Geophys. J. Int. (2009) **176,** 135–150 doi: 10.1111/j.1365-246X.2008.03943.x

Bayesian imaging of the 2000 Western Tottori (Japan) earthquake through fitting of strong motion and GPS data

D. Monelli, P. M. Mai, S. Jónsson and D. Giardini

Institute of Geophysics, ETH Zurich, Schafmattstr. 30, CH-8093 *Zurich, Switzerland. E-mail: monelli@erdw.ethz.ch*

Accepted 2008 August 18. Received 2008 August 18; in original form 2008 June 3

SUMMARY

We image the rupture process of the 2000 Western Tottori earthquake ($M_w = 6.6$) through fitting of strong motion and GPS data. We consider an observational network consisting of 18 strong motion and 16 GPS stations, located within three fault lengths from the epicentre. We assume a planar fault and compute Green's functions for a 1-D velocity model. The earthquake rupture is described as a shear dislocation parameterized in terms of peak slip velocity, rake angle, rupture time and rise time, defined on a regular grid of nodes on the fault surface and derived at inner points through bilinear interpolation.

Our inversion procedure is based on a Bayesian approach. The solution of the inverse problem is stated in terms of a posterior probability density function (pdf), representing the conjunction of prior information with information contained in the data and in the physical law relating model parameters with data. Inferences on model parameters are thus expressed in terms of posterior marginal pdfs. Due to the non-linearity of the problem, we use a Markov Chain Monte Carlo (MCMC) method, based on the Metropolis algorithm, to compute posterior marginals.

Except for a few cases posterior marginals do not show a Gaussian-like distribution. This prevents us from providing a mean model and from characterizing uncertainties in terms of standard deviations only. Resolution on each single parameter is analysed by looking at the difference between prior and posterior marginal pdfs.

Posterior marginals indicate that the best resolved feature is a major slip patch (peak value of 311 ± 140 cm), located between the hypocentre and the top edge of the fault, centered at a depth of 4.5 km. This shallow slip patch is triggered about 3 s after the earthquake nucleated and required about 4 s to reach its final slip value. The presence of this shallow slip patch is common to all previous studies. In contrast to some previous studies, we do not identify any significant slip (>1 m) at the bottom of the fault.

We also compare inferences from both strong motion and GPS data with inferences derived from strong motion data only. In both cases the shallow slip patch is identified. At other locations, the main effect of the GPS data is in reducing the probability associated with high values of slip. GPS data reduce the presence of spurious fault slip and therefore strongly influence the resulting final seismic moment.

Key words: Inverse theory, Earthquake source observations.

1 INTRODUCTION

The $M_w = 6.6$ Tottori earthquake struck southwestern Japan on 2000 October 6, at 04:30:17.75 UTC. The hypocentre was located at 35.275°N, 133.350°E at a depth of 9.6 km (Fukuyama *et al.* 2003). The best-fitting double-couple focal mechanism estimated by Fukuyama *et al.* (2003) indicates an almost pure left-lateral strike-slip event with a strike-angle of 150° and a dip of 85° (Fig. 1). No clear surface rupture was observed near the epicentre, although some cracks oriented parallel to the estimated fault were found on a paved road (Umeda 2002). Systematic displacement of 10–20 cm was also found in the concrete lining in a tunnel 200 m below the surface near the source region.

To reveal the details of the earthquake rupture process, a number of studies derived kinematic images. Using a linearized frequency-domain method and an initial slip model obtained through GPS data inversion, Semmane *et al.* (2005) inverted strong motion data to infer values of slip amplitude, rupture time and rise time. They

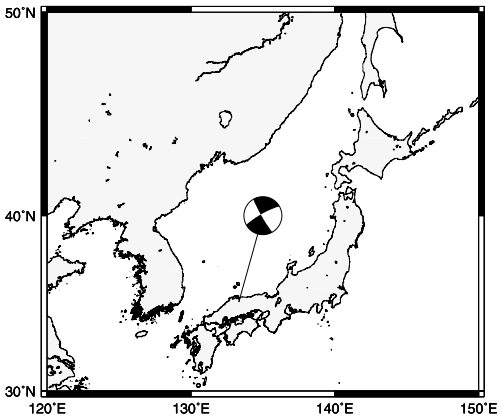


Figure 1. Location and focal mechanism for the 2000 Western Tottori earthquake (Fukuyama *et al.* 2003).

proposed different rupture models that all show a major slip patch located near the top edge of the fault (elongated towards southeast). Using strong motion data only and a backprojection method, Festa & Zollo (2006) inferred two major slip patches: one located above the hypocentre, close to the surface, extending southwards to the bottom of the fault and a second one located north of the hypocentre, at depths between 10 and 18 km. Fitting simultaneously strong motion and GPS data and using a direct search method based on a simulated annealing algorithm, Piatanesi *et al.* (2007) estimated peak slip-velocity, rise time, rupture time and rake angle. They confirm the presence of a major slip patch between the hypocentre and the surface, but also identify an additional slip patch (2–2.5 m) located at the bottom of the fault.

A dynamic model of the rupture process has also been derived for the Tottori earthquake. Assuming constant upper yield stress and uniform slip-weakening distance, and using a direct search method based on the neighbourhood algorithm, Peyrat & Olsen (2004) inferred the distribution of stress drop over the fault surface by fitting strong motion data. The resulting slip pattern again shows that most of the slip is concentrated in the uppermost part of the fault.

All these proposed images are similar in their general features—they all show a high slip patch near the surface. However, the presence of slip at the bottom of the fault is ambiguous: it has been recognized by Festa & Zollo (2006) and Piatanesi *et al.* (2007) but not by Semmane *et al.* (2005). Also the Peyrat & Olsen (2004) model does not require any slip at depth to fit the data, even though they consider a fault with a smaller depth extent compared with the ones used to obtain kinematic images.

One more aspect that has been investigated by both Semmane *et al.* (2005) and Piatanesi *et al.* (2007) is the rise time distribution on the fault surface. The model by Semmane *et al.* (2005) shows a highly heterogeneous pattern of rise time values that vary mostly between 0.5 and 2 s. Piatanesi *et al.*'s model shows a distribution that is instead more homogeneous (probably due to a coarser grid discretization and because they present a mean model) with rise time values varying mostly between 2.5 and 3.5 s. Clearly, these discrepancies can partially be due to the different approaches and parametrizations. However, no common features can be identified between the rise time distributions presented in these two studies, highlighting the intrinsic difficulty in imaging rise time in finite source inversions.

The Tottori earthquake is one of several examples, where multiple rupture models have been proposed to explain the observed data. All models are similar in some aspects but their obvious differences require a better understanding of where this variability comes from. Are these discrepancies in the source images only due to different approaches and modelling assumptions, or do they reveal some more fundamental lack of resolution?

Rupture-parameter estimates depend on how the inverse problem is stated, a well-known fact since the initial works of Olson & Apsel (1982) and Hartzell & Heaton (1983), who showed that results of linear slip inversions depend on the stabilization constraints and the data set used. More recently, considering the 2004 Parkfield earthquake, Custodio *et al.* (2005) analysed how kinematic rupture parameters depend on the chosen data set, while Hartzell *et al.* (2007) showed how source-inversion results may depend on the definition of the misfit function, the bounds on model parameters and the size of the model fault plane.

However, once a model parametrization, an inversion method and a data set are chosen, uncertainties on model parameters are determined by errors in data, modelling, and finite data coverage. All these factors influence the topology of the misfit function and therefore its minimum. Every minimum is characterized by a certain local topology, which determines the uncertainties on the corresponding model parameters. This is evident for the linear least-square problem, where the covariance matrix for model parameters is proportional to the inverse of the second derivative of the misfit function at the minimum (Menke 1989)—the sharper the minimum, the smaller the uncertainties. In case of non-linear problems, the minimization problem may even have multiple solutions, because the misfit function may have multiple (or degenerate) minima.

To estimate these uncertainties, some methods have been proposed. Emolo & Zollo (2005) used a genetic algorithm to search the model space and estimated resolution on the best-fitting model by defining a Gaussian probability density function, centred around it. For each model parameter they derived a marginal probability density function by computing the objective function in the neighbourhood of the best-fitting model, varying the parameter of interest but keeping all the remaining parameters fixed to their best-fitting values. With this approach, the posterior probability density function is forced to be Gaussian around the best-fitting model and, more importantly, the computed marginals do not take into account the correlation between different model parameters. Peyrat & Olsen (2004), Corish *et al.* (2007) and Piatanesi *et al.* (2007) derived uncertainty estimates by statistically analysing models generated by the optimization algorithm minimizing the misfit. The main drawback of this approach is that the statistical properties of a set of models produced by optimization do not necessarily represents the actual uncertainties (Sambridge 1999; Monelli & Mai 2008), but rather the tuning parameters and the operators adopted by the algorithm.

The aim of this paper is to investigate the rupture process of the Tottori earthquake, focusing on determining resolution on model parameters using a Bayesian approach (Mosegaard & Tarantola 1995; Tarantola 2005). A Bayesian approach allows one to estimate uncertainties, taking into account the non-linearity of the problem. It requires defining a posterior probability density function (pdf) on the model space, representing the conjunction of our prior information with information contained in the data (strong motion and GPS data in this case) and in the physical law relating model parameters with data. Inferences on model parameters are then expressed in terms of posterior marginal pdfs. Due to the non-linearity and large dimensionality of the problem, we use a Markov Chain Monte Carlo (MCMC) method, based on the Metropolis algorithm, to compute posterior marginals. Resolution on each model parameter is analysed by looking at the difference between the corresponding prior and posterior pdfs. With this approach, we can identify which

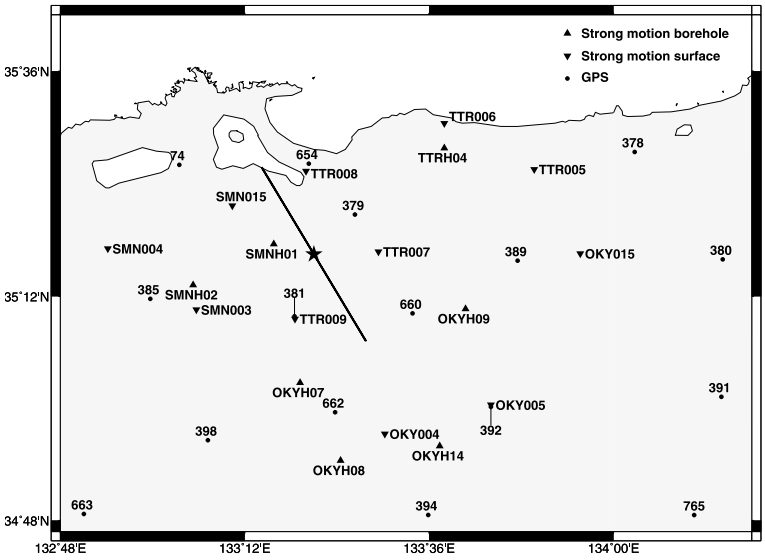


Figure 2. The observational network consists of 18 strong motion stations (black triangles) and 16 GPS stations (black dots) located within about 90 km from the epicentre (black star). We use seven borehole stations (upward-pointing triangles) and 11 surface stations (downward-pointing triangles). The black solid line indicates the assumed fault strike (150°).

regions of the fault surface are better illuminated by the data and which features of the rupture process can be considered well resolved.

2 THE OBSERVATIONAL NETWORK

The observational network we use consists of 18 strong motion and 16 GPS stations, located within about 90 km from the epicentre (Fig. 2). Among the strong motion stations, we use 11 K-net stations and 7 KiK-net borehole stations (SMNH01 and SMNH02 at 101 m depth, TTRH04 at 207 m depth, OKYH07, OKYH08, OKYH09, OKYH14 at 100 m depth).

The strong motion data (available at http://www.kik.bosai.go.jp/) come as raw accelerations with absolute time. We bandpass filter the waveforms in the frequency range 0.1–1 Hz using a first-order bandpass Butterworth filter, applied both in the forward and reverse directions to preserve phase. We then integrate the filtered waveforms to obtain ground velocities, which we resample to a sampling interval of 0.015 s. The horizontal components of station OKYH14 have been also rotated by 76° anticlockwise to correct for sensor misalignment. Each waveform lasts for 61.425 s and contains 4096 data points. Considering all components at all stations, the total number of waveform data points is therefore 221 184.

The GPS stations belong to the GEONET array, operated by the Geographical Survey Institute of Japan (Sagiya 2004). At each station, we define the coseismic static offset as the difference between the mean values of daily positions during the 5 days before and the 5 days after the earthquake. We also compute the corresponding standard deviations that we then propagate to compute the error on the final static displacement. For each station, we consider both the two horizontal components and the vertical component, resulting in a total number of GPS data points of 48.

3 THE FORWARD MODELLING

We adopt a 1-D piecewise-linear velocity–density–depth distribution, based on the velocity model used by Fukuyama *et al.* (2003) for the main shock location (Table 1). *S*-wave velocities are assumed to be $1/\sqrt{3}$ of the *P*-wave speed. Density values are deduced from *P*-wave velocities, using the Gardner's relationship (Gardner *et al.* 1974).

Table 1. Seismic velocity and density model for the Tottori region.

Depth (km)	V_P (km s^{-1})	V_S (km s^{-1})	ρ (g cm^{-3})
0.0	3.00	1.73	2.3
1.0	4.00	2.31	2.5
3.0	6.00	3.46	2.7
30.0	8.00	4.62	2.9

We represent the fault as a 40 km long and 20 km deep, vertically dipping plane surface, with a strike of 150°. The same strike and dip has been used by Peyrat & Olsen (2004), Festa & Zollo (2006) and Piatanesi *et al.* (2007). The fault's upper edge is at 0.5 km depth, because coseismic surface rupture was essentialy absent. On the fault surface, we define a regular grid of nodes,with a spacing of 4 km along-strike and along dip. The total number of nodes on the fault is therefore 66. At each node, we define four parameters: peak slip velocity, rise time, rupture time and rake angle.

We compute ground velocities using the frequency-domain representation theorem (Spudich & Archuleta 1987):

$$\dot{u}_m(\mathbf{y},\omega) = \int\int_{\Sigma} \dot{\mathbf{s}}(\mathbf{x},\omega)\cdot\mathbf{T^m}(\mathbf{x},\omega;\mathbf{y},\mathbf{0})\,\mathbf{d}\Sigma, \quad (1)$$

where $\dot{u}_m$ is the mth component of ground velocity at the receiver location $\mathbf{y}$, $\dot{\mathbf{s}}$ is the slip-velocity function, $\mathbf{T}^m$ is the traction exerted across the fault surface Σ at point $\mathbf{x}$, generated by an impulsive force applied in the mth direction at the receiver ($\omega = 2\pi f$: angular frequency). Tractions $\mathbf{T}^m$ are computed, up to a frequency of 1 Hz, using a Discrete Wavenumber/Finite Element method (Compsyn package, Spudich & Xu 2002), for a 1-D flat layered Earth model without attenuation. A trapezoidal-rule quadrature of the product $\dot{\mathbf{s}}\cdot\mathbf{T}^m$ is performed separately for each frequency, with the quadrature points being the sample points where $\mathbf{T}^m$ have been computed. Rupture-parameter values at integration points are derived through bilinear interpolation of values at surrounding grid nodes, similar to

the approach taken by Liu & Archuleta (2004) and Piatanesi *et al.* (2007).

In this study, we assume the slip-velocity function to be an isosceles triangle. With this parametrization, the peak slip velocity corresponds to the height of the triangle and the rise time, to the base length. Rupture time corresponds to the first point of the base segment. With this parametrization, rise and rupture times are nonlinearly related to ground velocity. Previous studies used different parametrizations, like a smooth ramp (Semmane *et al.* 2005) or a box-car function (Piatanesi *et al.* 2007).

Following eq. (1), we convolve tractions with the assumed slip-velocity function to compute ground velocity at the strong motion station locations. We compute GPS data predictions by integrating ground velocities to ground displacements and then selecting the final static offsets.

4 THE BAYESIAN APPROACH

In a Bayesian approach, inferences on model parameters (e.g. mean values, standard deviations, 1-D/2-D marginals) are derived from a posterior pdf defined on the model space. In Section 4.1, we introduce the general equations defining the posterior pdf. We then apply these equations to our specific case defining two different posteriors: one considering strong motion data only and the other considering both strong motion and GPS data. Our aim is to compare inferences from these two posteriors and analyse how GPS data influence the results. In Section 4.2, we define the model space. We pay special attention to defining a physically consistent model space to avoid considering unrealistic models. Finally, we present the numerical scheme used to derive inferences on the model parameters (Section 4.3).

4.1 The posterior pdf

We assume the M-dimensional model space and D-dimensional data space, $\mathbb{M}$ and $\mathbb{D}$, respectively, to be linear spaces. The prior probability density functions on model parameters and data are indicated with $\rho_M(\mathbf{m})$ and $\rho_D(\mathbf{d})$, respectively. $\theta(\mathbf{d}\,|\mathbf{m})$ denotes the conditional probability density representing the correlation between $\mathbf{d}$ and $\mathbf{m}$. The posterior pdf on the model space is given by (Tarantola 2005)

$$\sigma_M(\mathbf{m}) = k\rho_M(\mathbf{m})L(\mathbf{m}), \tag{2}$$

where k is a normalization constant and $L(\mathbf{m})$ is the likelihood function

$$L(\mathbf{m}) = \int_{\mathbb{D}} \mathrm{d}\mathbf{d}\ \rho_D(\mathbf{d})\theta(\mathbf{d}|\mathbf{m}), \tag{3}$$

which gives a measure of how well a model $\mathbf{m}$ explains the data.

In this study, we assume that our prior knowledge consists only of the information that each model parameter is strictly bounded by two values $m^\alpha_{\min}$ and $m^\alpha_{\max}$, where $\alpha \in I_M$, $I_M = \{1, \ldots, M\}$. We then write the prior pdf as

$$\rho(\mathbf{m}) = \prod_{\alpha_M \in I_M} \rho_\alpha(m^\alpha), \tag{4}$$

where

$$\rho_\alpha(m^\alpha) = \begin{cases} \frac{1}{m^\alpha_{\max} - m^\alpha_{\min}} & \text{for } m^\alpha_{\min} \le m^\alpha \le m^\alpha_{\max} \\ 0 & \text{otherwise} \end{cases}$$

is the prior marginal for each model parameter (i.e. a uniform probability density function [Monelli & Mai 2008]).

The common approach to define the likelihood function requires deriving a data covariance matrix for data uncertainties and a modelling covariance matrix for uncertainties in the forward modelling. Assuming Gaussian uncertainties the likelihood function takes a Gaussian functional form where the associated covariance matrix is the sum of the data and modelling covariance matrices (Gouveia & Scales 1998; Tarantola 2005).

Because we consider two different data sets (strong motion and GPS data), we define two distinct likelihood functions. For the strong motion data, we do not have a complete estimate of the associated uncertainties. Strong motion data represent a single measurement of the ground motion produced by an earthquake, and we therefore have a single realization of the data errors. A possible approach to still derive a data covariance matrix would be to analyse the portion before the P-wave arrival of each trace and to assume this portion to be representative of the seismic noise. More problematic is to derive the modelling covariance matrix, which would require knowing the uncertainties in the velocity and fault models (unknown in our case) and then propagating them into the Green's functions used to compute the predicted ground motion.

Due to the difficulty of deriving a realistic covariance matrix for strong motion data, we propose an alternative approach. First, we assume a 'perfect instrument' condition (Tarantola 2005). This assumption is valid if data uncertainties are negligible compared with modelling uncertainties. We propose this approach for the strong motion waveforms considered in this study, for which we find high signal-to-noise ratios thanks to the vicinity of the recording stations with respect to the source and the magnitude of the event. This assumption translates into the following condition:

$$\rho^{\mathrm{sm}}_D(\mathbf{d}) = \delta(\mathbf{d} - \mathbf{d}^{\mathrm{obs}}) \tag{5}$$

where $\rho^{\mathrm{sm}}_D(\mathbf{d})$ represents prior knowledge on strong motion data and $\mathbf{d}^{\mathrm{obs}}$ represents the observed data.

We define now the correlation $\theta(\mathbf{d} \mid \mathbf{m})$ between data and model parameters. Due to our lack of knowledge of the amplitude and type of uncertainites affecting our modelling, we cannot derive $\theta(\mathbf{d} \mid \mathbf{m})$ from a formal theory. We therefore propose an empirical formulation. Using an optimization algorithm, we examine which model produces the best fit, given the observed data. We then use this information to define a correlation function that assigns to each model $\mathbf{m}$, a correlation value that depends on how well it fits the data with respect to the level of fit produced by the best-fitting model. Models producing a level of fit close to the one of the best fitting models should then have a higher value of correlation than models producing a worse level of fit. Indicating with $\phi(\mathbf{d}, \mathbf{m})$ the percentage difference between the misfit produced by a model $\mathbf{m}$ and the misfit produced by the best-fitting model $\mathbf{m}^{\mathrm{best}}$ (which depends on the data $\mathbf{d}$), we obtain:

$$\phi(\mathbf{d}, \mathbf{m}) = \frac{S(\mathbf{m}) - S(\mathbf{m}^{\mathrm{best}}(\mathbf{d}))}{S(\mathbf{m}^{\mathrm{best}}(\mathbf{d}))} \cdot 100 \tag{6}$$

where S indicates the misfit function used and $\mathbf{m}^{\mathrm{best}}(\mathbf{d})$ represents the best-fitting model given data $\mathbf{d}$. We define the correlation between (strong motion) data and model parameter as

$$\theta^{\mathrm{sm}}(\mathbf{d}|\mathbf{m}) = \begin{cases} c & , \forall \mathbf{m} \in \mathbb{M} : \phi(\mathbf{d}, \mathbf{m}) < 0, \\ c\exp[-\phi(\mathbf{d}, \mathbf{m})] & , \forall \mathbf{m} \in \mathbb{M} : \phi(\mathbf{d}, \mathbf{m}) \ge 0, \end{cases} \tag{7}$$

where c is a normalization constant. Eq. (7) predicts that for all models producing a lower misit value than the best-fitting model, the correlation assumes its maximum value. This condition accounts for the possibility that the best-fitting model found during the optimization

process may not correspond to the absolute misfit minimum. For all other models, the value of the correlation decreases exponentially, depending on the percentage difference between the generated misfit and the minimum misfit associated with the best-fitting model. In writing eq. (7), we follow the analogy with a Guassian correlation function. When assuming Guassian modelling uncertainties, the correlation function $\theta(\mathbf{d}, \mathbf{m})$ assumes an exponential functional form where the argument is the L_2 norm of the data misfit weighted by the modelling covariance matrix. In our study we keep the exponential functional form, but we substitute the argument with eq. (6). Inserting eqs (5) and (7) into eq. (3), the integration yields

$$L^{\mathrm{sm}}(\mathbf{m}) = \begin{cases} c & , \forall \mathbf{m} \in \mathbb{M} : \phi(\mathbf{d}^{\mathrm{obs}}, \mathbf{m}) < 0, \\ c \exp[-\phi(\mathbf{d}^{\mathrm{obs}}, \mathbf{m})] & , \forall \mathbf{m} \in \mathbb{M} : \phi(\mathbf{d}^{\mathrm{obs}}, \mathbf{m}) \geq 0, \end{cases} \tag{8}$$

where $L^{\mathrm{sm}}(\mathbf{m})$ represents the likelihood function for strong motion data.

For GPS data, we define a data covariance matrix. As described in Section 2, we define the observed static offset as the difference between the mean values of daily positions during the 5 days before and after the earthquake. By computing the corresponding standard deviations, we can deduce the standard deviation on the final static displacement. Assuming uncorrelated uncertainties, we then define a covariance matrix for GPS data, which is a diagonal matrix of data variances. Assuming Gaussian uncertainties, the prior pdf on (GPS) data is

$$\rho_D^{\mathrm{gps}}(\mathbf{d}) = C \exp\left[-\frac{1}{2}(\mathbf{d} - \mathbf{d}^{\mathrm{obs}})^T \mathbf{C}_{d,\mathrm{gps}}^{-1}(\mathbf{d} - \mathbf{d}^{\mathrm{obs}})\right] \tag{9}$$

where C is a normalization constant and $\mathbf{C}_{d,\mathrm{gps}}$ is the data covariance matrix for GPS data.

As for the strong motion data, the modelling covariance matrix for uncertainties in the predicted GPS displacement requires knowing the uncertainties in the velocity and fault models. However, GPS data, measuring a static offset, reflect the zero frequency component of the wavefield, which is less sensitive to complexities in the velocity model. Also, GPS data seems to be well explained even using a simple planar fault (Piatanesi *et al.* 2007). We hence assume for GPS data to have neglible uncertainties in the forward modelling. This assumption translates into the following condition:

$$\theta^{\mathrm{gps}}(\mathbf{d}|\mathbf{m}) = \delta(\mathbf{d} - \mathbf{g}(\mathbf{m})), \tag{10}$$

where $\mathbf{g}(\mathbf{m})$ is the forward modelling operator. Inserting eqs (9) and (10) into eq. (3), the result of the integration is

$$L^{\mathrm{gps}}(\mathbf{m}) = C \exp\left[-\frac{1}{2}\mathbf{r}^T \mathbf{C}_{d,\mathrm{gps}}^{-1}\mathbf{r}\right], \tag{11}$$

where $L^{gps}(\mathbf{m})$ represents the likelihood function for GPS data and $\mathbf{r} = \mathbf{g}(\mathbf{m}) - \mathbf{d}^{\mathrm{obs}}$.

Considering eq. (2) we now define a posterior pdf representing the conjunction of our prior information with information contained in strong motion data

$$\sigma_M^{\mathrm{sm}}(\mathbf{m}) = k\rho_M(\mathbf{m})L^{\mathrm{sm}}(\mathbf{m}). \tag{12}$$

Eq. (12) can then be used as prior information to define a new posterior pdf for the model parameters, which also considers the GPS data:

$$\sigma_M^{\mathrm{sm,gps}}(\mathbf{m}) = k\rho_M(\mathbf{m})L^{sm}(\mathbf{m})L^{gps}(\mathbf{m}). \tag{13}$$

4.2 The model space

The posterior pdf is defined over the model space. Inferences on model parameters are therefore dependent on the chosen model space. A correct definition of the model space is of vital importance to avoid testing unrealistic models that make the inference process inefficent. We thus pay special attention to defining a physically consistent model space.

In our inversion, we assume the peak slip velocity (and therefore the slip) to be zero at the fault edges. Non-zero slip at the fault boundaries would constitute a discontinuity in slip that lead to unrealistically high values of stress change at the edges. This condition is assumed to be valid also for the top edge of the fault, because no surface rupture was reported for the Tottori earthquake. For the inner nodes, the peak slip velocity is allowed to vary between 0 and 400 cm s^{-1}. With these conditions, we generate peak slip velocity distributions, with non-zero values only inside the fault and tapered to zero at the edges.

The moment tensor solution for the Tottori earthquake indicates an almost pure left-lateral strike-slip event (Fukuyama *et al.* 2003); nevertheless we allow the rake angle to vary between -30° and $+30^\circ$ at each node. Positive angles indicate a downdip component whereas negative angles an updip component.

The range of rupture times at each grid node is defined as the time interval between the arrival times of two circular rupture fronts, propagating from the hypocentre (at 9.6 km depth) at two limiting rupture velocities: 1.5 and 4 km s^{-1}.

The range of possible values for rise time has been chosen according to the frequency band used in the inversion. Having bandpass filtered the waveforms in the frequency band 0.1–1 Hz, we consider 1 and 10 s as minimum and maximum values for rise time, respectively. However, from dynamic rupture simulations (Day 1982; Madariaga *et al.* 1998), it is known that when a rupture front reaches the unbreakable boundaries of a fault, it generates stopping phases that propagate inwardly and heal the slip process as they spread over the fault. As a consequence, the duration of slip at fault locations is influenced by the stopping phases emitted from the edges of the fault. In our case, the hypocentre is located approximately in the centre of the assumed fault plane; we may therefore expect that the inner portion of the fault will start slipping earlier and will be reached by the stopping phases, later than regions near the borders of the fault. For this reason, the minimum allowed rise time is assumed to be 1 s for each node, whereas the maximum allowed rise time is assumed to decrease from the maximum value (10 s), according to the following equation:

$$\tau_r^{\max,n} = \tau_r^{\min} + \left(\tau_r^{\max} - \tau_r^{\min}\right)\left(1 - \frac{d_{\mathrm{hyp}}^n}{d_{\mathrm{hyp}}^n + d_{\mathrm{bound}}^n}\right), \tag{14}$$

where $\tau_r^{\max,n}$ is the maximum rise time at the node n, $\tau_r^{\min}$ and $\tau_r^{\max}$ are the minimum and maximum rise time values allowed by the considered frequency range, respectively, d_{hyp}^n is the distance of the node n from the hypocentre and d_{bound}^n is the minimum distance of the node n from the boundaries of the fault. This equation predicts that the maximum allowed rise time is equal to 10 s, only for a node at the hypocentre ($d_{hyp}^n = 0$) and that for all the nodes on the boundaries ($d^n{}_{\mathrm{bound}} = 0$), the maximum rise time corresponds to the minimum allowed rise time. For all remaining nodes, the maximum rise time decreases as their distance from the boundary decreases (Fig. 13). For the nodes having the same minimum distance (e.g. nodes 14, 15, 16, 17), the maximum allowed rise time decreases with increasing distance from the hypocentre. Eq. (14) only predicts the maximum allowed rise time at each node and expresses the fact that long rise

times are not expected near the borders of the fault, simply because stopping phases are expected to reduce the duration of the slip process in these locations. The minimum rise time is 1 s everywhere. Between the minimum and maximum allowed rise time values, the prior pdf assumes uniform probability at each node. In other words, a crack-like rupture behaviour or a pulse-like propagation are assumed to be equally likely.

4.3 Sampling the posterior pdf

Once the posterior pdf and the model space are defined, information on each model parameter m^α can be quantified by computing the corresponding 1-D marginal posterior pdf:

$$M(m^\alpha) = \int \dots \int \sigma_M(\mathbf{m}) \prod_{k=1k\neq\alpha}^{M} \mathrm{d}m^k \tag{15}$$

Eq. (15) involves computing the integral of the posterior pdf over all dimensions of the model space, except the one corresponding to the parameter of interest. Due to the large dimensionality of the problem (204 model parameters) Eq (15) can be estimated only using Monte Carlo methods that generate models $\mathbf{m}$ as samples of the posterior pdf $\sigma_M(\mathbf{m})$. Once a large ensemble of such samples has been generated, the 1-D marginal of each parameter can be approximated by the histogram of the corresponding sampled values.

Among the different possible sampling algorithms (for a review see Tarantola 2005), we use a Markov Chain Monte Carlo (MCMC) method, based on the Metropolis algorithm (Martinez & Martinez 2002; Tarantola 2005). A Markov chain is a sequence of random variables $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_t$, such that the next value or state of the sequence $\mathbf{m}_{t+1}$ depends only on the previous one $\mathbf{m}_t$. An MCMC method, based on the Metropolis algorithm, generates a Markov chain where the state of the chain at $t+1$ is obtained by sampling a 'candidate point' $\tilde{\mathbf{m}}$ from a symmetric proposal distribution $q(.|\mathbf{m}_t)$. An example of a distribution like this is the normal distribution with mean $\mathbf{m}_t$ and fixed covariance. To generate variables that are samples of a given pdf P, the candidate point is accepted as the next state of the chain with a probability given by:

$$\alpha(\mathbf{m}_t, \tilde{\mathbf{m}}) = min\left\{1, \frac{P(\tilde{\mathbf{m}})}{P(\mathbf{m}_t)}\right\}. \tag{16}$$

This means that if $P(\tilde{\mathbf{m}}) \geq P(\mathbf{m}_t)$, that is, if the proposed model corresponds to an higher value of the target pdf, the move is always accepted because $\alpha(\mathbf{m}_t, \tilde{\mathbf{m}})$ will be equal to one. In the opposite case, if the move produces a lower value of the target pdf, the proposed model is accepted with probability given by $P(\tilde{\mathbf{m}})/P(\mathbf{m}_t)$. If the point $\tilde{\mathbf{m}}$ is not accepted, then the chain does not progress and $\mathbf{m}_{t+1} = \mathbf{m}_t$.

Our aim is to generate models that are samples of the posterior pdf. In our case the posterior pdf is given by the product of several pdfs [in case of $\sigma_M^{\mathrm{sm,gps}}(\mathbf{m})$, the prior and the likelihoods for strong motion and GPS data]. Using a general notation we write

$$\sigma_M(\mathbf{m}) = kP_1(\mathbf{m})P_2(\mathbf{m})P_3(\mathbf{m}). \tag{17}$$

To generate samples according to the posterior defined in eq. (17), we use the Cascaded Metropolis algorithm (Tarantola 2005). We start by defining a random walk that generates samples according to the first pdf. At a given step, the random walker is at point $\mathbf{m}_t$ (which is a sample of P_1). Using a proposal distribution, we generate a model $\tilde{\mathbf{m}}$. We accept the new model as a next step of the random walk, according to the following rules.

(a) If $P_2(\tilde{\mathbf{m}}) \geq P_2(\mathbf{m}_t)$, then go to step (c).

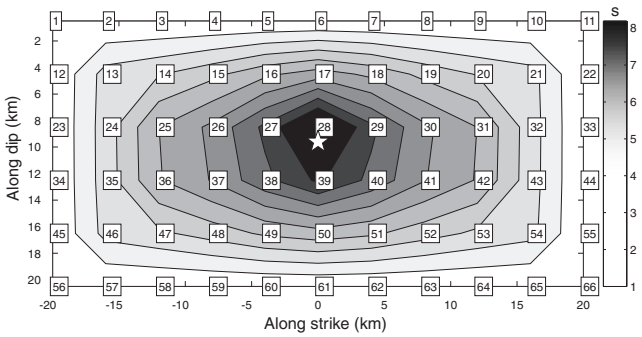


Figure 3. The maximum allowed rise time (s) on the fault surface. Numbered labels indicate node locations. The white star represents the hypocentre location.

(b) If $P_2(\tilde{\mathbf{m}}) < P_2(\mathbf{m}_t)$, then decide randomly to go to step (c) or to reject the proposed model, with a probability to go to step (c) given by $\alpha = P_2(\tilde{\mathbf{m}})/P_2(\mathbf{m}_t)$.

(c) If $P_3(\tilde{\mathbf{m}}) \geq P_3(\mathbf{m}_t)$, then accept the new model.

(d) If $P_3(\tilde{\mathbf{m}}) < P_3(\mathbf{m}_t)$, then decide randomly to accept a new model or to stay at $\mathbf{m}_t$, with a probability to accept the new model given by $\alpha = P_3(\tilde{\mathbf{m}})/P_3(\mathbf{m}_t)$.

5 RESULTS

In Section 4.1, we defined the posterior pdfs $\sigma_M^{\mathrm{sm}}(\mathbf{m})$ (eq. 12) for strong motion data and $\sigma_M^{\mathrm{sm,gps}}(\mathbf{m})$ (eq. 13) for both strong motion and GPS data. We now present the corresponding estimated maximum likelihood models, and compare their predictions with the observed data (Section 5.1). Then we compute the corresponding 1-D marginals and analyse how GPS data change inference results (Section 5.2). In Section 5.3, we finally compute 2-D marginals for a number of model parameters and investigate possible correlations.

5.1 The maximum likelihood models

The maximum likelihood model for $\sigma_M^{\mathrm{sm}}(\mathbf{m})$ corresponds to the model maximizing the likelihood function $L^{\mathrm{sm}}(\mathbf{m})$. By definition (eqs 6 and 7) the maximum is attained in correspondence with the best-fitting model. Given a model $\mathbf{m}$, we measure the level of fit with strong motion data, using a L2 norm of the misfit between observed and predicted waveforms in the time domain. We explore the model space to identify the best-fitting model, using a direct search method based on an evolutionary algorithm (Beyer 2001; Monelli & Mai 2008). An evolutionary algorithm is a population-based stochastic optimization method. According to this algorithm, the search of the model space starts with generating an initial set of models, which is obtained through a uniform random sampling of the model space. This initial population then evolves through the subsequent application of both stochastic and deterministic operators. Goal of these operators is to generate a new population of models that hopefully show better properties (i.e. lower misfit values). The creation of a new population is referred as a new generation.

We consider an initial population of 100 models from which we produce, at each generation, 2000 new models. The search lasts for 100 generations and the total number of models produced is therefore 200 100. The best objective function value for each generation versus the generation number is shown in Fig. 4. We can see that after the 40th generation, the level of fit reaches an approximately

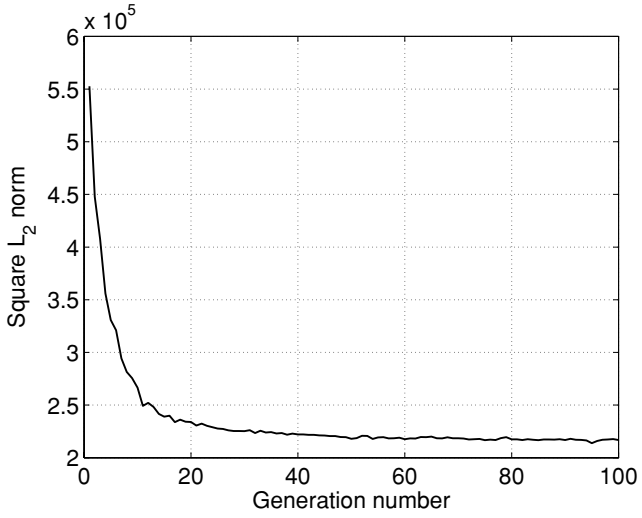


Figure 4. Misfit reduction during the search. After about the 40th generation, the level of fit reaches an approximately stationary level.

stationary level. The best-fitting model (generating the lowest misfit function value) is shown in Fig. 5 (left-hand panels).

The maximum likelihood model for $\sigma_M^{\mathrm{sm,gps}}$ (Fig. 5, right panels) corresponds to the model minimizing the sum of the exponents of the two likelihood functions, $L^{\mathrm{sm}}(\mathbf{m})$ and $L^{\mathrm{gps}}(\mathbf{m})$. We identify it among the models visited during the sampling process which we describe in detail in Section 5.2.

Comparing the two rupture models, we can see that both of them present several high slip-velocity patches. In both cases, we can identify a high slip-velocity patch between the hypocentre and the top edge of the fault (at 4.5 km depth). The maximum likelihood model for σ_M^{sm} presents significant peak slip velocity SE of the hypocentre, which is not observed in the maximum likelihood model for $\sigma_M^{\mathrm{sm,gps}}$. The latter presents also a low peak slip velocity region, NW of the hypocentre, which is also visible, but less extensive, in the maximum likelihood model for σ_M^{sm}.

In both cases, the rise time pattern shows higher values near the hypocentre and lower values near the borders, following approximately the pattern of the maximum allowed rise time.

In comparing the final slip distributions, we note in both cases a high slip patch (maximum value about 4 m) between the hypocentre and the top edge of the fault, with an elongation of the slip distribution towards SE. The major difference concerns the presence of deep slip. The maximum likelihood model for $\sigma_M^{\mathrm{sm,gps}}$ presents little slip at the bottom of the fault, especially in the NW, whereas the maximum likelihood model for σ_M^{sm} contains, instead, more deep slip.

The seismic moments of the maximum likelihood models for σ_M^{sm} and $\sigma_M^{\mathrm{sm,gps}}$ are 1.9×10^{19} and 1.6×10^{19} N m, respectively. Semmane *et al.* (2005) inferred values of seismic moment of

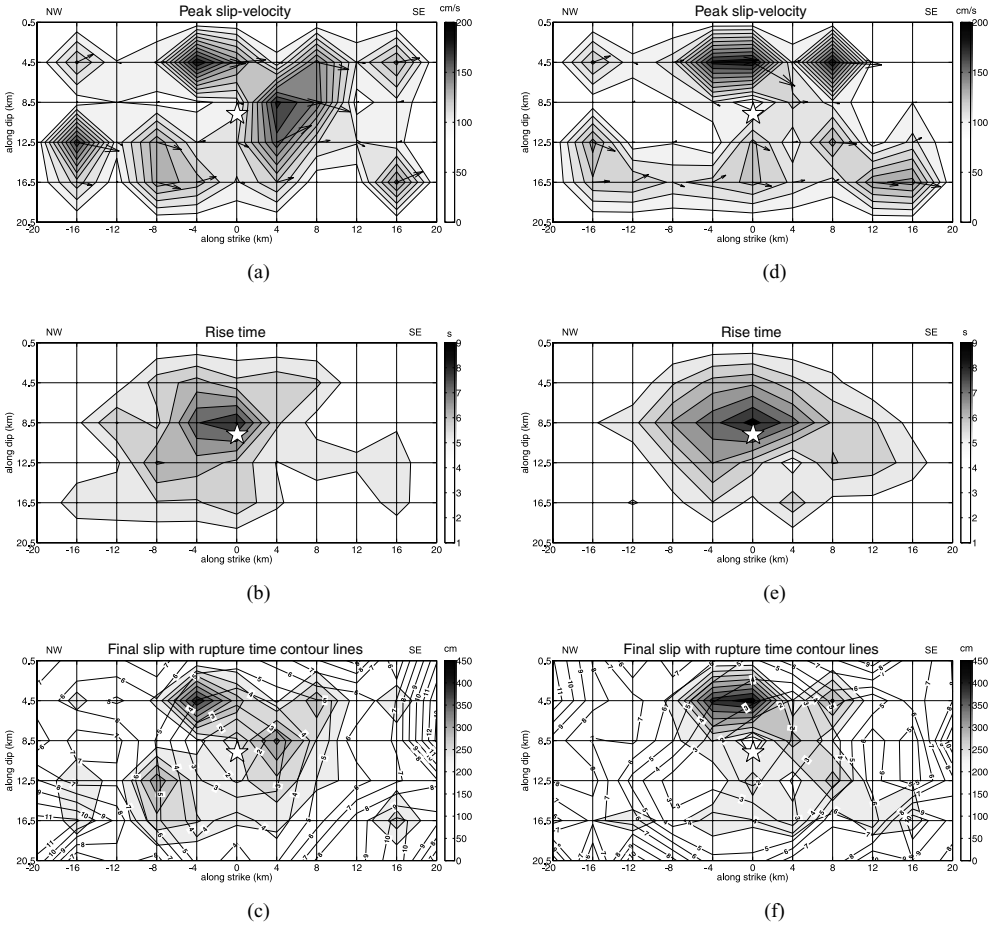


Figure 5. Peak slip velocity vector (cm s^{-1}), rise time (s) and final slip (cm) distributions (with rupture time contour lines every 1 s) of the maximum likelihood models for σ_M^{sm} (panels a, b, c, respectively) and for $\sigma_M^{\mathrm{sm,gps}}$ (panels d, e, f, respectively). The grid indicates the subfault discretization. The white star represents the hypocentre location.

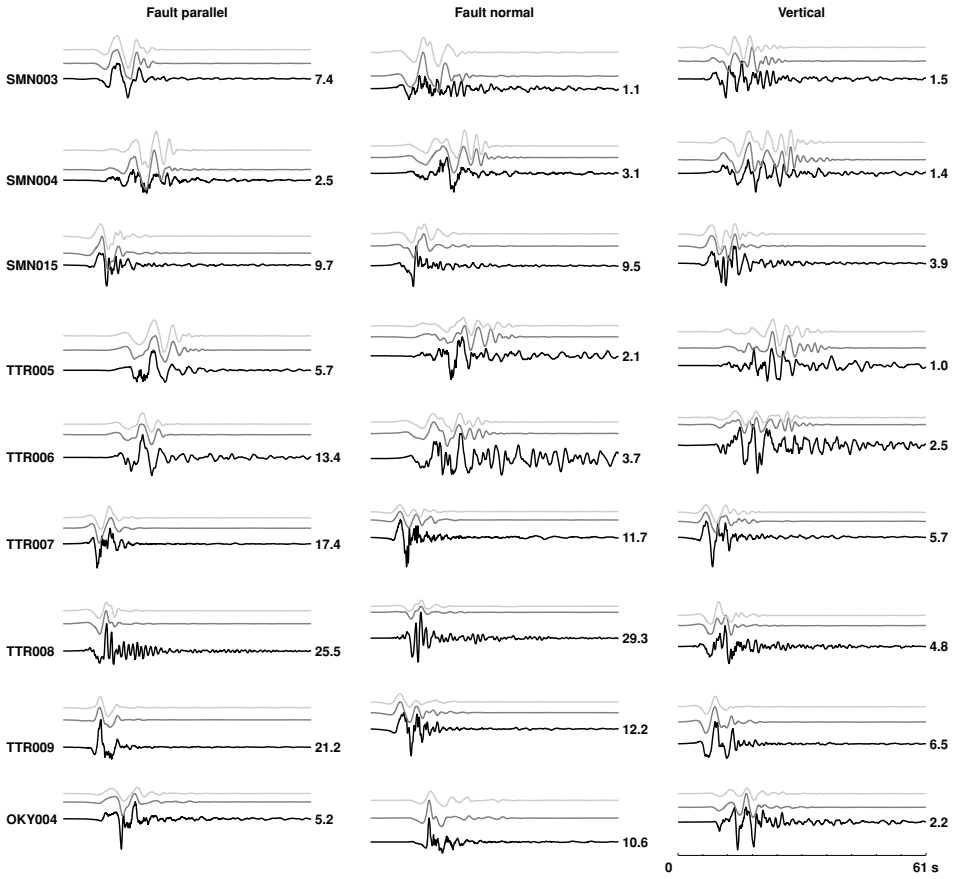


Figure 6. Level of fit produced by the maximum likelihood models for σ_M^{sm} (dark grey) and $\sigma_M^{\mathrm{sm,gps}}$ (light grey) with the observed ground motion (black). For each waveform the maximum observed ground velocity (cm s^{-1}) is shown at the end of each trace. Waveforms are not normalized. For each component, the vertical spacing between two subsequent traces is equal to the maximum positive amplitude of the lower trace.

1.5– 1.7 $\times$ 10^{19} N m, Festa & Zollo (2006), 2.6 $\times$ 10^{19} N m and Piatanesi *et al.* (2007) 1.7 $\times$ 10^{19} N m.

In Figs 6 and 7, we show the level of fit produced by both models with the observed strong motion data. For some components, both models reproduce the polarity of the first arrival and the amplitude and duration of the main phase (see fault parallel component at stations SMN003, SMN015, TTR005, SMNH01, SMNH02, TTRH04 for instance). For some other components, the forward modelling does not reproduce the observed complexity (see waveforms at station TTR008 for instance). Both models produce a similar level of fit. Without any uncertainty characterization, we cannot say which model is performing better in reproducing the observed strong motion data.

In Fig. 8, we compare the horizontal static displacement produced by both models with the one deduced from GPS data. Ellipses represent 95 per cent confidence level. We note that at some stations (74, 379, 660, 662, 381), the static displacement produced by the maximum likelihood model for σ_M^{sm} lies just on or is slightly outside the error ellipse. The maximum likelihood model for $\sigma_M^{\mathrm{sm,gps}}$, instead, reproduces the observed surface displacements within the estimated displacement error at all stations.

5.2 The 1-D marginals

According to Section 4.3, we express our inferences on the investigated rupture parameters in terms of marginal pdfs derived from the two posterior pdfs defined in eqs (12) and (13).

Following the algorithm described in Section 4.3, we simulated, for both cases, four random walks starting from different models obtained through uniform random sampling of the model space. Each random walk has a different seed value for the random number generator also. At each step, we generate a new model using a Gaussian probability distribution with fixed covariance matrix. We assume the covariance matrix to be diagonal, with standard deviations equal for parameters of the same type. After several trial random walks, we fix the standard deviations for peak slip velocity, rake angle, rupture time and rise time to be 5 cm s^{-1}, 2°, 0.1 s and 0.1 s, respectively. With these values, the acceptance rate of the Metropolis algorithm (ratio between accepted and generated models) is ∼50 per cent when sampling σ_M^{sm} and ∼30 per cent when sampling $\sigma_M^{\mathrm{sm,gps}}$. Tarantola (2005) suggests that the size of the perturbations in the model space should give an acceptance rate of ∼30–50 per cent.

Models produced by the Metropolis sampler are not independent samples of the posterior pdf, since each model depends on the previous one. However, the estimation of the integral in equation (15) requires independent samples. Only with n independent samples can eq. (15) be approximated as accurately as needed by increasing n (Martinez & Martinez 2002). After taking one sample, a possible strategy to generate a new independent sample is to wait a sufficient number of moves before collecting a new sample, such that the random walk has 'forgotten' the previous sample. Unfortunately, no general rule exists that helps to set the number of moves that should be done before collecting a new sample (Tarantola 2005).

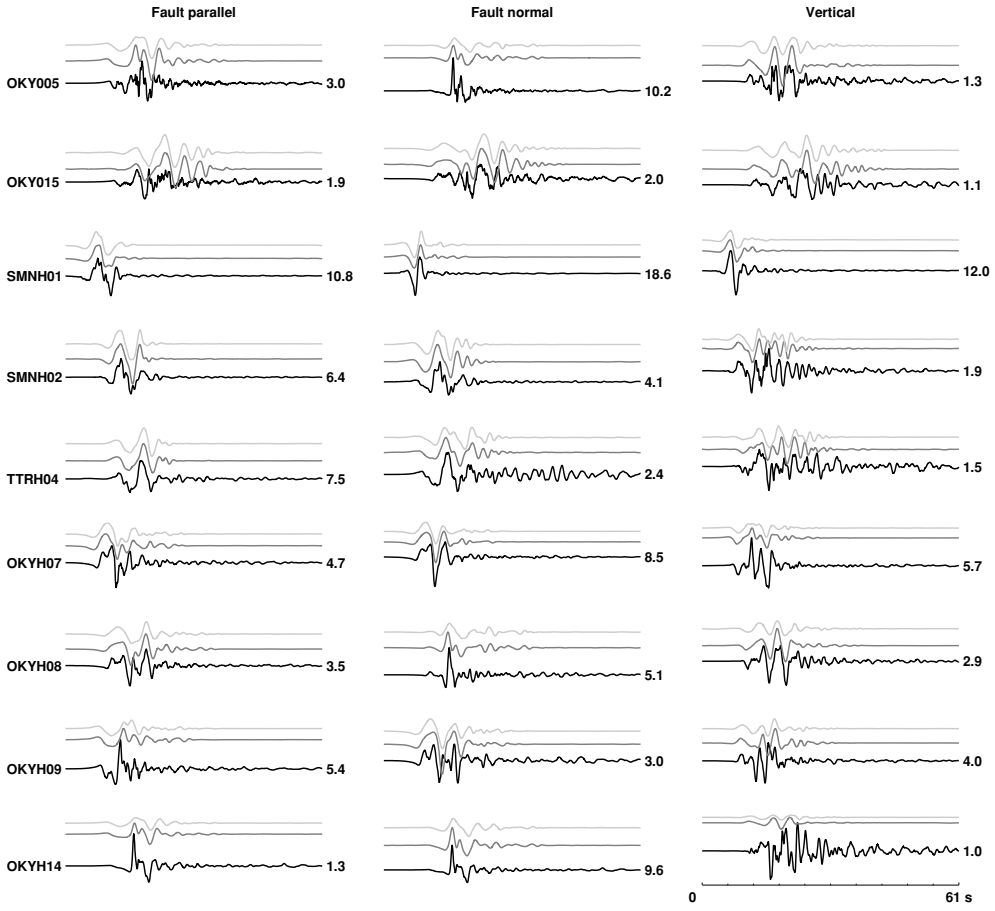


Figure 7. Continuation of Fig. 6.

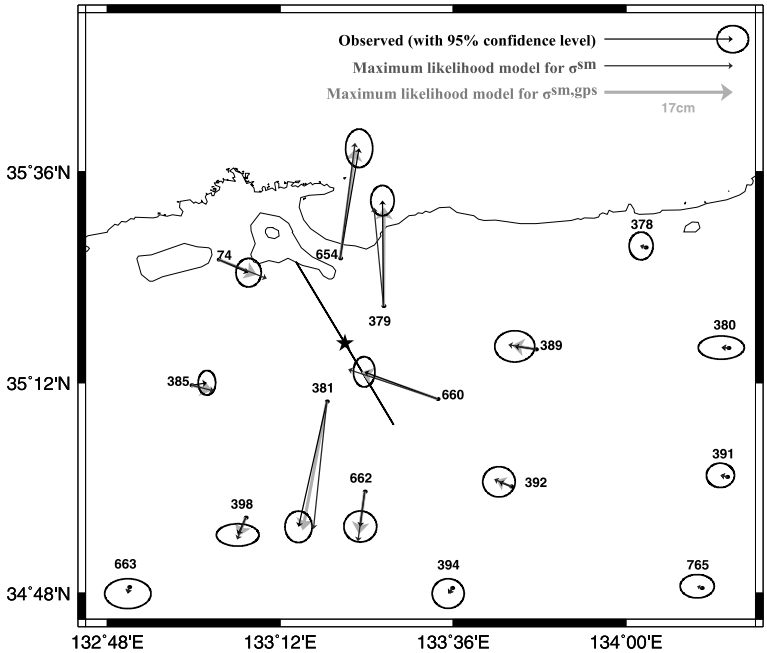


Figure 8. Horizontal static displacement predicted by the maximum likelihood models for σ_M^{sm} (thin dark grey) and $\sigma_M^{\mathrm{sm,gps}}$ (thick light grey) compared with the observed one (thin black). Ellipses represent 95 per cent confidence levels.

From a practical point of view, this parameter is also dependent on the computation time available. After some experimentation we decided to collect samples every 100 steps.

To generate samples according to σ_M^{sm}, we ran each random walk for 1 000 000 steps and collected samples every 100 moves. Each random walk produced therefore 10 000 approximately independent samples. We ran the four random walks in parallel, each of them requiring a single processor. The computation time needed was ∼40 days on a Linux cluster, based on AMD Opteron 64-bit CPUs. We then merged all ensembles produced by the different random walks into a single ensemble, which we finally used to estimate marginals.

To generate samples according to $\sigma_M^{\mathrm{sm,gps}}$, the sampling algorithm requires solving the forward modelling for the GPS data prediction. With this additional calculation, each random walk, produced 300 000 models in approximately the same computation time (∼35 days). From each random walk we extracted 3000 approximately independent samples that we then merged to estimate the corresponding marginals. Even with a smaller number of samples, we observed that each single random walk was able to produce approximately the same marginal, indicating therefore an acceptable convergence.

In Fig. 9(a), we present 1-D marginals for peak slip velocity at grid points, displaying only inner grid points, because on the fault plane boundaries, peak slip velocity is assumed to be zero (Section 4.2). For each node, we present the 1-D prior marginal, the posterior obtained from σ_M^{sm} and the one from $\sigma_M^{\mathrm{sm,gps}}$.

The most important feature to note is that the posteriors generally do not show a Gaussian shape, rather a skewed distribution. The only two posterior marginals with a Gaussian-like distribution corresponds to nodes number 16 and 17. For these two nodes, the posteriors from $\sigma_M^{\mathrm{sm,gps}}$ predict a peak slip velocity of 122 ± 57 and 140 ± 57 cm s^{-1}, respectively. The relative error for these two nodes is about 47 and 41 per cent, respectively. These two posteriors confirm the presence of a near-surface high slip-velocity patch as imaged in the maximum likelihood models (Fig. 5).

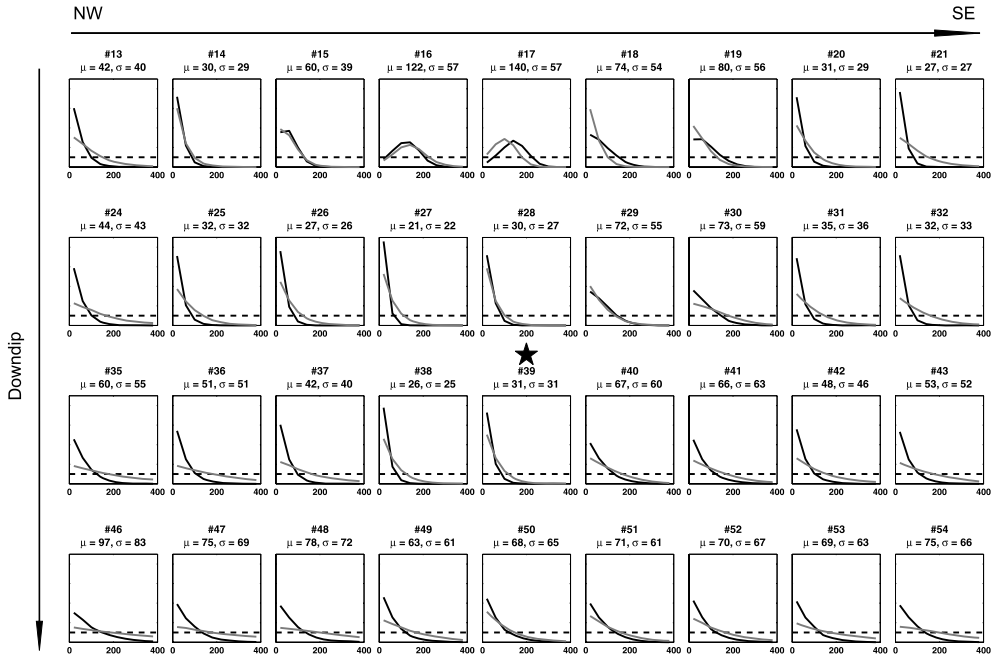
For all the remaining nodes, posteriors show a distribution skewed towards the minimum allowed peak slip velocity value (0 cm s^{-1}). Note that the skeweness depends on the node location. As a general trend, we find that the skeweness, and therefore the posterior peak, become less clear from the top edge of the fault towards the bottom (see subplots along the columns). This is particularly evident for posteriors from strong motion data only. This implies that the resolution power of the data sets (measured at each node by the difference between posterior and prior pdfs) follows the same trend and decreases with increasing depth.

Comparing posterior marginals obtained from σ_M^{sm} and $\sigma_M^{\mathrm{sm,gps}}$, we find that GPS data have a notable effect in constraining the peak slip velocity distribution. In fact, GPS data are sensitive to the final slip distribution. In our modelling, the final slip at each fault location is directly proportional to peak slip-velocity [assuming an isosceles triangle as source time function, final slip = (peak slip-velocity × rise time)/2]. Looking at nodes 16 and 17, we see that GPS data confirm the presence of a high slip-velocity patch. However, for node 17, GPS data suggest an even higher value of peak slip-velocity with respect to the one inferred using strong motion data only. In most of the remaining locations, GPS data have an effect in reducing the tail of the posteriors obtained from σ_M^{sm}. This is evident at nodes 21, 32, 43, for instance. It is also interesting to note that the GPS data used in this study have the same effect at the bottom of the fault (see nodes number 46, 47, 48). This shows that, at least in this case, GPS data bring useful information on the rupture process also for the deeper part of the fault.

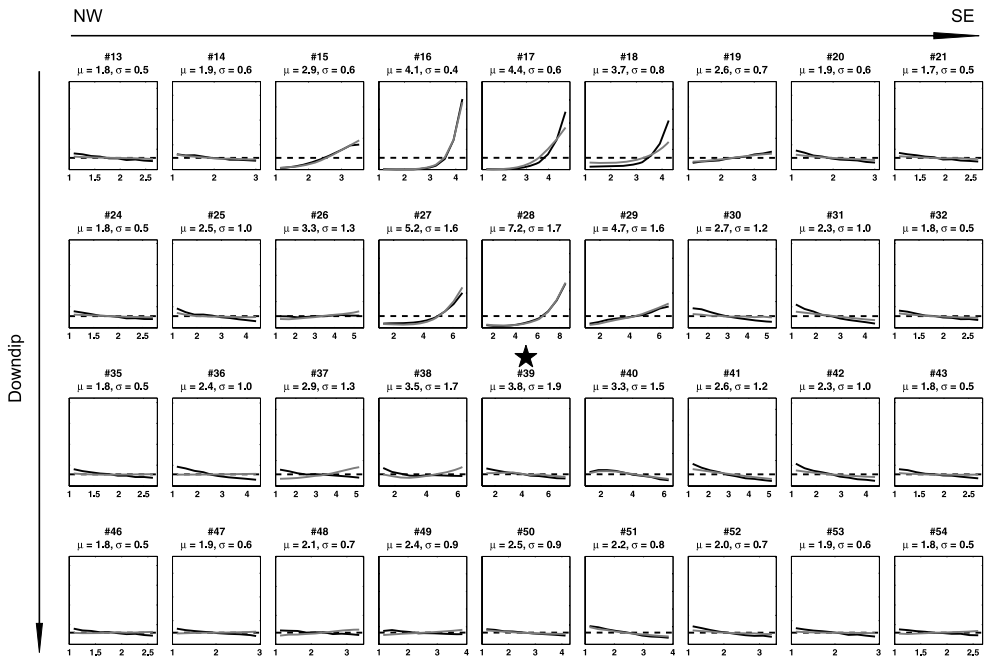
We show 1-D marginals for rise time in Fig. 9(b): the posterior marginals present a well-defined peak only for the nodes located near the high slip-velocity patch (nodes 15, 16, 17, 18 and 27, 28, 29). For all remaining nodes, posterior marginals present very little difference with respect to the prior uniform, indicating therefore very poor resolution for rise time. At node 17, corresponding to the highest inferred peak slip velocity value, the mean rise time is about 4.4 s. We also note that the maximum estimated mean rise time (7.2 s) corresponds to node 28, which is associated with low peak slip velocity values (see corresponding posterior in Fig. 9a). We could expect to have little resolution on rise time for a node associated with low slip-velocity. However, we recall that rupture parameters are defined on a coarse grid on the fault surface and then derived on a finer grid (where the actual integration is carried out) through bilinear interpolation. Even if a node is associated with a low value of peak slip velocity, its vicinity may not have low values if a neighbouring node is associated with an high value of peak slip velocity. Node 17, where the highest value of peak slip velocity is inferred, is a neighbouring node of node 28. This means that between these two nodes, significant peak slip-velocity may be present. In that case, the long rise time corresponding to node 28 is needed to describe the slip process in its neighbourhood. When comparing posteriors from σ_M^{sm} and $\sigma_M^{\mathrm{sm,gps}}$, we note the greatest differences only at nodes 17 and 18. For these nodes, GPS data increase the probability associated with larger values of rise time.

In Fig. 10(a) we show 1-D marginals for rupture time. In this case we consider also nodes located at the edges of the fault. Marginals from σ_M^{sm} and $\sigma_M^{\mathrm{sm,gps}}$ are very similar, since GPS data do not contribute information about rupture timing. Again, we find that posteriors present a well defined peak with respect to the prior marginals, only in the upper-most part of the fault (especially at nodes 4, 5, 6 and 15, 16, 17). Nodes 16 and 17 correspond to the nodes where the shallow high slip-velocity patch is located. Assuming mean values as estimates of the actual rupture times, the rupture front triggers the high slip-velocity patch located below the top edge of the fault (nodes 17) approximately 3.1 s after the rupture initiated. The average rupture velocity from the hypocentre in the updip direction is therefore 1.6 km s^{-1}, corresponding to 44 per cent of the average shear velocity in the involved depth range. For some nodes located on the boundary of the fault (4, 5 and 6 especially) the posterior pdfs show a clear peak, although for these nodes, the peak slip velocity is assumed to be zero. The fact that data are sensitive to these parameters, is an effect of the bilinear interpolation scheme. Even if these parameters correspond to nodes where the peak slip velocity is assumed to be zero, the rupture time defined on these nodes determines the rupture time in the neighbourhood points. Hence, if these neighbourhood points are associated with well-resolved slip, the rupture time in the neighbourhood nodes will also be well resolved.

Comparing 1-D marginals for the rake angle (Fig. 10b), with marginals for peak slip velocity, rise time (Fig. 9) and rupture time [Fig. 10(a)] we find that the rake angle is the least resolved parameter in the considered model space. Differences between priors and posteriors are generally less accentuated than for the other parameters. We also observe that GPS data have a notable effect in constraining the rake angle at some locations. This is evident at nodes 16, 17, 27, 28. In these locations, posterior marginals suggest that the high slip-velocity patch is associated with a positive rake angle, which implies a downdip movement in our modelling.



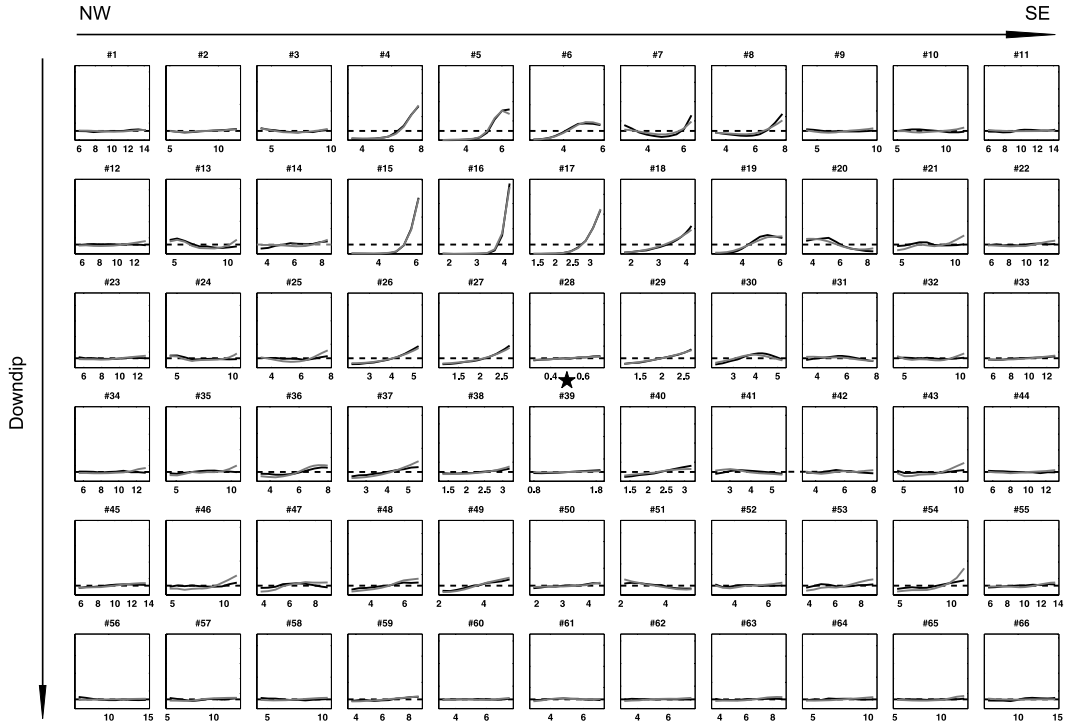
(a)



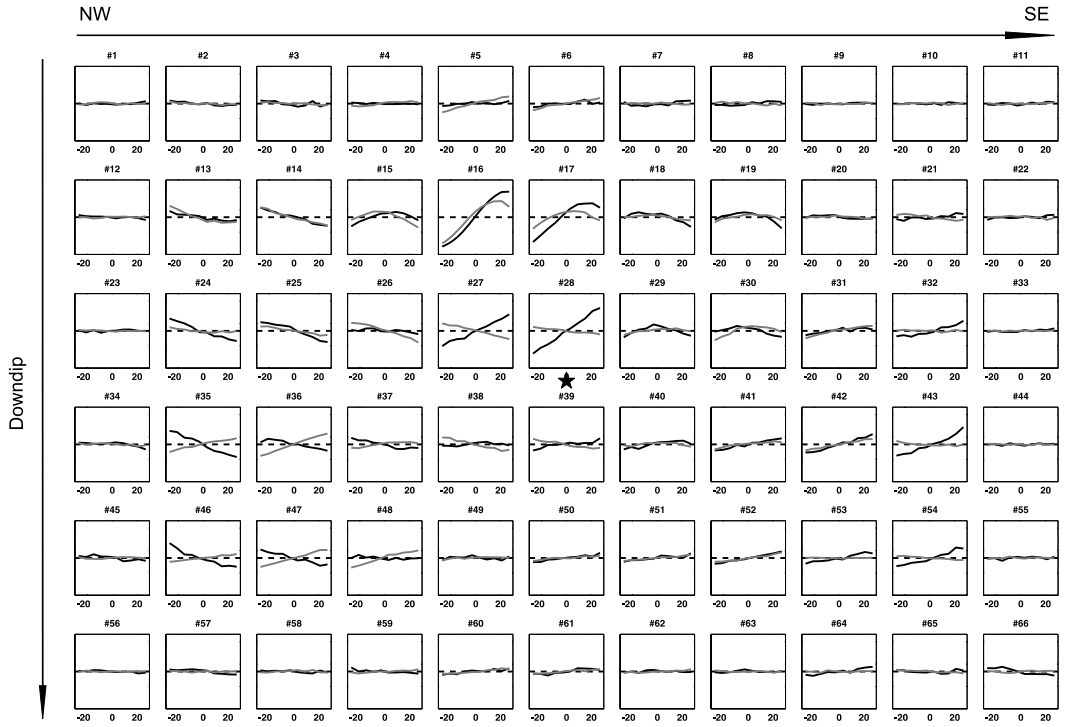
(b)

Figure 9. 1-D marginals for (a) peak slip velocity (cm s^{-1}) and (b) rise time (s). Each subplot corresponds to a node position. Dashed lines represent priors, solid lines posterior marginals (grey from σ_M^{sm}, black from $\sigma_M^{\mathrm{sm,gps}}$). For each subplot we indicate node number, posterior mean value (μ) and standard deviation (σ) of the posterior marginal obtained from $\sigma_M^{\mathrm{sm,gps}}$. The black star represents the hypocentre location.

Fig. 11 shows posterior marginals for final slip (derived from peak slip velocity and rise time values). Note that prior marginals are not uniform because they represent prior information on a combination of the original model parameters. Again, we find that posteriors show mostly a skewed distribution. Only posteriors at nodes 16 and 17 show a Gaussian-like shape. For these two nodes, posteriors predict a final slip of 250 ± 120 and 311 ± 140 cm, respectively. The relative error is ∼48 and 45 per cent, respectively. We also infer a low slip region NW of the hypocentre (nodes from 24 to 27 and 35 to 38). In these locations, 1-D marginals present a distribution, skewed towards the minimum allowed slip (0 cm) with standard deviations ∼50 cm. SE of the hypocentre, 1-D marginals present instead larger standard deviations: ∼100 cm at nodes 28 and 40 and ∼140 cm at node 29 indicating therefore a wider range of likely values. This feature may suggest an elongation of the slip distribution towards SE. The effect of GPS data in constraining the peak slip-velocity is



(a)



(b)

Figure 10. 1-D marginals for (a) rupture time (s) and (b) rake angle (degrees). The same notation of Fig. 9 is used. For clarity we do not report mean values and standard deviations.

reflected in the marginals for the final slip. GPS data have a notable effect in reducing the tail of the marginals (see nodes 21, 32, 43, for instance). They help also in constraining the shallow slip (see node 16 and 17).

These changes have a strong effect when computing the posterior marginal for seismic moment (Fig. 12). GPS data reduce the probability associated with high values of slip and produce a shift of the peak of the posterior towards lower values of

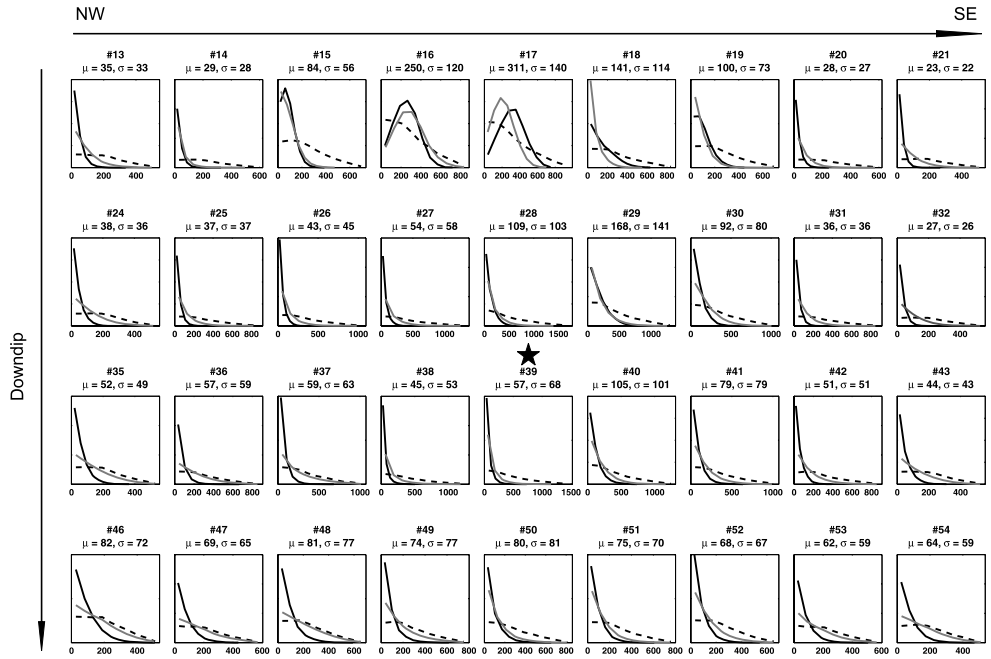


Figure 11. 1-D marginals for final slip (cm).

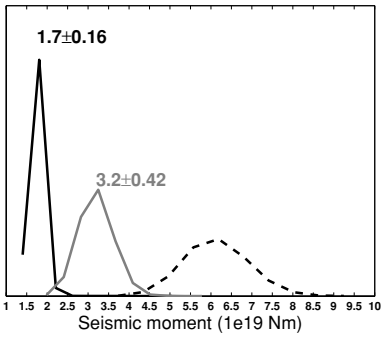


Figure 12. 1-D marginals for seismic moment as obtained from the prior pdf (dashed), σ_M^{sm} (grey) and $\sigma_M^{\mathrm{sm,gps}}$ (black).

seismic moment than obtained from σ_M^{sm}. From the posterior marginal from $\sigma_M^{\mathrm{sm,gps}}$ we infer a value of seimic moment equal to $1.7 \pm 0.16 \times 10^{19}$ N m. The corresponding relative error is about 10 per cent.

In Fig. 13 we present posterior 1-D marginals (derived from $\sigma_M^{\mathrm{sm,gps}}$ only) for seismic moment and moment rate, as they evolve in time. In other words, we compute moment and moment rate time histories for each sample of $\sigma_M^{\mathrm{sm,gps}}$ and then compute, at each time step, the corresponding 1-D marginal. In this way, we obtain a 'probabilistic' image of the moment and moment rate functions, where at each time step, we have not a single value but rather a distribution of values. From the seismic-moment time history, we see that most of the seismic moment starts to be released only after about 3 s from the origin time. This is consistent with the fact that the shallow slip patch is triggered, on an average, 3 s after the earthquake initiated. The moment rate function assumes its peak value at about 5 s. Again considering node 17, we infer a value of rise time of 4.4 s (average value deduced from posterior marginal for rise time, see Fig. 9b). In other words, at node 17, the slip-velocity reaches its peak value about 2 s after the rupture time, that is, at about 5 s. We therefore see a correlation between the peak of the moment rate function and the peak of the source time function at node 17, which is associated with the highest inferred slip.

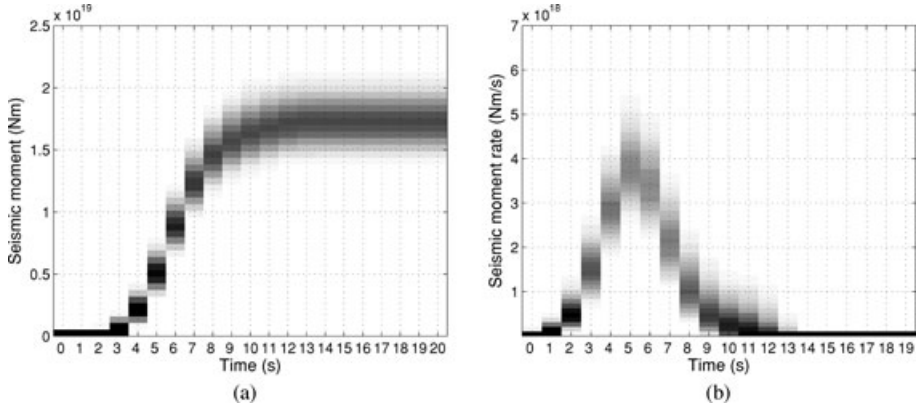


Figure 13. 1-D marginals for moment (a) and moment rate (b) (derived from $\sigma_M^{\mathrm{sm,gps}}$ only) as they evolve in time.

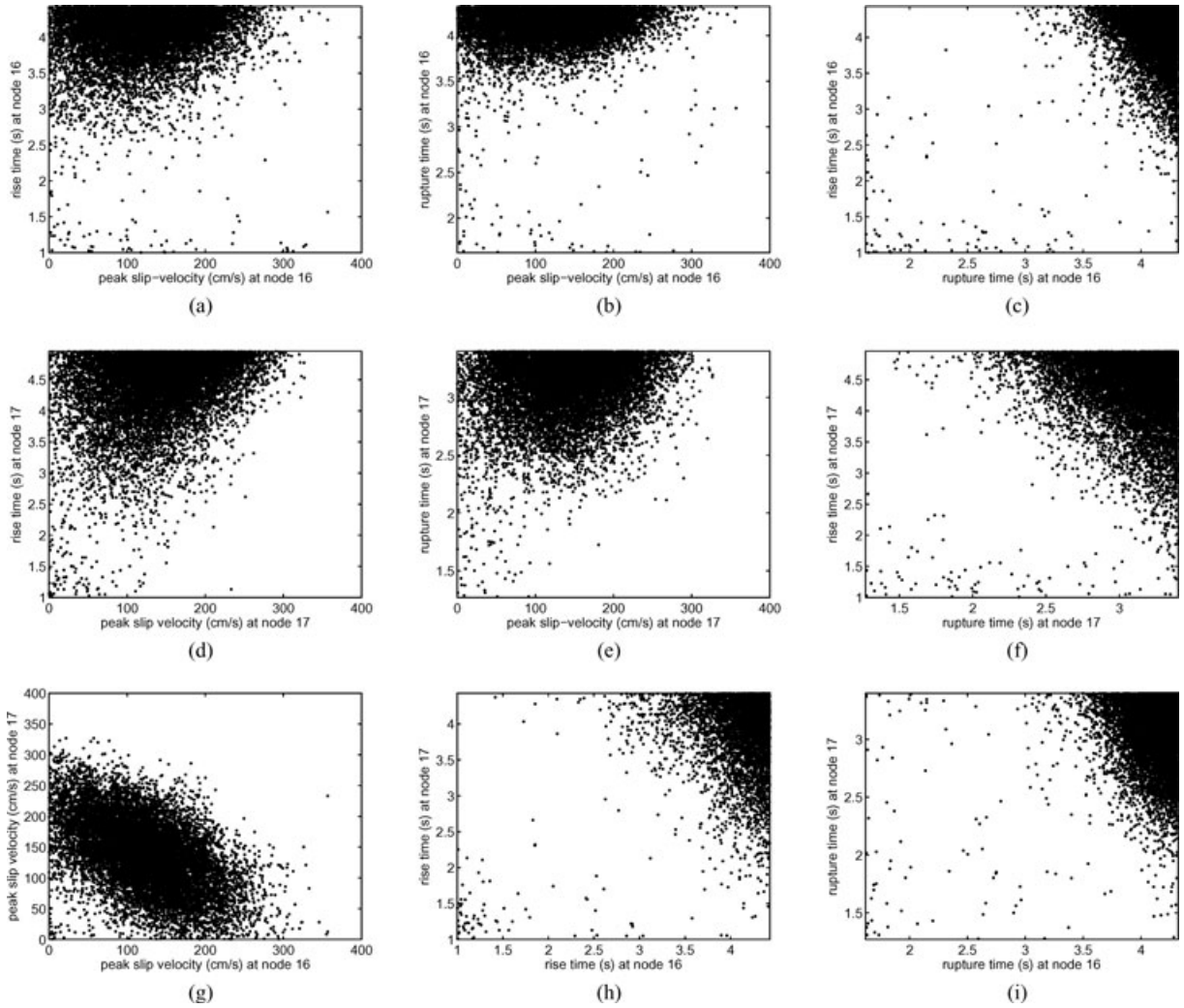


Figure 14. 2-D marginals between peak slip velocity, rise time and rupture time at node 16 and 17.

5.3 2-D marginals

1-D marginals represent all information we have on a single parameter. However, they do not contain any information about possible correlations with other parameters, which constitutes an integral part in any uncertainty analysis. If a pair of parameters is correlated, this implies that we cannot measure them independently. Correlations between pairs of different parameters can be analysed computing 2-D marginals.

Due to the large number of parameters (204 in this study) we did not explore all possible correlations. We focused our attention on the rupture parameters describing the shallow high slip patch, at nodes 16 and 17. We derived 2-D marginals from $\sigma_M^{\mathrm{sm,gps}}$ only because it considers all the data. We first computed 2-D marginals between rupture parameters (mainly peak slip velocity, rupture time and rise time) defined on the same node (Figs 14a–f). We do not identify any significant correlation between these parameters. In Figs 14(g)–(i), we instead present 2-D marginals between rupture parameters defined on different nodes. Here, we identify a strong anticorrelation between peak slip velocity values. In other words, if the peak slip velocity at node 16 increases, the peak slip velocity at node 17 will decrease, and vice versa.

6 DISCUSSION

From the analysis of the 1-D marginals computed from $\sigma_M^{\mathrm{sm,gps}}$ we identify the following main features in the rupture process of the 2000 Tottori earthquake.

(1) Between the hypocentre and the top edge of the fault, corresponding to a depth of 4.5 km (nodes 16, 17), we find a high slip-velocity patch. Posterior marginals show a Gaussian-like shape from which we deduce values of peak slip velocity of 122 ± 57 and 140 ± 57 cm s^{-1}.

(2) In correspondence to the high slip-velocity patch, the posterior marginals for rise time show a skewed distribution, with the maximum attained at the maximum allowed rise time. The mean values for rise time at nodes 16 and 17 are 4.1 and 4.4 s, respectively.

(3) Combining values of peak slip velocity and rise time, we infer for the shallow slip patch final displacements of 250 ± 120 and 311 ± 140 cm on nodes 16 and 17, respectively.

(4) 1-D marginals for rupture time indicate that the shallow slip patch is triggered about 3.1 s (mean value of posterior at node 17) after the rupture initiated at the hypocentre. We can therefore

estimate an average rupture velocity in the updip direction, of about 1.6 km s^{-1}.

(5) The rake angle is generally poorly resolved in the model space considered. Only on the shallow slip patch (nodes 16, 17), posterior marginals suggest that a positive angle (downdip component) is more likely than a negative one.

The presence of a high slip patch near the top edge of the fault has also been recognized in previous studies (Semmane *et al.* 2005; Festa & Zollo 2006; Piatanesi *et al.* 2007). Their models indicate a maximum value of slip of about 4 m, roughly in agreement with our estimates (311 $\pm$ 140 cm). We do not identify any significant slip at the bottom of the fault. For the deepest nodes (from node 46 to 54) posterior 1-D marginals of slip from $\sigma_M^{\mathrm{sm,gps}}$ exhibit a skewed distribution with maximum attained at the minimum allowed slip (0 cm, see Fig. 11). Assuming that standard deviations represent the range of most likely values, we infer for the deepest nodes, values of slip between 0 and $\sim$80 cm. Our inferences for the final slip distribution are therefore more consistent with the preferred model of Semmane *et al.* (2005), which does not show significant slip at the bottom, rather than with the models proposed by Festa & Zollo (2006) and Piatanesi *et al.* (2007), which suggest the presence of significant deep slip (up to 2.5 m).

Regarding the rupture timing, we infer a value of about 1.6 km s^{-1} for the rupture velocity in the updip direction. Festa & Zollo (2006) and Piatanesi *et al.* (2007) inferred values equal to 2.1 and 2.2 km s^{-1}, respectively. These higher values may be due to the deeper hypocentre assumed in these studies (13.5 km, Festa & Zollo 2006 and 12.5 km, Piatanesi *et al.* 2007) compated with the one we adopted (9.6 km).

Another difference from previous studies concerns the rise time pattern. Our results show that rise time values are well resolved only in the vicinity of the shallow high slip patch. At these locations (nodes 16, 17 for instance), the rise time values equal $\sim$4 s. Semmane *et al.*'s (2005)'s preferred model shows at the same locations lower values between 0.5 and 1.5 s. Piatanesi *et al.*'s 2007's average model shows instead more comparable values between 2.5 and 3 s.

As recognized in all studies (including this work), a peculiar feature of the Tottori earthquake is the presence of considerable slip at shallow depth (311 $\pm$ 140 cm at 4.5 km depth), without any evident surface rupture. Identifying the reasons why the slip did not reach the surface is beyond the scope of this paper and requires dynamic modelling of the earthquake rupture process. Qualitatively, we can imagine that possible reasons impeding slip propagation to the surface can be a velocity-strengthening behaviour of the shallow layers or low pre-stress in the uppermost part of the fault, or a combination of these two effects.

Also, the Tottori earthquake is not the only event showing shallow slip with no surface breaks. An earthquake showing similar behaviour is the 2003 $M_{\mathrm{W}} = 6.5$, Bam (Iran) earthquake. From the inversion of radar data, Fialko *et al.* (2005) showed how the Bam earthquake is characterized by right-lateral displacements having a maximum amplitude of about 2 m at a depth of 3–7 km. However both radar data and field investigations confirm lack of surface rupture associated with the faulting event.

Finally, we stress that all the results we show in this study depend and are limited by the chosen model space. For instance in our study we find that for some parameters (e.g. concerning rise time and rupture time), the posterior marginals are skewed towards the maximum allowed values, suggesting that the solution, for these parameters, is located beyond the upper bound of the considered range of values. We acknowledge therefore that a natural extension of this work would be considering a larger model space (e.g. by removing constraints on rise time) and checking if the inference results remain stable or if new solutions are found.

7 CONCLUSIONS

In this study, we investigate the rupture process of the 2000 Western Tottori earthquake through fitting of strong motion and GPS data. Our inversion methodology is based on a Bayesian approach. We state our inferences in terms of marginal pdfs derived from two distinct posterior pdfs: one that considers only strong motion data and the other that considers both strong motion and GPS data.

With both posteriors, we identify as a stable feature of the earthquake rupture process the presence of a high slip patch between the hypocentre and the top edge of the fault. This feature is common with previous studies. The analysis of the 1-D marginals for rupture time, rise time and rake angle indicates that these parameters are well resolved, only where this shallow slip patch is located, meaning that the signal emitted by this patch determines most of the wavefield that we fitted.

When using both strong motion and GPS data, we do not identify any significant slip (>1 m) at the bottom of the fault. For this aspect, our inference results disagree with some previous studies (Festa & Zollo 2006; Piatanesi *et al.* 2007).

We compare inferences obtained considering strong motion data only with ones derived considering both strong motion and GPS data. In our study, we note that the main effect of GPS data is in reducing the precence of spurious slip on the fault, which in turn, has a strong influence on the estimate of the final seismic moment.

A clear point in our analysis is that resolution on kinematic rupture parameters cannot be explained generally, using the Gaussian uncertainty hypothesis. In our study, most of the 1-D posterior marginals do not show a Gaussian distribution. Understanding the actual resolution requires taking into account the nonlinearity of the problem and therefore dealing with non-Gaussian distributions.

ACKNOWLEDGMENTS

We thank K-net and KiK-net for providing strong motion data. This study was founded through ETH-grant TH-16/05-1. Some figures were made using Generic Mapping Tools free software.

REFERENCES

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

Corish, S.M., Bradley, C.R. & Olsen, K.B., 2007. Assessment of a nonlinear dynamic rupture inversion technique applied to a synthetic earthquake, *Bull. seism. Soc. Am.,* **97**(3), 901–914.

Custodio, S., Liu, P. & Archuleta, R.J., 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.

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

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

Festa, G. & Zollo, A., 2006. Fault slip and rupture velocity inversion by isochrone backprojection, *Geophys. J. Int.,* **166,** 745–756.

Fialko, Y., Sandwell, D., Simons, M. & Rosen, P., 2005. Three-dimensional deformation caused by the Bam, Iran, earthquake and the origin of shallow slip deficit., *Nature,* **435,** 295–299.

Fukuyama, E., Ellsworth, W.L., Waldhauser, F. & Kubo, A., 2003. Detailed fault structure of the 2000 Western Tottori, Japan, earthquake sequence, *Bull. seism. Soc. Am.,* **93,** 1468–1478.

Gardner, G.H.F., Gardner, L.W. & Gregory, A.R., 1974. Formation velocity and density-the diagnostic basics for stratigraphic traps, *Geophysics,* **39**(6), 770–780.

Gouveia, W.P. & Scales, J.A., 1998. Bayesian seismic waveform inversion: parameter estimation and uncertainty analysis, *J. geophys. Res.,* **103,**(B2), 2759–2779.

Hartzell, H.S. & 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,** 1553–1583.

Hartzell, S., Liu, P., Mendoza, C., Ji, C. & Larson, K.M., 2007. Stability and Uncertainty of Finite-Fault Slip Inversions: application to the 2004 Parkfield, California, Earthquake, *Bull. seism. Soc. Am.,* **97**(6), 1911–1934.

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**(B02318), doi:10.1029/2003JB002625.

Madariaga, R., Olsen, K. & Archuleta, R., 1998. Modeling Dynamic Rupture in a 3D Earthquake Fault Model, *Bull. seism. Soc. Am.,* **88**(5), 1182–1197.

Martinez, W.L. & Martinez, A.R., 2002. *Computational Statistics Handbook with MATLAB*, Chapman and Hall/CRC, Boca Raton, Florida.

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

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.

Mosegaard, K. & Tarantola, A., 1995. Monte carlo sampling of solutions to inverse problems, *J. geophys. Res.,* **100**(B7), 12 431–12 447.

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

Peyrat, S. & Olsen, K.B., 2004. Nonlinear dynamic rupture inversion of the 2000 Western Tottori, Japan, earthquake, *Geophys. Res. Lett.,* **31**(L05604), doi:10.1029/2003GL019058.

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.

Sagiya, T., 2004. A decade of GEONET:1994-2003-The continuous GPS observation in Japan and its impact on earthquake studies-, *Earth Planets Space,* **56,** xxix–xli.

Sambridge, M., 1999. Geophysical inversion with a neighbourhood algorithm, II: appraising the ensemble, *Geophys. J. Int.,* **138,** 727–746.

Semmane, F., Cotton, F. & Campillo, M., 2005. The 2000 Tottori earthquake: a shallow earthquake with no surface rupture and slip properties controlled by depth, *J. geophys. Res.,* **110**(B03306), doi:10.1029/2004JB003194.

Spudich, P. & Archuleta, R., 1987. Techniques for earthquake ground motion calculation with applications to source parametrization of finite faults, in *Seismic Strong Motion Synthetics,* Vol. 37, pp. 205–265, ed. Bolt, B.A., Academic Press, Orlando, Florida.

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, San Diego, California.

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

Umeda, Y., 2002. The 2000 western Tottori earthquake, *Earth Planets Space,* **54,** iii–iv.