Bayesian inversion for finite fault earthquake source models - II: the 2011 great Tohoku-oki, Japan earthquake,

S.E. Minson, M. Simons, J. L. Beck, F. Ortega, J. Jiang, S.E. Owen, A.W. Moore, A. Inbal, A. Sladen

Geophysical Journal International

Geophys. J. Int. (2014) **198,** 922–940
GJI Seismology

doi: 10.1093/gji/ggu170

Bayesian inversion for finite fault earthquake source models – II: the 2011 great Tohoku-oki, Japan earthquake

S. E. Minson,[1] M. Simons,[1] J. L. Beck,[2] F. Ortega,[1] J. Jiang,[1] S. E. Owen,[3] A. W. Moore,[3] A. Inbal[1] and A. Sladen[4]

[1] *Seismological Laboratory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA* 91125, *USA. E-mail: minson@gps.caltech.edu*
[2] *Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA* 91125, *USA*
[3] *Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA* 91109, *USA*
[4] *Géoazur, Université de Nice, CNRS UMR* 7329, *IRD, UPMC, Observatoire de la Côte d Azur, Sophia-Antipolis, Valbonne, France*

Accepted 2014 May 7. Received 2014 May 6; in original form 2013 July 15

SUMMARY

We present a fully Bayesian inversion of kinematic rupture parameters for the 2011 M_w9 Tohoku-oki, Japan earthquake. Albeit computationally expensive, this approach to kinematic source modelling has the advantage of producing an ensemble of slip models that are consistent with physical *a priori* constraints, realistic data uncertainties, and realistic but simplistic uncertainties in the physics of the kinematic forward model, all without being biased by non-physical regularization constraints. Combining 1 Hz kinematic GPS, static GPS offsets, seafloor geodesy and near-field and far-field tsunami data into a massively parallel Monte Carlo simulation, we construct an ensemble of samples of the posterior probability density function describing the evolution of fault rupture. We find that most of the slip is concentrated in a depth range of 10–20 km from the trench, and that slip decreases towards the trench with significant displacements at the toe of wedge occurring in just a small region. Estimates of static stress drop and rupture velocity are ambiguous. Due to the spatial compactness of the fault rupture, the duration of the entire rupture was less than approximately 150 s.

Key words: Inverse theory; Probability distributions; Earthquake source observations; Computational seismology.

1 INTRODUCTION

Many finite fault rupture models have been produced and published in the time since the M_w 9 great Tohoku-oki earthquake of 2011 March 11. However, these models are usually based on traditional optimization approaches which produce a single model from the parameter space of all rupture scenarios that are consistent with the observed data. Inferring a finite fault earthquake rupture model is an ill-conditioned inverse problem, and so we expect the space of all possible and realistic models to be large. This non-uniqueness should be especially large for subduction zone earthquakes given that there are few observations available from the offshore region where the bulk of the fault slip occurs; but the non-unique characteristic of the inferred rupture model is lost when optimization is applied to solve the inverse problem. Compounding this shortcoming of traditional optimization, many optimization methods produce no error estimates, while the techniques that can estimate uncertainties often assume Gaussian errors and only provide variances of the inferred rupture parameters.

Each published rupture model for the Tohoku-oki earthquake can be thought of as one sample from the space of all plausible models, or rather they can be thought of as individual samples from a number of different parameter spaces since these studies often utilize different source parametrizations or constraints on those parameters (e.g. regularization). However to fully understand the range of possible rupture scenarios for the great Tohoku-oki earthquake, we should determine the complete distribution of all plausible source models including the trade-offs and uncertainties on each rupture parameter. Furthermore, the published body of work represents a limited set of models where the uncertainties of each model typically rely on conventional tests of resolution, all of which are of limited value. (For a discussion of this, see Lévêque *et al.* 1993.)

To improve our understanding of the 2011 great Tohoku-oki earthquake, we provide a complete Bayesian inversion that uses only physically based prior information. [We introduced this methodology in Minson *et al.* (2013), hereafter referred to as Paper I.] Paper I described and validated a fully Bayesian methodology for finite fault kinematic source models which yields the posterior probability density function (PDF) that describes all rupture scenarios which are consistent with observed data and three physical priors: each point on the fault is forbidden from slipping significantly in a direction opposite to the long-term relative fault motion, each point

on the fault is only allowed to rupture once during the earthquake using a prescribed source–time function (in this case, a triangular slip velocity function), and rupture propagates along the fault causally. We use the same inversion approach in this study to provide both static and kinematic models of the earthquake rupture constrained by 1 Hz kinematic GPS time-series (used in the kinematic inversion only), static GPS offsets, offshore geodetic data and tsunami data from both near-coast instruments and global DART (Deep-ocean Assessment and Reporting of Tsunamis) buoys. In addition, because we compute the Bayesian posterior PDF, our solutions allow for a more complete understanding of the uncertainties in our solution than can be derived from conventional resolution tests. While still not ideal, our error analysis is an improvement over current standard approaches to understanding the uncertainties in finite fault earthquake source models. However, as discussed towards the end of this paper, we still need further improvements to reach a more complete error analysis. Further, although our model is free of non-physical regularization, it is only as good as the physical model used to produce the predicted observations. Any errors or shortcomings in our forward model will lead to erroneous inferences about the source process, including the potential for seemingly well constrained yet erroneous conclusions.

2 INVERSION DESIGN

2.1 Data

Following the methodology presented in Paper I, we compute an ensemble of static slip models for the Tohoku-oki earthquake and then use that posterior distribution to inform the prior distribution of the complete joint kinematic-static rupture model. For the static inversion, we include static GPS offsets, offshore seafloor geodesy, tsunami records from near-coast pressure gauges and tsunami records from the global DART network of open-water buoys. (Our forward model for the tsunami data assumes an instantaneous rupture such that the tsunami forward model is purely a static deformation model.) The kinematic model uses all of these data as well as 1 Hz GPS time-series.

Information about the near-coast tsunami data, open-water tsunami records, locations of the six GPS stations used for 1 Hz kinematic modelling (out of the 738 GEONET stations whose static offsets were also incorporated), and the locations and observed static offsets at the 738 GEONET GPS stations and offshore GPS/acoustic instruments used in our inversions are tabulated in detail in the online Supporting Information. The seafloor geodesy data and their associated errors come from Kido *et al.* (2011) and Sato *et al.* (2011). The global DART data are the same as in Simons *et al.* (2011). The GPS positions were obtained using the same single-station bias-fixed processing as in Simons *et al.* (2011). (See the Simons *et al.* 2011 online Supporting Information for details.) The three-component static GPS data come from 738 GEONET permanent GPS stations and are derived from solutions with 30-s epochs so as to exclude as much post-seismic and aftershock-related deformation as possible. The 1 Hz kinematic GPS data are taken from the east and north components of six GPS stations which are a subset of the 738 GEONET stations. The six GPS stations used for kinematic modelling are located near the coast of Honshu in a transect that covers the along-strike extent of the rupture. We use high-rate data from a small subset of the GEONET array not because the data are not available or cannot be fit, but rather because it would be too computationally expensive to evaluate the misfit between the data and the predicted time-series for each of the $O(10^{10})$ models we evaluate during Monte Carlo simulation of the posterior distribution. As borne out by visual inspection of the data, we expect the observations over a spatially dense array of stations to be highly correlated such that the observed waveforms vary slowly from station to station, and thus a representative subset of observed time-series should be sufficient to describe the spatiotemporal variation of surface displacement while removing duplicate information. Therefore, in order to obtain the best possible data coverage to determine along-strike variations in the earthquake rupture, we selected a subset of stations that give a set of observed near-field displacement histories without duplication and which cover a wide range of near-field source–receiver azimuths. This source–receiver geometry is not ideal for imaging along-dip variations, but it is the best that is possible since there are no near-field observations available on the seaward side of the rupture and we do not expect that teleseismic data would provide additional resolution given the very small spatial extent of the rupture in the downdip direction.

We filter the 1 Hz GPS time-series and corresponding Green's functions between 0.005 and 0.1 Hz (200–10 s period) using a one-pass causal Butterworth band-pass filter. We adopt the low-pass filter to counteract the effects of potential high-frequency aliasing in the raw time-series. We use a high-pass filter because our kinematic and static GPS data come from identical stations, and our model predictions are made using Green's functions for identical Earth structures. Thus the 'static' GPS data and the final static offsets contained in the 'kinematic' GPS data should be identical (give or take the effects of different processing methods), and the predicted offsets from any source model will be identical for these two data sets. It would be in effect 'double counting' our observations to use both static offsets from the 30-s epoch GPS positions and the static offsets contained in the 1 Hz time-series. So we filter out the periods in our kinematic GPS data that are of the duration of the earthquake or longer. While the broadband kinematic GPS time-series are dominated by the near-field ramp-like accumulation of static offset, the filtered records emphasize the time-dependent and non-steady-state parts of the wavefield, the exact information which the high-rate GPS data can provide but that our other data sets cannot.

2.2 Model design

Our model for the geometry of the megathrust interface is an approximation of the Pacific Plate surface from Simons *et al.* (2011) (Figs S1 and S2) with the fault interface from the surface to approximately 30 km depth constrained by seismic refraction and reflection profiles (Iwasaki *et al.* 1994; Tsuru *et al.* 2000; Fujie *et al.* 2002; Miura *et al.* 2003; Ito *et al.* 2004; Takahashi *et al.* 2004; Miura *et al.* 2005), and the deeper fault surface constrained with seismic tomography (Zhang *et al.* 2004; Hasegawa *et al.* 2005; Yamamoto *et al.* 2006) and earthquake hypocentres (Katsumata *et al.* 2003). As described in Paper I, we use the Fast Sweeping algorithm (Zhao 2005) to calculate the initial rupture time at each source location given a random realization of a hypocentre location and a random distribution of rupture velocities. Because this algorithm is only applicable to solving the eikonal equation on a square grid, we are forced to tessellate our fault into square patches while Simons *et al.* (2011) used a complex triangular mesh. However, we have used the nearest regularly gridded approximation to the Simons *et al.* (2011) source geometry. Our non-planar gridded fault is curved in the downdip direction (Fig. 1). It is comprised of square patches,

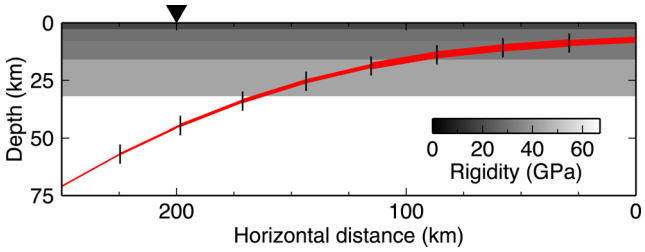

Figure 1. Profile of slab geometry as a function of depth. The fault interface is shown by a red line whose thickness is proportional to the mean integrated along-strike slip potency of the joint static-kinematic rupture model (Fig. 2). Assumed rigidity as a function of depth (Table 1) is shown by background colour. The coastline is located approximately 200 km from the trench (indicated with triangle).

Table 1. Depth extent, layer thickness, *P*-wave velocity, *S*-wave velocity and density used for calculating elastic Green's functions. The *P*-wave and *S*-wave quality factors used are 1000 and 500, respectively.

Depth (km)	Thickness (km)	α (km s^{-1})	β (km s^{-1})	ρ (kg m^{-3})
0.0–2.8	2.8	5.0	2.7	2500
2.8–7.8	5.0	5.7	3.3	2700
7.8–16.0	8.2	6.0	3.4	2750
16.0–32.0	16.0	6.7	3.9	2900
32.0–∞	∞	7.7	4.5	3300

29 km on a side, with a constant strike of 194° and dip that increases from 3° at the trench to 29° at depth. There are a total of 24 patches along-strike and nine patches downdip. For the kinematic model, we solve for two components of slip, slip duration, and rupture velocity at each of these 216 source locations. To produce smooth synthetics, our predicted seismograms are generated by first interpolating each rupture model onto a finer grid of point sources located less than 10 km apart.

We use the same 1-D layered elastic structure as Simons *et al.* (2011) which, in turn, is based on the seismic structure model of Takahashi *et al.* (2004) (Table 1). Dynamic Green's functions for this elastic structure were calculated using frequency–wavenumber integration (Zhu & Rivera 2002) with a 1 Hz sampling rate using wavenumbers sampled at an interval of 0.01 and applying a low-pass cosine filter with corner frequency equal to 70 per cent of the Nyquist frequency. The tsunami Green's functions were produced by first creating elastic Green's functions that map displacement on our fault surface to a dense grid of vertical seafloor displacement using the Earth structure in Table 1 vertically shifted 2.8 km to account for the ocean depth in the source region. A Gaussian spatial filter with a standard deviation of 15 km was then applied to these seafloor displacements so that our predicted seafloor deformation had the same spatial resolution as our original fault slip model but without discontinuities due to the discretization of the fault plane. The resulting seafloor displacements were then used as input into the COMCOT (COrnell Multi-grid COupled Tsunami model) tsunami propagation code (Liu *et al.* 1998). Using a shallow water approximation, we calculated predicted tsunami records for a unit step on a given fault patch assuming that the tsunami is excited not only by the resulting vertical motion of the seafloor but also by the change in seafloor height due to horizontal displacement of the sloping seafloor. We use an instantaneous rupture model for predicting the tsunami data, however we shift the tsunami Green's functions by 68 s so that the instantaneous rupture occurs at the centroid time rather than the hypocentre time.

The phase velocity, c, of a tsunami wave with wavelength, λ, travelling through water of depth, d, is $c = \sqrt{gd}$ for wavelengths much greater than d, where g is the gravitational acceleration. At shorter wavelengths, the phase velocity exhibits frequency-dependent dispersion, such that $c = \sqrt{\lambda g/2\pi}$. We should note that our tsunami Green's functions neglect this dispersion. However, we applied a low-pass filter with a corner frequency of 0.004 Hz (250 s period) to all tsunami data and Green's functions. This filtering is designed to accomplish two purposes: first, to counteract the lack of frequency-dependent dispersion in the tsunami Green's functions and, second, to remove complex reflections and other high-frequency features in the waveforms that are not reproducible given the limited resolution of the available bathymetric data. Comparisons of Green's function with and without frequency-dependent dispersion show that the effects of dispersion are removed by this filter.

Even if dispersion were non-negligible, tsunami arrival times are controlled to the first order by water depth such that propagation models that assume a shallower water depth will have later arriving tsunamis than those that assume deeper water. Thus it is critical to have an accurate bathymetric model in order to properly model the tsunami traveltimes. To this end, we have used the highest resolution bathymetry available to us. In the near-field, we use the J-EGG500 500-meter bathymetry data set (http://www.jodc.go.jp/data_set/jodc/jegg_intro.html). Deep water propagation is modelled using the ETOPO1 1-min bathymetric grid. For the far-field DART tsunami data, we empirically time-shift the tsunami waveforms to remove systematic errors in tsunami arrival times due to the shortcomings of using the shallow water approximation to predict global tsunami propagation (see e.g. Tsai *et al.* 2013; Watada 2013; Watada *et al.* 2014).

The resulting forward model for the tsunami generated by the earthquake rupture contains four simplifications. We have neglected, first, the propagation delay between the fault plane and the seafloor, second, the propagation delay and attenuation between seafloor deformation and changes in ocean surface height, third, frequency-dependent dispersion of tsunami waves, and, fourth, the finite duration of the earthquake rupture. Neglecting both propagation delays is valid since elastic wave speeds are much larger than tsunami speeds. The effects of attenuation between the seafloor and the sea surface can be ignored because we have applied a Gaussian spatial filter to the seafloor displacement field. (Neglecting the propagation delay and attenuation between seafloor deformation and changes in sea surface height is also justified, respectively, by the fact that the water depth is small and water can be considered incompressible.) The effect of frequency-dependent dispersion can be ignored because we have lowpass-filtered both the tsunami data and Green's functions. Finally, we compared predictions of an instantaneous tsunami rupture model and one made with a propagating source and found that the instantaneous rupture approximation is valid for the Tohoku-oki earthquake which is both spatially compact and of short duration.

2.3 Bayes' theorem

A thorough introduction to the use of Bayesian inference in inverse problems was given in Paper I. The reader is referred to that paper for additional background as the discussion is too long to be repeated here. To briefly summarize, a generic inverse problem can be solved using Bayes' theorem (e.g. Gelman *et al.* 2004; Tarantola 2005),

$$\mathrm{p}(\theta|\mathbf{d}) \propto \mathrm{p}(\mathbf{d}|\theta)\,\mathrm{p}(\theta)\,, \qquad (1)$$

Table 2. Prior distributions.

Source parameter	Prior PDF
Strike-slip displacement	$U_{\perp} \sim \mathcal{N}(0\text{ m}, 10\text{ m})$
Dip-slip displacement	$U_{\parallel} \sim \mathcal{U}(-10\text{ m}, \infty)$
Slip duration	$T_r \sim \mathcal{U}(12\text{ s}, 42\text{ s})$
Rupture velocity	$V_r \sim \mathcal{U}(0\text{ km s}^{-1}, 7.7\text{ km s}^{-1})$
Hypocentre[a]	$H_0 \sim \mathcal{N}(0\text{ km}, 50\text{ km})$

[a]Hypocentre location is measured in distance along fault plane from a point at the latitude and longitude of the Chu *et al.* (2011) epicentre (38.19°N, 142.68°E).

where $\mathrm{p}(\theta|\mathbf{d})$ is the *a posteriori* or posterior PDF representing the solution to our inverse problem, that is the conditional probability of a set of model parameters, θ, given a set of observations, $\mathbf{d}$. It is proportional to the *a priori* or prior PDF, $\mathrm{p}(\theta)$, which represents our *a priori* beliefs about the relative plausibility of different potential values of the model parameters, and the data likelihood, $\mathrm{p}(\mathbf{d}|\theta)$, which describes the misfit between our observed data and the forward predictions of a model whose parameters are given by θ. Next we describe our choice of prior distribution and data likelihood for the modelling presented in this paper.

2.4 Prior distribution

We require the specification of a prior distribution for each model parameter. We use the same families of prior PDFs as in Paper I. Specifically, we use a zero-mean Gaussian prior on the strike-slip component of slip (since we believe that, on average, the slip direction was thrust and not strike-slip) but consider all possible values of thrust motion equally likely with the constraint that we forbid back-slip (i.e. normal slip) in excess of 10 m. Slip duration and rupture velocity at each point are both assigned uniform prior PDFs broad enough to encompass all plausible values for this earthquake. The lower bound on rupture velocity is 0 km s^{-1} while the high bound is equal to the highest *P*-wave velocity in our elastic structure thus eliminating the possibility of supersonic rupture speeds. The hypocentre location is given a Gaussian prior centred on the location of the Chu *et al.* (2011) epicentre. These prior distributions are summarized in Table 2.

2.5 Data likelihood

For all data sets we use a Gaussian PDF for our data likelihood,

$$\begin{aligned}\mathrm{p}(\mathbf{d}|\theta) &= \mathcal{N}\left(\mathbf{d}|\mathbf{G}(\theta), \mathbf{C}_{\chi}\right)\\ &= \frac{1}{(2\pi)^{n/2}|\mathbf{C}_{\chi}|^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}[\mathbf{d}-\mathbf{G}(\theta)]^{T}\cdot\mathbf{C}_{\chi}{}^{-1}\cdot[\mathbf{d}-\mathbf{G}(\theta)]},\end{aligned} \tag{2}$$

where $\mathbf{d}$ is a vector of n data points, $\mathbf{G}(\theta)$ is the output of a forward model defined by a vector of model parameters θ, and $\mathbf{C}_{\chi}$ is a covariance matrix of the misfits between the observations and our predictions. As explored in Paper I, the errors in our assumed Earth structure and fault model can cause errors in our predictions that are orders of magnitude greater than the formal observational uncertainties. We refer to this class of errors (the differences between the actual observed quantities and what our forward model produces) as the model prediction error.

Let us define $\mathbf{C_d}$ as the covariance matrix describing our formal data uncertainties and $\mathbf{C_p}$ as the covariance matrix for our model prediction errors. Under the assumption that $\mathbf{C_d}$ and $\mathbf{C_p}$ are independent and Gaussian, we can rewrite $\mathbf{C}_{\chi}$ as,

$$\mathbf{C}_{\chi} = \mathbf{C_d} + \mathbf{C_p}. \tag{3}$$

Our preliminary investigations found that ignoring $\mathbf{C_p}$ results in enormous overfitting of the data. This behaviour is expected since even a small percentage of error in the Green's functions would, when multiplied by a source of strength M_{w} 9.0, result in large absolute errors in the predicted near-field surface displacements. Instead, we introduce into the inversion process an error source representing an error in our assumed Earth structure, parametrized as,

$$\mathbf{C_p} = \alpha^2 \mathrm{diag}\left(d_1^2, \ldots, d_n^2\right), \tag{4}$$

where α is the percentage error on our Green's functions. For waveform data (tsunami data and 1 Hz kinematic GPS time-series) whose amplitude varies as a function of time, we replace $\mathrm{diag}(d_1^2, \ldots, d_n^2)$ with $\max(d^2)\mathbf{I_n}$, where $\mathbf{I_n}$ is the n-dimensional identity matrix. This model prediction error model is insensitive to the moment magnitude of the earthquake because it scales with the size of the observed surface displacements, but it has two main simplifying assumptions: that the model prediction error is Gaussian and that the errors are independent. Although the errors resulting from an imperfect elastic structure are not generally Gaussian, our performance tests in Paper I show that this form of the model prediction error still improves recovery of the true slip model and posterior uncertainties on the model parameters.

In Paper I, we solved for the value of α as part of our Bayesian inversion. However, by using a diagonalized covariance matrix, we lose the spatial covariance between different observation locations and the temporal correlations within waveform data such as the 1 Hz kinematic GPS time-series and tsunami data. For these latter data sets, it may be preferable to use a methodology which includes temporal covariance such as the approach of Yagi & Fukahata (2011). Since spatially dense data are expected to have highly covarying prediction errors, sampling for the correlation-free model prediction error associated with each data set tends to return results that favour fitting the spatially dense (and thus highly correlated) GEONET GPS data at the expense of the sparse offshore geodetic data. (We further explore the systematic effects of errors in the assumed elastic structure in Section 4.2.) Instead, for all of the results presented in this paper, we fix $\mathbf{C_p}$ using eq. (4) with α set to 10 per cent. This choice reflects a very conservative estimate of the quality of our 1-D elastic structure. Given the large magnitude of this earthquake, the model prediction error ($\mathbf{C_p}$) not the formal observational error ($\mathbf{C_d}$) is the dominant component of the total misfit covariance matrix ($\mathbf{C}_{\chi}$) in eq. (3).

2.6 Simulating the posterior PDFs

As described in Paper I, we employ the cascading approach in which we first compute a static slip distribution using only the static GPS data,

$$\begin{aligned}\mathrm{p}(\theta_{\mathrm{s}}|\mathbf{d}_{\mathrm{s}}) &\propto \mathrm{p}(\mathbf{d}_{\mathrm{s}}|\theta_{\mathrm{s}})\mathrm{p}(\theta_{\mathrm{s}})\\ &= \mathrm{p}(\mathbf{d_{GPS}}|\theta_{\mathrm{s}})\mathrm{p}(\mathbf{d_{tsunami}}|\theta_{\mathrm{s}})\\ &\quad\cdot\mathrm{p}(\mathbf{d_{SO}}|\theta_{\mathrm{s}})\mathrm{p}(\theta_{\mathrm{s}})\\ &= \mathcal{N}(\mathbf{d_{GPS}}|\mathbf{G}(\theta_{\mathrm{s}}), \mathbf{C_{GPS}})\\ &\quad\cdot\mathcal{N}(\mathbf{d_{SO}}|\mathbf{G}(\theta_{\mathrm{s}}), \mathbf{C_{SO}})\\ &\quad\cdot\mathcal{N}(\mathbf{d_{tsunami}}|\mathbf{G}(\theta_{\mathrm{s}}), \mathbf{C_{tsunami}})\\ &\quad\cdot\mathrm{p}(U_{\perp})\mathrm{p}(U_{\parallel}),\end{aligned} \tag{5}$$

where $\mathbf{d_{GPS}}$, $\mathbf{d_{SO}}$ and $\mathbf{d_{tsunami}}$ are the observed onshore GPS offsets, seafloor offsets (from GPS/acoustic measurements) and tsunami waveforms, respectively, which together comprise our observations of static deformation, $\mathbf{d_s}$. $\mathbf{C_{GPS}}$, $\mathbf{C_{SO}}$ and $\mathbf{C_{tsunami}}$ are the total misfit covariance matrices for the respective data sets. θ_s is the vector of parameters for the static slip model consisting of the two components of slip on each fault patch in coordinates parallel to the rake direction, $U_{\parallel}$, and perpendicular to the rake direction, $U_{\perp}$.

We use the posterior PDF from the static slip model to inform the posterior distribution of the joint kinematic-static rupture model,

$$\begin{aligned} p(\theta|\mathbf{d}) &= p(\theta_k, \theta_s|\mathbf{d_k}, \mathbf{d_s}) \\ &\propto p(\mathbf{d_k}|\theta_k, \theta_s)\, p(\mathbf{d_s}|\theta_s)\, p(\theta_s) p(\theta_k) \\ &\propto p(\theta_s|\mathbf{d_s})\, p(\mathbf{d_k}|\theta_k, \theta_s)\, p(\theta_k) \\ &= p(\theta_s|\mathbf{d_s})\, p(\mathbf{d_{1Hz}}|\theta_k, \theta_s)\, p(T_r)\, p(V_r)\, p(\mathbf{H_0}) \\ &= p(\theta_s|\mathbf{d_s})\, \mathcal{N}(\mathbf{d_{1Hz}}|\mathbf{G}(\theta_k, \theta_s), \mathbf{C_{1Hz}}) \cdot \\ &\quad p(T_r)\, p(V_r)\, p(\mathbf{H_0}), \end{aligned} \tag{6}$$

where $\mathbf{d_{1Hz}}$ is a vector of high-rate kinematic GPS time-series with associated covariance matrix $\mathbf{C_{1Hz}}$, and θ_k is the vector of kinematic-only rupture parameters consisting of the duration of slip, T_r, and rupture velocity, V_r, on each fault patch as well as the point on the fault plane at which the rupture initiates, $\mathbf{H_0}$. Thus the complete vector of parameters characterizing the kinematic rupture model is $\theta = (\theta_k, \theta_s)$.

The posterior PDFs in eqs (5) and (6) were evaluated via Markov Chain Monte Carlo (MCMC) simulation using the Cascading Adaptive Transitional Metropolis In Parallel (CATMIP) algorithm introduced in Paper I. CATMIP is an embarassingly parallel algorithm that shares information between each parallel process in order to dynamically optimize sampling, making it tractable to simulate high-dimensional PDFs.

The PDFs in eqs (5) and (6) are high-dimensional. As described earlier, we have 216 fault patches in our source model. Thus the static slip model (eq. 5) which solves for two components of slip on each fault patch requires simulation of a 432-dimensional posterior PDF. The full kinematic slip model (eq. 6) also includes slip duration and rupture velocity on each fault patch and two additional parameters to describe the location of the rupture initiation point (hypocentre) on the fault plane, yielding an 866-dimensional posterior PDF.

As discussed in Paper I, we gain computational efficiency by initializing our MCMC simulation of the posterior PDF with random slip models which have moment magnitudes that compare to the actual magnitude of the earthquake. This approach is not a constraint on the inversion nor is it part of the prior PDF. We are simply choosing not to use seeds for our random walks which we know *a priori* are highly unlikely. The inversion algorithm is free to propose and keep models with any seismic moment.

To generate random slip values for each fault patch that sum to a specified seismic moment, we draw samples from the Dirichlet distribution (e.g. Gelman *et al.* 2004). However, we assume that we do not know the magnitude of the earthquake perfectly *a priori*. So we first draw a set of random moment magnitudes from a normal distribution which, for the Tohoku-oki earthquake, we assign a mean of M_w 9.0 and standard deviation of ± 0.5 magnitude units. Then, for each of these random realizations of moment magnitude, we assign random slips to each fault patch using the Dirichlet distribution scaled to match our selected moment magnitude. These random slips comprise one of our initial seed slip models for our MCMC simulation.

CATMIP computes a sequence of transitional PDFs which moves the simulated PDF from a known initial state (the prior PDF) to the final posterior PDF. Each transitional PDF is simulated via random walks (Markov chains) using multiple parallel instances of the Metropolis algorithm (Metropolis *et al.* 1953). At run time, the user chooses the number of Markov chains (random walks) that will be run at each transitioning stage and the number of steps in each random walk. All other parameters defining the Monte Carlo simulation are dynamically optimized by CATMIP as it runs. Each transitional PDF for both the static slip and kinematic rupture models was explored with 1 000 000 Markov chains (random walks) each with a length of 500 steps. The total number of model evaluations made over the lifetime of the CATMIP algorithm is $N_{\text{chains}} + M \cdot N_{\text{chains}} \cdot (N_{\text{steps}} - 1)$, where N_{chains} is the number of Markov chains, N_{steps} is the number of steps in each Markov chain and M is the number of intermediate PDFs, which is chosen dynamically by CATMIP. For the static slip inversion, $M = 65$, meaning that $O(10^{10})$ models were evaluated. For the kinematic model with a fixed hypocentre, $M = 52$, requiring the evaluation of $O(10^{10})$ full kinematic rupture scenarios. Simulation of the full joint kinematic-static posterior PDF is only computationally tractable by 'cascading' from the static-only posterior. While the ensemble of samples of the posterior PDF are comprised of only the last sample from each of the N_{chains} at the Mth transitioning step, it required evaluation of almost 60 billion potential source models to compute the final ensemble of fully Bayesian rupture models presented here.

3 RESULTS

The Bayesian approach to seismic source inversion yields a posterior PDF describing the ensemble of all plausible values for a set of model parameters given a set of observations, *a priori* knowledge of the physics of the rupture process, and a model for the distribution of the residuals between the observations and the predictions of the forward model. Although there is significant computational expense in simulating the posterior PDF, once the posterior PDF has been obtained, it can be analysed to determine what is known and not known about the rupture process. The following discussion of our inversion results represents just an initial exploration of some of the potential avenues for analysis of the Tohoku-oki earthquake rupture.

The most obvious model to consider first is the mean of the posterior PDF. This solution is the average of the entire ensemble of plausible models. The posterior mean slip for the static inversion (of GPS offsets, seafloor geodesy and both near-coast and deep-water tsunami records) is plotted in Fig. 2. The posterior mean slip, slip duration, and spatial distribution of initial rupture time for the kinematic inversion (which includes all static data sets as well as 1 Hz kinematic GPS time-series) are plotted in Figs 2 and S3. (The kinematic inversion includes the hypocentre location as an additional unknown.) The static and kinematic slip models (Fig. 2) are almost identical, indicating that the kinematic GPS time-series and static deformation data are very compatible. These two slip models are, in turn, similar to the slip model of Simons *et al.* (2011) which was computed on a complex triangular mesh. This similarity gives us confidence that our slip model is not overly sensitive to our choice of fault geometry and discretization scheme.

Note that the peak slip occurs well downdip of the trench. This does not mean that there was not significant slip at the trench. In fact, the mean posterior slip at the trench has a maximum of 47 m. However, the peak slip in the mean posterior slip model is 74 m, located at a depth of about 16 km.

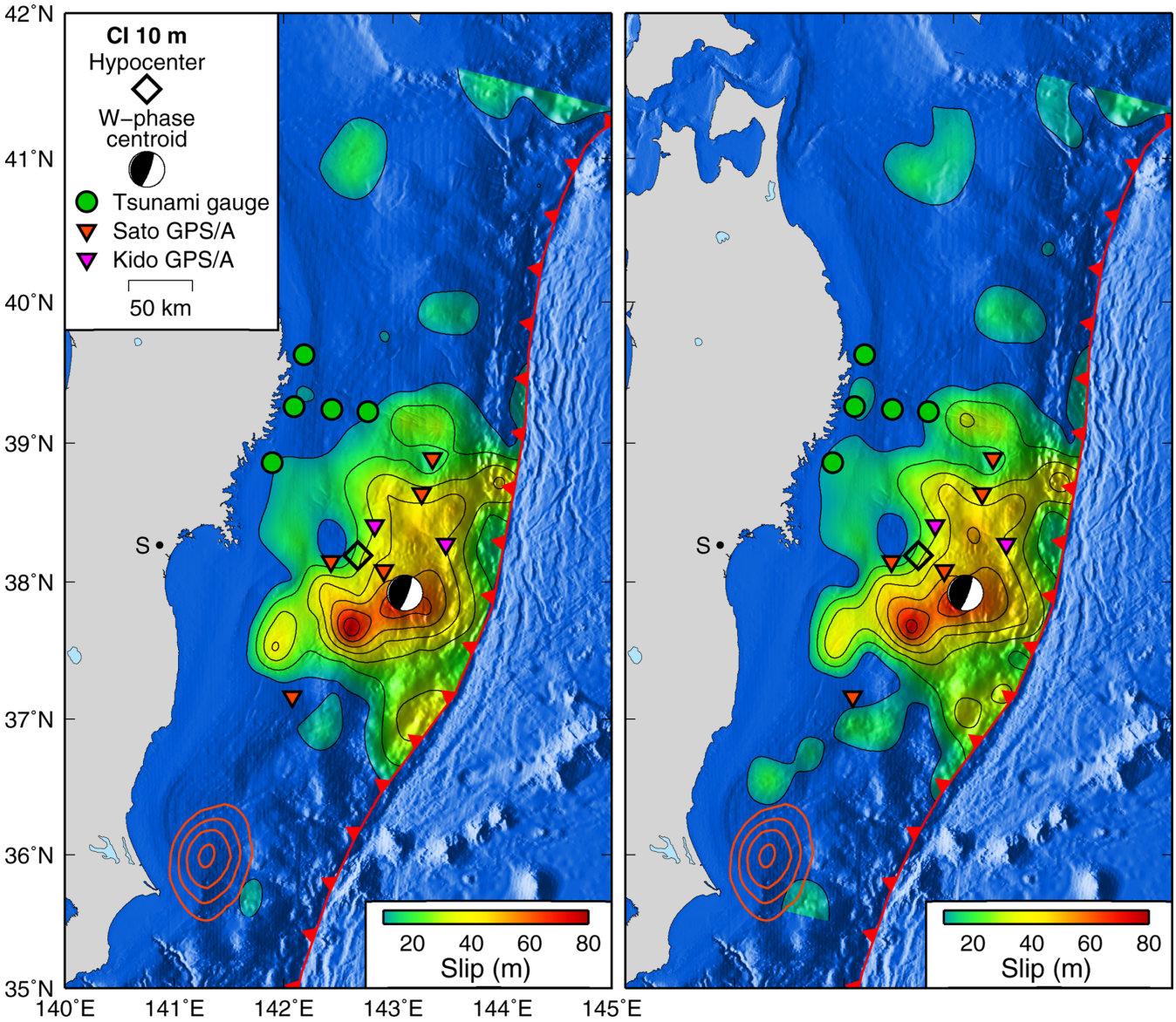

Figure 2. Mean posterior slip distribution from static inversion (left-hand panel) and joint kinematic-static inversion (right-hand panel). The mean of the posterior PDF for the static displacements on the fault plane is shown. Green circles, red triangles, and pink triangles mark the locations of seafloor geodesy and near-field tsunami records. The black diamond is the location of the Chu *et al.* (2011) hypocentre. The W-phase focal mechanism of Duputel *et al.* (2011) is shown at its centroid location. Red contour lines denote slip model for largest aftershock from Simons *et al.* (2011) shown with a contour interval of 1 m. The location of the city of Sendai is marked by the letter S.

The rupture appears to have begun by slowly propagating through the region of largest slip (Fig. S3). If we look at the uncertainties of our model parameters as measured by the standard deviation of the posterior samples (Fig. S4), we see that rupture velocity is much better constrained in the region of high slip than it is elsewhere. Thus our best interpretation of Fig. S3 is that rupture velocity is robustly determined to be low in the main asperity and is not well constrained in the rest of the model. It is self-evident that the values of the parameters that define the time evolution of slip (slip duration and rupture velocity) must be arbitrary (or, if not arbitrary, controlled by the prior PDF) for patches with zero slip. However based on the rake direction in the posterior mean rupture model (Fig. 3), which does not appear to be well recovered for magnitudes of slip less than 20 m, it could be argued that we should not make any inferences about the rupture process of the Tohoku earthquake in regions which experienced less than approximately 20 m of slip. We can consider a slip-weighted average rupture velocity,

$$\bar{V}_r = \frac{\Sigma_i U_i \cdot V_{ri}}{\Sigma_i U_i}, \tag{7}$$

where V_{ri} and U_i are the rupture velocity and magnitude of slip on the ith patch, respectively. $\bar{V}_r$ is 3.0 km s^{-1} for the posterior mean kinematic rupture model if we consider all patches and 2.2 km s^{-1} if we only include contributions from patches with more than 20 m of slip. The slip-weighted average rupture velocity decreases when we only include contributions from patches with significant slip, however this should not be construed to imply that rupture velocity is negatively correlated with the amount of slip. In fact, they are not correlated (Fig. S5).

The goodness of fit between the data and the predictions of the mean of our posterior PDF for the kinematic inversion are shown in Figs 4–9. The fits to the static GPS offsets are, in general, excellent

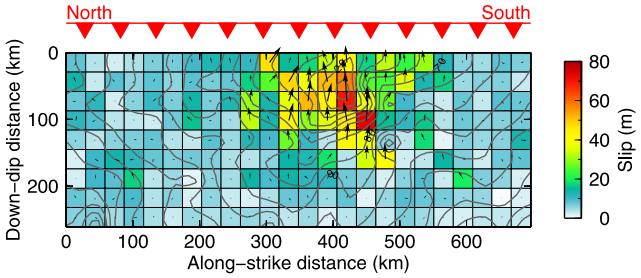

Figure 3. Posterior mean rupture model for kinematic inversion. Background colour is magnitude of the mean scalar slip. Contour lines denote the location of the rupture front in intervals of 10 s. The black vectors show the direction of slip on each patch.

(Fig. 4). The residuals of the fits to the vertical static GPS offsets have a random pattern, indicating that we are not missing any meaningful signals (Fig. 5). However, there is a systematic overprediction of the horizontal displacements at the stations immediately adjacent to the coastline and a systematic underprediction of the horizontal displacements at the more inland stations. The source of this misfit is unclear, but slip on a secondary structure not included in our source model, incompatibility between the GPS data and the near-coast tsunami data, and 3-D complexities in the elastic structure of the Earth (e.g. Hsu *et al.* 2011) are three possible explanations. The misfit to the seafloor geodesy is perhaps larger than would be preferred (Fig. 6), however the residuals associated with our posterior mean slip model (Fig. 7) show a complex spreading signal that may indicate secondary faulting and deformation within the wedge that we would never be able to reproduce with our simple single fault interface model. (These residuals could also possibly be caused by unknown variations in the 3-D elastic structure.) The vertical seafloor deformation is overpredicted at a few stations, but these misfits are within the assumed errors. The low-frequency content of the deep-water tsunami data is also very well fit (Fig. 8), although our tsunami propagation model is clearly not sophisticated enough to reproduce the complex high-frequency signal in the DART data. Finally, the predictions of the mean of the posterior PDF do an excellent job of fitting the near-coast tsunami data (Fig. 8) and 1 Hz GPS data (Fig. 9).

An open question in seismology is whether the high-frequency component of seismic radiation comes from large asperities or (possibly due to anelastic processes) mostly originates from the fringes of the rupture. The location and amplitude of high-frequency radiators are shown in Fig. 10. Consistent with Simons *et al.* (2011) and Meng *et al.* (2011), we again find that the sources of high-frequency energy are systematically located downdip of the regions of significant slip rather than within the asperities themselves. This result agrees with a number of other studies (e.g. Koper *et al.* 2011a,b).

Contours of rupture time are also plotted in Fig. 10, illustrating the effects of the slower rupture velocity in the main asperity. Interestingly, Fig. 11 hints that the rupture first propagated in the along-dip direction, filling the region between the trench and the bottom of the main asperity before expanding significantly along-strike. The rupture propagation proceeded slowly at first, building up slip near the hypocentre region, before expanding bilaterally

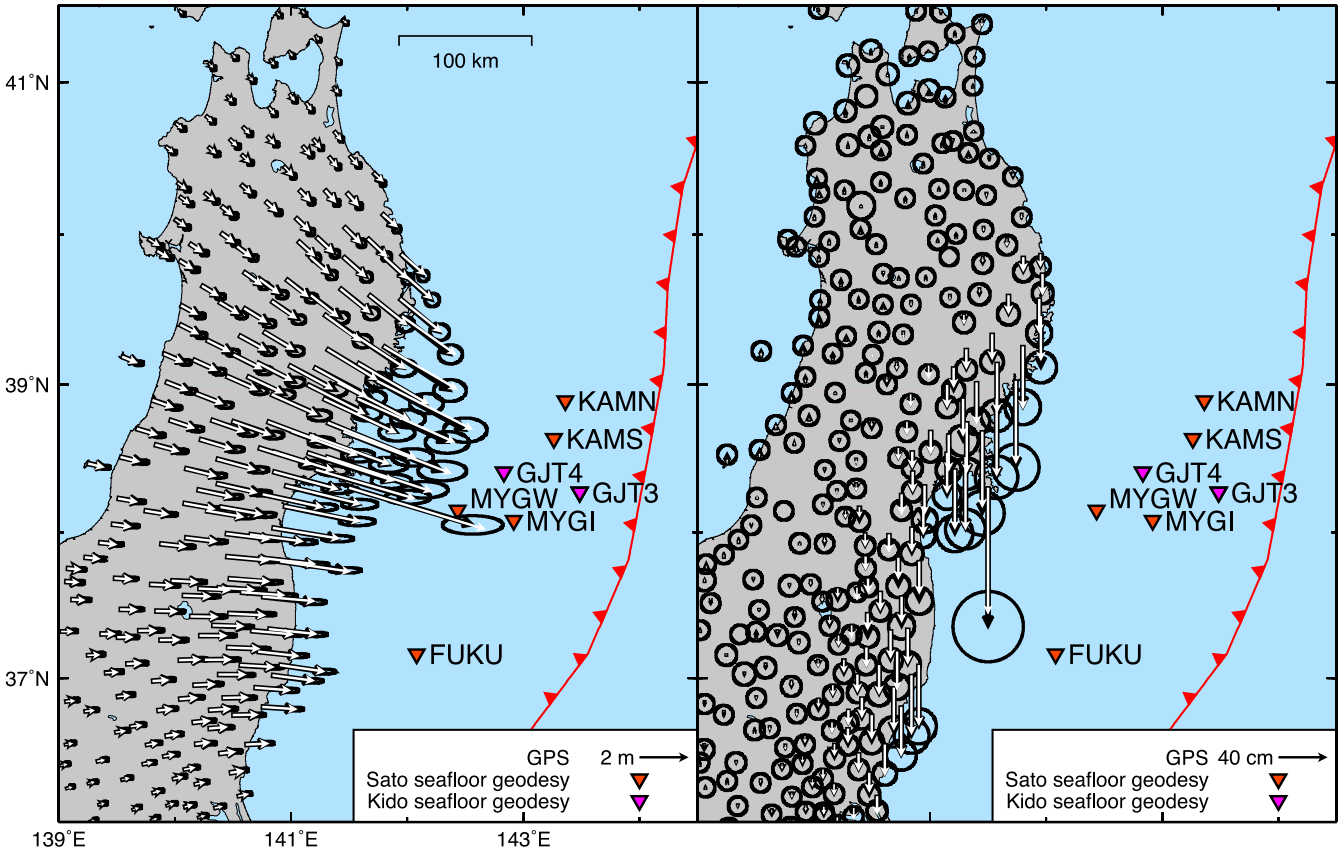

Figure 4. Comparison of observed GPS offsets and predictions from the mean of the posterior PDF for the kinematic inversion. Data are shown in black and predictions are shown in white for horizontal (left-hand panel) and vertical (right-hand panel) deformation. The uncertainty ellipses correspond to 95 per cent confidence bounds of $\mathbf{C}_{\chi}$ (not $\mathbf{C}_{\mathbf{d}}$).

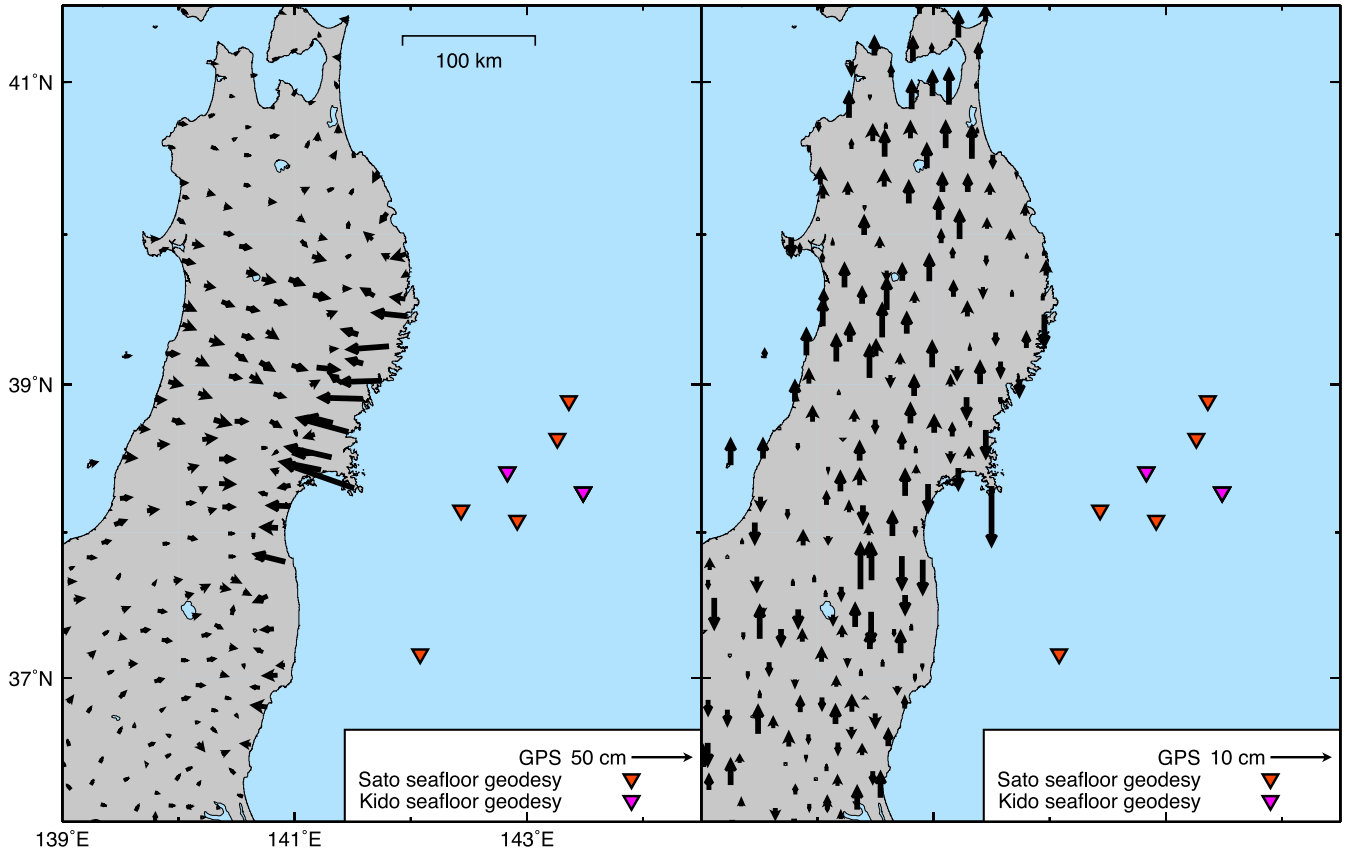

Figure 5. Residuals between observed and predicted displacements for GPS offsets. Left-hand panel: horizontal residuals. Right-hand panel: vertical residuals.

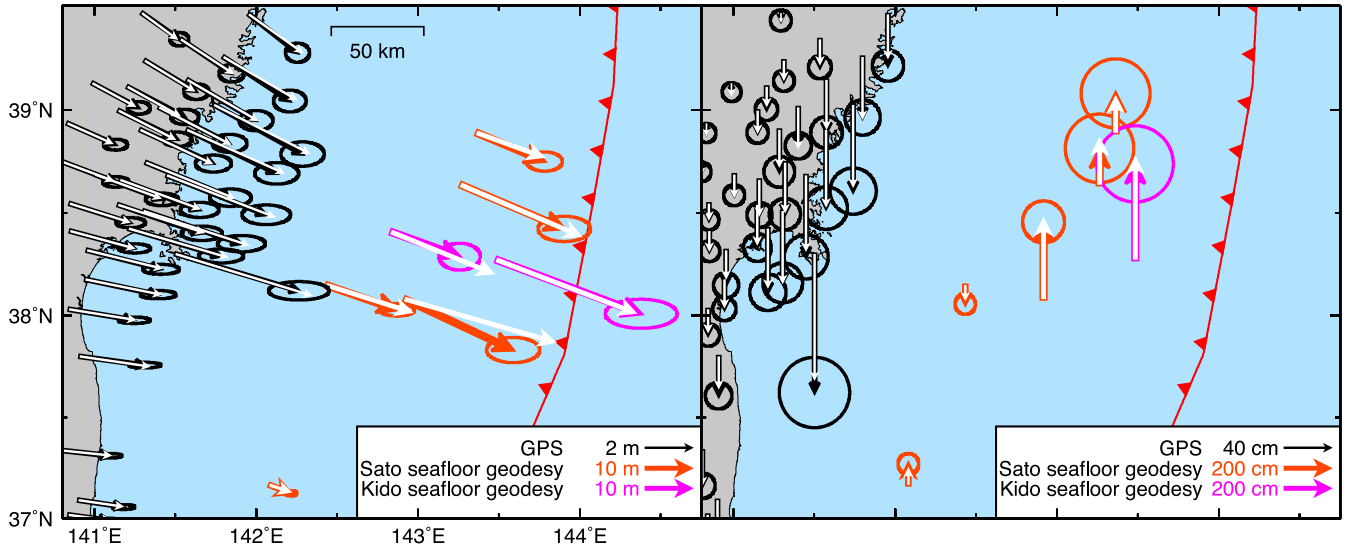

Figure 6. Comparison of observed seafloor geodesy and predictions from the mean of the posterior PDF for the kinematic inversion. Data are shown in colour and predictions are shown in white for horizontal (left-hand panel) and vertical (right-hand panel) deformation. The uncertainty ellipses correspond to 95 per cent confidence bounds of $\mathbf{C}_\chi$ (not $\mathbf{C}_\mathbf{d}$).

along-strike (Fig. 11). From 50 s after the origin time onward, slip proceeds like a bilaterally propagating slip pulse (Heaton 1990). In fact, the later parts of the rupture may be significantly more pulse-like than Fig. 11 suggests because each subplot represents slip integrated over 10 s and our spatial resolution is fundamentally limited by the size of our patches. Thus, at any instant, a much smaller fraction of the fault could be rupturing. In the first 50 s, given our limited spatial resolution, we cannot distinguish whether the rupture behaves like a crack or is simply a slip pulse squeezed into a small spatial extent.

Fig. 12 shows the static three-component seafloor deformation field predicted by the mean of the posterior PDF from the kinematic inversion. The direction and magnitude of our predicted deformation field are consistent with displacements observed by seafloor

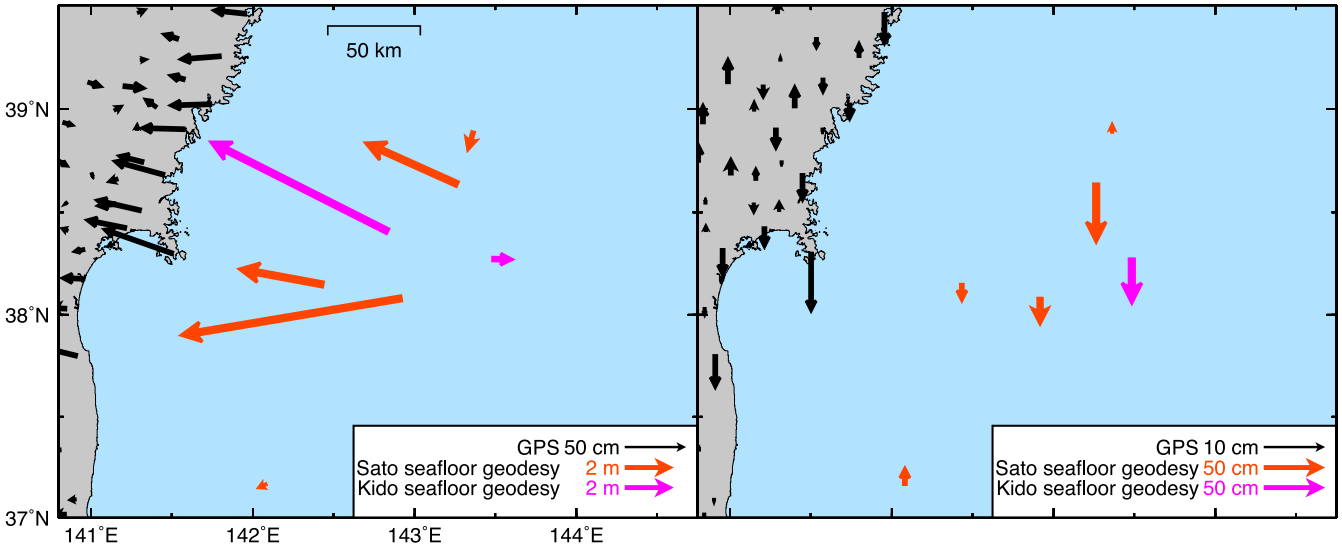

Figure 7. Residuals between observed and predicted displacements for offsets from seafloor geodesy. Left-hand panel: horizontal residuals. Right-hand panel: vertical residuals.

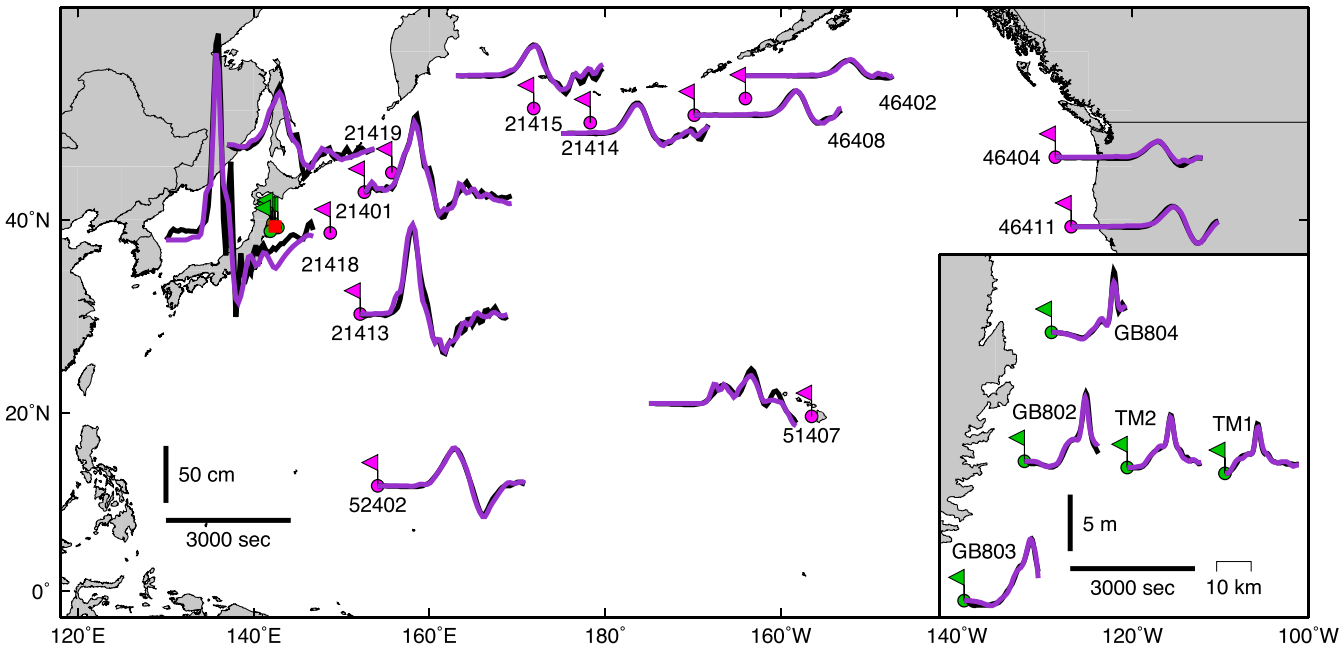

Figure 8. Comparison of observed tsunami records and predictions from the mean of the posterior PDF for the kinematic inversion. Main plot shows fits to DART data (magenta stations) with red square demarcating the region of the inset map with fits to near-coast tsunami records (green stations). Data are shown in black, predictions in purple.

geodesy. (This is as expected because our inversion includes those same data.) The region of greatest seafloor uplift does not extend to the trench, and uplift of any significant magnitude is confined to one small region. However we must note that this seafloor deformation pattern is the result of one slip model (the posterior mean slip) from our posterior PDF.

In Fig. 13, we consider the full posterior ensemble of models by plotting the mean and standard deviation of all samples from all patches as a function of depth to the fault interface. (This is equivalent to plotting the mean and standard deviation for all samples on all patches in a given row of Fig. 3. However, we only include in this calculation contributions from patches with more than 10 m of slip.) The decrease in slip near the trench is a robust feature. Given our limited model resolution, there are no robust variations in slip duration as a function of depth, however rupture velocity is possibly reduced at the depths at which the greatest slip occurred. Although there are large uncertainties for the slip on each patch, the slip integrated over larger spatial areas is robustly constrained. In Fig. 13, we also show the mean and standard deviation of the slip potency integrated along-strike. (This is equivalent to plotting the mean and standard deviation for the sum of all samples on all patches with slip greater than 10 m in a given row of Fig. 3.) The uncertainties associated with the total slip potency integrated along-strike are much smaller than the uncertainties associated with all slip values on all patches along-strike. Thus, if we consider the total slip or slip potency integrated along-strike, the decrease in the total slip near the trench is even more significant. In fact, the posterior PDF completely eliminates the possibility of maximum

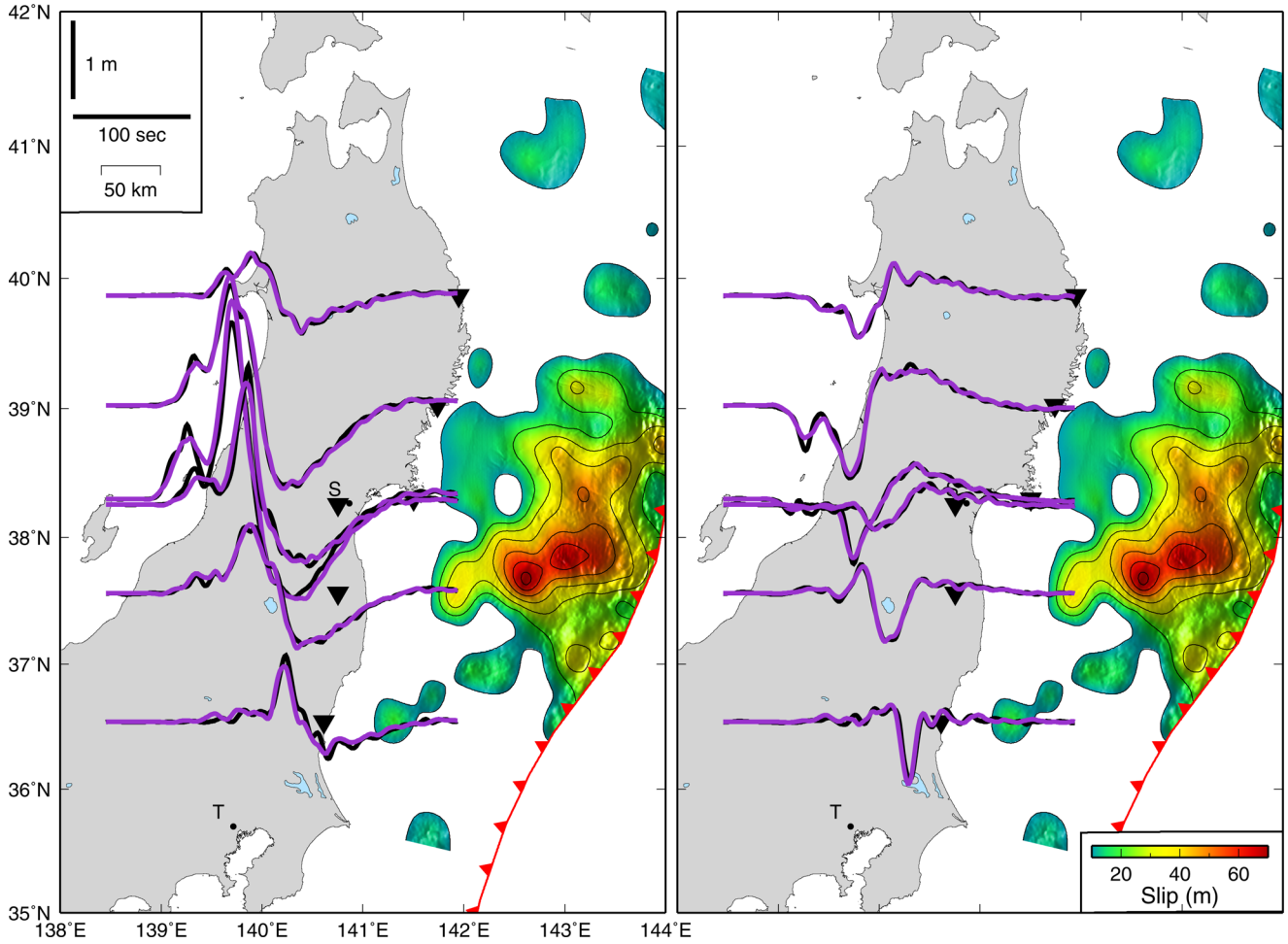

Figure 9. Comparison of observed 1 Hz GPS time-series and prediction from the mean of the posterior PDF for kinematic inversion. Data are plotted in black, predictions in purple. The filtered east and north component of motion is shown in the left and right panels, respectively. Station locations are shown with triangles. The locations of Sendai and Tokyo are marked by S and T, respectively.

slip occurring at the trench with greater than 95 per cent confidence (Fig. 14).

Instead of one optimized slip model we have instead calculated an ensemble of plausible slip models. Thus instead of obtaining one value for scalar seismic moment, stress drop, etc., we can calculate PDFs describing the plausible values for each of these quantities (Fig. 15). It should be noted that merely by including a fault patch or potential deformation source in a model you are allowing for the possibility of slip at that location and, except in the rare cases where a sample includes identically zero slip at that location, it will act to bias our estimates of rupture area upwards, and thus bias our estimates of seismic moment upwards and static stress drop downwards. Therefore, we also compute the posterior PDFs we obtain from only including contributions to each model in the posterior PDF from patches with at least 10 per cent or 20 per cent of the maximum slip in that posterior sample.

While it is straightforward to precisely calculate the slip potency or seismic moment associated with a particular kinematic rupture model, static stress drop is a cruder estimate made by assuming uniform slip on a dip-slip fault of length, L, and width, W, embedded in a Poisson solid,

$$\Delta\sigma = \frac{8}{3\pi}\frac{M_0}{W^2 L}, \tag{8}$$

where M_0 is the scalar seismic moment. Eq. (8) can be rewritten as,

$$\Delta\sigma = \frac{8}{3\pi} M_0 \sqrt{\frac{R}{A^3}}, \tag{9}$$

where A is the rupture area and R is the aspect ratio of the fault ($R = L/W$), which we assume to be 2.

Eq. (9) is valid for faults which reach the surface. If we instead assume that the faulting occurred at great depth, the theoretical static stress drop doubles (Parsons *et al.* 1988) so that,

$$\Delta\sigma = \frac{16}{3\pi} M_0 \sqrt{\frac{R}{A^3}}. \tag{10}$$

We evaluated the maximum slip amplitude, rupture area, slip potency, seismic moment and the static stress drop (using both eqs 9 and 10) for each member of our ensemble of rupture models using the 0, 10 and 20 per cent slip thresholds (Fig. 15). (Here we should note that, when trimming the fault area in accordance with a minimum slip threshold, we only consider the moment released from the trimmed fault and not the total scalar seismic moment of the full rupture model. This choice has particular significance to the calculation of static stress drop.) The maximum slip is fairly well constrained, and we can say with 95 per cent confidence that the

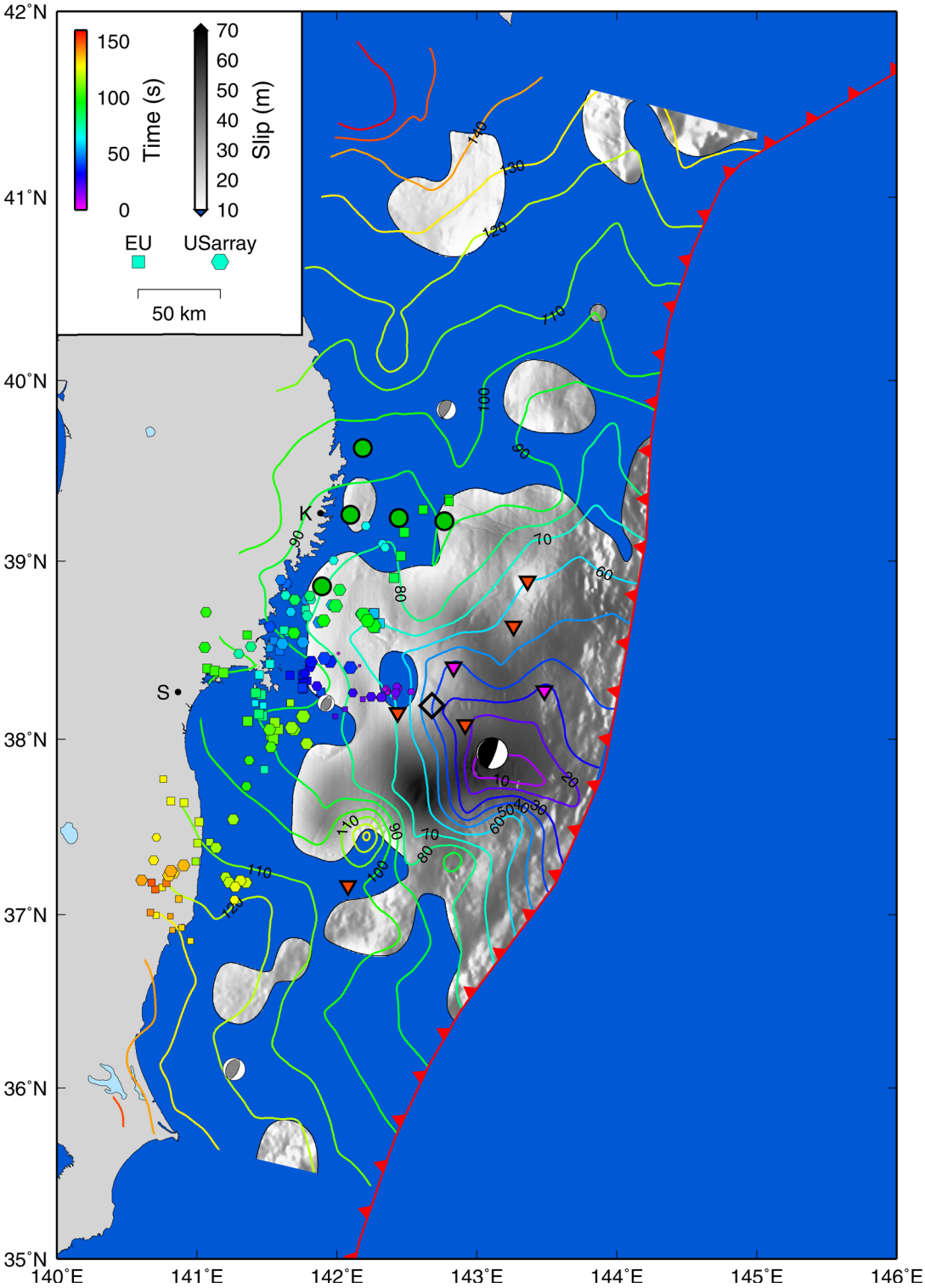

Figure 10. Plot of slip and rupture evolution. Contours of initial rupture time (in seconds after origin time) are plotted in colour above final accumulation of slip plotted in gray. In addition to the W-phase focal mechanism which is shown in black (Duputel *et al.* 2011), F-net broadband seismograph network focal mechanisms for significant thrust aftershocks are plotted in gray. The colored hexagons and squares show the location and magnitude of high-frequency radiators identified with data from USarray and European Union seismic arrays, respectively (Meng *et al.* 2011). All other symbols are as in Fig. 2.

fault slipped in excess of 70 m at its peak. However, our slip values represent slip averaged over the area of one of our fault patches, 841 km^2. For example, if we had used patches which were half this size, we might have found that the maximum slip actually consisted of an area with no slip and an area with twice as much slip. Thus, our maximum slip values should be considered the minimum possible peak slip.

The rupture area, slip potency, and seismic moment are well constrained: there is very little uncertainty associated with a given method for calculating these quantities, although the inferred slip

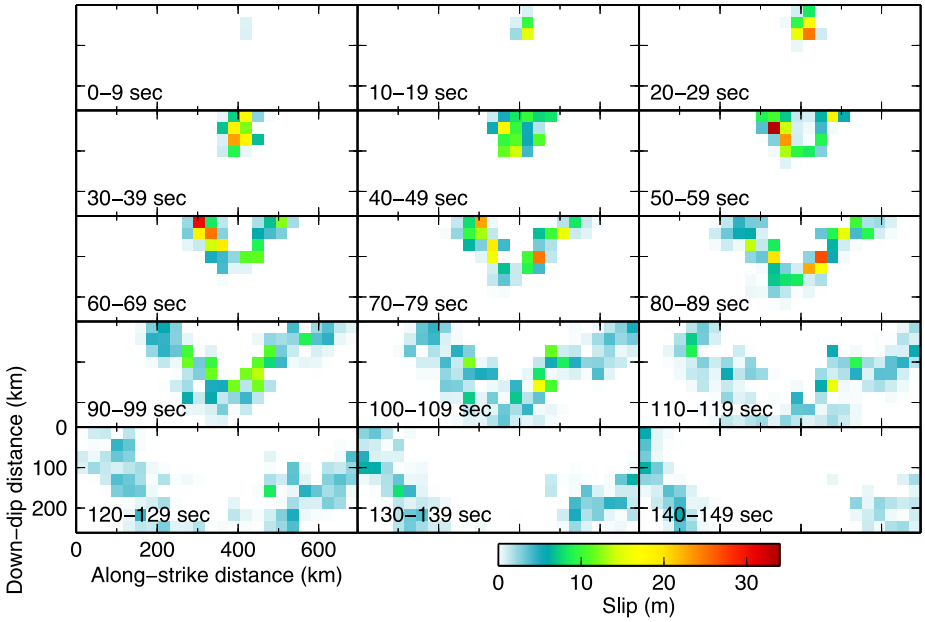

Figure 11. Snapshots of slip on the fault plane from the posterior mean of the joint kinematic-static inversion. All times are relative to the origin time.

potency and seismic moment can change by up to a factor of two if we use scalar sums instead of vector sums. However, given the large uncertainties in how to estimate the scalar stress drop (e.g. eq. 9 versus eq. 10) and the relatively small difference between the stress changes associated with low stress drop versus high stress drop events (only a change of 9 MPa between a 1 MPa low stress drop earthquake and a 10 MPa high stress drop earthquake), it is impossible to conclude whether the Tohoku-oki earthquake was a low stress drop or high stress drop earthquake based on its kinematic rupture model. We emphasize that the uncertainty in static stress drop is not due to uncertainty in the rupture model because the scalar seismic moment and rupture area are both well constrained. (For example, the moment magnitude estimate using vector slip with the 0 per cent of maximum slip threshold is 9.08 ± 0.00089.) Instead, the uncertainty in the scalar stress drop is entirely due to the wide variation in these simplistic and physically unrealistic models for static stress drop (eqs 9 and 10).

The temporal evolution of moment release and its associated uncertainties are presented in Fig. 16. The uncertainties on the source–time function are small and thus we can robustly determine that this earthquake had a very short duration source process. Within the context of the models presented here, we can say with confidence that the entire rupture was over within 150 s. This duration is almost identical to the result in Lay & Kanamori (2011), in which the authors concluded that the great Tohoku-oki earthquake had a much shorter source duration than other magnitude 9+ events such as the 1964 Alaska earthquake or the 2004 Sumatra–Andaman earthquake, although the duration of the Tohoku-oki earthquake is comparable to that of the 2010 M_w 8.8 Maule, Chile earthquake.

In addition to slip, slip duration, and rupture velocity, our kinematic inversion also solves for the hypocentre location constrained to live on the pre-determined fault plane. The prior PDF on the initiation point is a 2-D normal distribution with mean location based on Chu *et al.* (2011) and standard deviations of 50 km each in the along-strike and downdip directions. The location of the rupture initiation point is then sampled along with the rest of the kinematic rupture parameters (Fig. 17). The mean posterior hypocentre location is located near the W-phase centroid (Duputel *et al.* 2011) up-dip of our *a priori* hypocentre from Chu *et al.* (2011) and the relocated hypocentre of Zhao *et al.* (2011). Our hypocentre is calculated in the context of a finite fault kinematic rupture model and thus represents the place at which non-negligible slip first began to accumulate. In contrast, high frequency seismic hypocentres represent the source of the earliest high frequency radiation. Fig. 10 shows that the location of high frequency radiators are located systematically downdip of the regions of major slip. So perhaps it is not surprising that the point of rupture initiation associated with the kinematic rupture model is located updip of the high frequency radiators and seismic hypocentres. We also should note that, since we do not solve for the origin time but instead assume that the origin time associated with the seismic hypocentre is correct, our hypocentre location search might be absorbing errors in the origin time. In Fig. 16, it appears that there is some delay between the origin time and when significant moment is first released. This might indicate that the high frequency seismic origin time is not entirely appropriate to describe the lower frequency kinematic rupture model.

An even fuller understanding of the resolution of each model parameter can be made by comparing the posterior PDF for each model parameter to its prior PDF. As discussed in Paper I, the posterior PDF is only meaningful in the context of the prior PDF. If the data are completely uninformative, then the posterior PDF will be the same as the prior PDF. But as the data add information about the relative plausibility of different models, the posterior PDF diverges from the prior PDF. Thus, a model parameter is only well constrained if its posterior PDF is both highly peaked (in an absolute sense) and more peaked than the prior PDF. From Fig. S6, we see that the data provide significant information about the source process. The posterior PDFs of slip are well constrained in the region of highest slip. Outside of this region, the PDFs on the thrust component of motion ($U_{\|}$) are often highly peaked near zero. But this does not necessarily mean that slip on these patches with small slips are individually well constrained. One side effect of our use of a positivity constraint is that it eliminates the ability of neighbouring patches to trade-off large positive slips against large negative slips. If the average slip over multiple patches is low, then the slip

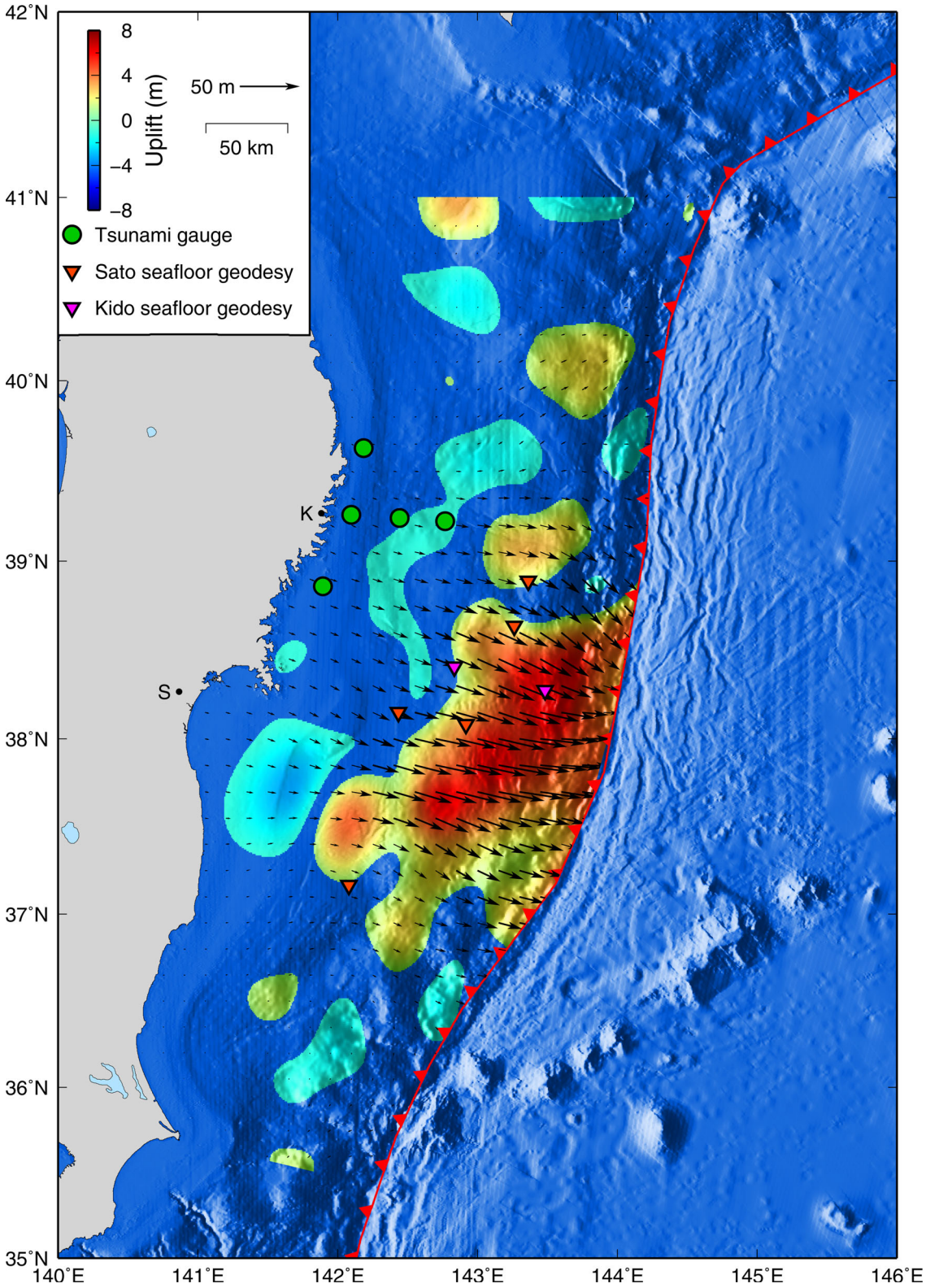

Figure 12. Plot of predicted coseismic seafloor deformation. The static seafloor displacements predicted by the mean of the posterior PDF for the kinematic rupture inversion are shown. Vertical offsets are mapped in colour. Horizontal offsets are shown as a black vector field. The locations of offshore geodetic and tsunami observations are also shown.

on each individual patch will be forced to be near zero so that the total slip across those patches remains small. In fact, the posterior PDFs resemble a Dirichlet distribution (the distribution of random values for one of a set of variables which sum to a particular value) and, as such, the average of slip across these patches should be accurate and robustly determined. We also find that rupture velocity is best resolved in the region of high slip (bottom subplot of Fig. S6).

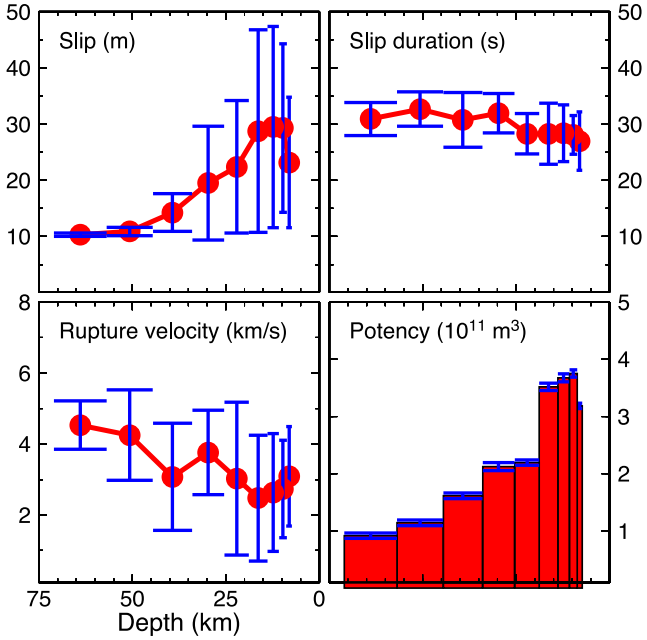

Figure 13. Slip, slip duration, rupture velocity, and slip potency as a function depth. Red symbols and blue error bars are the mean and standard deviation, respectively, of the samples of the posterior PDF from the kinematic inversion. (The ensemble of slip potency is calculated from the total slip summed across all patches along-strike.) Only contributions from patches with more than 10 m of slip were included in calculating the mean and standard deviation.

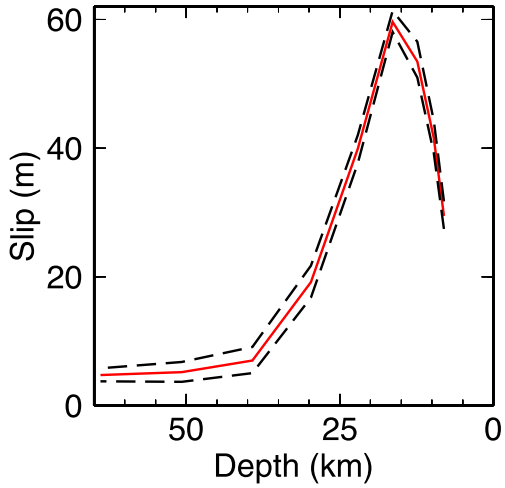

Figure 14. Profile through the largest asperity of slip as a function of depth. Values are averaged over the two columns of patches with highest mean posterior slip. The resulting mean posterior slip is shown in red. The black dashed lines represent 95 per cent credibility intervals calculated from the posterior PDF.

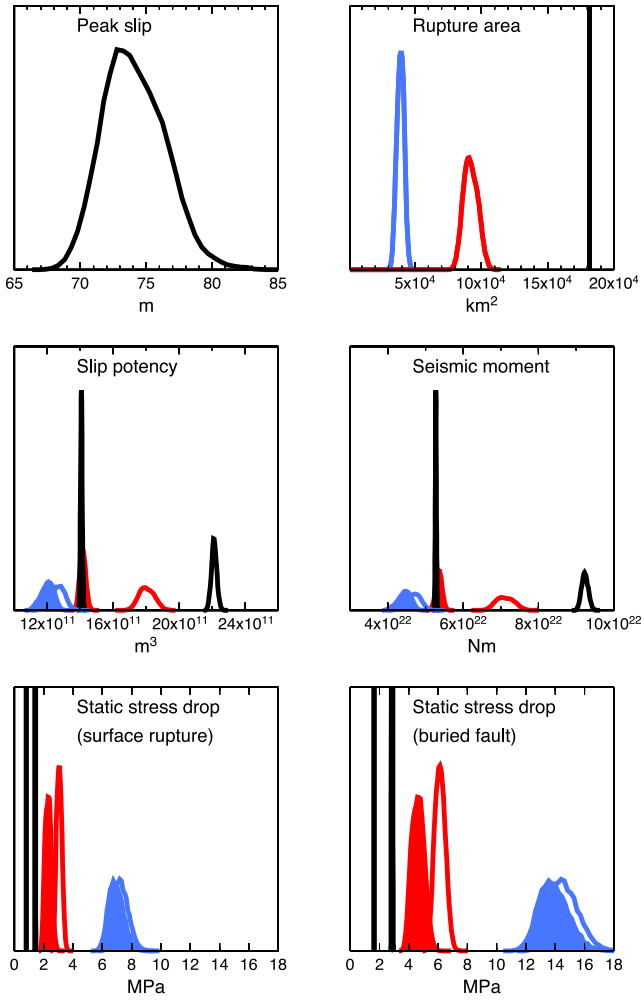

Figure 15. Derived rupture characteristics. The black histograms represent the posterior PDFs of maximum slip, rupture area, slip potency, scalar seismic moment, and static stress drop for the Tohoku-oki earthquake. The red histograms depict the posterior histograms if, for each sample of the posterior PDF, we only include information from patches with slip exceeding 10 per cent of the maximum slip in that rupture scenario. The blue histograms use a threshold of 20 per cent of the maximum slip. Filled histograms represent PDFs computed using the vector sum of slip on each patch; open lines are PDFs derived from the sum of the scalar magnitude of slip on each patch.

4 DISCUSSION

4.1 Spatial extent of peak slip

Our posterior distribution of slip models indicates that the region of maximum slip does not occur at the trench but instead occurs 40–100 km downdip (at a depth of about 10-20 km), tapering updip and downdip, and practically excludes the possibility of peak slip occurring at the trench. This is in contrast to several published slip models which also have slip tapering from the region of highest slip towards the coast, but which differ from our inferences by having slip increase towards the trench with maximum displacements at the trench itself (e.g. Fujii *et al.* 2011; Lay *et al.* 2011; Yamazaki *et al.* 2011; Yoshida *et al.* 2011; Yue & Lay 2011). There are several factors which could contribute to the differences between those models and our inferred slip distribution.

The first issue is regularization. Most of the available regional data for the Tohoku-oki earthquake are located on land. For these data, the downdip edge of the fault is best resolved and the slip towards the trench is not as well constrained. So, in this scenario, when traditional optimization methods with spatial smoothing are used, the optimization might fit the well-constrained low slip under Honshu, fit the higher slips offshore, and then, due to the regularization, extrapolate high slip out to the trench. This behaviour may be what we are seeing in some of the published models for the

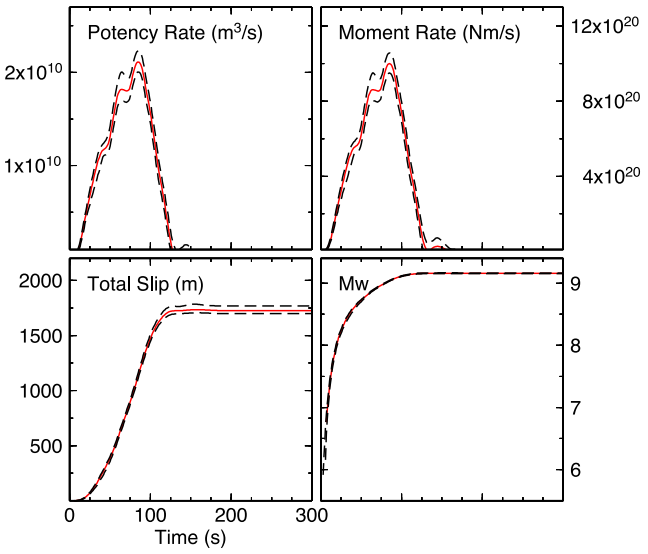

Figure 16. History of moment release. The mean potency rate function, moment rate function, total slip accumulated, and total moment magnitude released are plotted as a function of time with red lines. The black dashed lines represent 95 per cent credibility intervals calculated from the posterior PDF.

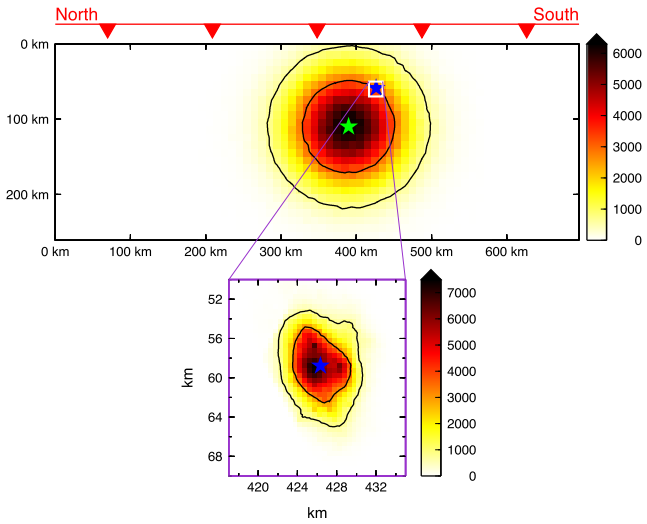

Figure 17. Prior and posterior PDFs on hypocentre location. (Top) 2-D histogram of samples of the prior PDF for the point of rupture initiation on the fault plane expressed as kilometres along-strike and kilometres downdip. The prior PDF is a 2-D normal distribution with mean location based on Chu *et al.* (2011) (green star) and a standard deviation of 50 km in both directions. The mean posterior hypocentre is plotted with a blue star. Contours containing 50 per cent and 90 per cent of the prior samples are shown for reference. (Bottom) 2-D histogram of samples of the posterior PDF for the point of rupture initiation. Symbols are the same as in the top plot. Because the posterior PDF is so much narrower than the prior PDF, the whole fault is shown in the top plot, but only a small section of the fault is shown in the bottom plot. The spatial extent of this zoomed-in region is shown in the top plot by a white box. Background colour indicates the frequency of samples in each 2-D histogram bin. Contours containing 50 and 90 per cent of the posterior samples are shown for reference.

Tohoku-oki earthquake which use traditional optimization methods to solve the inverse problem.

Second, we must consider the effects of errors in the assumed elastic structure. These errors can be enormous and systematic. Previous work has found that homogeneous elastic half-space models tend to put more slip closer to the trench (Lin *et al.* 2013). In contrast, if the elastic moduli were reduced in the toe of the wedge, which may be the case in the Tohoku region (e.g. Zhao *et al.* 2011), then more slip in the shallow fault zone would be required to produce the same surface displacements. In this case, if the Green's functions used in the inversion failed to capture the low elastic moduli, the inversion would tend to infer less slip on the fault surface than actually occurred. Of course, these biases would affect all source modelling, not just ours.

Third, the data utilized to constrain the inversion along with the assumed observational errors and choice of weighting between data sets all shape the inversion's null space. For example, we conducted a test in which we forward predicted the observed data if we extended the high slip region all the way out to the trench. We found that only the DART stations located in the updip direction from the source were significantly affected. (Specifically, additional updip slip predicted waveforms at those stations that were much more complex than the observed waveforms, and the fits to the observations were significantly degraded.)

Not all data equally influence all parts of the source model. When it comes to inferring the shallow slip, we expect the tsunami data to be the strongest influence. Given that our inversion is free from artefacts due to regularization and the fact that the tsunami data (especially the near-field tsunami data) have a high signal-to-noise ratio, we, in turn, expect the greatest errors or uncertainties associated with our inferred shallow slip to be due to errors in our tsunami propagation model.

Using tsunami data requires a host of different choices to be made regarding the propagation model, and in many cases there is not yet consensus in the modelling community as to what the optimal choices are. The tsunami source is the static vertical seafloor displacement (including changes in seafloor height due to horizontal motion of seafloor topography), meaning that the observed tsunami is sensitive to not only the effects of bathymetric variation on tsunami wave propagation but also the 3-D elastic response of the seafloor to deformation on the fault surface. Further, for earthquake ruptures that have a long duration, the tsunami source may need to be staggered in time to reflect the finite time over which the static dislocation of the fault occurs. Like many other modellers, we represent the tsunami source as a static and instantaneous offset. This assumption may be less than optimal (e.g. Satake *et al.* 2013). But, given the relatively short duration and spatially compact nature of the Tohoku-oki earthquake, this is probably the type of earthquake rupture to be least biased by using an instantaneous tsunami source model.

Tsunami propagation models consist of two parts: propagating slip on the fault surface to seafloor deformation, and propagating seafloor deformation through the water. Unlike some models based on using a simple homogeneous elastic half-space model to predict the seafloor deformation, we use a layered elastic space. Although it is unclear whether errors in the elastic structure are a dominant source of error in tsunami propagation models, following the argument of Lin *et al.* (2013), we do know there are significant and systematic variations in the inferred slip model between inversions with and without vertically varying elastic structure, and we know that the mapping between surface deformation and fault slip is a key component of the tsunami modelling process. Since we used a layered elastic structure, we feel confident that the computation of static seafloor displacements as part of our tsunami propagation

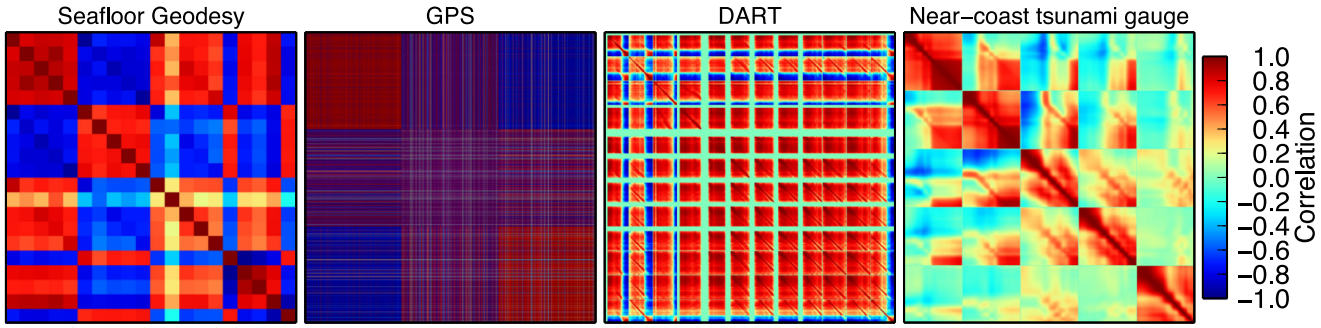

Figure 18. Correlations of data predictions. 500 000 random slip models were produced with the same methodology as is used to seed the CATMIP inversion. The forward predictions from each of these models were then calculated, and the correlation matrices for these 500 000 realizations of random predicted observations are shown.

model is reasonable and probably at least as good as much of the published literature. Further, it should be noted that while tsunamis are excited by changes in seafloor depth, this does not mean that only the vertical part of the 3-D deformation of the seafloor contributes to tsunami generation. Rather, in a megathrust earthquake, the horizontal motion on the shallowly dipping fault interface acts to move the accretionary wedge seaward, changing the bathymetric depth significantly. Satake *et al.* (2013) claim that the horizontal displacement of the accretionary wedge accounts for 20–40 per cent of the tsunami generation during the Tohoku-oki earthquake. Our tsunami propagation model includes these effects. Given our good quality elastic deformation model and inclusion of the effects of horizontal as well as vertical motion to changes in seafloor height, it is, therefore, likely that the dominant source of potential errors lies in the second part of our tsunami propagation model: propagating through the water the gravitational waves due to the static seafloor deformation.

Although our propagation model for the tsunami contains several simplifications as discussed in Section 2.2, we have compensated for the impacts of most of these simplifications by filtering our tsunami Green's functions both in space and time. Thus the largest source of error probably lies in incorrect arrival times due to poorly known bathymetry. However, here we have used the highest resolution bathymetry data set available to us, again minimizing any potential error and helping to make the tsunami propagation model more accurate.

The most direct way to resolve the amount of shallow slip would be to collect data near the trench. There are some observations of seafloor deformation suggesting large near-trench slip, most notably the seafloor geodesy measurements of Ito *et al.* (2011), the change in bathymetry reported by Fujiwara *et al.* (2011), and the seismic imaging of Kodaira *et al.* (2012). However, the Ito *et al.* (2011) data show extreme changes in displacement direction over extremely short distances and at least one station may have been affected by landsliding (Ito, personal communication, 2011). Similarly, the Kodaira *et al.* (2012) data were collected only near the toe of the wedge and these data show substantial landsliding. The Fujiwara *et al.* (2011) bathymetry change data show significant displacement on at least one secondary fault. Thus all of these data appear to be influenced by secondary deformation in the wedge such as landsliding and splay faulting, making it difficult to reliably infer deformation on the megathrust from these observations. Nonetheless, our predicted maximum seafloor displacements in the region of the Fujiwara *et al.* (2011) bathymetric data (about 40 m horizontal displacement and 6 m vertical displacement) are fairly consistent with the Fujiwara *et al.* (2011) estimates of ∼50 m horizontal and ∼7–10 m vertical motion given the uncertainties they reported.

4.2 Empirical effects of errors in the assumed elastic structure

In Paper I and in Section 2.5 of this paper, we discussed the importance of including model prediction errors in the inversion design, and methods for estimating the model prediction error as part of the inversion. Yet in this study we have opted to assign a model prediction error of a specified magnitude rather than determine the optimal strength of the model prediction error as part of the inversion. This was not our original preference. But we found that both our data and our data predictions were so highly correlated that we could not accurately solve for the model prediction error using the formulation in eq. (4). To illustrate this issue, we generated 500 000 random slip models whose strike-slip motions were zero-mean Gaussian (using our prior PDF on $U_{\perp}$) and thrust motions randomly drawn from the Dirichlet distribution (using the methodology discussed in Section 2.6). This sampling produces a set of completely random, spatially rough slip models, where the average slip direction of each model is in the thrust direction. We then computed the predicted observations for each of these slip models, and then calculated the correlations between these 500 000 predicted data sets. The resulting correlation matrices are shown in Fig. 18. These correlation matrices can be in some ways thought of as an empirical Bayesian equivalent to a data resolution matrix (e.g. Menke 2012). Most of the data are highly correlated (or highly anticorrelated) with almost all the rest of the data. The tsunami records, which are waveform data, also have a very strong autoregressive correlation which is evidenced by the fact that the elements of the first few subdiagonals of the correlation matrix are nearly 1. For the five near-coast tsunami records (right-most subplot in Fig. 18), we see what is effectively a 5×5 grid of submatrices. The submatrices along the diagonal represent the temporal correlations within each tsunami record, which are significant. The off-diagonal submatrices, particularly for the first two stations, look very similar to the diagonal correlation submatrices but shifted back in time (shifted off of the diagonal of the submatrix). This type of correlation is to be expected since the records at these stations are nearly identical to each other but shifted in time.

The GPS correlations are a bit harder to interpret in matrix form. If we instead map out the correlation between one component of displacement at one station and all other components of motion at all other stations (Fig. 19), we see that the east components of displacement are highly correlated over a large spatial extent. Similarly, the vertical components of displacement are either highly correlated or highly anticorrelated depending on whether a given station is on the same or the opposite side of the hinge-line as our reference station. This analysis implies that once a fault geometry

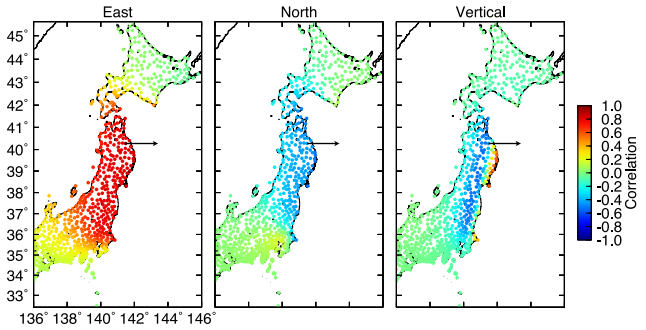

Figure 19. Spatial correlations of GPS data predictions. A subset of the GPS data prediction correlation matrix of Fig. 18 is shown. The coloured circles in the three subplots show the magnitude of the correlation between the east, north, and vertical components (respectively) of predicted displacement at each station with the predicted east component of motion at GEONET station 0903. The observed east component of displacement at station 0903 is plotted with a vector.

and elastic structure are chosen, the shape of the surface deformation field produced is largely known regardless of what slip model you input into that model design. From the standpoint of inverse theory, given the high spatial density of the available data and our source–receiver geometry, our source model will not be able to independently predict the motions at each receiver location. (The equivalent statement in the traditional least-squares approach is to say that our data resolution matrix is far from the identity matrix.) Our data predictions are clearly not independent. Similarly, we can conclude that the prediction errors due to errors in our elastic structure will be highly spatially correlated as well. Therefore, solving for α in eq. (4) will fail or give spurious results because that parametrization of $\mathbf{C_p}$ is missing the spatial correlation between the predictions of our models, and thus inverting for α will result in α being greatly underestimated. This tendency to misestimate the prediction error (due to the lack of spatial covariance in our parametrization of the prediction error) is why we fixed α to 10 per cent. Fixing the model prediction error *a priori* is a poor substitute for accurately estimating the model prediction error with all of its complexity and correlations. However, we at least can force the inversion to recognize that there is a large additional source of error which scales with the strength of the source (and thus the magnitude of the observed surface displacement) and thus the inversion should not fit the data more than is merited given the existence of that additional error source. Future work should explore more sophisticated approaches to estimating $\mathbf{C_p}$, such as the approach proposed by Duputel *et al.* (2014). For now, we limit ourselves to observing that errors in our Earth structure are assumed to be the single largest error source for large earthquakes.

5 CONCLUSIONS

We have added an ensemble of rupture models for the 2011 great Tohoku-oki earthquake to the large and ever-growing literature on this topic. The use of Bayesian analysis produces a distribution of plausible source models which fit the data and our *a priori* knowledge of the physics of the rupture process. There are no non-physical constraints or regularization on the earthquake source process in our inversion. We also include a crude estimate of the effects of errors in our Green's functions to ensure a more reasonable level of fit to the data than would be obtained if only the formal observation uncertainties were included in the inversion design. And, finally, because we used Bayesian inference to solve the inverse problem, we have also obtained a relatively complete description of uncertainties associated with our model parameters.

The possible variation of rupture velocity with depth in Fig 13 is quite intriguing. First, rupture velocity tends to decrease with decreasing depth. This makes intuitive sense since rupture velocity is expected to be tied to the elastic constants of the medium which decrease with shallowing depth. Second, where rupture velocity is well constrained (i.e. in the region of largest slip), it appears to be quite low. This could be misleading since Uchide (2013) suggests that the complexity of the early rupture process might result in the rupture velocity appearing low if the propagation speed is inferred from low-frequency data. However, if our results are correct, then rupture velocity as a function of depth is almost anti-correlated with slip as a function depth. Further, the low rupture velocity in the main asperity gives the Tohoku-oki earthquake many of the features of a classic tsunamigenic earthquake, specifically it is an earthquake rupture with slowly propagating slip concentrated in the upper part of the subduction zone.

Despite the limits of onshore and sparse offshore data to resolve subduction zone earthquakes, there are nonetheless several interesting features of this earthquake rupture which are well constrained. For example, although we cannot distinguish whether the initial phase of the rupture propagated as a slip pulse rather than an expanding crack, the later evolution of the rupture was definitely pulse-like. But perhaps the single most interesting feature of our rupture model is that the location of peak slip was located at depth and not at the trench.

We can only draw conclusions in the context of the data, error structure, and model parametrization we utilized. We would expect any other inversion which either neglects the data from the DART stations located updip of the source or which strongly regularizes the slip distribution (especially through the use of Laplacian smoothing) would tend to overestimate the shallow slip relative to our results. A different parametrization of the model prediction error, one which included spatial and temporal covariances, could have yielded a significantly different posterior PDF. (However, most traditional optimization approaches ignore the model prediction error entirely, which almost inevitably leads to an overly complex model due to overfitting of the data and/or underestimating the uncertainties associated with the inferred model.) We also must acknowledge that, if the elastic moduli near the trench are lower than what is included in our Earth structure, then our inversion might be underestimating the shallow slip on the fault. However, we can say that using all available data with the exception of the Ito *et al.* (2011) and Kodaira *et al.* (2012) data that may have been affected by landsliding and the Fujiwara *et al.* (2011) data that suggest secondary faulting, in concert with the 1-D Earth structure in Table 1 and even allowing for a 10 per cent error in the Green's functions calculated from that elastic structure, we robustly infer that the maximum slip occurred downdip of the trench.

There have been several important contributions made to seismic source studies using Bayesian or approximately Bayesian methods, (e.g. Fukuda & Johnson 2008; Monelli & Mai 2008; Hooper *et al.* 2013). However, this study is the first to use Bayesian methods to determine the complete posterior PDF for a kinematic finite fault model of an actual earthquake without any non-physical regularization constraints. Our solution represents the ensemble of plausible source models which fit a wide-variety of geophysical observations given realistic uncertainties on those observations, constraints based on our *a priori* knowledge of the physics of the rupture process, and a model for the errors in our forward model. This last point is perhaps the most important.

By removing all non-physical regularization, we can be assured that our rupture models are free from any artefacts due to regularization. An artefact is any feature of the source model that appears to be well constrained but is due to constraints of the inversion process and are not required by either the data or the physical *a priori* constraints on the source model. Thus we must note that, if our representation of the uncertainties in our observations and forward model is flawed, we can still have spurious features in our source model or, equivalently, have some parts of the source model appear to be better constrained than they are. Despite our rather simplistic parametrization of the error in our forward model ($\mathbf{C_p}$), we infer significant uncertainty in the rupture process, something that we intuitively know must exist when we attempt to constrain the spatial and temporal evolution of offshore slip given mostly land-based observations. Despite these uncertainties, there are many parts of the rupture process that are well constrained including the maximum slip (which must be large given that our maximum slip value is a minimum constraint based on averaging slip over more than 800 km^2), the total slip and rupture area (and thus, in turn, the slip potency, scalar seismic moment, and moment magnitude are all well constrained), and the fact that the slip appears to peak downdip of the trench.

ACKNOWLEDGEMENTS

The CATMIP MCMC simulations were executed on NASA's Pleiades supercomputer. This work is supported by the National Science Foundation through grant number EAR-0941374 and is Caltech Seismological Laboratory contribution 10087.

REFERENCES

Chu, R., Wei, S., Helmberger, D. & Zhan, Z., 2011. Initiation of the great Mw 9.0 Tohoku–Oki earthquake, *Earth planet. Sci. Lett.*, **308**, 277–283.

Duputel, Z., Rivera, L., Kanamori, H., Hayes, G., Hirshorn, B. & Weinstein, S., 2011. Real-time W phase inversion during the 2011 off the Pacific coast of Tohoku earthquake, *Earth, Planets Space*, **63**(7), 535–539.

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**(1), 464–482.

Fujie, G., Kasahara, J., Hino, R., Sato, T., Shinohara, M. & Suyehiro, K., 2002. A significant relation between seismic activities and reflection intensities in the Japan Trench region, *Geophys. Res. Lett.*, **29**(7), 1100, doi:10.1029/2001GL013764.

Fujii, Y., Satake, K., Sakai, S., Shinohara, M. & Kanazawa, T., 2011. Tsunami source of the 2011 off the Pacific coast of Tohoku earthquake, *Earth, Planets space*, **63**(7), 815–820.

Fujiwara, T., Kodaira, S., No, T., Kaiho, Y., Takahashi, N. & Kaneda, Y., 2011. The 2011 Tohoku-Oki earthquake: displacement reaching the trench axis, *Science*, **334**(6060), 1240–1240.

Fukuda, J. & Johnson, K., 2008. A fully Bayesian inversion for spatial distribution of fault slip with objective smoothing, *Bull. seism. Soc. Am.*, **98**(3), 1128–1146.

Gelman, A., Carlin, J., Stern, H. & Rubin, D., 2004. *Bayesian Data Analysis*, Chapman and Hall.

Hasegawa, A., Nakajima, J., Umino, N. & Miura, S., 2005. Deep structure of the northeastern Japan arc and its implications for crustal deformation and shallow seismic activity, *Tectonophysics*, **403**(1), 59–75.

Heaton, T.H., 1990. Evidence for and implications of self-healing pulses of slip in earthquake rupture, *Phys. Earth planet. Inter.*, **64**(1), 1–20.

Hooper, A. *et al.*, 2013. Importance of horizontal seafloor motion on tsunami height for the 2011 Mw = 9.0 Tohoku-Oki earthquake, *Earth planet. Sci. Lett.*, **361**, 469–479.

Hsu, Y.-J., Simons, M., Williams, C. & Casarotti, E., 2011. Three-dimensional FEM derived elastic green's functions for the coseismic deformation of the 2005 Mw 8.7 Nias-Simeulue, Sumatra earthquake, *Geochem. Geophys. Geosyst.*, **12**(7), Q07013, doi:10.1029/2011GC003553.

Ito, A., Fujie, G., Tsuru, T., Kodaira, S., Nakanishi, A. & Kaneda, Y., 2004. Fault plane geometry in the source region of the 1994 Sanriku-oki earthquake, *Earth planet. Sci. Lett.*, **223**(1–2), 163–175.

Ito, Y. *et al.*, 2011. Frontal wedge deformation near the source region of the 2011 Tohoku-Oki earthquake, *Geophys. Res. Lett.*, **38**(15), L00G05, doi:10.1029/2011GL048355.

Iwasaki, T. *et al.*, 1994. Precise p and s wave velocity structures in the Kitakami massif, northern Honshu, Japan, from a seismic refraction experiment, *J. geophys. Res.*, **99**(B11), 22 187–22 204.

Katsumata, K., Wada, N. & Kasahara, M., 2003. Newly imaged shape of the deep seismic zone within the subducting Pacific plate beneath the Hokkaido corner, Japan-Kurile arc-arc junction, *J. geophys. Res.*, **108**(B12), 2565, doi:10.1029/2002JB002175.

Kido, M., Osada, Y., Fujimoto, H., Hino, R. & Ito, Y., 2011. Trench-normal variation in observed seafloor displacements associated with the 2011 Tohoku-Oki earthquake, *Geophys. Res. Lett.*, **38**(24), L24303, doi:10.1029/2011GL050057.

Kodaira, S. *et al.*, 2012. Coseismic fault rupture at the trench axis during the 2011 Tohoku-oki earthquake, *Nat. Geosci.*, **5**(9), 646–650.

Koper, K. D., Hutko, A. R. & Lay, T., 2011a. Along-dip variation of teleseismic short-period radiation from the 11 March 2011 Tohoku earthquake (Mw 9.0), *Geophys. Res. Lett.*, **38**, L21309, doi:10.1029/2011GL049689.

Koper, K.D., Hutko, A.R., Lay, T., Ammon, C.J. & Kanamori, H., 2011b. Frequency-dependent rupture process of the 2011 Mw 9.0 Tohoku earthquake: comparison of short-period P wave backprojection images and broadband seismic rupture models, *Earth, Planets Space*, **63**(7), 599–602.

Lay, T. & Kanamori, H., 2011. Insights from the great 2011 Japan earthquake, *Phys. Today*, **64**(12), 33–39.

Lay, T., Ammon, C.J., Kanamori, H., Xue, L. & Kim, M.J., 2011. Possible large near-trench slip during the 2011 Mw 9.0 off the Pacific coast of Tohoku earthquake, *Earth, Planets Space*, **63**(7), 687–692.

Lévěque, J., Rivera, L. & Wittlinger, G., 1993. On the use of the checker-board test to assess the resolution of tomographic inversions, *Geophys. J. Int.*, **115**(1), 313–318.

Lin, Y.-n.N. *et al.*, 2013. Coseismic and postseismic slip associated with the 2010 maule earthquake, chile: characterizing the arauco peninsula barrier effect, *J. geophys. Res.: Solid Earth*, **118**(6), 3142–3159.

Liu, P.L.-F., Woo, S.-B. & Cho, Y.-S., 1998. *Computer programs for tsunami propagation and inundation*, Tech. rep., Cornell University.

Meng, L., Inbal, A. & Ampuero, J.-P., 2011. A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-Oki earthquake, *Geophys. Res. Lett.*, **38**(7), L00G07, doi:10.1029/2011GL048118.

Menke, W., 2012. *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & Teller, E., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.*, **21**(6), 1087–1092.

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.

Miura, S., Kodaira, S., Nakanishi, A., Tsuru, T., Takahashi, N., Hirata, N. & Kaneda, Y., 2003. Structural characteristics controlling the seismicity crustal structure of southern Japan trench fore-arc region, revealed by ocean bottom seismographic data, *Tectonophysics*, **363**(1), 79–102.

Miura, S., Takahashi, N., Nakanishi, A., Tsuru, T., Kodaira, S. & Kaneda, Y., 2005. Structural characteristics off miyagi forearc region, the Japan trench seismogenic zone, deduced from a wide-angle reflection and refraction study, *Tectonophysics*, **407**(3), 165–188.

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

Parsons, I., Hall, J. & Lyzenga, G., 1988. Relationships between the average offset and the stress drop for two- and three-dimensional faults, *Bull. seism. Soc. Am.*, **78**(2), 931–915.

Satake, K., Fujii, Y., Harada, T. & Namegaya, Y., 2013. Time and space distribution of coseismic slip of the 2011 Tohoku earthquake as inferred from tsunami waveform data, *Bull. seism. Soc. Am.*, **103**(2B), 1473–1492.

Sato, M., Ishikawa, T., Ujihara, N., Yoshida, S., Fujita, M., Mochizuki, M. & Asada, A., 2011. Displacement above the hypocenter of the 2011 Tohoku-Oki earthquake, *Science*, **332**(6036), 1395.

Simons, M. *et al.*, 2011. The 2011 magnitude 9.0 Tohoku-Oki earthquake: Mosaicking the megathrust from seconds to centuries, *Science*, **332**(6036), 1421–1425.

Takahashi, N. *et al.*, 2004. Seismic structure and seismogenesis off Sanriku region, northeastern Japan, *Geophys. J. Int.*, **159**(1), 129–145.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM.

Tsai, V.C., Ampuero, J.-P., Kanamori, H. & Stevenson, D.J., 2013. Estimating the effect of Earth elasticity and variable water density on tsunami speeds, *Geophys. Res. Lett.*, **40**(3), 492–496.

Tsuru, T., Park, J.-O., Takahashi, N., Kodaira, S., Kido, Y., Kaneda, Y. & Kono, Y., 2000. Tectonic features of the Japan trench convergent margin off Sanriku, northeastern Japan, revealed by multichannel seismic reflection data, *J. geophys. Res.: Solid Earth*, **105**(B7), 16 403–16 413.

Uchide, T., 2013. High-speed rupture in the first 20 s of the 2011 Tohoku earthquake, Japan, *Geophys. Res. Lett.*, **40**(12), 2993–2997.

Watada, S., 2013. Tsunami speed variations in density-stratified compressible global oceans, *Geophys. Res. Lett.*, **40**(15), 4001–4006.

Watada, S., Kusumoto, S. & Satake, K., 2014. Traveltime delay and initial phase reversal of distant tsunamis coupled with the self-gravitating elastic Earth, *J. geophys. Res.: Solid Earth*, **119,** doi:10.1002/2013JB010841.

Yagi, Y. & Fukahata, Y., 2011. Introduction of uncertainty of Green's function into waveform inversion for seismic source processes, *Geophys. J. Int.*, **186**(2), 711–720.

Yamamoto, Y., Hino, R., Nishino, M., Yamada, T., Kanazawa, T., Hashimoto, T. & Aoki, G., 2006. Three-dimensional seismic velocity structure around the focal area of the 1978 Miyagi-Oki earthquake, *Geophys. Res. Lett.*, **33**(10), L10308, doi:10.1029/2005GL025619.

Yamazaki, Y., Lay, T., Cheung, K., Yue, H. & Kanamori, H., 2011. Modeling near-field tsunami observations to improve finite-fault slip models for the 11 March 2011 Tohoku earthquake, *Geophys. Res. Lett.*, **38**(7), L00G15, doi:10.1029/2011GL049130.

Yoshida, K., Miyakoshi, K. & Irikura, K., 2011. Source process of the 2011 off the Pacific coast of Tohoku earthquake inferred from waveform inversion with long-period strong-motion records, *Earth, Planets Space*, **63**(7), 577–582.

Yue, H. & Lay, T., 2011. Inversion of high-rate (1 sps) gps data for rupture process of the 11 March 2011 Tohoku earthquake (Mw 9.1), *Geophys. Res. Lett.*, **38**, L00G09, doi:10.1029/2011GL048700.

Zhang, H., Thurber, C.H., Shelly, D., Ide, S., Beroza, G.C. & Hasegawa, A., 2004. High-resolution subducting-slab structure beneath northern Honshu, Japan, revealed by double-difference tomography, *Geology*, **32**(4), 361–364.

Zhao, D., Huang, Z., Umino, N., Hasegawa, A. & Kanamori, H., 2011. Structural heterogeneity in the megathrust zone and mechanism of the 2011 Tohoku-oki earthquake (Mw 9.0), *Geophys. Res. Lett.*, **38**(17), L17308, doi:10.1029/2011GL048408.

Zhao, H., 2005. A fast sweeping method for eikonal equations, *Math. Comput.*, **74**(250), 603–628.

Zhu, L. & Rivera, L., 2002. A note on the dynamic and static displacements from a point source in multilayered media, *Geophys. J. Int.*, **148**(3), 619–627.

SUPPORTING INFORMATION

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

Supplementary figures and a description of the supplementary tables and files containing the rupture model. Zip archive containing the supplementary tables and rupture model files as described in the PDF. (http://gji.oxfordjournals.org/lookup/suppl/doi:10.1093/gji/ggu170/-/DC1)

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.