Geophysical Journal International

Geophys. J. Int. (2015) **200,** 1–16
GJI Seismology

doi: 10.1093/gji/ggu252

Evaluating the effect of network density and geometric distribution on kinematic source inversion models

Youbing Zhang,[1] Luis A. Dalguer,[1,*] Seok Goo Song,[1,†] John Clinton[1] and Domenico Giardini[2]

[1]*Swiss Seismological Service (SED), ETH Zurich, Zurich, Switzerland. E-mail: youbing.zhang@sed.ethz.ch*
[2]*Institute of Geophysics, ETH Zurich, Zurich, Switzerland*

Accepted 2014 June 30. Received 2014 June 25; in original form 2013 November 22

SUMMARY

The effect of network density and geometric distribution on kinematic non-linear source inversion is investigated by inverting synthetic ground motions from a buried strike-slip fault (M_w 6.5), that have been generated by dynamic spontaneous rupture modelling. For the inversion, we use a physics-based regularized Yoffe function as slip velocity function. We test three different cases of station network geometry: (i) single station, varying azimuth and epicentral distance; (ii) multistation circular configurations, that is stations at similar distances from the fault, and regularly spaced around the fault; (iii) irregular multistation configurations using different numbers of stations. Our results show: (1) single station tests suggest that it may be possible to obtain a relatively good source model even using a single station. The best source model using a single station is obtained with stations at which amplitude ratios between three components are not large. We infer that both azimuthal angle and source-to-station distance play an important role in the design of optimal seismic network for source inversion. (2) Multistation tests show that the quality of the inverted source systematically correlates neither with the number of stations, nor with waveform misfit. (3) Waveform misfit has a direct correlation with the number of stations, resulting in overfitting the observed data without any systematic improvement of the source. It suggests that the best source model is not necessarily derived from the model with minimum waveform misfit. (4) A seismic network with a small number of well-spaced stations around the fault may be sufficient to obtain acceptable source inversion.

Key words: Inverse theory; Earthquake dynamics; Computational seismology.

1 INTRODUCTION

Detailed imaging of the spatio- and temporal distributions of the source parameters from natural earthquakes is one of the principal goals for seismology and geophysics. Kinematic inversion models are obtained routinely by inverting seismic and geodetic data for many moderate-to-large earthquakes nowadays, which improve our understanding of the physical source processes governing dynamic rupture propagation and seismic-wave generation (e.g. Hartzell & Heaton 1983; Fukuyama & Irikura 1986; Wald & Heaton 1994; Cotton & Campillo 1995; Bouchon *et al.* 2002; Sekiguchi & Iwata 2002; Custódio *et al.* 2009, among many others). Kinematic source inversion is an efficient tool to retrieve heterogeneous distributions of source parameters on the finite fault, such as slip, peak slip velocity (maximum slip rate), rake angle (slip direction), rupture time (time that slip starts) and risetime (slip duration), etc.

However, an earthquake source inversion is often highly ill-posed and produces many non-unique solutions even though they fit observed data equally well. Although many methods have been developed to solve ill-posed inverse problems (e.g. Monelli & Mai 2008, and others), the inherent uncertainties in resulting earthquake source models are poorly understood. Many factors affect the estimated source image, such as incompletely known Earth structure, simplifications in assumed fault geometry, data processing steps, type of data sets used, Green's function calculation, and parameterization of the inversion Beresnev (2003). One 'Blind Test' on source inversion within the framework of the European commission project SPICE in 2006 shows that the results are often not useful, '4 out of 9 inversion results are, statistically speaking, not better than a random model with somehow correlated slip!' (Mai *et al.* 2007). Hence, there is an inevitable question of whether we should trust any source inversion results for real earthquakes? (Shao

*Now at: Swissnuclear, Aarauerstrasse 55, 4601 Olten, Switzerland.
†Now at: Earthquake Research Center, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124 Gwahang-no, Yuseong-gu, Daejeon 305-350, South Korea.

& Ji 2012). Some researchers have introduced approaches aimed at quantifying the errors on the source kinematic images retrieved from real data inversion (e.g. Piatanesi *et al.* 2007; Lucca *et al.* 2012; Toraldo Serra *et al.* 2013a,b). One concern is that researchers often cannot directly evaluate their estimated models and quantitatively measure the uncertainty in the inversion of real earthquakes.

In order to understand the uncertainty and reliability of source inversions, based on observed data from natural earthquakes, researchers start to pay attention to synthetic source inversion tests based on the prior information about the source. Spontaneous dynamic rupture models that are based on the physics of the causative rupture and wave propagation, incorporating conservation laws of continuum mechanics, frictional sliding, and the state of stress in the crust, simulate physically consistent earthquakes rupture solutions (e.g. Andrews 1976; Das & Aki 1977; Day 1982; Olsen *et al.* 1997; Dalguer *et al.* 2001; Peyrat *et al.* 2001; Day *et al.* 2008; Gabriel *et al.* 2012) and have the potential to expand our understanding of both source- and propagation-dominated ground motion phenomena (e.g. Dalguer *et al.* 2008; Dunham & Bhat 2008; Ripperger *et al.* 2008; Pitarka *et al.* 2009; Duan & Day 2010; Song & Dalguer 2013). Thus, spontaneous dynamic rupture simulations enable us to test our knowledge about earthquake rupture and wave propagation (proved and/or still in assumption) against observations (Harris *et al.* 2009). Therefore, testing kinematic source inversion methods by inverting the synthetic ground motions generated by dynamic rupture simulations is a rigorous way of evaluating the suitability of different source inversion methods (e.g. Konca *et al.* 2013) for exploring the physics of the real earthquake source.

Beresnev (2003) points out that the goodness of kinematic slip inversion solution depends on the network geometry. Generally, seismologists expect to obtain better source images by improving observational data (e.g. more stations and high quality data). However, recent studies show that improving the surface station density alone does not significantly affect the source inversion results. As the number and density of available strong-motion and high-rate GPS stations increase, the evaluation of the reliability of the inverted source model with respect to the number and array geometry of the strong-motion stations has gained importance. Particularly important is the ability of a set of near-field records to evaluate reliable location and characteristics of asperities. Custódio *et al.* (2005) examined the dependence of the kinematic inversion solutions on different combinations of data sets by using 12 equivalent data subsets for the 2004 Parkfield, California, earthquake, and suggest that similar inversion results can be archived using less than all 43 stations. Asano & Iwata (2009) point out that 12 near-field stations are sufficient to obtain a stable source image of the 2004 Chuetsu earthquake. Furthermore, Sarao *et al.* (1998) point out that the azimuthal distribution and the station location to the direction of rupture propagation are more important than simply the number of stations. Miyatake *et al.* (1986) and Iida *et al.* (1990) discuss optimal strong-motion array configuration for source inversion. They also point out that the circular distributed station network with well-spaced azimuthal coverage and certain epicentral distance to fault length play an important role. Overall, it is not clear yet how the azimuthal and epicentral distance distribution of station networks affect the source inversion, which we think still deserves further investigation.

In this paper, we investigate how the density of stations and the network geometry influence the quality of the inverted source and its corresponding predicted ground motion. For that purpose we use a kinematic non-linear source inversion method, in which we consider a physics-based regularized Yoffe function as slip velocity function. As a case study we use a synthetic vertical strike-slip earthquake simulated by spontaneous dynamic rupture modelling (Dalguer & Mai 2011). We investigate three cases of network geometry: (i) single stations with varying azimuth and epicentral distance; (ii) three cases of multistation circular configurations with similar distances from the fault and regularly spaced around the fault and (iii) five cases of irregular spaced station configurations on the free surface (5, 10, 20, 40 and 168 stations) with good azimuthal and epicentral distance coverage around the fault. We find that the quality of model estimates is not significantly improved by simply increasing the number of stations and that a good azimuthal and distance distribution of the stations is more important than the number of stations. We also find that a single station can be sufficient to estimate reliable source depending on the location of the station used. Our synthetic inversion tests are performed in a relatively ideal situation, that is no errors in the velocity structure and fault geometry, and no noise on the synthetic data. However our findings may provide some insights into the potential effect of network distribution on source model estimation, and serve as a guideline for network design in the field.

2 SOURCE INVERSION METHOD

2.1 Kinematic source inversion approach

The fault rupture process can be kinematically described as a shear dislocation propagating along a discontinuous surface within an elastic medium. Each point on the fault is usually parameterized in terms of slip, peak slip velocity (maximum slip rate), rake angle (slip direction), rupture time (time that slip starts) and risetime (slip duration), etc. The kinematic description of the earthquake processes is represented by the so-called slip velocity function (SVF, henceforth) or source time function, which is a proxy for the true slip velocity evolution during rupture propagation on the assumed fault. Several simplified versions of SVFs have been used in source inversion such as boxcar and triangle functions. Recently more dynamically compatible SVFs (Tinti *et al.* 2005; Bizzarri 2012) have been introduced.

Two main methods are used for kinematic source inversion: the linear multitime window method and the non-linear single time window method. For the former, the SVF is not prescribed a priori, but is expanded into a number of basis functions such as isosceles triangles (e.g. Olson & Apsel 1982; Sekiguchi *et al.* 2000). In general, 3–5 windows are assigned for each subfault patch. The width of each window and interval between adjacent windows are predetermined, and the onset time of the first window is also predetermined with an assumed rupture velocity. Thus only the amplitude for each window needs to be determined in the inversion. By fixing the risetime and the rupture time on each grid node on the fault, the inversion problem becomes linear. So the solution can be obtained using the linear least-square method. Since most source inversion problems are underdetermined and ill conditioned, regularization schemes, that is smoothing, minimum norm, non-negative constraint, are often applied to get a stable solution.

In the non-linear single time window approach the slip velocity function is predefined, like a triangular function (Hartzell & Heaton 1983), a boxcar (Emolo & Zollo 2005) and some more complex forms used, for instance, trigonometric (Hartzell *et al.* 1996) or power (Liu & Archuleta 2004) functions. In general both rupture time and risetime are model parameters and estimated in the inversion, which make the inversion problem non-linear. One of the

advantages in the non-linear inversion is that we can implement a physically realistic slip velocity function to describe the slip evolution on grid node. The main drawbacks of the non-linear approach are that inversion results may depend on starting models, it requires large computational costs, and estimated solutions may be trapped in a local minimum.

Due to the non-uniqueness of kinematic source inversion, recent work uses Bayesian inference to recover the searching parameters and their uncertainties (Monelli & Mai 2008; Minson *et al.* 2013). In the Bayesian framework, the prior of source parameters and accuracy of the posterior are encapsulated as probability distributions, and the physics-based prior information could be considered as regularization in inversion problem (Tarantola & Valette 1982; Tarantola 2005). Thus the Bayesian approach appraises the ensemble of models that are consistent with the data rather than finding one single best solution.

2.2 Slip velocity function

The SVFs obtained from the solution of dynamic rupture models with heterogeneous source properties have usually complex spatio-temporal forms. Fig. 1 shows some examples of the diversity of complexities of SVFs on a fault for a pure strike-slip, buried and normal stress depth-dependent dynamic rupture model (Dalguer & Mai 2011), which is used in this study. It is obvious that the slip evolution on the fault is complex and variable, and not consistent with the traditionally used boxcar or triangular function. Therefore, a more sophisticated single SVF function is necessary to closely represent the dominant forms, such as those from Figs 1(a) and (b), but forms as in Fig. 1(c) need multi-SVF functions. Fig. 1(d) is a proxy of a single SVF used in our inversion, as explained in the next paragraph.

We build our kinematic models using the regularized Yoffe function proposed by Tinti *et al.* (2005) as the SVF. This function is derived by convolving the Yoffe function (Yoffe 1951, with duration τ_R), which is self-similar, self-healing and compatible with evidence for pulse-like rupture propagation, with a triangular function (half-duration is τ_s). The resultant function is not singular both at the rupture onset and at the healing time. The regularized Yoffe function is characterized by three independent kinematic parameters: $T_{\rm acc}$, $\tau_R^{\rm eff}$ and $D_{\rm max}$ (as Fig. 1d). $T_{\rm acc}$ is the duration of the positive slip acceleration time, $\tau_R^{\rm eff}$ is the local duration of slip velocity (risetime) and $D_{\rm max}$ is the final slip. The peak slip velocity ($V_{\rm peak}$) is related to the above three parameters through the following asymptotic relation.

$$V_{\rm peak} = 1.04 \cdot \frac{D_{\rm max}}{(T_{\rm acc})^{0.54}(\tau_R)^{0.47}}. \quad (1)$$

Detailed analytical representation of this function is shown in Tinti *et al.* (2005), and eqs (A.7) and (A.18) are corrected by Bizzarri (2012). According to the definition, the risetime $\tau_R^{\rm eff}$ should be about 2.5 times larger than the $T_{\rm acc}$. The regularized Yoffe function is consistent with the self-similar solution of the elasto-dynamic equation, and with spontaneous dynamic models governed by the slip weakening traction evolution of spontaneous-crack models (Tinti *et al.* 2005; Bizzarri 2012). In addition, the variable risetime of this function is consistent with the character of a local healing process on a fault from laboratory experiments (Ohnaka & Yamashita 1989). This function can be easily implemented in non-linear kinematic source inversion.

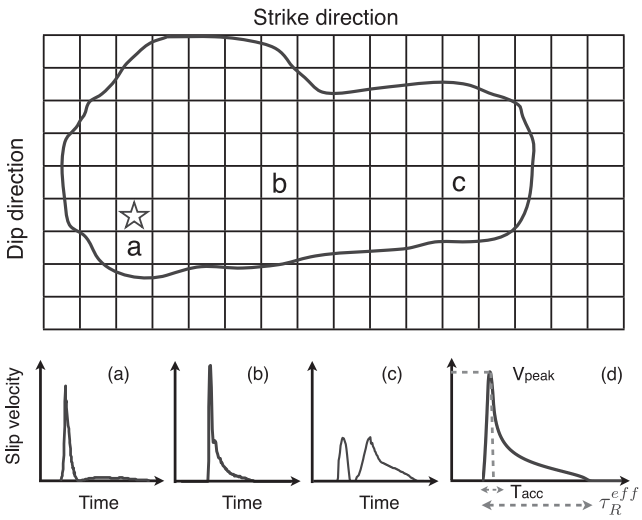

Figure 1. Schematic diagram of the slip velocity function on finite dynamic rupture fault (Dalguer & Mai 2011), contour shows the entire rupture area; Star indicates the hypocentre location. Panels (a), (b) and (c) indicate the slip evolution of patch a, b, and c on fault, respectively, amplitude and width of the SVFs are not identical; panel (d) shows the regularized Yoffe function.

2.3 Source representation

The source representation theory used in this study has been discussed in previous works (Monelli & Mai 2008). For the convenience of further discussions, we repeat the basic formulation here. We compute ground velocity waveforms using the representation theorem in frequency-domain (Spudich & Archuleta 1987).

$$\dot{u}_m(y,\omega) = \iint_\Sigma \dot{s}(x,\omega) \cdot T^m(x,\omega;y,0)\mathrm{d}\Sigma, \quad (2)$$

where $\dot{u}_m$ is the mth component of ground velocity at observer location y and angular frequency ω ($\omega = 2 \cdot \pi f$), $\dot{s}$ describes the slip evolution at point x on fault, traction T^m at point x is caused by a point impulsive force in the mth direction at the observer location y and Σ is fault surface.

2.4 Inversion framework

We use the non-linear kinematic inversion code (KISS, Monelli & Mai 2008) in which its original version uses boxcar and triangular functions as SVF. For the purpose of this paper we implemented the regularized Yoffe function in the code. The Discrete Wavenumber/Finite Element method (Compsyn package, Spudich & Xu 2002) was used to calculate the Green's function and forward generate the velocity waveforms in a 1-D flat layered earth model without attenuation. Rupture-parameter values at integration points are derived through bi-linear interpolation of values at surrounding grid nodes. And at least four samples within one wavelength are needed to generate a precise approximation of rupture propagation (e.g. Spudich & Archuleta 1987), which is similar to the approach taken by Liu & Archuleta (2004). We calculate the following objective function:

$$O(m) = \|g(m) - \mathrm{d}\|_2^2 + \alpha^2 \|Lm\|_2^2. \quad (3)$$

The $\|g(m) - \mathrm{d}\|_2^2$ is the misfit measure between the synthetic and observed waveforms using $L2$ norm; $\|Lm\|_2^2$ is the model regularization term; and α^2 controls the relative weighting between the two terms. In this study, $\|Lm\|_2^2$ is to smooth the final slip on fault to stabilize the inversion results. To cope with the ill-posed nature of

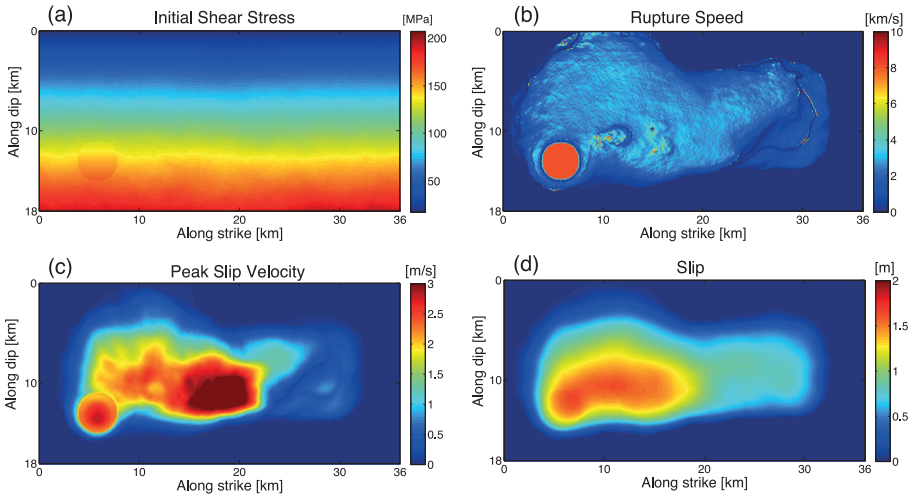

Figure 2. Source parameters of dynamic rupture strike-faulting model. (a) Initial heterogeneous shear stress distribution; (b) rupture speed; (c) peak slip velocity and (d) final slip on fault.

the problem, we restrict our solutions to smooth slip models rather than randomly distributed final slip models even though the latter could fit the waveform better. We include the spatial slip distribution smoothing using a four-neighbours Laplacian filter matrix.

$$L = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

It is an important, but challenging task to choose the value of α^2 for the final inversion result. We test the α^2 value by trial-and-error, selecting the 'kink' in the well known L-curve (see Fig. S1, and take the value $\alpha^2 = 5 \times 10^{-4}$ as reference). The computational time for the non-linear kinematic source inversion is quite large in our study, especially for the tests that require the inversion of a large number of waveforms (up to a maximum of 168 stations), and the total number of investigated models in our study is large. Hence, we did not calculate the optimal α^2 values for every test. We simply assume that α^2 is linearly proportional to the number of stations and so scale the reference value with the number of stations.

The optimization algorithm we use to explore the model space is an evolutionary algorithm (EA, Beyer 2001), which is a population-based stochastic optimization method, involving a series of mechanisms like recombination, mutation, and selection to improve the characteristics (fitness). The EA requires a certain number of parameters to be tuned, e.g. number of parents (μ) and offsprings (λ), respectively, and the standard deviations of source parameters for the mutation operator. Additional details about these parameters are described in Monelli & Mai (2008). Unfortunately, no general theory is available that helps to choose optimal values for these parameters. Additional trial and error procedure is usually required, and we choose these parameters in this study according to previous work (e.g. Monelli & Mai 2008).

3 NUMERICAL TESTS

3.1 Dynamic rupture model

We choose a vertical purely strike-slip spontaneous dynamic model from the database of Dalguer & Mai (2011), which was calculated using the Support Operator Rupture Dynamics code (Ely *et al.* 2008, 2010) based on a generalized Finite Difference Scheme. This vertical fault model is 36 km along strike and 18 km along dip, buried 2 km from surface. The stress parameterization assumes depth-dependent normal stress distribution that results also in a depth-dependent initial shear stress distribution as shown in Fig. 2(a). The model has a grid resolution that can solve waveforms up to 3 Hz on the surface. More details on the dynamic parameterization can be found in Dalguer & Mai (2011). Baumann & Dalguer (2014) evaluated the compatibility of the ground motion generated by the rupture models of Dalguer & Mai (2011) with empirical ground motion prediction equations (GMPEs). In the model used in this paper, rupture initiates spontaneously at a circular patch of 4 km diameter located at 14.9 km depth. The dynamic rupture solution is represented by the rupture speed, peak slip velocity and slip distribution shown in Figs 2(b)–(d). The final rupture has a heterogeneous slip distribution, including one large asperity, centred between 5 and 15 km along strike, at 10 km depth. A notable feature is that the area with peak slip velocity does not completely overlap with the slip asperity. In fact, the slip is relatively small where the peak slip velocity is the largest. The 1-D seismic velocity structure and density model shown in Fig. 3 are used for both dynamic rupture model simulation and kinematic source inversion in this study.

3.2 Model setups

For the purpose of the inversion the fault plane is subdivided into 2 km × 2 km subfaults, resulting in 10 patches along dip and 19 patches along strike direction, with a total number of 190 subfaults. There are four kinematic source parameters to be determined of each grid node: peak slip velocity ($V_{\rm peak}$), rake angle, rupture time ($T_{\rm rup}$) and risetime ($\tau_R^{\rm eff}$). Tables 1 and 2 show the allowable range for source parameters and the various tuning parameters within the inversion code. At the boundary grid nodes, we set both the $V_{\rm peak}$ and slip as zero. We constrain the rake angle to be 180°, therefore only pure right-lateral strike-slip faulting is investigated in this study. Due to the heterogeneous rupture speed distribution, we consider a wide range of variation from subshear to supershear rupture speed. We set $V_s = 3.5\,{\rm km\,s^{-1}}$, and rupture speed V_r ranges from 0.5 to 1.1 V_s. Then the rupture initiation time is bounded between $D/\max(V_r)$ and $D/\min(V_r)$ for every grid node, where D is the on-fault distance to the hypocentre. We set $T_{\rm acc}$ to a constant value for every subfault patch via trial and error, and we select $T_{\rm acc} = 0.15$ s. Mena *et al.* (2012) sets it as 0.2 s. Moreover, Cirella *et al.* (2009) point out

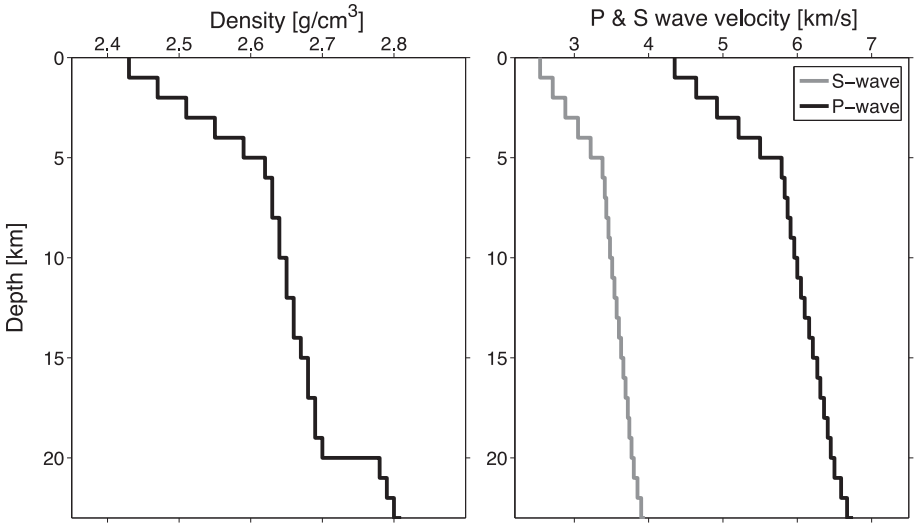

Figure 3. 1-D flat layered seismic velocity and density model for the test.

Table 1. Ranges of kinematic source parameters.

V_{peak} (cm s^{-1})	Rake (°)	V_{rup} (km s^{-1})	T_{rise} (s)
0–400	180	0.5–1.1 × Vs^a	0.4–8.0

[a] Vs means the average shear velocity value from depth 2–20 km in Fig. 3

Table 2. Tuning parameters for the Evolutionary Algorithm.

No. of parents and offsprings		Standard deviation σ		
μ	λ	V_{peak} (cm s^{-1})	T_{rise} (s)	T_{rup} (s)
480	19 200	2	0.2	0.2

that different T_{acc} does not substantially affect the imaged slip and rupture time patterns. The risetime should be two times larger than τ_s ($\tau_s = T_{acc}/1.27$), approximately 0.4 s. Therefore we limit the value of risetime from 0.4 to 8.0 s, and set it to 0.4 s at the boundary grid nodes. The standard deviations of V_{peak}, τ_R^{eff} and T_{rup} for the mutation operator is 2 cm s^{-1}, 0.2 s and 0.2 s. Consequently the total number of the unknown source parameters is 462 (190 for rupture time, 136 for V_{peak} and 136 for τ_R^{eff}). Following the setting for the Evolutionary Algorithm in Monelli & Mai (2008), we also set ratio of μ and λ as $\frac{1}{40}$, and searching generation as 100. According to the computational capacity in our lab, we set the $\mu = 480$, $\lambda = 19\,200$, so the total number of the estimated models is 1 920 480. The smoothing parameter α^2 is set as 0.0001 for every station set. We do not expect these settings to be optimal (in terms of rendering the search the most efficient), as even if some guidelines are available, trial and error work is usually required to set these parameters.

To quantitatively assess the overall resolution of the source parameter distribution obtained from the inversion tests, we compare them with the source model obtained by the dynamic rupture solution, hereafter called 'true-model', and calculate the rms error. The effective fault area for comparison has been defined as the subfaults in which the slip is larger than the mean slip, corresponding for both the true-model and inverted source model.

3.3 Single station test

Here we investigate the effect of azimuthal and epicentral distance distribution on the quality of a kinematic source inversion, by inverting three components of velocity ground motion at only one single station. Due to symmetric ground motion distribution on either side of a pure strike-slip fault in a 1-D velocity model as considered in this study, we select 31 stations located on one side only as shown in Fig. 4(a). The fault discretization, Green's function calculation and estimated source parameters are discussed in Sections 2 and 3.2. We restrict the inversion to a low frequency band (0.01–1 Hz).

For each of the 31 single station tests, we select the model that produces the minimum waveform misfit after 100 iterations. Fig. 5 shows the model misfit distribution of estimated source parameters to true model for all 31 single station tests. The misfit values are normalized by the maximum and minimum rms error. A general summary of our result, as seen in Fig. 5, is that the errors of rupture time, slip, peak slip velocity and risetime are smaller for models with stations located at the quadrant of forward directivity. A close analysis of the error distribution of slip and peak slip velocity is that larger errors correspond to stations located near the *P*-wave nodal planes.

In order to assess the capability of our estimated source models to generate ground motion for each of the 31 best models, we compute synthetic ground velocity waveforms for all 168 stations distributed across the fault (see Fig. 4c). These waveforms are then compared with the waveforms generated by the true model. Fig. 6 shows the distribution of rms three-component velocity waveform fitting error for all 31 single station tests. Note that the rms misfit is computed for all 168 stations although only three-component waveforms of a single station are used in each inversion. Error patterns are similar to those of the source errors observed in Fig. 5, but the pattern is sharper for the waveform misfit, in which the best fitting models are from stations located midway between the two -wave nodal planes and the fit is better at the forward directivity nodal plane than backward directivity direction. The choice of objective function to calculate the waveform fitting is arbitrary to some extent. The stations with large amplitude waveforms have relatively large weighting in *L2* norm calculation in eq. (3). We also calculate the waveform fitting error using the normalized rms $O(m) = \sum \|g_i(m) - \mathrm{d}_i\|_2^2 / \|\mathrm{d}_i\|_2^2$ (d_i: waveform for ith station). As we show in Fig. S2, normalizing the amplitudes of the waveforms does not significantly impact the pattern comparing to Fig. 6. Although the patterns of waveform fit error in Fig. 6 and Fig. S2 are not similar in the region of backward directivity

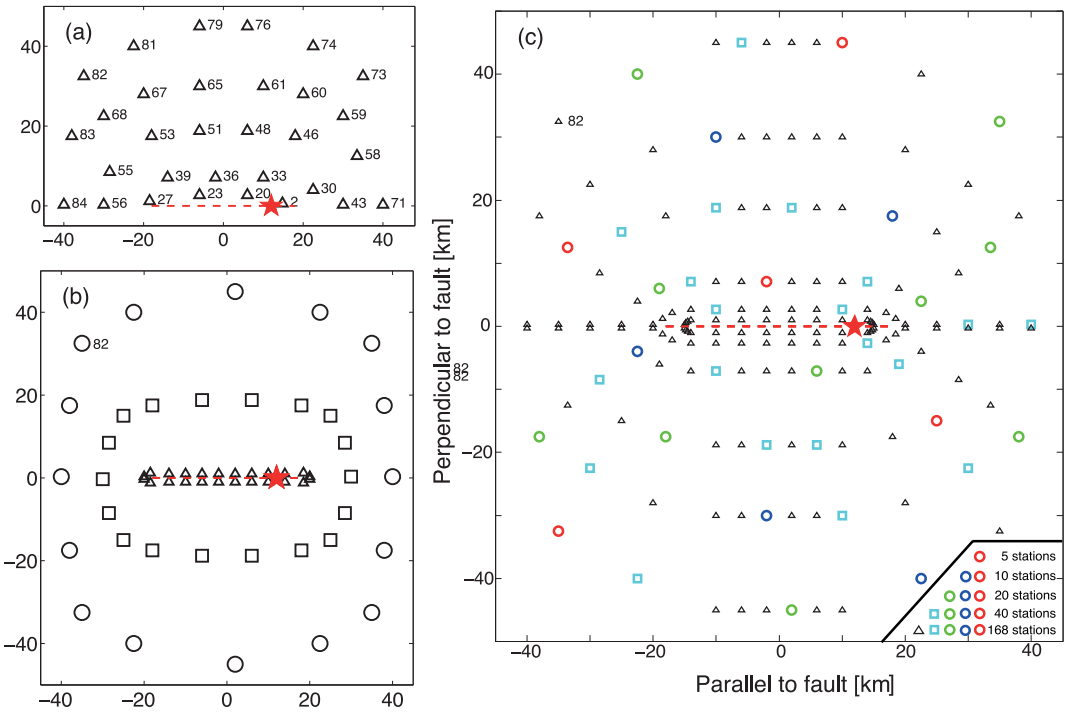

Figure 4. Network geometry tests: (a) single station, the 31 single station (triangles) selected; (b) circular distributed station networks (triangle, square and circle indicate C1 with stations near fault, C2 with stations moderate distance to fault and C3 with stations far distance to fault); (c) multistation, five well-spaced station networks (5, 10, 20, 40 all 168 stations). Red dashed line indicates the fault mapping on the surface, and red star indicates the epicentre location.

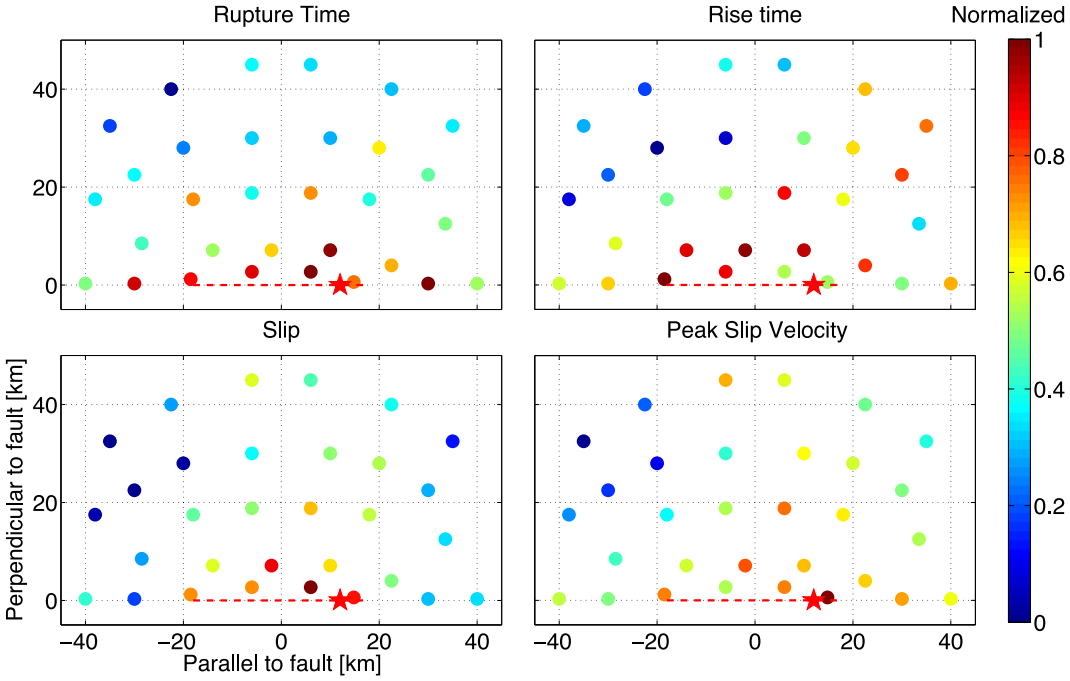

Figure 5. Comparison of source parameters estimated using single station inversion. The values are normalized by the the minimal and maximal rms error of these source parameters. Red dashed line indicates the fault mapping on the surface and red star indicates the epicentre location.

direction, both figures indicate that the best-fitting models are mainly from the stations located midway between the two *P*-wave nodal planes in the forward directivity region.

The best single station model, with the lowest errors in both source model and waveform, is from station Nr 82 (see Fig. 4a). Fig. 7 compares the three components velocity waveforms recorded at 84 stations (upper side of the fault shown in Fig. 4c) generated by this model (red line) with those from the true-model (black line) at low frequencies (<1.0 Hz), and the maximum values (cm s^{-1}) of the velocity waveforms from dynamic modelling are shown in the right of each trace. The qualitative comparison is very good for all the 84 stations with small discrepancies at short periods. This is surprising since only one station (Nr. 82) has been used for the inversion. We further evaluate five source models (using station 27, 43, 48, 73

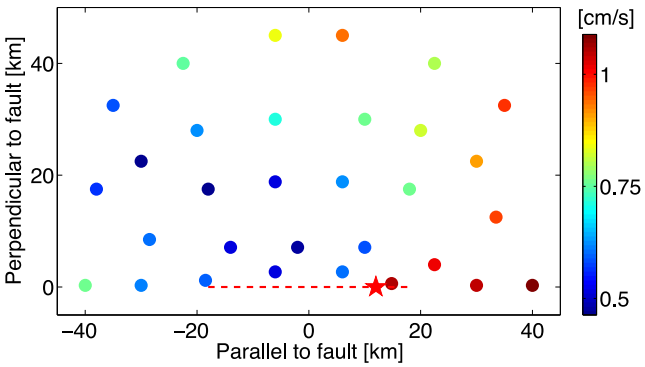

Figure 6. The rms three-component velocity waveforms fitting errors between the dynamic model and the estimated source models basing on various single station combinations. The waveform misfit is calculated for all 168 stations, although a single station is used in inversion.

and 82) as representative examples, named as model S27, S43, S48, S73 and S82 (S means single station test, the numbers indicate the single station number used as shown in Fig. 4a). Fig. 8(a) shows the slip, peak slip velocity and rupture time distribution for each model. The corresponding figures for the true-model are plotted at the top of this figure for comparison. For the slip, S82 and S27 (the model with the lowest waveform misfit, Fig. 6) show good consistency with the true-model, capturing both the location of the asperity and the amplitude of slip. S73 (with poor waveform misfit) and S48 (with good waveform misfit) capture half of the asperity, but do not image the rest. Model S43 (with poor waveform misfit) poorly matches the asperity, and does not match the amplitude. The stations located in the direction of rupture propagation (27 and 82) better resolve slip characteristics than those located in the backward direction (43 and 73). For the peak slip velocity (V_{peak}), all models fail to match the location of largest V_{peak}. S43 and S73 are quite similar in terms of amplitude and location of the largest slip area. Overall, all models follow the same general pattern, with the largest asperity area (defined by the slip) at a similar position to the true-model. Regarding the rupture time, all estimated models show similar patterns to the true-model during the first 6 s, and fail to capture the pattern toward the end of the rupture.

3.4 Circular distributed network test

Iida *et al.* (1990) investigate circularly distributed network geometries, and suggest that this type of strong-motion array is optimal for source inversion if the radius is 0.75–2.0 times the fault length. Hence, we investigate three uniformly spaced, circularly distributed station networks at similar distances from the fault (Fig. 4b). We call these networks as C1 (C means circular distributed network, shown in Fig. 4b) with 24 stations extremely close to the fault (Joyner–Boore distance, $R_{\mathrm{jb}} = 0.3$ km), C2 with 18 stations with intermediate Joyner–Boore distance (R_{jb} is about 10–20 km), and C3 with 16 stations with largest Joyner–Boore distance to fault (R_{jb} is about 22–50 km), about 0.6–1.4 times the fault length (36 km). We investigate whether the circular distributed station network could improve the source inversion results.

We use the same set of model parameters as in the single station test. We assume the value of α^2 is proportional to the number of station used, and then set the value as 0.0024, 0.0018 and 0.0016 for the three station configurations. The best estimated source models for each station network are named as model C1, C2 and C3. Fig. 8(b) compares the estimated slip, peak slip velocity and rupture time with the true model. All the estimated models by the three circular distributed network match the location of slip patch despite having variable peak slip. However, they all fail to capture the patch of V_{peak}, similar to the single station tests. Fig. 9 shows the quantitative estimation of errors for all the models used in this study (and is discussed later in more detail). For the three circular geometries, C3 with the largest distance to fault, produces the lowest source misfit and largest waveform misfit, while C1 is exactly opposite, that is, it has the lowest waveform misfit, but is the worst model for matching the source, with the stations closest to the fault.

3.5 Multistation test

Sarao *et al.* (1998) investigate the effect of non-uniform station coverage on the inversion for earthquake rupture history using a simple and classic Haskell-type source model, and show that even four stations located at forward direction of the rupture propagation are sufficient to retrieve the earthquake image. Although this model is simple, the conclusion is consistent with our general findings using single station (Fig. 5). Although here we show that only one station may be sufficient to capture the key elements of source and have good waveform fitting, however, Asano & Iwata (2009) study the 2004 Chuetsu earthquake (Japan) using dense strong-motion data, and point out that more than 12 stations are sufficient to obtain a stable solution. In the same vein, here we test different combinations of non-uniform station distributions that do have good azimuthal coverage and epicentral distance. We use five cases with 5, 10, 20, 40 and 168 stations (Fig. 4c) to investigate how station number and geometry affect the source inversion results.

The same set of model parameters are estimated as in the previous tests, setting the value of α^2 as 0.0005, 0.001, 0.002, 0.004 and 0.017 for these station configurations (5, 10, 20, 40 and 168 stations). We select the best source models in 100 iterations as before. Models are named as M5, M10, M20, M40 and M168 (M means multistation, the numbers indicate stations used as shown in Fig. 4c). Fig. S3 shows the relationship between waveform data misfit and the generation number for the multistation test M10. The maximum generations in this case is 200. After about the 40th generation, the misfit reaches an approximately stationary level. It is noticeable that the convergence speed is related to the parameters in optimization algorithm used (e.g. number of parents and offsprings in each generation in the Evolutionary Algorithm), and the size of earthquake itself (e.g. the number of unknown source parameters and the amount of observed data sets). Fig. 8(c) shows slip, peak slip velocity and rupture time distribution for these models. For the final slip, all five estimated models match the location of target asperity, where the peak amplitude of estimated final slip is slightly larger than the target value. However, all the source models fail to capture the details of small slip patches along rupture direction. For the peak slip velocity (V_{peak}), the five estimated models fail to capture the location of the major rupture area and amplitude. The models suggest large V_{peak} is located close to epicentre, however the true model shows V_{peak} is in the centre of fault plane. Contour lines of rupture time show that the estimated models show a similar pattern of rupture time as to the 'true' model for first few seconds (0–6 s), and fail to capture the pattern after this, as was the case for single and circular station models (Figs 8a and b). However, the estimated rupture front is not as smooth as the target, for example the contour of rupture time at 4 s are obviously zig-zag shaped among model M5, M20 and M40.

Nevertheless, a quantitative estimate of the misfit of source parameters for all the cases studied here, as shown in Fig. 9, suggests

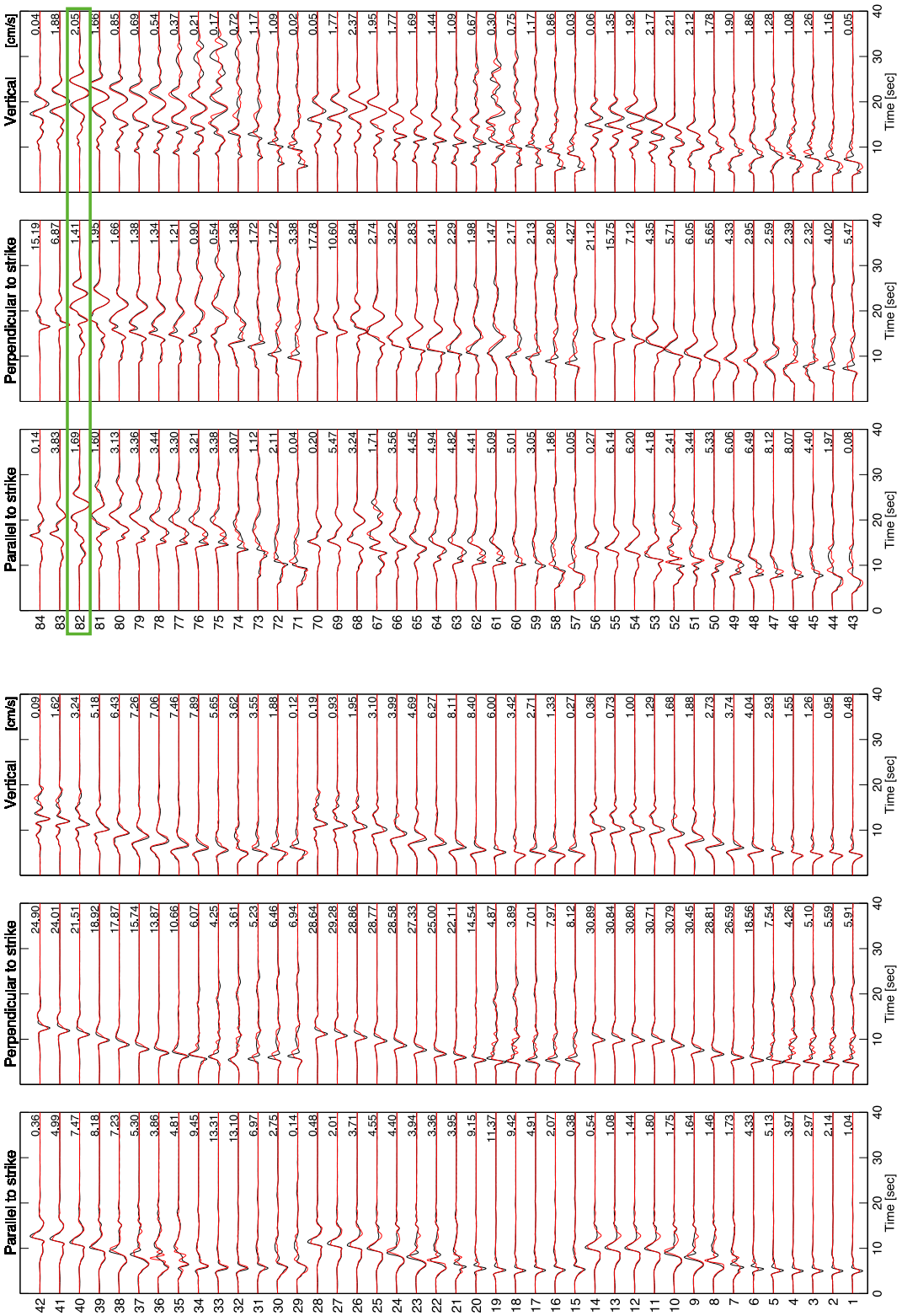

Figure 7. Comparison of the three-component synthetic velocity waveforms (red lines) of estimated model S82 using station 82 indicated in green box and target waveforms generated by dynamic model (black lines) at low frequency (<1 Hz). The maximum values (cm s^{-1}) of the target waveforms are shown in the right of each trace. All 84 stations on upper side of the fault are shown in Fig. 4(c).

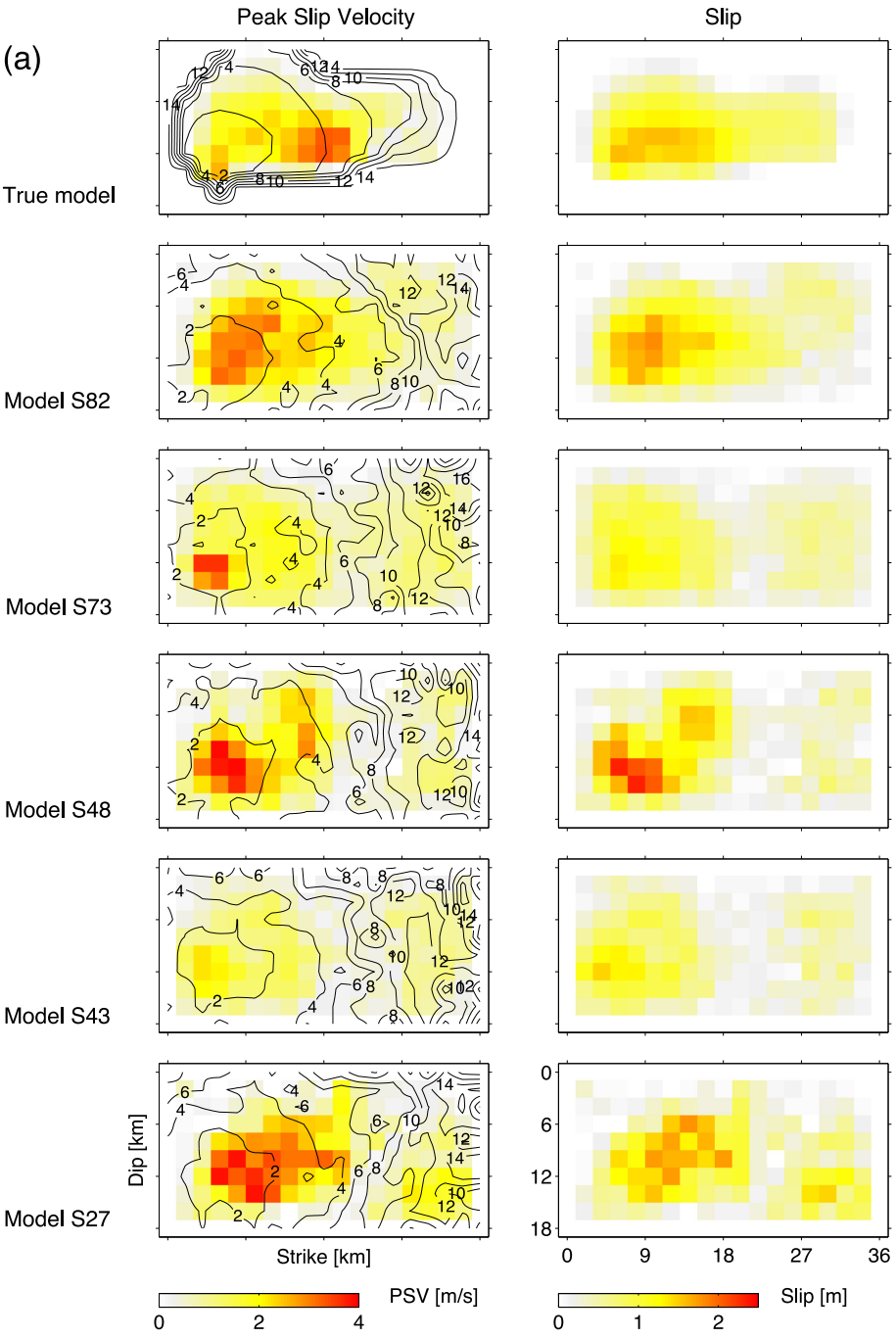

Figure 8. Comparison of estimated kinematic source parameters in three tests with true values derived from dynamic rupture model. (a) Single station estimated models S82, S73, S48, S43 and S27. Station locations are shown in Fig. 3(a). (b) Circular geometry models C1(stations near fault), C2(stations moderate distance to fault), C3(stations far distance to fault). Station locations are shown in Fig. 3(b). (c) Multistation variable geometry models basing on five stations (M5), 10 stations (M10), 20 stations (M20), 40 stations (M40) and 168 stations (M168). Station locations are shown in Fig. 3(c). The left-hand column indicates the peak slip velocity (contours indicate the estimated rupture time). The right-hand column indicates the slip.

that there is no a systematic improvement in source modelling when increasing the number of stations. The source misfit from multistation models fluctuates around the range of errors obtained from single station models. To calculate the waveform misfit, we proceed as the previous tests, that is, the estimated source models are used to generate velocity waveforms at 168 stations, and compare them with the waveforms generated by the true model at low frequency (<1 Hz). In this case, the waveform misfit systematically reduces with the number of stations, as seen in Fig. 9.

4 DISCUSSION AND CONCLUSIONS

4.1 Source model and waveform misfit

In order to evaluate source models estimated from source inversion, we quantitatively compare estimated source models derived from various network geometries with the true model and evaluate the capability of each estimated source model to generate ground motion at stations outside the network, that is including stations not

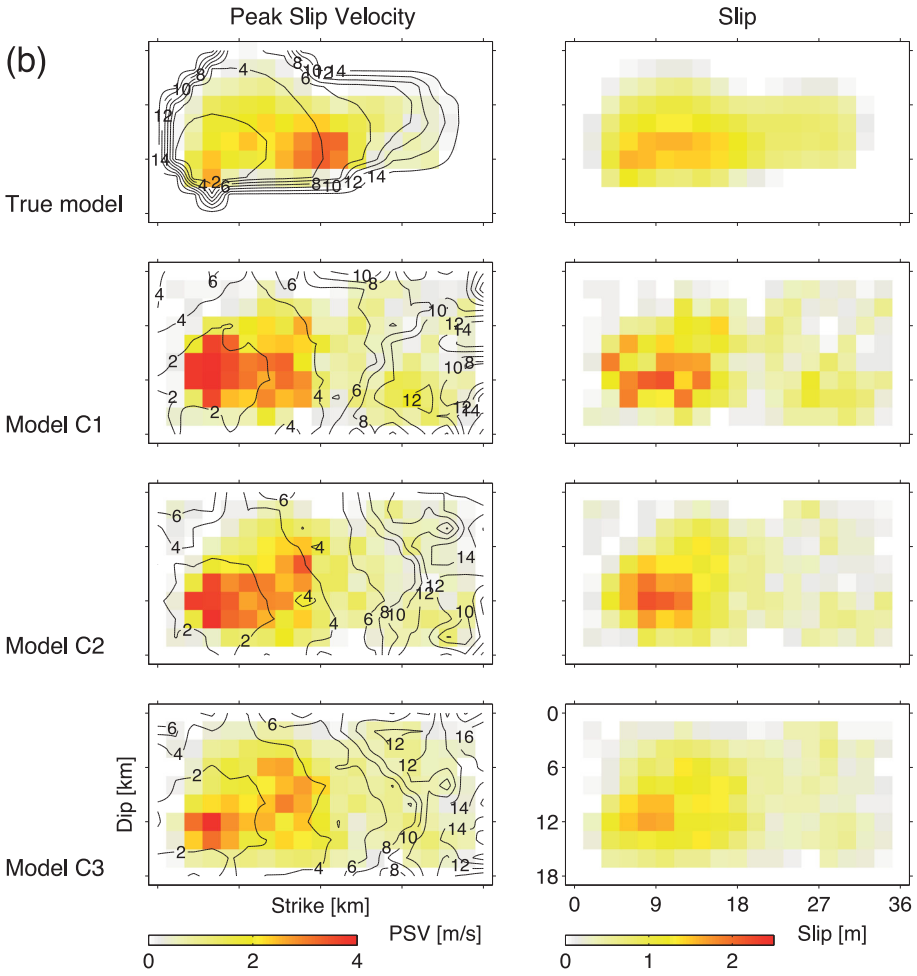

Figure 8. *(Continued.)*

used in the inversion. A first order understanding of the effect of network geometry on source inversion without investigation, likely be: (1) the best kinematic source model generates the minimum waveform misfit; (2) the larger the number of available stations, the more reliable the source; (3) a dense array of stations very near the fault would best capture the source detail and (4) a source model obtained by inverting only one station may not be sufficient to generate reliable ground motion and the estimated source should not contain sufficient details to describe the true source. Nevertheless, the summary of the source and waveform misfits from all the source models we have investigate, as shown in Fig. 9, reveals that these first order expectations may not be correct. Figs 9(a)–(c) show the source rms errors, for the full trace of slip velocity function (see eq. 1), slip, and peak slip velocity; and Fig. 9(d) shows the rms waveform misfit of ground motion for all 168 stations. All estimated source models are linearly interpolated to 0.1 km × 0.1 km and directly compared to true dynamic rupture model. The SVF gives the best source error quantification as it uses the complete waveform in time (Fig. 9a), but these errors have similar pattern as those obtained from the slip (Fig. 9b) and peak slip velocity (Fig. 9c). As shown in Fig. 9, the source error from multistation models has the same range of errors seen in many single station models. In fact, the source model with the lowest model misfit correspond to a single station, but this model does produce larger waveform errors than those with multistation models. If we consider models derived from multistation geometries only, the trend is similar. Model C3 that uses 16 stations regularly spaced around the fault has in general (considering the three metrics) the lowest source error but the largest waveform error. There is no systematic correlation of the quality of source estimates with the number of stations and with the waveform misfit, but there is a systematic improvement of the waveform misfit (computed for all 168 stations) with respect to the number of stations as shown in Fig. 9(d).

Our study suggests that the source inversion algorithm is an efficient tool to find source estimates that produce the best waveform misfit, but not necessarily the best (or true) source model. We think that this may be linked to the fact that we simplified our model space both in space and time in the inversion, that is coarse gridding (2 km × 2 km patch) and a simplified SVF (regularized Yoffe function). As shown in Fig. 1(d), the regularized Yoffe function does not represent all the variations of the true model in time, but also in space due to the coarse gridding of the fault. Given the fact that we have deviations from the true model in model parameterization in the inversion, the best model may not necessarily produce the best waveform data fitting even in a low frequency band (<1 Hz). On the other hand, the non-linear inversion algorithm used may automatically find model estimates that produce the best waveform misfit, but not necessarily the best source estimate. This procedure of finding a model that fit data may add non-existing terms to the source that may result in overfitting the waveforms. Yagi & Fukahata (2008) also mention the overfitting phenomenon if the covariance components of waveforms are neglected. This overfitting is shown

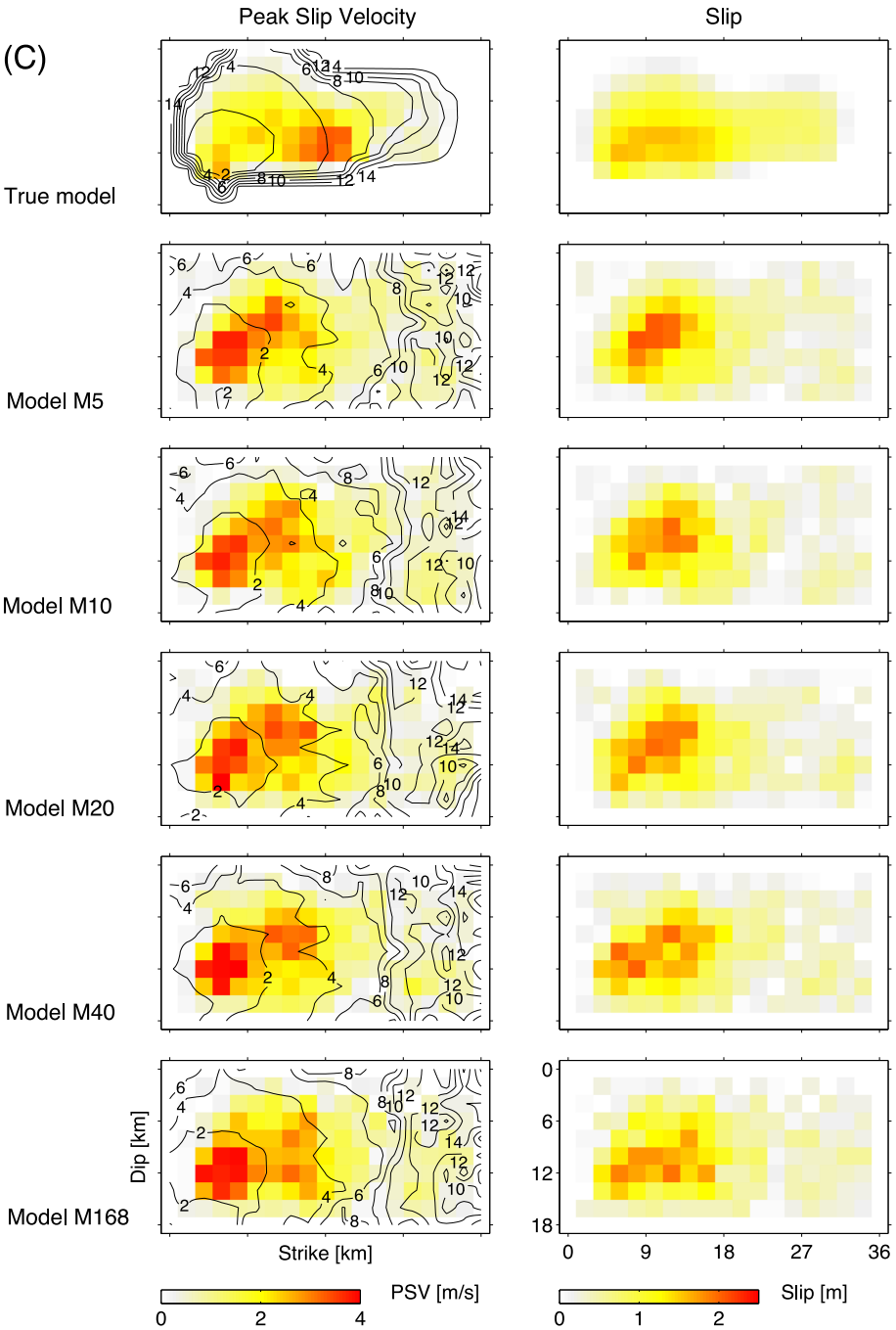

Figure 8. *(Continued.)*

in Fig. 9(d), the source models obtained from the multistation tests produce waveform errors smaller than the minimum expected misfit obtained from the best kinematic source models we can get (horizontal lines). These best expected kinematic source models are derived directly from the true model (without any inversion) and characterized by the SVF defined by eq. (1). The regularized Yoffe function has four parameters (V_{peak}, T_{acc}, slip and risetime). We calculate one source parameter using eq. (1) while fixing the other three. T_{acc} is always determined through dynamic model within 0.1–0.35 s. Hence we build three kinematic models named best kinematic model K1 (risetime), K2 (psv) and K3 (slip), which are, respectively the red, blue, and black horizontal lines shown in Fig. 9. The source parameters in brackets are calculated in each model while fixing the other with values extracted from the dynamic model. As mentioned before, these three models are the best kinematic source model we can obtain within the limitations of the kinematic representation, as shown in Figs 9(a)–(c).

A dense array of stations located very close to the fault (C1 in Fig. 9) produces the worst source model within the circular distributed network tests, but has the best waveform misfit. The best source model corresponds to model C3 that has a circular station distribution of stations around the fault located at larger distance (about 0.6–1.4 times the fault length) than model C1. This is consistent with the work from Iida *et al.* (1990) who point out that the appropriate circular distributed network to find a good source has a radius around 0.75–2.0 times the fault length. Nevertheless,

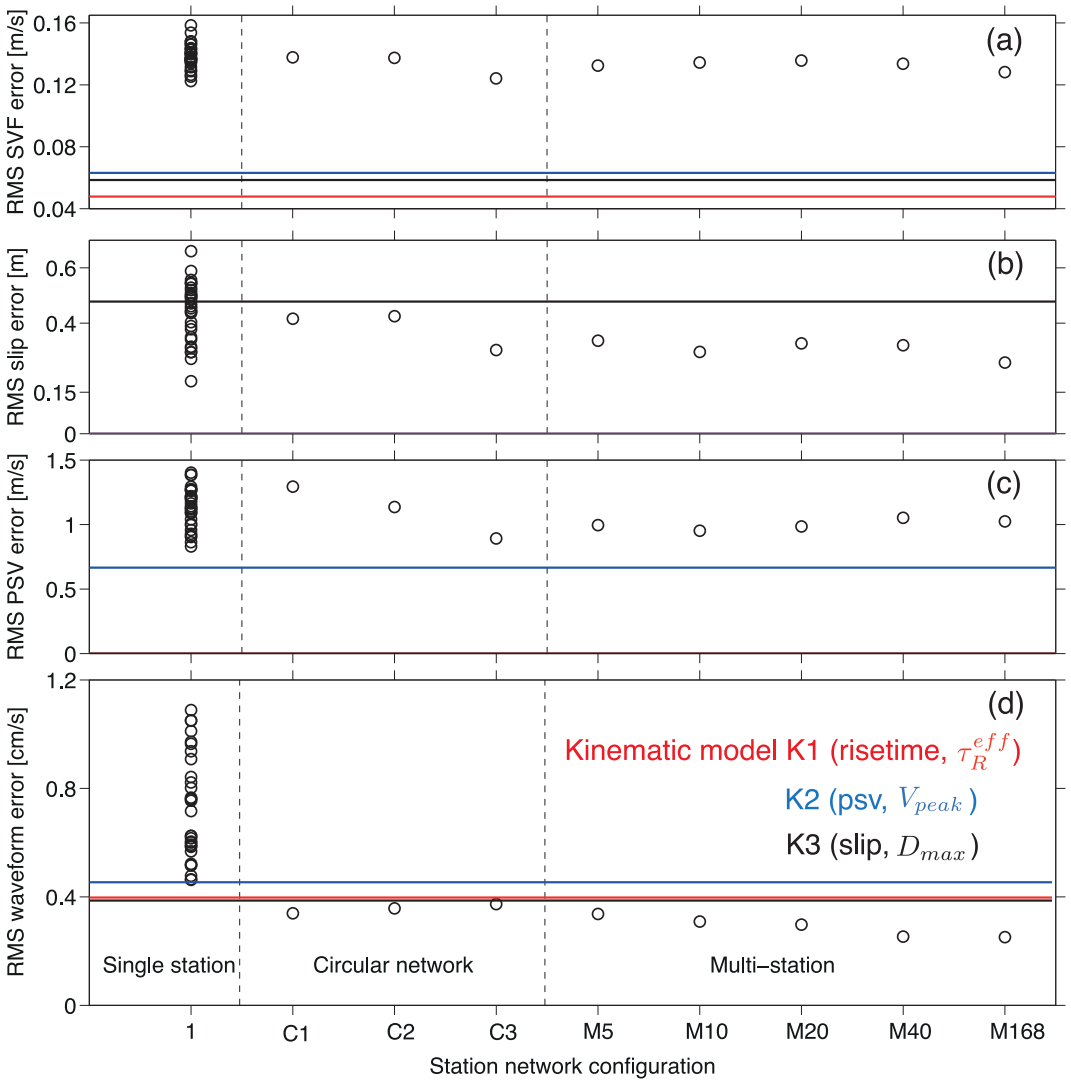

Figure 9. The rms error of estimated source parameters, (a) slip velocity function; (b) slip; (c) peak slip velocity; (d) all 168 stations three-component velocity waveforms fitting. Red, blue, and black horizontal lines indicate the three best kinematic models (K1, K2 and K3) derived from target dynamic model by varying one source parameter in parentheses while fixing the rest according the eq. (1).

this model results in the largest waveform misfit when compared to C1 (very close to the fault) and C2 (intermediate distance). These three models from the circular distributed network tests have the common feature that the waveform misfits are about the same as the kinematic model K1, K2 and K3.

4.2 Seismic wave radiation pattern: Is one single station enough to retrieve the source?

The inversion test using one single station reveals surprising results. As shown in Figs 7–9, the source and waveforms obtained by selected single stations are comparable to those obtained by multistation tests. The source error can be even lower than those obtained by the best source model with multistations for some cases (Figs 9a–c), and the lowest waveform error has about the same error level obtained from the best kinematic source models (Fig. 9d). The qualitative comparison of synthetic waveforms generated by the best model in the single station test with waveforms generated by dynamic modelling is excellent at most stations (Fig. 7). This suggests that the source model obtained from the inversion of a single station can contain the necessary features to generate accurate estimated ground motions at stations not used in the inversion. Figs 5 and 6, showing the patterns of source and waveform misfit respectively, indicate there is a well-defined pattern of station locations where source and waveform misfit are lower. The better source estimates correspond to stations located at the quadrant of forward directivity. In a similar pattern, better waveform misfits corresponds to models obtained from stations located midway between the two *P*-waves nodal planes. Interestingly, along the horizontal nodal plane, the stations located at the forward directivity produce better waveform errors than those models located at the backward directivity. It seems that seismic radiation pattern plays a role in defining the optimum location of stations used to retrieve the source. In order to understand the contribution of the three components of ground motion in the inversions that uses a single station, we calculate the normalized three components peak velocity amplitude at each station and represent them in circles as shown in Fig. 10. Each colour represents one component (blue = fault parallel, green = fault perpendicular, red = vertical). A large circle indicates one component is dominant, and three smaller circles about

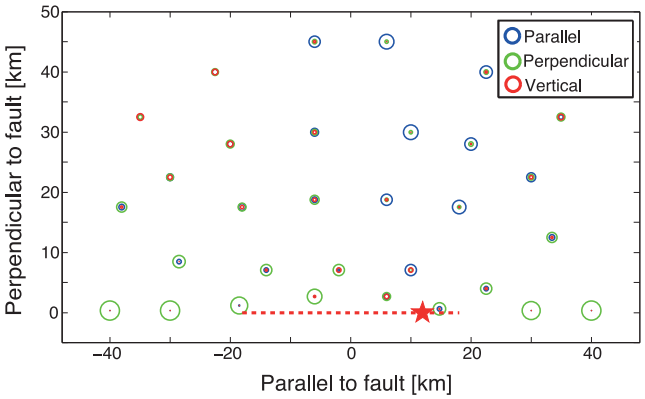

Figure 10. Comparison of the three components weighting for each single station.

the same size indicate that the peak velocity amplitude of the three components are similar. Comparing Figs 10 and 5, we observe the location of the smaller circles correspond to starting with smaller source error distribution, except risetime, and the larger circles correspond to the larger source errors. The stations located at the -wave nodal plane are dominated by the large circles, that is dominated by only one component. At these nodal planes locations, the waveform and source errors are larger in general. This analysis suggests that the best location of a station in order to retrieve a reliable source is in a position where the amplitude ratios between the three components are not large and at a distance about 0.7–1.2 the fault length. It indicates that the seismic wave radiation pattern should play an important role in selection of the stations.

If we were to consider realistic conditions such as large uncertainty in fault location, velocity structure, site effect, noise in observed ground motions, and other factors that may adversely affect the inversion results, the use of only one station for the source inversion of real earthquakes may be too optimistic. Therefore, we argue that in a realistic case, a relative small number of stations (5–10) with high quality records of the event, and with good azimuthal and epicentral distance should be sufficient to retrieve a reliable solution because increasing the number of stations do not improve the source (Fig. 9). This conjecture is consistent with other studies that use kinematic source inversion techniques for real earthquake (e.g. Asano & Iwata 2009), but it is also consistent with other studies that use different methods to retrieve sources, such as the study of Jakka *et al.* (2010) that uses the back projection method to study synthetic near-source ground motions. This author has demonstrated the utility of the method to identify slip asperities and their associated intensities, with a small number of stations (less than 5) with good azimuthal and epicentral coverage around the source in the ideal case.

4.3 Uncertainty in source inversion

The uncertainty in the derived source is an issue that must be addressed, especially for non-linear kinematic source inversion. Duputel *et al.* (2014) summarize that two kinds of uncertainties should be considered in fault slip inverse problem. First, the observational error is induced by imperfect measurements. Second, the prediction error is the uncertainty due to imperfect forward modelling. Inaccurate underground velocity structure could reduce the resolution of source image (Graves & Wald 2001; Razafindrakoto & Mai 2014), and broaden the posterior probability density function (PDF) of estimated source parameters (Razafindrakoto & Mai 2014), though we note other authors report minor inaccuracies in the Green's function do not significantly affect source inversion results (Konca *et al.* 2013).

In this study, the velocity waveforms that we inverted are generated by a scenario dynamic rupture model, and we do not add noise to them. Therefore there is no observational error in this study. Additionally, we investigate a well-known scenario earthquake with a known underground velocity structure, that means there is no uncertainty on underground velocity structure, fault geometry and hypocentre location, etc. We do not consider the anelastic attenuation in both dynamic rupture simulation and synthetic source inversion in this study. So there is no prediction error on anelastic behaviour of Earth structure (e.g. seismic wave energy attenuation). The predicted ground motion could be different for elastic and anelastic crustal model given the same source image. Furthermore, anelastic attenuation of the Earth could lead to temporal correlation of densely sampled seismic waveform data (Yagi & Fukahata 2008).

However, the major uncertainty of prediction comes from the parameterization of earthquake source. Fig. 1 shows that the evolution of slip rate on fault varies significantly along the rupture direction. The pulse-like slip velocity evolution (as Figs 1a and b) which could be well fitted by using single window SVF, for example regularized Yoffe function (Tinti *et al.* 2005) used in this study (as Fig. 1d). However Fig. 1(c) shows slip reactivation which is more suitable to fit by multiwindow than single window SVF. This is a disadvantage of non-linear kinematic source inversion for these complicated earthquake source ruptures. Overall, in this study, the prediction uncertainty mainly comes from the unmatchable evolution of slip rate on subfaults and the pre-assumed regularized Yoffe function, and the uncertainty is systematic for all the three group tests.

4.4 Macro scale representation of the source: moment rate

The source rupture evolution in time can be also characterized by a single macro scale parameter such as the moment rate function. Its time integral is the seismic moment that is a measure of the size of the earthquake. We estimated the moment rate from all the source models estimated by the inversion and compare them with the one from the true model. Fig. 11 shows this comparison. The discrepancy of amplitude with respect to the true model is obvious. All the models fluctuate around the true solution. However, the shape and total rupture time closely follow the true model in all cases. At this scale of qualitative comparison, all the models, independent of the number of stations and geometrical distributions, are consistent with the true solution.

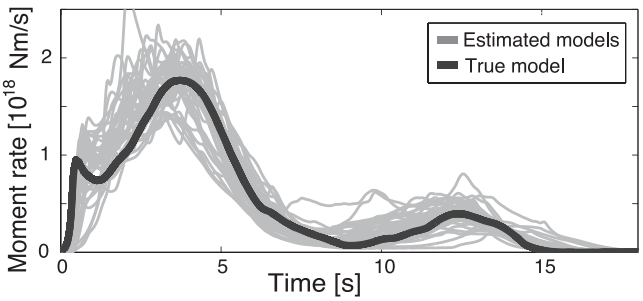

Figure 11. Comparison of seismic moment rate function of true model (dynamic rupture model) and estimated models including all the best estimated models in three tests.

4.5 Future direction: pseudo-dynamic source inversion

As demonstrated in the paper, the quality of model estimates in source inversion does not systematically correlate with the quantity of data. Although the waveform misfit metric for all 168 stations decreases as we include more stations in the inversion, the model estimates are not significantly improved. This also leads to the issue of data over-fitting in the inversion. If it is difficult to constrain the model space stably by simply increasing the number of stations, we may need to constrain it with some other means in addition to observational data. We adopted a smoothing operator for final slip in this study to regularize the model space. But the smoothing operator for a certain model parameter does not have any physical basis, and it is even somewhat arbitrary to determine the appropriate level of smoothing for each inversion problem. We believe that an alternative solution to this problem is to impose physical constraints in the source inversion and regularize the model space with our prior knowledge about earthquake source. Pseudo-dynamic source modelling methods have been developed to produce a number of rupture scenarios for simulating ground motions for future events (e.g. Guatteri *et al.* 2004; Schmedes *et al.* 2010; Song & Somerville 2010; Mena *et al.* 2012; Song & Dalguer 2013; Trugman & Dunham 2014). This approach aims to derive simple relationships between kinematic source parameters obtained from physics-based dynamic rupture models. The pseudo-dynamic source models are considered physically self-consistent since they are constrained by both rupture dynamics and past events. This suggests that we may also be able to utilize the concept of pseudo-dynamic source modelling in constraining the model space in kinematic source inversion. We would call it 'pseudo-dynamic source inversion' and it may have the potential to obtain better model estimates in source inversion by combining data constraints with physical constraints from the pseudo-dynamic source models.

ACKNOWLEDGEMENTS

The authors thank Ralph Archuleta, Susana Custodio, Jean-Paul Ampuero, and Frantisek Gallovic for helpful and enjoyable discussions during the development of this study. We also thank Damiano Monelli for providing the kinematic source inversion code (KISS) and helpful discussion. This project is funded by Swiss National Science Foundation (SNF), SNF Grant 200021-130061. Simulations were done at the Swiss National Supercomputing Center (CSCS), under the production projects, 'Development of Dynamic Rupture Models to Study the Physics of Earthquakes and Near-Source Ground Motion' and 'Development of a Database of Physics-Based Synthetic Earthquakes for Ground Motion Prediction'.

REFERENCES

Andrews, D.J., 1976. Rupture velocity of plane strain shear cracks, *J. geophys. Res.*, **81**(32), 5679–5687.

Asano, K. & Iwata, T., 2009. Source rupture process of the 2004 Chuetsu, Mid-Niigata prefecture, Japan, earthquake inferred from waveform inversion with dense strong-motion data, *Bull. seism. Soc. Am.*, **99**(1), 123–140.

Baumann, C. & Dalguer, L.A., 2014. Evaluating the compatibility of dynamic rupture-based synthetic ground motion with empirical ground-motion prediction equation, *Bull. seism. Soc. Am.*, **104**(2), doi:10.1785/0120130077.

Beresnev, I.A., 2003. Uncertainties in finite-fault slip inversions: to what extent to believe? (a critical review), *Bull. seism. Soc. Am.*, **93**(6), 2445–2458.

Beyer, H.G., 2001. *The Theory of Evolution Strategies*, Springer.

Bizzarri, A., 2012. Analytical representation of the fault slip velocity from spontaneous dynamic earthquake models, *J. geophys. Res.*, **117**(B6), doi:10.1029/2011JB009097.

Bouchon, M., Toksöz, M.N., Karabulut, H., Bouin, M.-P., Dietrich, M., Aktar, M. & Edie, M., 2002. Space and time evolution of rupture and faulting during the 1999 Izmit (Turkey) earthquake, *Bull. seism. Soc. Am.*, **92**(1), 256–266.

Cirella, A., Piatanesi, A., Cocco, M., Tinti, E., Scognamiglio, L., Michelini, A., Lomax, A. & Boschi, E., 2009. Rupture history of the 2009 L'Aquila (Italy) earthquake from non-linear joint inversion of strong motion and GPS data, *Geophys. Res. Lett.*, **36**(19), doi:10.1029/2009GL039795.

Cotton, F. & Campillo, M., 1995. Frequency domain inversion of strong motions: application to the 1992 Landers earthquake, *J. geophys. Res.*, **100**(B3), 3961–3975.

Custódio, S., Liu, P. & Archuleta, R.J., 2005. The 2004 M w6.0 Parkfield, California, earthquake: inversion of near-source ground motion using multiple data sets, *Geophys. Res. Lett.*, **32**(23), L23312, doi:10.1029/2005GL024417.

Custódio, S., Page, M.T. & Archuleta, R.J., 2009. Constraining earthquake source inversions with GPS data: 2. A two-step approach to combine seismic and geodetic data sets, *J. geophys. Res.*, **114**(B1), doi:10.1029/2008JB005746.

Dalguer, L.A. & Mai, P.M., 2011. Near-source ground-motion variability from M 6.5 dynamic rupture simulations, in *Proceedings of the 4th IASPEI/IAEE International Symposium on Effects of Surface Geology on Seismic Motion*, August 23–26, University of California Santa Barbara, CA, USA.

Dalguer, L.A., Irikura, K., Riera, J.D. & Chiu, H.C., 2001. The importance of the dynamic source effects on strong ground motion during the 1999 Chi-Chi, Taiwan, earthquake: brief interpretation of the damage distribution on buildings, *Bull. seism. Soc. Am.*, **91**(5), 1112–1127.

Dalguer, L.A., Miyake, H., Day, S.M. & Irikura, K., 2008. Surface rupturing and buried dynamic-rupture models calibrated with statistical observations of past earthquakes, *Bull. seism. Soc. Am.*, **98**(3), 1147–1161.

Das, S. & Aki, K., 1977. Fault plane with barriers: a versatile earthquake model, *J. geophys. Res.: Solid Earth (1978–2012)*, **82**(36), 5658–5670.

Day, S.M., 1982. Three-dimensional simulation of spontaneous rupture: the effect of nonuniform prestress, *Bull. seism. Soc. Am.*, **72**(6A), 1881–1902.

Day, S.M., Gonzalez, S.H., Anooshehpoor, R. & Brune, J.N., 2008. Scale-model and numerical simulations of near-fault seismic directivity, *Bull. seism. Soc. Am.*, **98**(3), 1186–1206.

Duan, B. & Day, S.M., 2010. Sensitivity study of physical limits on ground motion at Yucca Mountain, *Bull. seism. Soc. Am.*, **100**(6), 2996–3019.

Dunham, E.M. & Bhat, H.S., 2008. Attenuation of radiated ground motion and stresses from three-dimensional supershear ruptures, *J. geophys. Res.*, **113**(B8), B08319, doi:10.1029/2007JB005182.

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

Ely, G.P., Day, S.M. & Minster, J.-B., 2008. A support-operator method for viscoelastic wave modelling in 3-D heterogeneous media, *Geophys. J. Int.*, **172**(1), 331–344.

Ely, G.P., Day, S.M. & Minster, J.B., 2010. Dynamic Rupture Models for the Southern San Andreas Fault, *Bull. seism. Soc. Am.*, **100**(1), 131–150.

Emolo, A. & Zollo, A., 2005. Kinematic source parameters for the 1989 Loma Prieta earthquake from the nonlinear inversion of accelerograms, *Bull. seism. Soc. Am.*, **95**(3), 981–994.

Fukuyama, E. & Irikura, K., 1986. Rupture process of the 1983 Japan Sea (Akita-Oki) earthquake using a waveform inversion method, *Bull. seism. Soc. Am.*, **76**(6), 1623–1640.

Gabriel, A.A., Ampuero, J.P., Dalguer, L.A. & Mai, P.M., 2012. The transition of dynamic rupture styles in elastic media under velocity-weakening friction, *J. geophys. Res.*, **117**(B9), B09311, doi:10.1029/2012JB009468.

Graves, R.W. & Wald, D.J., 2001. Resolution analysis of finite fault source inversion using one and three dimensional Green's functions: 1. Strong motions, *J. geophys. Res.: Solid Earth (1978–2012)*, **106**(B5), 8745–8766.

Guatteri, M., Mai, P. & Beroza, G., 2004. A pseudo-dynamic approximation to dynamic rupture models for strong ground motion prediction, *Bull. seism. Soc. Am.*, **94**(6), 2051–2063.

Harris, R.A. *et al.*, 2009. The SCEC/USGS dynamic earthquake rupture code verification exercise, *Seismol. Res. Lett.*, **80**(1), 119–126.

Hartzell, S., Liu, P. & Mendoza, C., 1996. The 1994 Northridge, California, earthquake: investigation of rupture velocity, risetime, and high-frequency radiation, *J. geophys. Res.*, **101**(B9), 20 091–20 108.

Hartzell, S.H. & Heaton, T.H., 1983. Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake, *Bull. seism. Soc. Am.*, **73**(6A), 1553–1583.

Iida, M., Miyatake, T. & Shimazaki, K., 1990. Relationship between strong-motion array parameters and the accuracy of source inversion and physical waves, *Bull. seism. Soc. Am.*, **80**(6A), 1533–1552.

Jakka, R.S., Cochran, E.S. & Lawrence, J.F., 2010. Earthquake source characterization by the isochrone back projection method using near-source ground motions, *Geophys. J. Int.*, **182**(2), 1058–1072.

Konca, A.O., Kaneko, Y., Lapusta, N. & Avouac, J.P., 2013. Kinematic inversion of physically plausible earthquake source models obtained from dynamic rupture simulations, *Bull. seism. Soc. Am.*, **103**, 2621–2644.

Liu, P. & Archuleta, R.J., 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**(B2), B02318, doi:10.1029/2003JB002625.

Lucca, E., Festa, G. & Emolo, A., 2012. Kinematic inversion of strong-motion data using a gaussian parameterization for the slip: application to the 2008 Iwate-Miyagi, Japan, earthquake, *Bull. seism. Soc. Am.*, **102**(6), 2685–2703.

Mai, P., Burjanek, J., Delouis, B., Festa, G., Francois-Holden, C., Monelli, D., Uchide, T. & Zahradnik, J., 2007. Earthquake source inversion blindtest: initial results and further developments, *AGU Fall Meeting Abstracts*, p. C8.

Mena, B., Dalguer, L.A. & Mai, P.M., 2012. Pseudodynamic source characterization for strike-slip faulting including stress heterogeneity and supershear ruptures, *Bull. seism. Soc. Am.*, **102**(4), 1654–1680.

Minson, S.E., Simons, M. & Beck, J.L., 2013. Bayesian inversion for finite fault earthquake source models I—theory and algorithm, *Geophys. J. Int.*, **194**(3), 1701–1726.

Miyatake, T., Iida, M. & Shimazaki, K., 1986. The effect of strong-motion array configuration on source inversion, *Bull. seism. Soc. Am.*, **76**(5), 1173–1185.

Monelli, D. & Mai, P. M., 2008. Bayesian inference of kinematic earthquake rupture parameters through fitting of strong motion data, *Geophys. J. Int.*, **173**(1), 220–232.

Ohnaka, M. & Yamashita, T., 1989. A cohesive zone model for dynamic shear faulting based on experimentally inferred constitutive relation and strong motion source parameters, *J. geophys. Res.: Solid Earth (1978–2012)*, **94**(B4), 4089–4104.

Olsen, K. B., Madariaga, R. & Archuleta, R.J., 1997. Three-dimensional dynamic simulation of the 1992 Landers earthquake, *Science*, **278**(5339), 834–838.

Olson, A.H. & Apsel, R.J., 1982. Finite faults and inverse theory with applications to the 1979 Imperial Valley earthquake, *Bull. seism. Soc. Am.*, **72**(6A), 1969–2001.

Peyrat, S., Olsen, K. & Madariaga, R., 2001. Dynamic modeling of the 1992 Landers earthquake, *J. geophys. Res.*, **106**(11; SECT 2), 26 467–26 482.

Piatanesi, A., Cirella, A., Spudich, P. & Cocco, M., 2007. A global search inversion for earthquake kinematic rupture history: application to the 2000 western Tottori, Japan earthquake, *J. geophys. Res.*, **112**, B07314, doi:10.1029/2006JB004821.

Pitarka, A., Dalguer, L.A., Day, S.M., Somerville, P.G. & Dan, K., 2009. Numerical study of ground-motion differences between buried-rupturing and surface-rupturing earthquakes, *Bull. seism. Soc. Am.*, **99**(3), 1521–1537.

Razafindrakoto, H.N.T. & Mai, P.M., 2014. Uncertainty in earthquake source imaging due to variations in source time function and earth structure, *Bull. seism. Soc. Am.*, **104**(2), 855–874.

Ripperger, J., Mai, P.M. & Ampuero, J.P., 2008. Variability of near-field ground motion from dynamic earthquake rupture simulations, *Bull. seism. Soc. Am.*, **98**(3), 1207–1228.

Sarao, A., Das, S. & Suhadolc, P., 1998. Effect of non-uniform station coverage on the inversion for earthquake rupture history for a Haskell-type source model, *J. Seismol.*, **2**(1), 1–25.

Schmedes, J., Archuleta, R.J. & Lavallée, D., 2010. Correlation of earthquake source parameters inferred from dynamic rupture simulations, *J. geophys. Res.*, **115**(B3), doi:10.1029/2009JB006689.

Sekiguchi, H. & Iwata, T., 2002. Rupture process of the 1999 Kocaeli, Turkey, earthquake estimated from strong-motion waveforms, *Bull. seism. Soc. Am.*, **92**(1), 300–311.

Sekiguchi, H., Irikura, K. & Iwata, T., 2000. Fault geometry at the rupture termination of the 1995 Hyogo-ken Nanbu earthquake, *Bull. seism. Soc. Am.*, **90**(1), 117–133.

Shao, G. & Ji, C., 2012. What the exercise of the SPICE source inversion validation BlindTest 1 did not tell you, *Geophys. J. Int.*, **189**(1), 569–590.

Song, S.G. & Dalguer, L.A., 2013. Importance of 1-point statistics in earthquake source modelling for ground motion simulation, *Geophys. J. Int.*, **192**(3), 1255–1270.

Song, S.G. & Somerville, P., 2010. Physics-based earthquake source characterization and modeling with geostatistics, *Bull. seism. Soc. Am.*, **100**(2), 482–496.

Spudich, P. & Archuleta, R.J., 1987. Techniques for earthquake ground motion calculation with applications to source parameterization of finite faults, *Seismic Strong Motion Synthetics*, pp. 205–265, ed. Bolt, B.R., Academic Press.

Spudich, P. & Xu, L., 2002. Documentation of Software Package Compsyn sxv3. 11: Programs for Earthquake Ground Motion Calculation Using Complete 1-D Green's Functions, Academic Press.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics.

Tarantola, A. & Valette, B., 1982. Inverse problems = quest for information, *J. Geophys.*, **50**(3), 159–170.

Tinti, E., Fukuyama, E., Piatanesi, A. & Cocco, M., 2005. A kinematic source-time function compatible with earthquake dynamics, *Bull. seism. Soc. Am.*, **95**(4), 1211–1223.

Toraldo Serra, E.M., Delouis, B., Emolo, A. & Zollo, A., 2013a. Combining strong-motion, InSAR and GPS data to refine the fault geometry and source kinematics of the 2011, Mw 6.2, Christchurch earthquake (New Zealand), *Geophys. J. Int.*, **194**(3), 1760–1777.

Toraldo Serra, E.M., Emolo, A., Orefice, A. & Zollo, A., 2013b. Earthquake source kinematics of moderate earthquakes from the inversion of apparent source time functions, *Geophys. J. Int.*, **194**(2), 1180–1189.

Trugman, D.T. & Dunham, E.M., 2014. A 2D Pseudodynamic rupture model generator for earthquakes on geometrically complex faults, *Bull. seism. Soc. Am.*, **104**(1), 95–112.

Wald, D.J. & Heaton, T.H., 1994. Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake, *Bull. seism. Soc. Am.*, **84**(3), 668–691.

Yagi, Y. & Fukahata, Y., 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**(1), 215–221.

Yoffe, E.H., 1951. LXXV. The moving griffith crack, *Philos. Mag.*, **42**(330), 739–750.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. (a) Trade-off curve of waveforms fitting error (abscissa) versus model roughness for the multistation case test M5. It shows five smoothing parameters a^2. (b) The slip distributions with increasing model smoothness from top to bottom. The selected model is highlighted by the red rectangle with $a^2 = 0.0005$.

Figure S2. Normalized rms three-component velocity waveforms fitting errors between the dynamic model and the estimated source models basing on various single station combinations. The waveform misfit is calculated for all 168 stations, although a single station is used in inversion.

Figure S3. Misfit reduction during the search for the multistations case test M10. The best-fitness function value for each generation versus generation number is shown. After about the 40th generation the misfit, reaches an approximately stationary level. (http://gji.oxfordjournals.org/lookup/suppl/doi:10.1093/gji/ggu252/-/DC1).