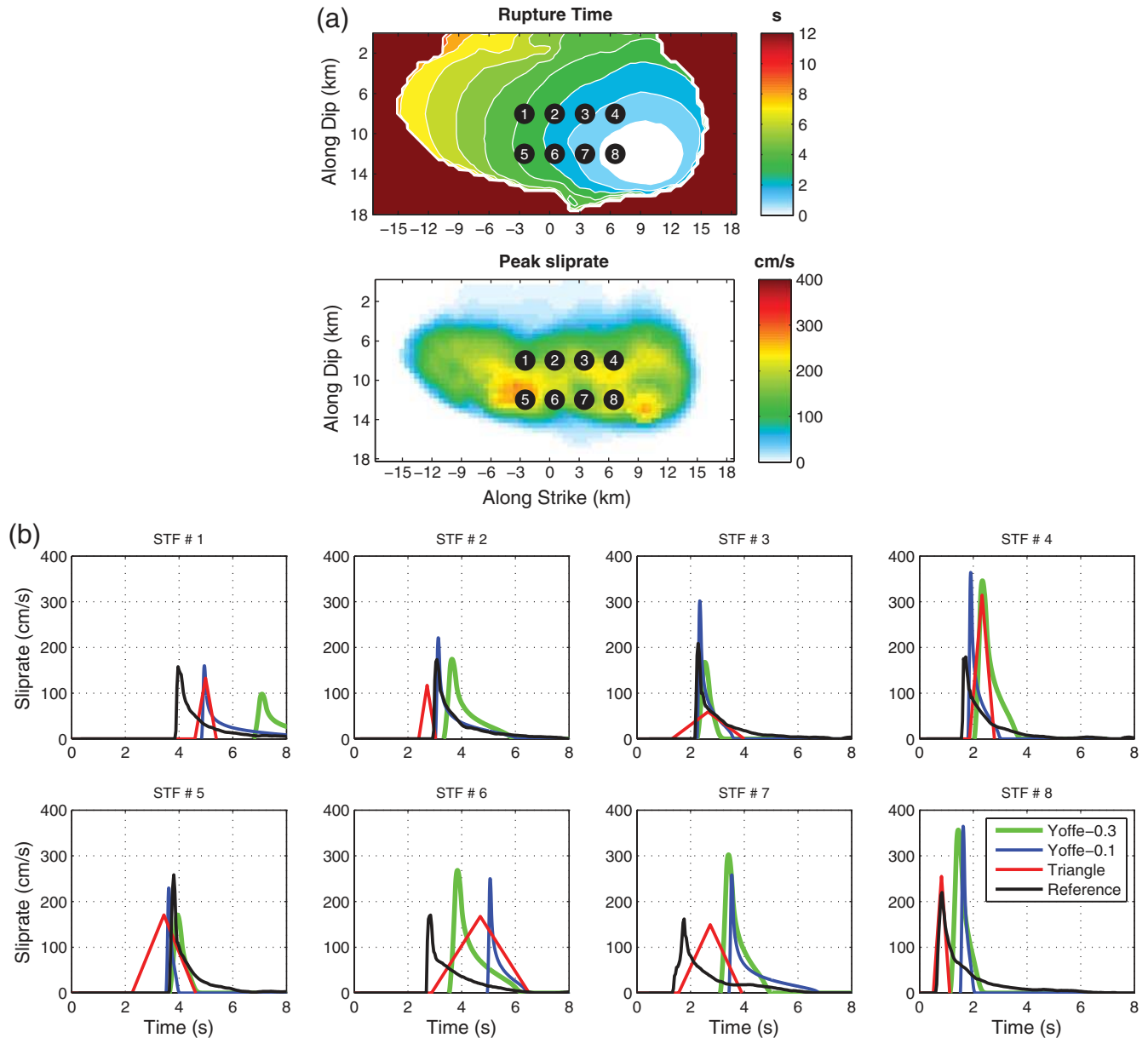
**Figure 4.** Amplitude differences between the reference and inverted model parameters (peak slip rate, rise time, and slip amplitude). The black contour line limits the area of the fault used in the error analysis.

STFs at eight selected points (Fig. 5). When peak slip rate is low (less than 2 m/s) in the reference model, the imaged slip-rate pulse occurs later for all STFs, with a delay of 0.5–2 s. Conversely, in cases when peak slip rate is greater than 2 m/s, the peaks of the three STFs occur at around the same time as in the reference model. We attribute this observation to the difficulty of resolving the temporal rupture pattern, demonstrating the nonuniqueness of the problem, as these rupture times are all acceptable values when examining the posterior PDFs at these points on the fault. Part of this uncertainty originates from the trade-off between neighboring points of the fault. Including spatial smoothing between neighboring points, with a smoothing parameter  $\alpha^2$  of 20, generally reduces this time difference of the peaks, particularly for Yoffe-0.1 close to the hypocenter (see Appendix).

Figure 6 presents the levels of fit of the synthetics obtained from the best-fitting models with the data at a selection of 12 stations. At all stations, the predictions fit the data well, with an average variance reduction of 93.9%, 94.2%, and 94.1% for the triangle, Yoffe-0.1, and Yoffe-0.3, respectively. To check the dependence of these best models on the initial state of the random model parameters, we run the evolutionary algorithm using different initial states and find that the high slip and slip-rate patches around the hypocenter are maintained for the generated rupture models. © The results of this test are available in the electronic supplement to this article (Figs. S1, S2, and S3).

Figure 7 shows posterior PDFs for peak slip rate, rise time, rupture time, and the calculated slip, at two points of the fault (points 4 and 5 in Fig. 5) located at 5 and 12 km from the hypocenter, for three different STFs. They reveal that the characteristics of the uncertainty vary from point to point, and they generally do not follow a Gaussian distribution due to the nonlinearity of the kinematic source imaging problem. The peak slip rate PDFs close to the hypocenter (point 4 in Fig. 5), for example, show distributions that are skewed toward higher values for the three STFs, with a larger uncertainty for Yoffe-0.3. Farther from the hypocenter, on the other hand, the PDFs become broader, following nearly uniform distributions. Therefore, the PDFs at those two points suggest the accuracy of source parameters decreases with increasing distance from the hypocenter.

Another interesting aspect of these PDFs is that the rupture for the triangle propagates faster than for the Yoffe (see PDFs of rupture time, Fig. 7). This feature of rupture behavior originates from the sensitivity of ground motion to peak slip rate. In this case, to obtain the correct timing for the peak slip rate, the rupture time needs to be earlier due to the symmetric form of this STF. The slip-rate functions on points 3 and 5 (Fig. 5) nicely illustrate this process. At these specific points, the rupture times are artificially earlier by about 0.8 s (or about 25%) compared with the Yoffe function, and the target model (Fig. 7b). This incorrect timing has the impact of overestimating the rupture speed, which is one of the key



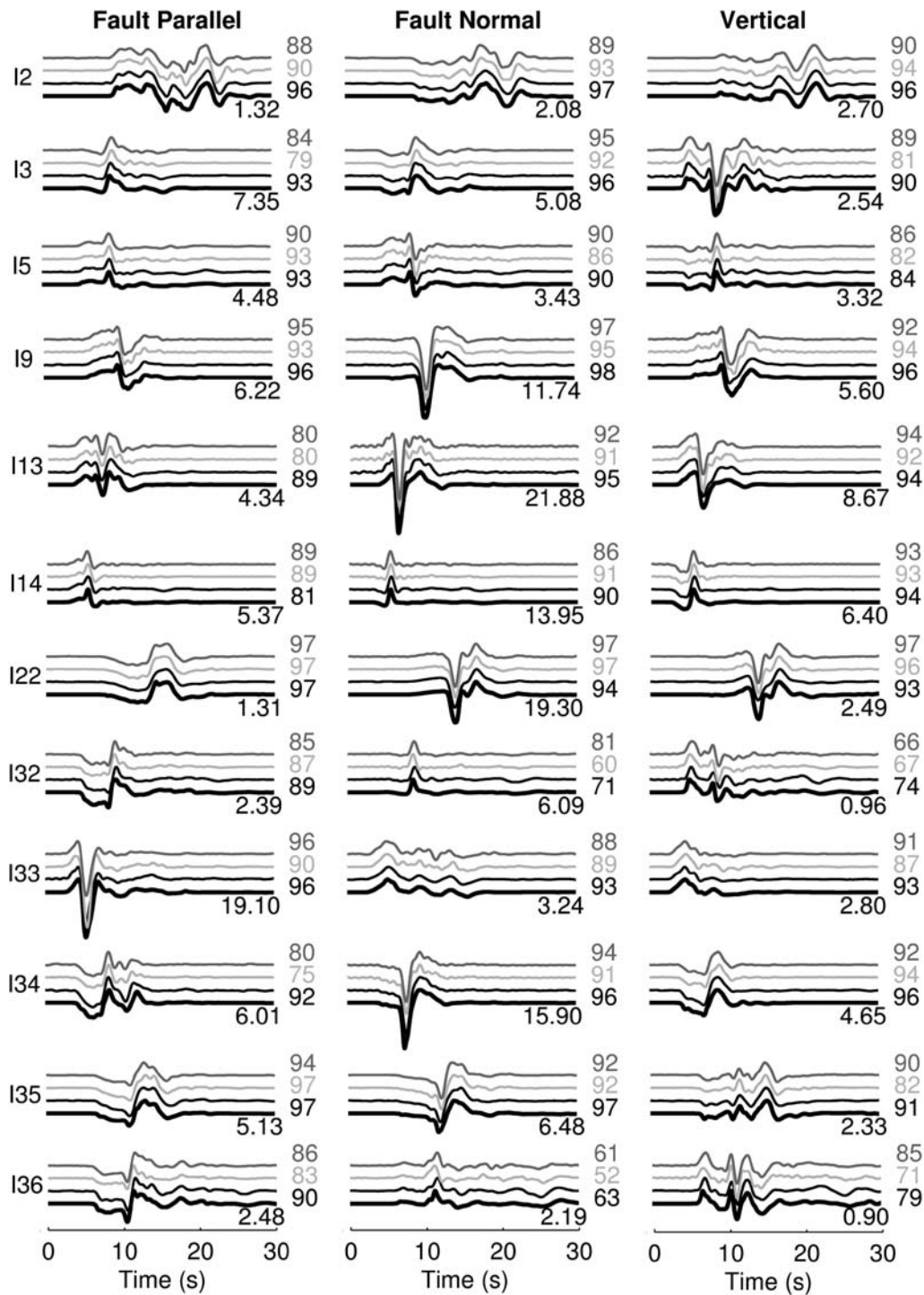
**Figure 5.** (a) Reference rupture model with eight selected points on the fault. (b) STF at eight points on the fault, for the reference model (black), and from the inversion results, using as STF the isosceles triangle (red), Yoffe-0.1 (blue), and Yoffe-0.3 (green).

parameters needed to better understand earthquake source physics.

Figures 8 and 9 present the maximum, mean, and median posterior models for the triangle and Yoffe-0.1, respectively. These statistical estimates are extracted from the posterior PDFs. At each node, we find the kinematic source parameters according to the maximum, mean, and median independently of the other node. We then assemble the composite to obtain the model estimates in Figures 8 and 9. These figures show that all models capture the main pattern of the rupture, characterized by high slip and slip-rate patches around the hypocenter (compare to Fig. 3a). However, the roughness of the models and the extension of the patches change between the various estimates. In addition,

the rupture time PDFs are skewed in the case of the triangle STF solutions, such that the inferred rupture times for the maximum extracted from the posterior arrive earlier than for the mean/median, implying faster rupture propagation.

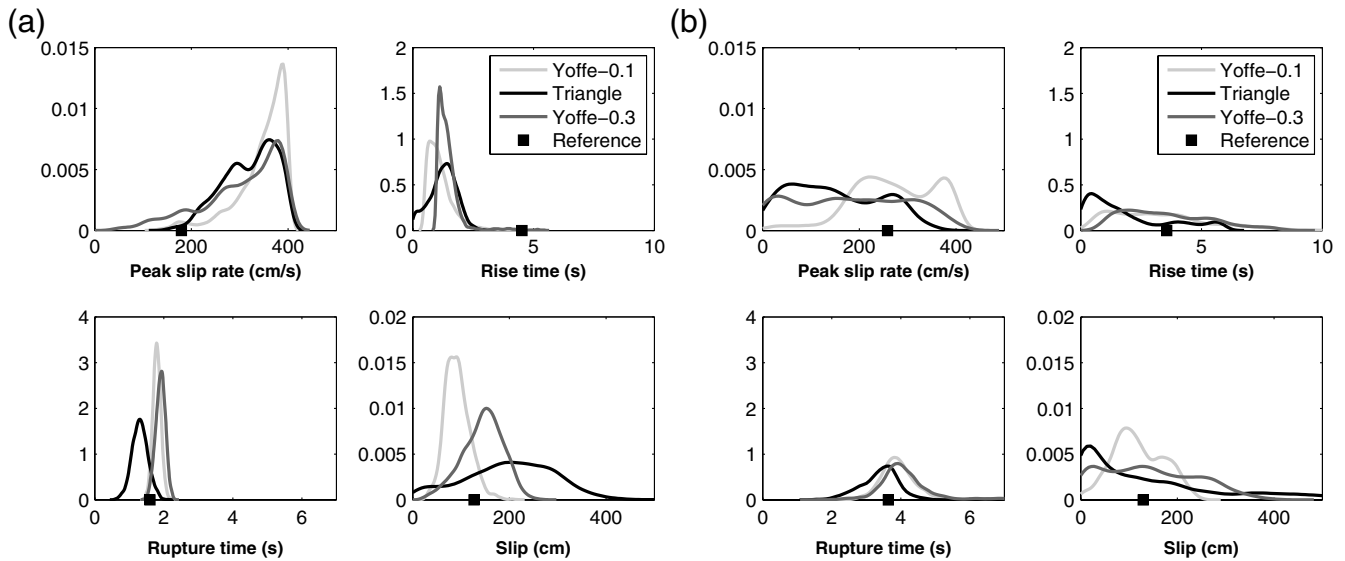
In addition, we statistically compare the estimated models with the reference model computing the 2D spatial cross correlation between source parameters (Mai *et al.*, 2007). As listed in Table 3, we find that for the triangle STF, the maximum posterior model has the lowest cross correlation for the peak slip rate, slip, and rise time (0.72, 0.61, and 0.55, respectively). On the other hand, the mean and median models have comparable cross-correlation values. They both have a correlation of 0.77 for the peak slip rate, 0.64 and 0.67 for the slip, and 0.73 and 0.76 for the rise time and have higher



**Figure 6.** Levels of fit produced by the best model for the isosceles triangle (thin black line), Yoffe-0.1 (light gray), and Yoffe-0.3 (dark gray) with the synthetic data (thick black line). The maximum velocity for the synthetic record (cm/s) for each waveform is shown below each black trace, and the variance reduction for each modeled seismogram is given at the end of the traces.

cross-correlation value compared with the maximum model. For Yoffe-0.1 on the other hand, the three model estimates (maximum, mean, and median) have similar ability to represent the reference model (see Table 3), and with larger values than for the triangle STF. Therefore, the Yoffe STF is a better representation of the dynamic rupture behavior.

As we assess the levels of fit of the synthetics with respect to the synthetic data, we find that among these estimates, the median posterior for the triangle STF explains the data best, with a variance reduction of 93% compared with the maximum and mean, which give 66% and 87%, respectively. The seismograms of the maximum posterior



**Figure 7.** The marginal posterior density for peak slip rate, rise time, rupture onset time, and slip at two points on the fault located (a) 5 km from the hypocenter and (b) 12 km from the hypocenter (indicated as points 4 and 5, respectively, in Figure 5). We assume as STF the isosceles triangle (black), Yoffe-0.1 (light gray), and Yoffe-0.3 (dark gray).

show differences in timing, due to the fast rupture. For the Yoffe function (Fig. 9) on the other hand, the overall pattern of the source models are similar, and the fits vary from 89% to 92%. The levels of fit is consistent with the cross-correlation comparison, as the higher the variance reduction, the higher the cross-correlation coefficient. Therefore, both analyses show that the Yoffe STF more accurately captures the dynamic rupture behavior.

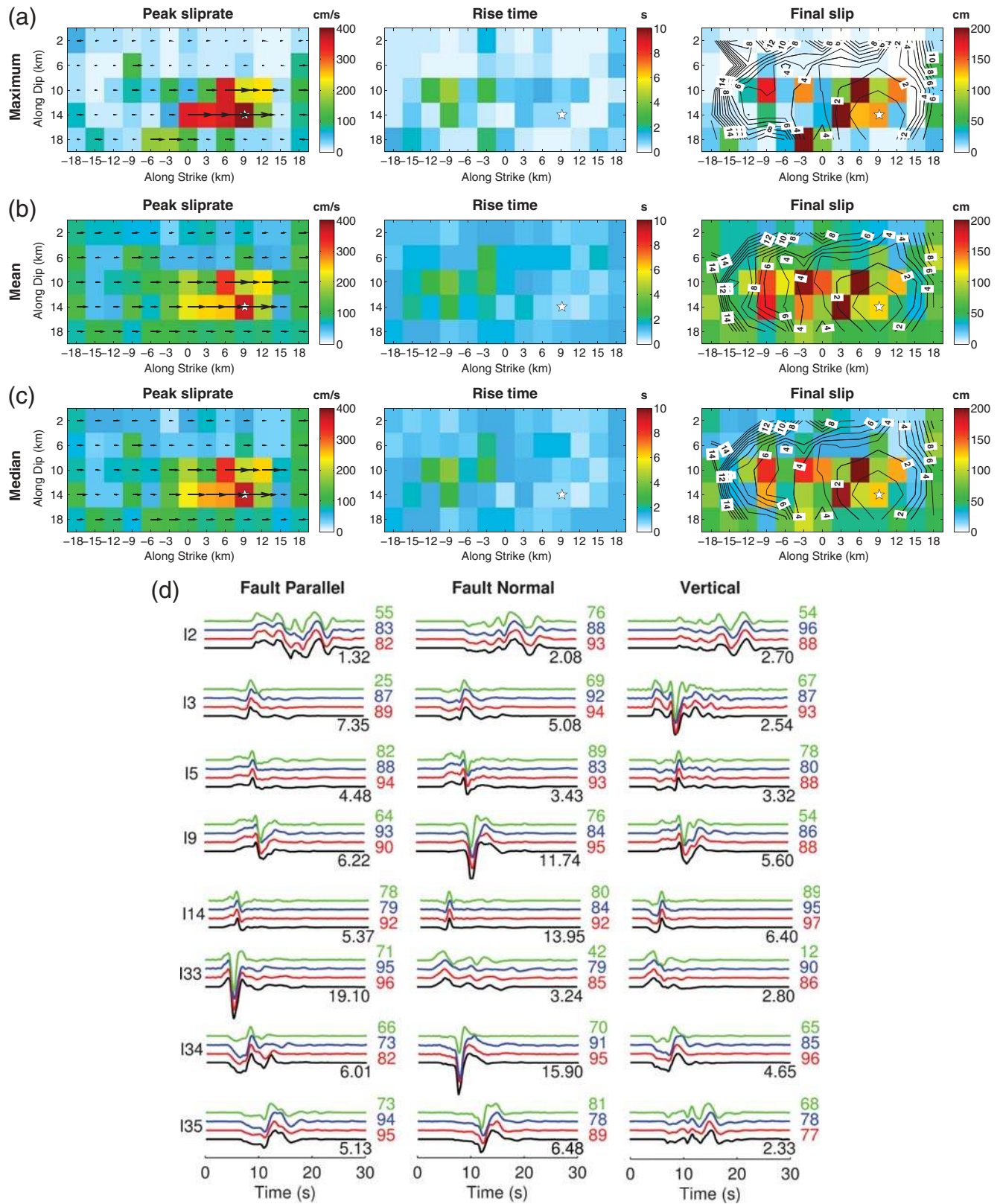
Figure 10 displays the marginal distribution of the seismic moment generated from the posterior source models and considering the three different STFs. The maximum values of the seismic moment marginals for isosceles triangle, Yoffe-0.1, and Yoffe-0.3 are about  $1.1 \times 10^{19}$ ,  $1.3 \times 10^{19}$ , and  $1.6 \times 10^{19}$  N·m, respectively. The triangle STF best fits the seismic moment of the reference model, while the Yoffe-0.1 and Yoffe-0.3 overestimate the seismic moment by about 18% and 40%.

In terms of 2D marginal distributions, Figure 11 presents the joint PDFs of rise time and rupture time for the isosceles triangle and Yoffe STF at the same eight points shown in Figure 5. In general, the farther away from the hypocenter the node is located, the stronger the trade-off between parameters. This can be understood from the fact that the problem is less ill-posed near the hypocenter as rupture is constrained to be close to the hypocenter at the beginning of the earthquake. Therefore, the inversion converges more easily toward the true model. Later in the rupture, increasing possibilities to fit the data with various parameter combinations complicates the inversion process and decreases the accuracy of the source parameters.

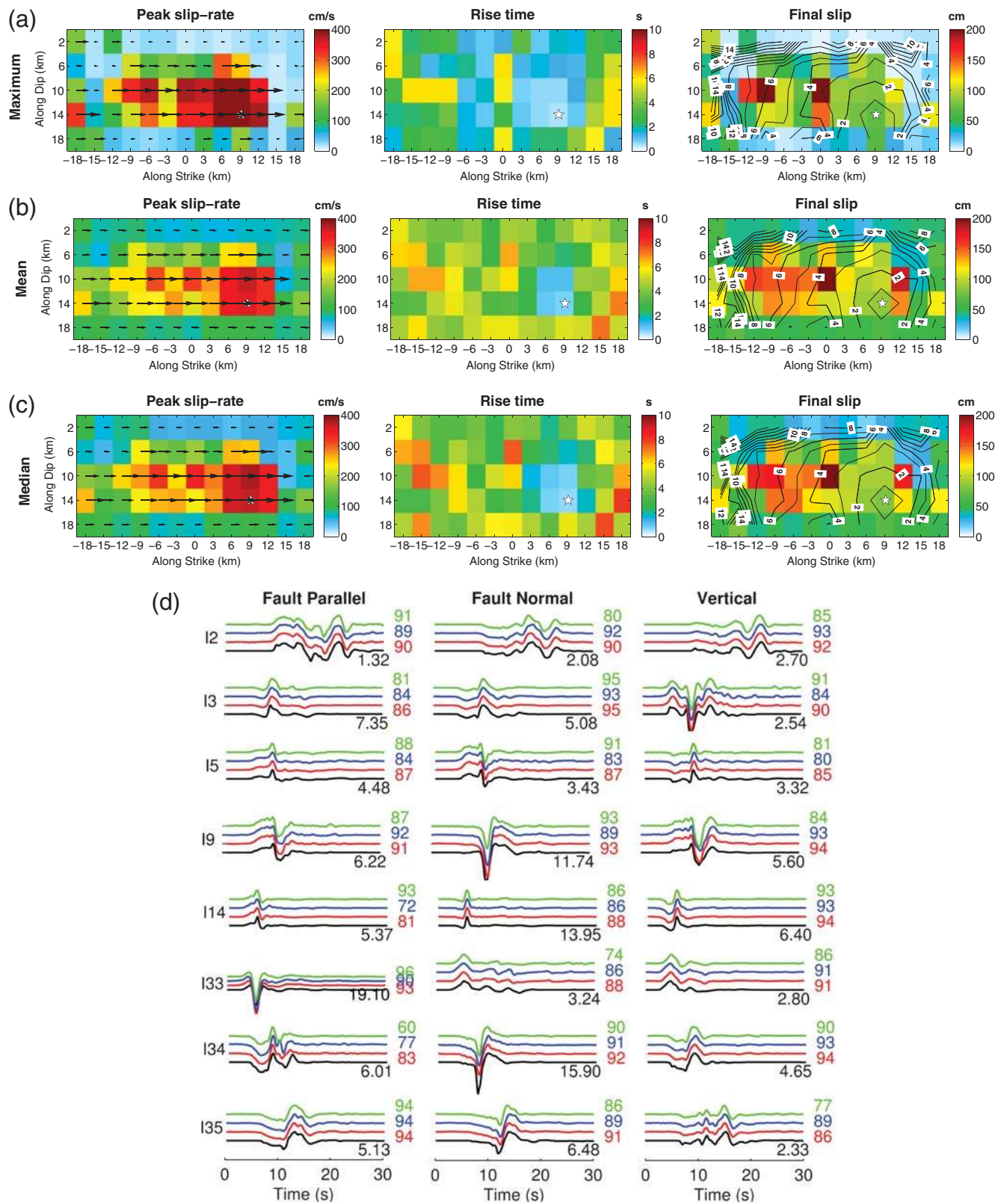
A possible explanation for the behavior of the inversion in this test case can be given using the isochrone concept

(Bernard and Madariaga, 1984; Spudich and Frazer, 1984). The isochrone consists of all points on the fault from which the radiated seismic energy arrives at a specific station at the same time. The radiated energy is significant in areas with large slip, high isochrone velocity, and/or a strong change in isochrone velocity. Therefore, large slip near the hypocenter would radiate more strongly than the same amount of slip away from the hypocenter (Schmedes and Archuleta, 2008). In our test case, the energy is released mainly from the area around the hypocenter, where the slip and the isochrone velocity are highest. Moreover, the isochrones around the hypocenter are in general quite different from station to station. Hence, the source parameters in this area can be well constrained. In contrast, away from the hypocenter, the isochrones become shorter with roughly the same spacing (velocity), and they do not change significantly from station to station (Schmedes and Archuleta, 2008). These lead to a wider range of plausible values for the source parameters and hence to a decrease in the accuracy when inferring these source parameters, in particular when slip is low farther away from the hypocenter, as in this example.

We also observe an interesting feature of the 2D marginals for the triangle STF (Fig. 11a). The rupture time and rise time are negatively correlated; that is, larger values of rupture time are expected for smaller rise times. Therefore, by constraining the rupture time for a triangle STF, the data also constrain the rise time. To the best of our knowledge, this implicit constraint of temporal parameters for the triangle STF has never been described before. This originates from the fact that the ground motions are most sensitive to the time of the peak slip velocity. This sensitivity was also reported by Goto and Sawada (2010) and Oglesby and Mai (2012), who



**Figure 8.** (a) Maximum, (b) mean, and (c) median estimates extracted from the posterior for an isosceles triangle STF. (d) Levels of fit produced by these three estimates, shown at eight stations for the maximum (green), mean (blue), and median (red). Synthetic data are shown in black lines.



**Figure 9.** (a) Maximum, (b) mean, and (c) median estimates extracted from the posterior for Yoffe-0.1. (d) The levels of fit produced by these three estimates, shown at eight stations for the maximum (green), mean (blue), and median (red). Synthetic data are shown in black lines.

Table 3  
Spatial Cross-Correlation Coefficient between the Reference and the Model Estimates

Model	Triangle			Yoffe-0.1*		
	Peak Slip Rate	Slip	Rise Time	Peak Slip Rate	Slip	Rise Time
Maximum	0.72	0.61	0.55	0.88	0.78	0.80
Mean	0.77	0.67	0.76	0.85	0.77	0.85
Median	0.77	0.64	0.73	0.86	0.77	0.84

\*Yoffe source time function with acceleration time of 0.1 s.

analyzed slip rates and waveforms from rupture dynamics and related ground motion. This negative correlation between temporal parameters for the triangle STF still appears when considering the spatial correlation between neighboring points. In contrast, for the Yoffe function this trade-off is not observed because of the antisymmetric shape of this function.

### Crustal Model Variability

In this section, we analyze the impact of 1D crustal structure variability on the rupture model uncertainty. In Figure 12, we show the posterior PDFs of the source parameters at two points of the fault generated using a single Earth model and the two types of crustal model uncertainties (Fig. 2), with uncertainties in wavespeeds only and with variability in both wavespeeds and layer depths. A significant change appears in the resolution of the parameters, including that velocity uncertainty broadens the posterior PDFs with some shift in peak location. This is consistent with the findings of Graves and Wald (2001), which show how the use of

the incorrect Earth model reduces the resolution of the source image. Our results quantify this loss of resolution (blurring effect) in terms of changes in the PDFs of kinematic parameters. Near the hypocenter, the posterior PDFs for the slip and peak slip rate becomes broader but maintains the peak value. The PDFs for the rise time, on the other hand, is nearly uniform. For the rupture time, the distribution appears to be shifted; the rupture tends to propagate faster if the Earth model uncertainties are included. Because the arrival time of each phase of the seismogram is related to both travel time and rupture time, the distortion of the travel time due to the changes in Earth structure ultimately affects the estimated rupture time. In addition, because most of the variations in shear-wave velocity occur in the third layer (where the fault is located) and are mostly smaller than the reference wavespeed, the resulting wave propagation is slower. Therefore, to compensate for this slower shear-wavespeed effect, the rupture velocity needs to be higher leading to earlier rupture arrival. On the other hand, away from the hypocenter, the PDFs of the resulting source parameters are mostly uniform, hence we cannot constrain the source parameters over this area of the fault.

Figure 13 displays the spatial variations of source parameters in terms of the posterior median for the two crustal model uncertainties. The result shows that despite the uncertainties in the Earth model, the inversion still retrieves the high slip and slip-rate features near the hypocenter, with values comparable to the reference model (Fig. 3a). However, the patch is more extended toward the surface for both types of crustal model uncertainty. In addition, slip on the fault boundary, which generally is not well constrained, appears to be larger than in previous results. Comparing the cross-correlation coefficient for the median model with and without Earth structure variability (see Table 4), we find that the correlation becomes lower, by 20% and 10% for the peak slip rate and slip, respectively, as we incorporate the crustal variability.

Comparing the results for the two crustal model variabilities, we do not find significant discrepancies for the posterior PDFs and the extracted source parameters (with a comparable value of correlation, as displayed in the Table 4), mainly because the major differences between the two crustal model variabilities appears at a depth greater than 15 km. Only a small fraction of the rupture is located in this region,

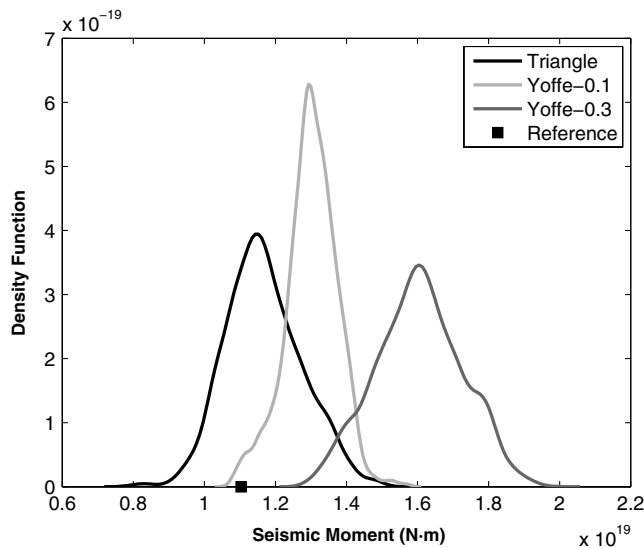
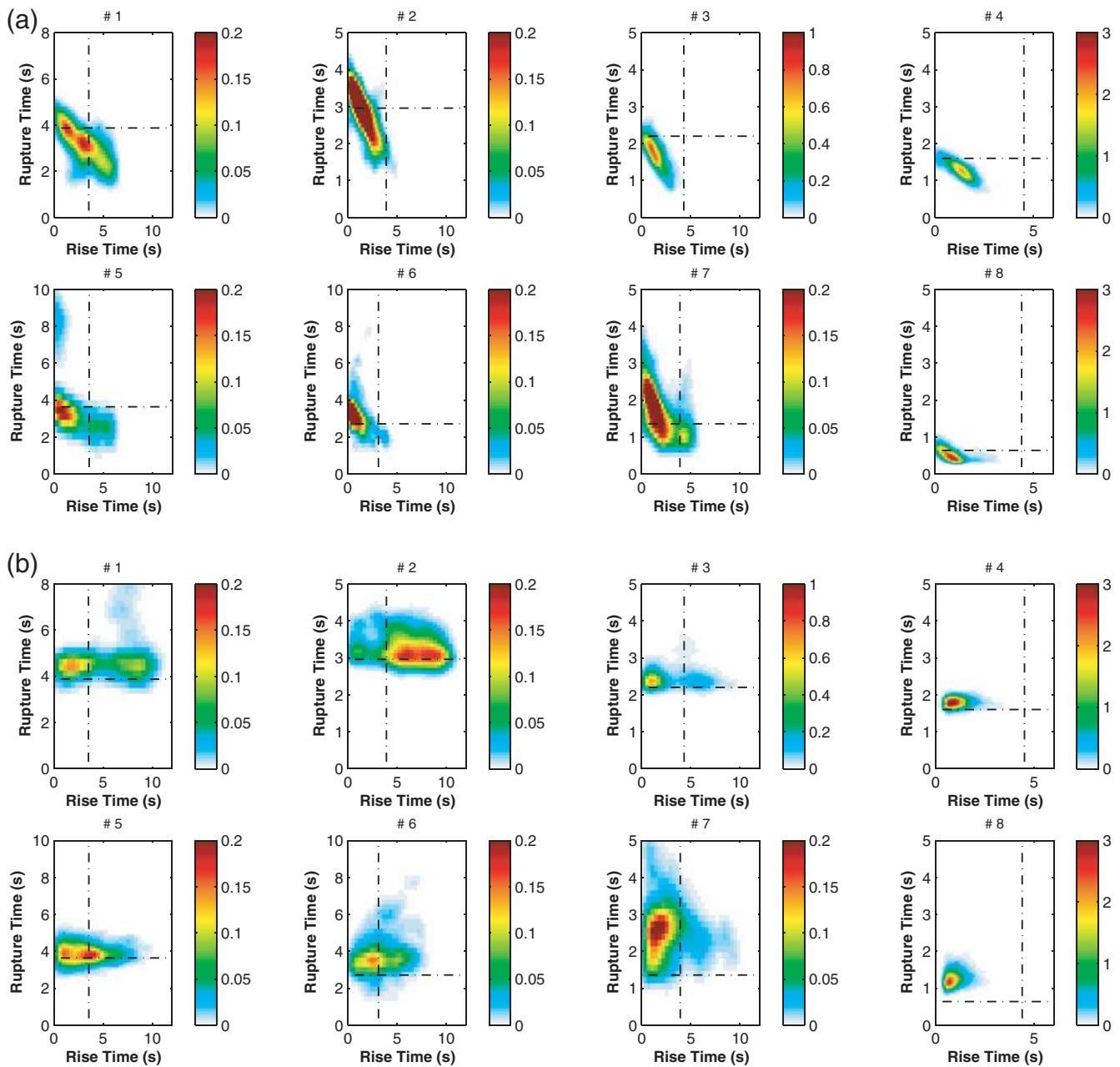


Figure 10. Marginal posterior PDFs for the seismic moment, assuming as STFs the isosceles triangle (black), Yoffe-0.1 (light gray), and Yoffe-0.3 (dark gray). The black square marks the seismic moment of the reference solution.



**Figure 11.** 2D marginal distributions between rise time and rupture time on the eight points shown in Figure 5 for (a) an isosceles triangle STF and (b) the Yoffe-0.1 STF.

and thus it has only a small effect on the inferred source models.

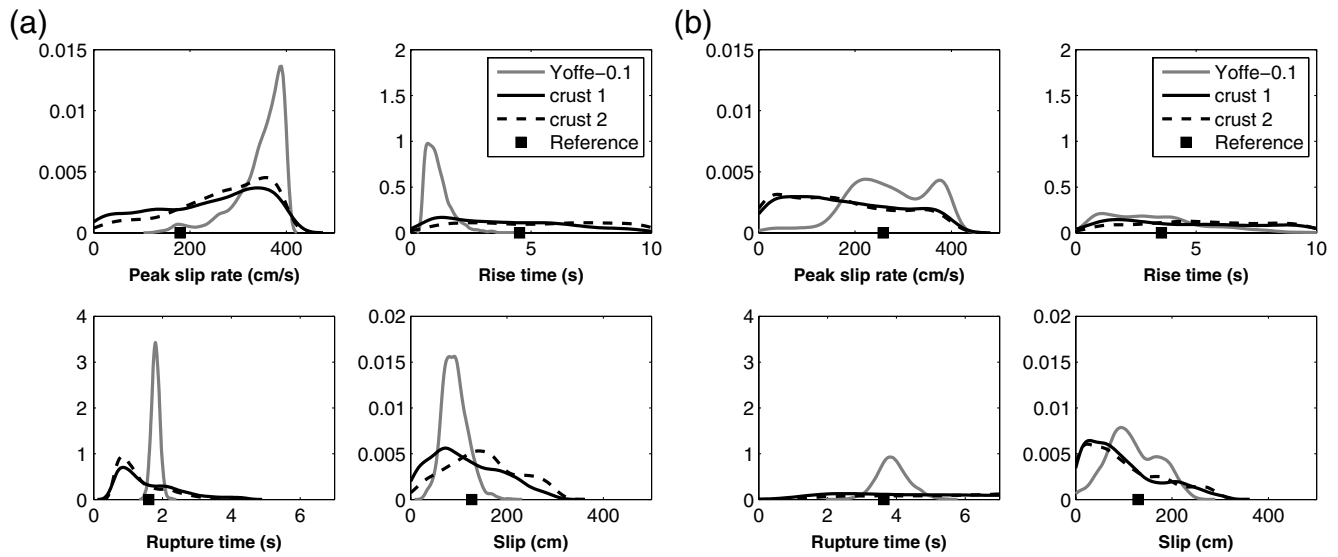
### Discussion

In this study, we apply a Bayesian inference approach to obtain a kinematic rupture model from synthetic near-fault ground motion, generated from a spontaneous dynamic rupture calculation. The principal advantage of this technique compared to using only an optimization technique, is its ability to not only solve the nonlinear inversion, but also to provide a quantitative assessment of the uncertainty. Instead

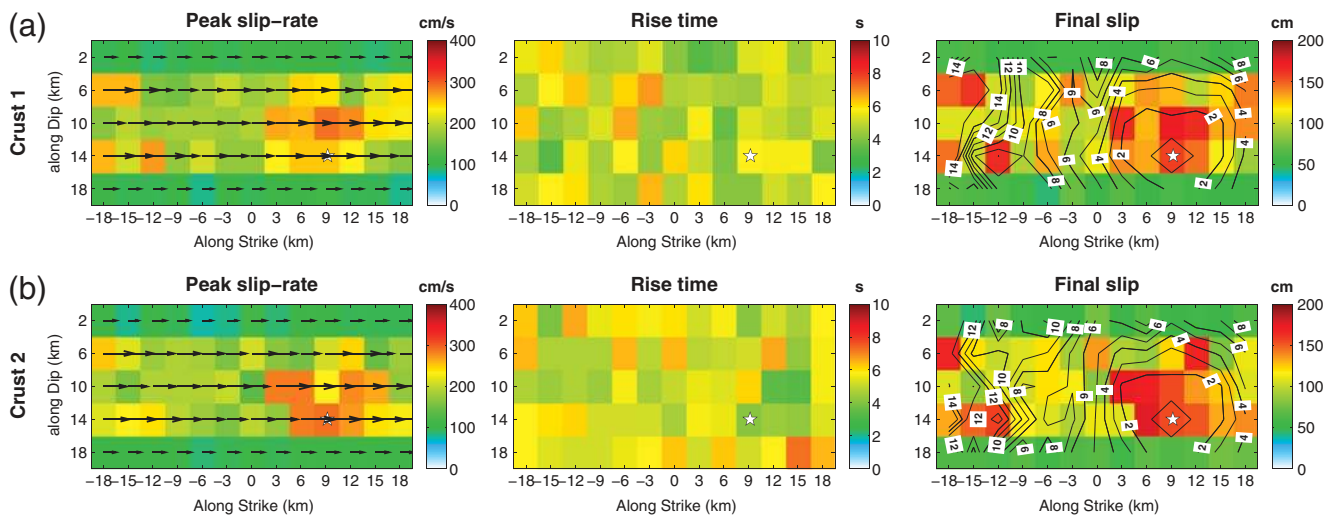
of obtaining one single optimal model, it yields an ensemble of plausible models represented in terms of posterior PDF. We can then assess the resolution of rupture models by analyzing the posterior distributions. Broad posterior distributions mean, for example, that the source parameters are not well constrained. In addition, to learn more about the ensemble of models, we can extract statistical information regarding the posterior distribution at each point on the fault independently.

In Figure 8, for example, we present three models, maximum, mean, and median of the posterior PDFs from isosceles triangle. These estimates are not always well representative





**Figure 12.** Posterior probability density function of slip rate, rise time, rupture onset time, and slip at two points on the fault located, (a) 5 km from the hypocenter and (b) 12 km from the hypocenter. Results are shown without crustal structure variability (gray), with variations in wavespeed (black dashed line), and with simultaneous variations in wavespeed and layer depth (black line).



**Figure 13.** Median extracted from the posterior PDFs for Yoffe-0.1 STF and including variability in the 1D crustal structure. (a) Wave-speed variability only; (b) wavespeed and layer depth variabilities.

of the full solution space as they are just representations of a statistical model. They are calculated by combining the estimated values from the posterior distribution at each individual point of the fault. Hence, the combination of these

selected source parameters does not in general represent a single rupture model that has been tested during the inversion stage and found to fit the data. Instead, these are independently generated rupture models that may not fit the data as well as any of the inverted solutions. We may expect these three estimates to be similar, reflecting the center of the distribution, only if the posterior distributions are Gaussian. However, if the distribution is broad, skewed, or contains several peaks, the maximum posterior is not preferable, as the peak is not aligned with the bulk of the distribution. For distributions that are not Gaussian, the mean and median do not well represent the central tendency of the distribution, which explains the high values on the boundary of the fault. In this part of the fault, the source parameters are not well

**Table 4**  
Spatial 2D Cross-Correlation Coefficient between the Reference and the Median Posterior Models for Yoffe-0.1, Including 1D Crustal Structure Variability

Case	Peak Slip Rate	Slip	Rise Time
Without uncertainty	0.86	0.77	0.84
Wavespeed only	0.68	0.70	0.85
Wavespeed and layer depth	0.69	0.70	0.85

resolved, and posterior distributions are nearly uniform. These characteristics of the extracted statistical models from the posterior PDFs are also reported in [Minson \*et al.\* \(2013\)](#).

We also test the effects on the inversion when using different STFs, an isosceles triangle and two regularized Yoffe functions with different acceleration times. The fits to the data of the synthetics produced by these rupture models give a variance reduction around 94%, lower than reported by [Shao and Ji \(2012\)](#), who reported a variance reduction of about 99% for the initial blind test of the source inversion validation exercise. The reason for this difference is that the true model in our case is based on a dynamic rupture simulation, whose temporal properties are difficult to recover in detail by a kinematic source inversion ([Konca \*et al.\*, 2013](#)); the initial blind test featured a simple kinematic rupture as reference rupture model. We also find that the use of an isosceles triangle causes two main issues. First, it generates an artificial rupture acceleration, which has important implications for earthquake physics, as it would overestimate the rupture speed. Second, the use of an isosceles triangle generates artificial linear correlation between rupture time and rise time, by constraining the rupture time for a triangle STF, the data also constrain the rise time. The former problem could be remedied by a multitime window inversion, whereas it is not clear at present how the rise time/rupture time correlation would manifest itself in this case.

Finally, away from the hypocenter, the distribution for all kinematic source parameters becomes nearly uniform (which means the accuracy of the parameter decreases) as the point is located away from the hypocenter. Investigating the rupture process of the 2000 Tottori earthquake, [Monelli \*et al.\* \(2009\)](#) also find that the source parameters are well resolved around the high slip patch located above the hypocenter and poorly resolved elsewhere. These two observations suggest the source parameter resolution is related to the amount of seismic energy released at different parts of the fault and to the position of these strongly radiating regions with respect to the hypocenter and the observational network. The stronger the seismic radiation is, the better resolved the source parameters should be at this specific area of the fault. The radiated energy from the slip patches close to the hypocenter should be large ([Schmedes and Archuleta, 2008](#)), hence the source parameters in this area will be more accurately resolvable than the same slip patch located far away from the hypocenter. The type of data being used also controls the resolution of the rupture patch, as various data sets are sensitive to different aspects of the rupture process.

One crucial element to achieve a robust rupture model is adopting a realistic Earth model. In this study, the Green's function was first calculated using a single 1D Earth model assuming perfectly known Earth structure. We then include the variability in Earth structure as part of the modeling error in the Bayesian inference, to mimic the realistic case in which we do not know the Earth structure precisely. This allows us to capture the uncertainty in wave propagation that in turn affects the resolution of the kinematic model. Our re-

sult ([Fig. 12](#)) shows that the use of inaccurate Earth structure significantly reduces the resolution of the rupture model parameters, consistent with the study of [Graves and Wald \(2001\)](#). This approach to include Earth structure variability allows the capture of uncertainty of the Green's function ([Yagi and Fukahata, 2011](#)). However, these are still approximate representations of the Earth structure. Therefore, an extensive analysis of the rupture model uncertainty adopting a 3D Earth model that includes the site effects and topographic scattering deserves an extensive exploration. This will particularly help in improving the rupture model resolution, which has important implications for near-fault ground-motion prediction and seismic-hazard assessment.

Another factor that controls the source inversion result is the fault parameterization. In this study we adopt a planar fault of 36 km by 16 km, discretized based on the choice of maximum frequency. This particular choice is reasonable for the synthetic case considered. However, for real earthquakes, different modelers may use different fault geometry, which are likely to affect the result. For some cases, too-simplistic fault geometry might not be well representative of the local tectonics. On the other hand, using very complex fault geometry may induce other complications for the source parameter resolution due to the increase of free parameters and rupture complexity at fault segment boundaries. It is therefore worthwhile to extend this Bayesian approach by including uncertainty in the fault parameterization, as part of the modeling error. This choice of fault parameterization is even more challenging for an earthquake early warning purpose, as we may not have any prior information about the fault geometry. Therefore, the source geometry needs to be solved simultaneously with the rupture models. As an example, [S. Minson \*et al.\* \(unpublished report, 2013\)](#) recently use Bayesian inference to find the optimal fault geometry and the distribution of possible slip models for that geometry, and then apply the approach to the 2011 Tohoku-Oki and the 2003 Tokachi-Oki earthquakes.

## Conclusions

Kinematic source models are obtained considering the full nonlinearity of the inversion problem. We use a Bayesian technique to obtain the posterior probability function of each source parameter on a regular grid representing the fault. We find that a symmetric triangle STF compensates missing dynamic constraints by adding an artificial correlation between rise time and rupture time. Incorporating spatial smoothing reduces the skewness of the posterior distribution, as this constrain reduces the trade-off of the parameters at neighboring nodes. To avoid biasing the result due to the assumption on an Earth model, we propose incorporating Earth model variability in the inference of source parameters. Our results indicate that including the uncertainty in Earth structure broadens the PDFs of the source parameters and may shift the location of the peak rupture time. Rupture time is the

parameter that is most strongly affected both by variations of the STF and crustal structure.

To further explore the effect of Earth structure variability, one should (1) consider the implication of the choice of the inverted data frequency range on source model uncertainty and (2) analyze the use of 3D Earth structure, which includes site effects and topographic scattering. In addition, our tests of different acceleration time  $T_{acc}$  for the regularized Yoffe STF showed that an inappropriate choice on  $T_{acc}$  may create artifacts in the source models (Tinti *et al.*, 2009). The inference and effects of this parameter need further exploration. Additional work on exploring the uncertainty due to assumption of fault plane complexity and fault parameterization is also warranted, as in general they are chosen differently by different research teams. Several studies propose different ways of choosing grid spacing (Bernard *et al.*, 1996; Emolo and Zollo, 2005; Page *et al.*, 2009). A natural extension of our work would consider a transdimensional inversion (Bodin *et al.*, 2012) in which the data itself are used to define the model parameterization and hence dictate the overall uncertainties.

### Data and Resources

All data used in this study are generated in the context of the source inversion validation (SIV) project and are available at <http://equake-rc.info/sivdb/> (last accessed September 2013).

### Acknowledgments

We would like to thank Damiano Monelli for providing his code and for helpful discussions. We also thank two anonymous reviewers and Associate Editor Heather DeShon for their thoughtful comments and suggestions that helped improve this paper. This study was supported by the King Abdullah University of Science and Technology (KAUST).

### References

- Abrahamson, N. A., and J. J. Bommer (2005). Probability and uncertainty in seismic hazard analysis, *Earthq. Spectra* **21**, no. 2, 603–607.
- Arakawa, A., H. Iwasaki, and K. Anada (2009). Investigation of Gibbs sampling conditions to estimate variance components from Japanese Black carcass field data, *Animal Sci. J.* **80**, no. 5, 491–497.
- Archuleta, R. (1984). A fault model for the 1979 Imperial Valley earthquake, *J. Geophys. Res.* **89**, 4559–4585.
- Bernard, P., and R. Madariaga (1984). A new asymptotic method for the modeling of near-field accelerograms, *Bull. Seismol. Soc. Am.* **74**, 539–558.
- Bernard, P., A. Herrero, and C. Berge (1996). Modeling directivity of heterogeneous earthquake ruptures, *Bull. Seismol. Soc. Am.* **86**, 1149–1160.
- Beyer, H. G. (2001). *The Theory of Evolution Strategies*, Springer, Berlin, Heidelberg, 143–256.
- Bodin, T., M. Sambridge, N. Rawlinson, and P. Arroucau (2012). Transdimensional tomography with unknown data noise, *Geophys. J. Int.* **189**, 1536–1556.
- Bouchon, M., N. Toksöz, H. Karabulut, M. Bouin, M. Dietrich, M. Aktar, and M. E. Tübitak (2000). Seismic imaging of the 1999 Izmit (Turkey) rupture inferred from the near-fault recordings, *Geophys. Res. Lett.* **27**, no. 18, 3013–3016.
- Clévéde, E., M.-P. Bouin, B. Bukchin, A. Mostinskiy, and G. Patau (2004). New constraints on the rupture process of the 1999 August 17 Izmit earthquake deduced from estimates of stress glut rate moments earthquake deduced from estimates of stress glut rate moments, *Geophys. J. Int.* **159**, 931–942.
- Cohee, B., and G. Beroza (1994). A comparison of two methods for earthquake source inversion using strong motion seismograms, *Ann. Geophys.* **37**, 1515–1538.
- Custodio, S., P. Liu, and R. Archuleta (2005). The 2004  $M_w$  6.0 Parkfield, California, earthquake: Inversion of near-source ground motion using multiple data sets, *Geophys. Res. Lett.* **32**, L23312, doi: [10.1029/2005GL024417](https://doi.org/10.1029/2005GL024417).
- Delouis, B., D. Giardini, P. Lundgren, and J. Salichon (2002). Joint inversion of InSAR, GPS, teleseismic, and strong-motion data for the spatial and temporal distribution of earthquake slip, *Bull. Seismol. Soc. Am.* **92**, 278–299.
- Dunham, E., and R. Archuleta (2004). Evidence for a supershear transient during the 2002 Denali fault earthquake, *Bull. Seismol. Soc. Am.* **94**, 256–268.
- Duputel, Z., L. Rivera, Y. Fukahata, and H. Kanamori (2012). Uncertainty estimations for seismic source inversions, *Geophys. J. Int.* **190**, no. 2, 1243–1256.
- Emolo, A., and A. Zollo (2005). Kinematic source parameters for the 1989 Loma Prieta earthquake from the nonlinear inversion of accelerograms, *Bull. Seismol. Soc. Am.* **95**, 981–994.
- Gelman, A., J. Carlin, H. Stern, and D. Rubin (1995). *Bayesian Data Analysis*, Chapman Hall/CRC Texts in Statistical Science, London, 143–256.
- Goto, H., and S. Sawada (2010). Trade-offs among dynamic parameters inferred from results of dynamic source inversion, *Bull. Seismol. Soc. Am.* **100**, 910–922.
- Gouveia, W. P., and J. A. Scales (1998). Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis, *J. Geophys. Res.* **103**, no. B2, 2759–2779.
- Graves, R., and D. Wald (2001). Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions 1. Strong motions, *J. Geophys. Res.* **106**, no. B5, 8745–8766.
- Hartzell, S., 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. Liu, and C. Mendoza (1996). The 1994 Northridge, California, earthquake: Investigation of rupture velocity, risetime, and high-frequency radiation, *J. Geophys. Res.* **101**, no. B9, 20,091–20,108.
- Hartzell, S., P. Liu, C. Mendoza, C. Ji, and K. Larson (2007). Stability and uncertainty of finite-fault slip inversions: Application to the 2004 Parkfield, California, earthquake, *Bull. Seismol. Soc. Am.* **97**, 1911–1934.
- Hartzell, S., G. Stewart, and C. Mendoza (1991). Comparison of L1 and L2 norms in a teleseismic waveform inversion for the slip history of the Loma Prieta, California, earthquake, *Bull. Seismol. Soc. Am.* **81**, 1518–1539.
- Hoff, P. (2009). *A First Course in Bayesian Statistical Methods*, Springer, Dordrecht, Heidelberg, London, New York, 89–104.
- Jónsson, S., H. Zebker, P. Segall, and F. Amelung (2002). Fault slip distribution of the 1999  $M_w$  7.1 Hector Mine, California, earthquake, estimated from satellite radar and GPS measurements, *Bull. Seismol. Soc. Am.* **92**, 1377–1389.
- Konca, A., Y. Kaneko, N. Lapusta, and J. Avouac (2013). Kinematic inversion of physically plausible earthquake source models obtained from dynamic rupture simulations, *Bull. Seismol. Soc. Am.* **103**, no. 5, 2621–2644.
- Lee, S.-J., B.-S. Huang, M. Ando, H.-C. Chiu, and J.-H. Wang (2011). Evidence of large scale repeating slip during the 2011 Tohoku-Oki earthquake, *Geophys. Res. Lett.* **38**, L19306, doi: [10.1029/2011GL049580](https://doi.org/10.1029/2011GL049580).
- Liu, P., 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**, no. B02318, doi: [10.1029/2003JB002625](https://doi.org/10.1029/2003JB002625).

- Mai, P. M., J. Burjanek, B. Delouis, G. Festa, C. Holden, D. Monelli, and T. Uchide (2007). Earthquake source inversion blindtest: Initial results and further developments, *Eos Trans. AGU* **88**, 855–877.
- Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953). Equation of state calculations by fast computing machines, *J. Chem. Phys.* **21**, 1087–1092.
- Minson, S., M. Simons, and J. Beck (2013). Bayesian inversion for finite fault earthquake source models I—Theory and algorithm, *Geophys. J. Int.* **194**, 1701–1726.
- Monelli, D., and P. M. Mai (2008). Bayesian inference of kinematic earthquake rupture parameters through fitting of strong motion data, *Geophys. J. Int.* **173**, no. 1, 220–232.
- Monelli, D., P. M. Mai, S. Jónsson, and D. Giardini (2009). Bayesian imaging of the 2000 Western Tottori (Japan) earthquake through fitting of strong motion and GPS data, *Geophys. J. Int.* **176**, 135–150.
- Mooney, W. (1989). Seismic methods for determining earthquake source parameters and lithospheric structure, in *Geophysical Framework of the Continental United States*, L. C. Pakiser and W. Mooney, Geol. Soc. Am. Memoir, Boulder, Colorado, Vol. 172, 11–34.
- Mosegaard, K., and A. Tarantola (1995). Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.* **100**, 12,431–12,447.
- Oglesby, D. D., and P. Mai (2012). Fault geometry, rupture dynamics and ground motion from potential earthquakes on the north Anatolian fault under the Sea of Marmara, *Geophys. J. Int.* **188**, 1071–1087.
- Olson, A., and R. Apsel (1982). Finite faults and inverse theory with applications to the 1979 Imperial Valley earthquake, *Bull. Seismol. Soc. Am.* **72**, 1969–2001.
- Page, M., S. Custodio, R. Archuleta, and J. Carlson (2009). Constraining earthquake source inversions with GPS data: 1. Resolution-based removal of artifacts, *J. Geophys. Res.* **114**, no. B01314, doi: [10.1029/2007JB005449](https://doi.org/10.1029/2007JB005449).
- Page, M., P. Mai, and D. Schorlemmer (2011). Testing earthquake source inversion methodologies, *Eos Trans. AGU* **92**, no. 9, 75.
- 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**, no. B07314, doi: [10.1029/2006JB004821](https://doi.org/10.1029/2006JB004821).
- Schmedes, J., and R. Archuleta (2008). Near-source ground motion along strike-slip faults: Insights into magnitude saturation of PGV and PGA, *Bull. Seismol. Soc. Am.* **98**, no. 5, 2278–2290.
- Sekiguchi, H., and T. Itawa (2002). Rupture process of the 1999 Kocaeli, Turkey, earthquake estimated from strong-motion waveforms, *Bull. Seismol. Soc. Am.* **92**, 300–311.
- 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**, no. 1, 117–133.
- Shao, G., and C. Ji (2012). What the exercise of the SPICE source inversion validation BlindTest1 did not tell you, *Geophys. J. Int.* **189**, 569–590.
- Spudich, P., and R. Archuleta (1987). Techniques for earthquake ground motion calculation with applications to source parametrization of finite faults, in *Seismic Strong Motion Synthetics*, Vol. 37, Academic Press, Orlando, Florida, 205–265.
- Spudich, P., and L. Frazer (1984). Use of ray theory to calculate high-frequency radiation from earthquake sources having spatially variable rupture velocity and stress drop, *Bull. Seismol. Soc. Am.* **74**, 2061–2082.
- Spudich, P., and L. Xu (2002). *Documentation of Software Package Comp-syn svx3.11: Programs for Earthquake Ground Motion Calculation Using Complete 1-D Green's Functions*, Academic Press, San Diego, California.
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.
- Tinti, E., M. Cocco, E. Fukuyama, and A. Piatanesi (2009). Dependence of slip weakening distance ( $D_c$ ) on final slip during dynamic rupture of earthquakes, *Geophys. J. Int.* **177**, 1205–1220.
- Tinti, E., E. Fukuyama, A. Piatanesi, and M. Cocco (2005). A kinematic source time function compatible with earthquake dynamics, *Bull. Seismol. Soc. Am.* **95**, 1211–1223.
- Tinti, E., P. Spudich, and M. Cocco (2005). Earthquake fracture energy inferred from kinematic rupture models on extended faults, *J. Geophys. Res.* **101**, 1–25.
- Vallée, M., and M. Bouchon (2004). Imaging coseismic rupture in far field by slip patches, *Geophys. J. Int.* **156**, 615–630.
- Wald, D., and R. Graves (2001). Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions 2. Combining seismic and geodetic data, *J. Geophys. Res.* **106**, no. B5, 8767–8788.
- Wald, D. J., and T. H. Heaton (1994). Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake, *Bull. Seismol. Soc. Am.* **84**, 668–691.
- Yagi, Y., and Y. Fukahata (2008). Importance of covariance components in inversion analyses of densely sampled observed data: an application to waveform data inversion for seismic source processes, *Geophys. J. Int.* **175**, 215–221.
- Yagi, Y., and Y. Fukahata (2011). Introduction of uncertainty of Green's function into waveform inversion for seismic source processes, *Geophys. J. Int.* **186**, 711–720.
- Zhou, S., K. Irikura, and X. Chen (2004). Analysis of the reliability and resolution of the earthquake source history inferred from waveforms, taking the Chi-Chi earthquake as an example, *Geophys. J. Int.* **157**, 1217–1232.

## Appendix

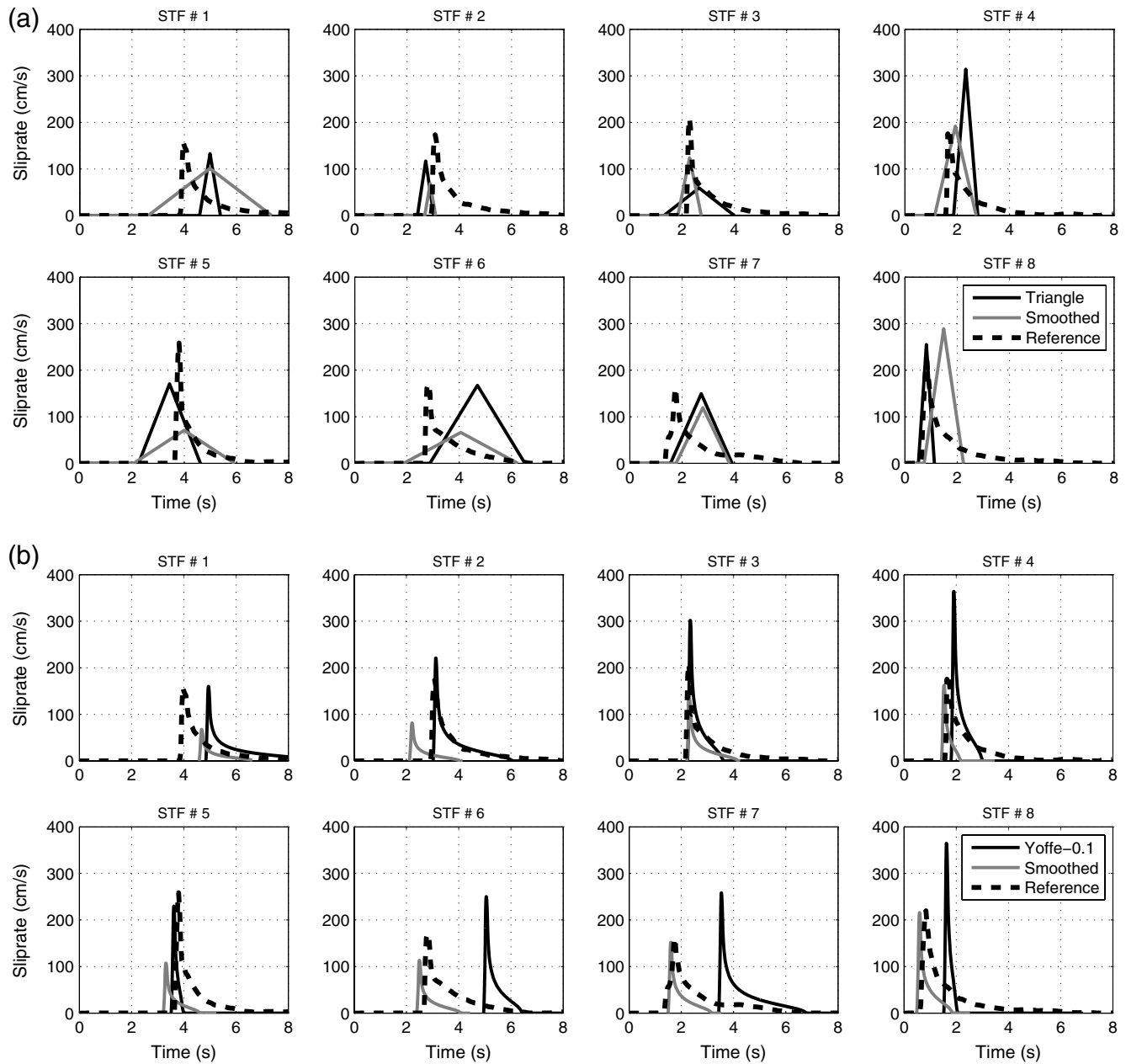
### Effect of Spatial Smoothing in Uncertainty Analysis

To minimize the trade-off between rupture parameters at neighboring points, we include the spatial smoothing using an eight-neighbors Laplacian filter  $D$ ,

$$D = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

respecting the distance between patches and the boundary.

Figure A1 shows that including spatial smoothing between neighboring points, with a smoothing parameter  $\alpha^2$  of 20, generally reduces the time difference between the peaks of the slip-rate pulse for the reference model and the inversion solutions, particularly for Yoffe-0.1 close to the hypocenter. However, for triangle STF, the improvement is only apparent for a few points on the fault. To analyze this effect on the parameter uncertainties in more detail, we constrain the inversion following the penalized likelihood (equation 10) in which the smoothing factor is chosen deterministically. In (E) Figure S4, available in the electronic supplement, we present the kinematic source models generated from an evolutionary algorithm (Beyer, 2001) applying different smoothing factors  $\alpha$  (equation 10). For  $\alpha^2 = 0.1$  and  $\alpha^2 = 20$ , we identify high slip rate and slip around the hypocenter, which is not observed for  $\alpha^2 = 200$  due to oversmoothing of the parameters. The variability of misfits to the seismic data with respect to the model roughness allows us to deterministically choose the preferred model. We choose a smoothing factor



**Figure A1.** STFs on the same eight points as in Figure 5 for the reference model (dashed line) and different inversion solutions: (a) using triangle slip-rate function and including spatial smoothing and (b) using Yoffe-0.1 STF and including spatial smoothing.

$\alpha^2 = 20$ , based on the curvature of the trade-off function. Therefore, we use this factor in Figure A1. Sekiguchi *et al.* (2000) also discuss a way to optimally choose this smoothing value. In  $\text{\textcircled{E}}$  Figure S5, available in the electronic supplement, we also show the resulting posterior PDFs at two selected points on the fault (point 4 and 5 in Fig. 5) for different smoothing factors. The area around the hypocenter shows better constrained slip and slip rate by incorporating the spatial smoothing. The PDFs become less skewed as this constraint reduces the trade-off between neighboring points that partially leads to this skewness.

King Abdullah University of Science and Technology (KAUST)  
 Division of Physical Sciences and Engineering  
 Thuwal 23955-6900, Saudi Arabia  
 hoby.razafindrakoto@kaust.edu.sa  
 martin.mai@kaust.edu.sa