Geophysical Journal International

Geophys J. Int. (2014) **196,** 1564–1579
Advance Access publication 2014 January 7

doi: 10.1093/gji/ggt406

A method for the joint inversion of geodetic and seismic waveform data using ABIC: application to the 1997 Manyi, Tibet, earthquake

Gareth J. Funning,[1] Yukitoshi Fukahata,[2] Yuji Yagi[3] and Barry Parsons[4]

[1] *Department of Earth Sciences, University of California, Riverside, CA* 92521, *USA. E-mail: gareth@ucr.edu*
[2] *Disaster Prevention Research Institute, Kyoto University, Uji, Kyoto,* 611-0011, *Japan*
[3] *Graduate School of Life and Environmental Sciences, University of Tsukuba, Tennodai* 1-1-1, *Tsukuba* 305-0006, *Japan*
[4] *COMET, Department of Earth Sciences, University of Oxford, Parks Road, Oxford,* OX1 3PR, *UK*

Accepted 2013 October 2. Received 2013 October 1; in original form 2013 March 25

SUMMARY

Geodetic imaging data and seismic waveform data have complementary strengths when considering the modelling of earthquakes. The former, particularly modern space geodetic techniques such as Interferometric Synthetic Aperture Radar (InSAR), permit high spatial density of observation and thus fine resolution of the spatial pattern of fault slip; the latter provide precise and accurate timing information, and thus the ability to resolve how that fault slip varies over time. In order to harness these complementary strengths, we propose a method through which the two data types can be combined in a joint inverse model for the evolution of slip on a specified fault geometry. We present here a derivation of Akaike's Bayesian Information Criterion (ABIC) for the joint inversion of multiple data sets that explicitly deals with the problem of objectively estimating the relative weighting between data sets, as well as the optimal influence of model smoothness constraints in space and time.

We demonstrate our ABIC inversion scheme by inverting InSAR displacements and teleseismic waveform data for the 1997 Manyi, Tibet, earthquake. We test, using a simplified fault geometry, three cases—InSAR data inverted alone, vertical component teleseismic broad-band waveform data inverted alone and a joint inversion of both data sets. The InSAR-only model and seismic-only model differ significantly in the distribution of slip on the fault plane that they predict. The joint-inversion model, however, has not only a similar distribution of slip and fit to the InSAR data in the InSAR-only model, suggesting that those data provide the stronger control on the pattern of slip, but is also able to fit the seismic data at a minimal degradation of fit when compared with the seismic-only model. The rupture history of the preferred, joint-inversion model, indicates bilateral rupture for the first 20 s of the earthquake, followed by a further 25 s of westward unilateral rupture afterwards, with slip peaking at 7 m in the upper 6 km of the fault. This joint-inversion approach is thus shown to be a viable method for the study of large shallow continental earthquakes, and may be of particular benefit in cases where near-field seismic observations are not available.

Key words: Inverse theory; Satellite geodesy; Radar interferometry; Earthquake source observations; Body waves.

1 INTRODUCTION

Space geodetic imaging [e.g. Interferometric Synthetic Aperture Radar (InSAR) and satellite image correlation] can provide precise and spatially dense observations of the displacement of the surface during earthquakes. Such observations place strong constraints on elastic dislocation models of the earthquake source, both in terms of the geometry of the fault responsible (e.g. from the orientation and degree of asymmetry in the surface displacement pattern), and the distribution of slip on that fault (e.g. from the amplitude and location of peaks in the deformation signal).

Where such data can often be lacking is in their temporal resolution—the ability to distinguish the deformation due to individual events out of a sequence. In most cases and in most areas, the repeat interval between satellite image acquisitions is several weeks at a minimum, and in a worst case scenario, for example, where regular background acquisitions have not been tasked, can be as long as several years. A major limitation of geodetic imaging methods for the study of earthquakes is that they measure the total displacement that occurs between observation epochs. We can never know from such data alone the time-evolution of earthquake slip (which occurs on timescales of seconds to minutes), or the order of

occurrence of subevents of an earthquake sequence, only the total slip distribution, and so the amount of information gained about earthquake dynamics is necessarily limited.

Seismic waveform data collected in the aftermath of an earthquake, on the other hand, have excellent temporal resolution and accuracy—at Global Seismographic Network (GSN) stations, for example, the timing of phase arrivals is designed to be accurate to 0.01 s, at least an order of magnitude smaller than the periods of interest. If the full waveforms are modelled for a globally distributed network of stations, it is possible to place constraints on fault geometry and slip, and in particular, the temporal evolution of that slip.

Conversely, although the ability of seismic observations to constrain temporal changes is greatly superior to that given by geodesy, there are reasons to consider estimates of the total slip occurring in an earthquake from seismic methods as more uncertain—foremost that in seismic data inversions, the observed data, which are ground displacement or velocity measurements, relate directly to slip velocities or accelerations on the fault, which must then be integrated in order to obtain slip distance. This can result in non-uniqueness in the model solution (e.g. Clévédé *et al.* 2004; Vallée & Bouchon 2004). Another area where seismic data, particularly teleseismic data, may be lacking is in their spatial resolution. Given the large source–receiver distances involved and the possibility of unmodelled mantle heterogeneity along seismic ray paths, it is perhaps not surprising that the spatial resolving power of teleseismic data is generally lower than that of space geodetic data. In finite-fault slip inversions that employ body wave data, near-field strong motion seismic observations are often used to place stronger constraints on the spatial slip pattern (e.g. Yagi *et al.* 2004). In many cases, however, particularly in more remote, uninhabited and/or inaccessible areas, such near-field data are not available.

Thus geodetic imaging data and seismic data both give information on earthquake slip, in terms of its magnitude, its location and the geometry of the fault it occurs on, but each data type has different advantages—geodetic imaging data being more sensitive to fault geometry and the spatial pattern of total fault slip, seismic data more sensitive to the temporal evolution of slip during an earthquake. Given these complementary strengths, it naturally follows that one means of mitigating such weaknesses could be to invert both data sets simultaneously. In so doing, an earthquake model could be produced that has both a strong control on the spatial pattern of slip, but also contains information on the propagation of the earthquake rupture with time.

A number of joint inversions of various types of seismic and geodetic data have been presented in the literature, for example, GPS, teleseismic body wave and strong motion data (Wald & Heaton 1994; Wu *et al.* 2001), GPS and regional seismic waveform data (Rolandone *et al.* 2006), InSAR and teleseismic waveform data (Salichon *et al.* 2003), InSAR, GPS and strong motion data (Kim & Dreger 2008), and InSAR, GPS, teleseismic body wave and strong motion data (Delouis *et al.* 2002). Such 'true' joint inversions, where two or more data sets are inverted simultaneously and are each permitted to influence the final solution, have several advantages over methods where geodetic data are used as some form of constraint on a single data set inversion, for example, by solving first for a slip model using the geodetic data and using this as a constraint on a seismic waveform inversion (e.g. Hernandez *et al.* 1999), by using geodetic data as a means to select between a series of slip models that fit seismic waveform data approximately equally well (e.g. Wen & Ma 2010), or by employing GPS data as an explicit prior constraint on a waveform inversion of near-field strong motion data (e.g. Asano *et al.* 2005). By forcing matching of one data type or the result of an inversion of one data set on an inversion of another, only a restricted portion of model parameter space is searched, and the resulting model may not be the best explanation of both data sets when considered together.

In each of the true joint inverse modelling studies mentioned, the relative weighting assigned to each data set strongly affects the final solution. The methods used for estimating such relative weights in these studies range from trial and error, to assignment of uniform weights, to observations of the change to a given misfit statistic as each weighting parameter is varied systematically, methods which may be limited in terms of objectivity, or in their consideration of trade-offs between the relative weights of various data sets. Given the influence that these parameters have on the final model solution, it is clear that an objective criterion for the estimation of relative weighting is both desirable and preferable. In this study, we present a new derivation of one such objective method for the relative weighting of two data sets in a joint inversion, based upon minimization of Akaike's Bayesian Information Criterion (ABIC). We apply it, using a simplified, fixed fault geometry, to InSAR and broad-band teleseismic waveform data sets from a large strike-slip earthquake in northern Tibet, the 1997 Manyi event, for which no near-field seismic, GPS or field geological data are available.

2 MATHEMATICAL FORMULATION

We present the derivation for our joint-inversion scheme. Note that while this has many similarities with the methods used for the individual data set inversions that we later perform to compare with the joint inversion (Sections 3.2 and 3.3), there are some differences in detail (e.g. in the definitions of data covariance, errors, prior information and hyperparameters). For details on individual data set inversions, the reader is referred to the relevant previous studies (Yabuki & Matsu'ura 1992; Fukahata *et al.* 2003, 2004; Funning 2005).

2.1 Model setup and the observation equations

We consider first the simple case in which geodetic observations are used to infer the distribution of the total slip that occurred in an earthquake on a fixed rectangular fault geometry subdivided into smaller subfault patches, each of which slips in the same, fixed direction (i.e. rake is fixed). With I along-strike and J downdip divisions, there will be a total of $M_{\mathrm{geod}} = IJ$ patches (Fig. 1a). We will assume that slip on these fault patches can adequately account for the surface deformation associated with the earthquake. We further assume that such deformation can be measured geodetically, for instance as InSAR range-change measurements in the line-of-sight of a radar satellite. Let $\mathbf{d_{geod}}$ represent a column vector of N_{geod} such geodetic displacement observations, and let us assume that these displacements can be related to slip on fault patches by the relationship:

$$\mathbf{d}_{\mathrm{geod}} = \mathbf{H}_{\mathrm{geod}}\mathbf{a}_{\mathbf{total}} + \mathbf{e}_{\mathrm{geod}}, \tag{1}$$

where $\mathbf{H_{geod}}$ is a $N_{\mathrm{geod}} \times M_{\mathrm{geod}}$ kernel, or matrix of Green's functions, which relates the slip of each individual fault patch to the predicted displacements due to that slip at each observation location, typically by an elastic dislocation model calculation, $\mathbf{a_{total}}$ is a vector containing the estimated total slip during the earthquake of each fault patch and $\mathbf{e_{geod}}$ is a vector comprising the uncertainties of the terms in $\mathbf{d_{geod}}$. We assume that the terms in $\mathbf{e_{geod}}$ follow

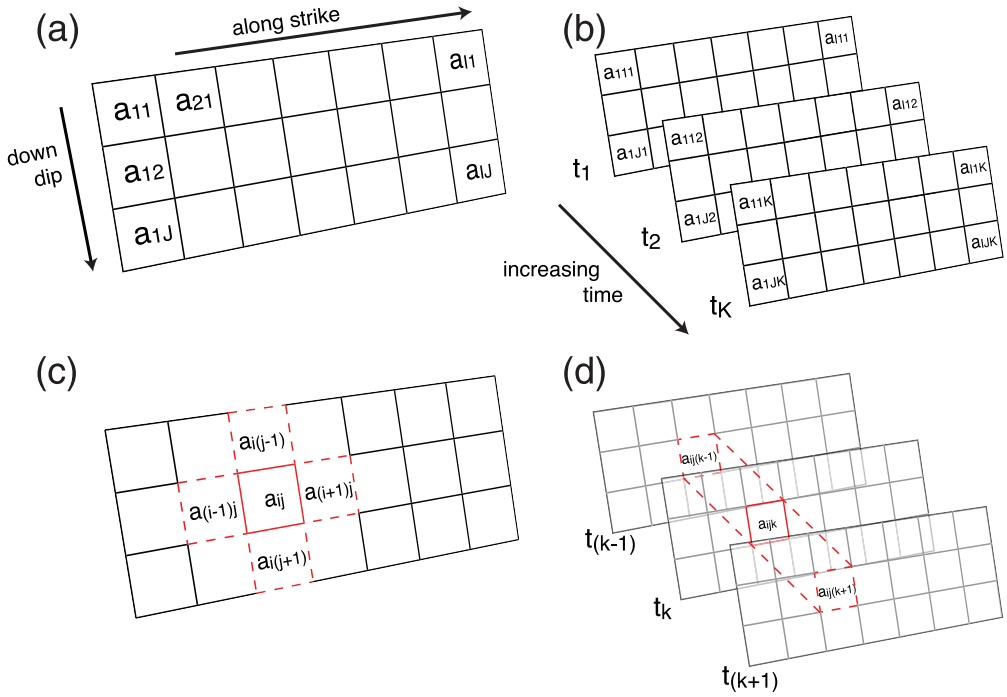

Figure 1. Schematic description of model layout and smoothing. (a) Rectangular fault divided into $I \times J$ patches; a_{ij} represents slip on the i-th patch along strike and j-th patch downdip. (b) Rectangular fault with K time steps; a_{ijk} represents the slip on the same patch as described, but at the k-th time step. (c) 2-D spatial smoothing, relating slip on a patch to that of its neighbours. (d) 1-D temporal smoothing, relating slip on a patch at a given time step, to the time steps before and after. Further details given in the main text.

a Gaussian distribution with zero mean and covariance $\sigma^2_{\rm geod}\mathbf{E}_{\rm geod}$, where $\sigma^2_{\rm geod}$ is an unknown scale factor, and $\mathbf{E}_{\mathbf{geod}}$ will be given later. While the uncertainties and covariances in our geodetic observations can in principle be estimated in advance, we must also account for 'modelling errors'—errors due to mismodelling of data (e.g. Tarantola 2005; Yagi & Fukahata 2008)—which are accounted for here by $\sigma^2_{\rm geod}$. This system of normal equations can be solved by standard least-squares methods, although for large $M_{\rm geod}$ it is usually necessary to introduce prior information on the smoothness of the spatial pattern of slip, as we will explain later.

Next we consider the case where we have information on the time-dependent evolution of an earthquake, for example, from seismic observations. Over the course of an earthquake, the distribution of slip will vary over the fault surface with respect to time. Thus, in our model, each of the fault patches is permitted to slip at each of K time steps (Fig. 1b), giving a total of $M = IJK$ parameters describing the behaviour of the fault during the earthquake. Let $\mathbf{d}_{\mathbf{seis}}$ represent a column vector of $N_{\rm seis}$ seismic waveform observations (for instance, a series of regularly sampled waveforms from different seismic stations concatenated together), with $\mathbf{e}_{\mathbf{seis}}$ being the uncertainties in those observations, the terms in $\mathbf{e}_{\mathbf{seis}}$ again following a Gaussian distribution with zero mean and covariance $\sigma^2_{\rm seis}\mathbf{E}_{\mathbf{seis}}$, where $\sigma^2_{\rm seis}$ is an unknown scale factor, again accounting for modelling errors, and $\mathbf{E}_{\mathbf{seis}}$ will be given later. If $\mathbf{a}^{\rm T} = (a_{111}, a_{211}, \ldots, a_{I11}, a_{121}, \ldots, a_{IJ1}, a_{112}, \ldots, a_{IJK})$, describes the slip of each fault patch at each time step, then we can relate our observations to the model parameters by

$$\mathbf{d}_{\mathbf{seis}} = \mathbf{H}_{\mathbf{seis}}\mathbf{a} + \mathbf{e}_{\mathbf{seis}}. \tag{2}$$

Here, $\mathbf{H}_{\mathbf{seis}}$ is a $N_{\rm seis} \times IJK$ kernel matrix linearly relating the slip of each individual fault patch at each time step to predicted seismograms at each seismic station location.

Our intention here is to combine both geodetic and seismic data into a joint inversion; this can be achieved by combining eqs (1) and (2):

$$\begin{pmatrix} \mathbf{d}_{\rm geod} \\ \mathbf{d}_{\mathbf{seis}} \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{\rm geod} \ldots \mathbf{H}_{\rm geod} \\ \mathbf{H}_{\mathbf{seis}} \end{pmatrix}\mathbf{a} + \begin{pmatrix} \mathbf{e}_{\rm geod} \\ \mathbf{e}_{\mathbf{seis}} \end{pmatrix}, \tag{3}$$

where the $\mathbf{H}_{\mathbf{geod}} \ldots \mathbf{H}_{\mathbf{geod}}$ notation here represents K concatenations of the geodetic data kernel matrix, and given $a^{\rm total}_{ij} = \sum_{t=1}^{K} a_{ijt}$. This is equivalent to the standard linear algebra relation

$$\mathbf{d} = \mathbf{H}\mathbf{a} + \mathbf{e}, \tag{4}$$

where $\mathbf{d}^{\rm T} = (\mathbf{d}^{\rm T}_{\rm geod}\ \mathbf{d}^{\rm T}_{\mathbf{seis}})$, $\mathbf{e}^{\rm T} = (\mathbf{e}^{\rm T}_{\rm geod}\ \mathbf{e}^{\rm T}_{\mathbf{seis}})$ and

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{\rm geod} \ldots \mathbf{H}_{\rm geod} \\ \mathbf{H}_{\mathbf{seis}}. \end{pmatrix}. \tag{5}$$

By assuming no correlation between the errors of the geodetic and seismic data, the covariance matrix of the combined data set is written as $\sigma^2_{\rm geod}\mathbf{E}(\gamma^2)$, where

$$\mathbf{E}(\gamma^2) = \begin{pmatrix} \mathbf{E}_{\rm geod} & 0 \\ 0 & \gamma^2\mathbf{E}_{\mathbf{seis}} \end{pmatrix}, \tag{6}$$

and $\gamma^2 = \sigma^2_{\rm seis}/\sigma^2_{\rm geod}$ is a measure of the relative variance of the geodetic and seismic data sets. Hence we obtain the following stochastic relation between the data $\mathbf{d}$ and model parameters $\mathbf{a}$:

$$p(\mathbf{d}|\mathbf{a}, \sigma^2_{\rm geod}, \gamma^2) = (2\pi\sigma^2_{\rm geod})^{-N/2}|\mathbf{E}(\gamma^2)|^{-\frac{1}{2}} \times \exp\left[-\frac{1}{2\sigma^2_{\rm geod}}(\mathbf{d} - \mathbf{H}\mathbf{a})^T\mathbf{E}(\gamma^2)^{-1}(\mathbf{d} - \mathbf{H}\mathbf{a})\right], \tag{7}$$

where $|\mathbf{E}(\gamma^2)|$ represents the determinant of $\mathbf{E}(\gamma^2)$, and $N = N_{\text{geod}} + N_{\text{seis}}$.

We can regard γ^2 as a hyperparameter controlling the relative weight of the two data sets in the probability distribution, with larger values of γ^2 giving increased influence to the geodetic data and vice versa.

2.2 Prior information

In order to further restrict the suite of possible fault slip models $\mathbf{a}$, we use prior information on the form of the fault slip distribution in space and time (Yabuki & Matsu'ura 1992). In the absence of a spatial smoothing constraint, for instance, it is often the case that the best-fitting solution can show large spatial oscillations in slip. Such oscillations imply large localized stresses on the fault that can exceed the inferred mechanical strength of the crust. Similarly, oscillatory slip with respect to time implies repeat loading and unloading of the fault segment with time; this is physically implausible given our knowledge of fault loading, and can be mitigated by applying smoothing criteria on slip between neighbouring time steps for every fault patch.

Therefore, two smoothing constraints are defined. The first, a spatial smoothing constraint, is a 2-D finite difference Laplacian approximation that relates the total slip on each fault patch, $\sum_k a_{ijk}$, to the total slip of its neighbours by

$$\sum_k (a_{(i-1)jk} + a_{(i+1)jk} + a_{i(j-1)k} + a_{i(j+1)k} - 4a_{ijk}) + e_s = 0, \quad (8)$$

where we assume the values of e_s follow a Gaussian distribution. In this study we assume slip to be zero beyond the edges of the fault (i.e. $a_{0jk} = a_{(I+1)jk} = a_{i0k} = a_{i(J+1)k} = 0$), but other slip boundary conditions may be implemented here if desired. We can rewrite the spatial smoothing constraint in matrix notation as

$$\mathbf{Sa} + \mathbf{e_s} = \mathbf{0}, \quad (9)$$

where $\mathbf{S}$ has dimensions $M \times M$, and $\mathbf{e_s}$ is a vector of errors that follow a Gaussian distribution with zero mean and covariance $\rho_s^2\mathbf{I}$, where $\mathbf{I}$ represents a unit matrix.

The second, temporal, smoothing constraint is a 1-D finite difference Laplacian approximation that relates slip on a fault patch at a given time step, a_{ijk} to that on the same fault patch at the previous and subsequent time steps, that is,

$$a_{ij(k-1)} - 2a_{ijk} + a_{ij(k+1)} + e_t = 0, \quad (10)$$

where, again, e_t has a Gaussian distribution. A zero-slip boundary condition is also enforced here, with $a_{ij0} = a_{ij(K+1)} = 0$. In vector form, this is

$$\mathbf{Ta} + \mathbf{e_t} = \mathbf{0}, \quad (11)$$

where, similarly to the spatial smoothing case, $\mathbf{T}$ has dimensions $M \times M$, and $\mathbf{e_t}$ is a vector of errors that follow a Gaussian distribution with zero mean and covariance $\rho_t^2\mathbf{I}$. As for the observation equations, the smoothing constraints can be written as probability density functions (pdfs), which we call here the 'prior distributions':

$$p\left(\mathbf{a}; \rho_s^2\right) = \left(2\pi\rho_s^2\right)^{-M/2} |\mathbf{G_s}|^{\frac{1}{2}} \exp\left(-\frac{1}{2\rho_s^2}\mathbf{a}^{\mathrm{T}}\mathbf{G_s}\mathbf{a}\right) \quad (12)$$

and

$$p\left(\mathbf{a}; \rho_t^2\right) = \left(2\pi\rho_t^2\right)^{-M/2} |\mathbf{G_t}|^{\frac{1}{2}} \exp\left(-\frac{1}{2\rho_t^2}\mathbf{a}^{\mathrm{T}}\mathbf{G_t}\mathbf{a}\right). \quad (13)$$

Here $\mathbf{G_s} = \mathbf{S}^{\mathrm{T}}\mathbf{S}$, $\mathbf{G_t} = \mathbf{T}^{\mathrm{T}}\mathbf{T}$ and ρ_s^2 and ρ_t^2 are hyperparameters controlling the distribution of model parameters with respect to spatial and temporal smoothing, respectively. Following Fukahata *et al.* (2003, 2004), these two pdfs can be combined into a single prior distribution:

$$p(\mathbf{a}; \rho_s^2, \rho_t^2) = (2\pi)^{-M/2} \left|\frac{1}{\rho_s^2}\mathbf{G_s} + \frac{1}{\rho_t^2}\mathbf{G_t}\right|^{\frac{1}{2}} \times \exp\left[-\mathbf{a}^{\mathrm{T}}\left(\frac{1}{2\rho_S^2}\mathbf{G_s} + \frac{1}{2\rho_t^2}\mathbf{G_t}\right)\mathbf{a}\right], \quad (14)$$

where $|\mathbf{G_s}/\rho_s^2 + \mathbf{G_t}/\rho_t^2|$ represents the determinant of the full-rank $M \times M$ matrix $(\mathbf{G_s}/\rho_s^2 + \mathbf{G_t}/\rho_t^2)$.

2.3 The likelihood function and ABIC

With the application of prior information to our model, a critical question arises—how are the competing influences of individual data sets and prior information accounted for? In other words, how do we give each information source an appropriate level of weight in the model, and how can we do this objectively?

Akaike (1980) proposed a statistical scheme for the objective determination of relative weighting of observations and prior information in such problems. The method proposed by Akaike was introduced to geophysical problems by Yabuki & Matsu'ura (1992), who presented a suitable form for geodetic and seismic data inversion. The solution lies with combining both the observations $p(\mathbf{d}|\mathbf{a}, \sigma_{\text{geod}}^2, \gamma^2)$ and the prior information $p(\mathbf{a}; \rho_s^2, \rho_t^2)$ to create a flexible, Bayesian model. The 'posterior' probability given these two quantities is obtained by applying Bayes' theorem (Bayes 1763):

$$p\left(\mathbf{a}; \sigma_{\text{geod}}^2, \gamma^2, \rho_s^2, \rho_t^2|\mathbf{d}\right) = cp\left(\mathbf{d}|\mathbf{a}; \sigma_{\text{geod}}^2, \gamma^2\right) p\left(\mathbf{a}; \rho_s^2, \rho_t^2\right), \quad (15)$$

where c is a normalizing factor independent of the model parameters $\mathbf{a}$ and the hyperparameters σ_{geod}^2, γ^2, ρ_s^2 and ρ_t^2.

Substituting eqs (7) and (14) into (15) gives

$$p\left(\mathbf{a}; \sigma_{\text{geod}}^2, \alpha^2, \beta^2, \gamma^2|\mathbf{d}\right) = c\left(2\pi\sigma_{\text{geod}}^2\right)^{-(M+N)/2} |\mathbf{E}(\gamma^2)|^{-\frac{1}{2}} |\alpha^2\mathbf{G_s} + \beta^2\mathbf{G_t}|^{\frac{1}{2}} \exp\left(-\frac{1}{2\sigma_{\text{geod}}^2}s(\mathbf{a})\right), \quad (16)$$

where $\alpha^2 = \sigma_{\text{geod}}^2/\rho_s^2$ and $\beta^2 = \sigma_{\text{geod}}^2/\rho_t^2$ are hyperparameters controlling the relative weighting of the spatial and temporal smoothing constraints (respectively) with respect to the data, and $s(\mathbf{a})$ is a measure of data misfit given by

$$s(\mathbf{a}) = (\mathbf{d} - \mathbf{Ha})^{\mathrm{T}}\mathbf{E}(\gamma^2)^{-1}(\mathbf{d} - \mathbf{Ha}) + \mathbf{a}^{\mathrm{T}}\left(\alpha^2\mathbf{G_s} + \beta^2\mathbf{G_t}\right)\mathbf{a}. \quad (17)$$

The optimal values of $\mathbf{a}$, σ_{geod}^2, α^2, β^2 and γ^2 can be obtained by maximizing the posterior pdf given in eq. (16). Therefore, if the hyperparameters σ_{geod}^2, α^2, β^2 and γ^2 are fixed, the problem becomes one of simply minimizing $s(\mathbf{a})$; the set of best-fitting model parameters $\mathbf{a}^*$ can therefore be obtained by solving the least-squares problem

$$\mathbf{a}^* = [\mathbf{H}^{\mathrm{T}}\mathbf{E}(\gamma^2)^{-1}\mathbf{H} + \alpha^2\mathbf{G_s} + \beta^2\mathbf{G_t}]^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{E}(\gamma^2)^{-1}\mathbf{d}. \quad (18)$$

In the general case, we wish to obtain the best estimates of the hyperparameters σ_{geod}^2, α^2, β^2 and γ^2. This is achieved through the minimization of ABIC (Akaike 1980), defined as

$$\text{ABIC} = -2\log L\left(\sigma_{\text{geod}}^2, \alpha^2, \beta^2, \gamma^2|\mathbf{d}\right), \quad (19)$$

where $L(\sigma^2_{\rm geod}, \alpha^2, \beta^2, \gamma^2|\mathbf{d})$ is called the 'marginal likelihood' of the hyperparameters for given data $\mathbf{d}$, and is given by

$$L\left(\sigma^2_{\rm geod}, \alpha^2, \beta^2, \gamma^2|\mathbf{d}\right) = \int p\left(\mathbf{a}; \sigma^2_{\rm geod}, \alpha^2, \beta^2, \gamma^2|\mathbf{d}\right) \mathrm{d}\mathbf{a}. \quad (20)$$

In minimizing ABIC, we are maximizing the information entropy, a measure of the closeness between two distributions—a hypothetical (model) distribution and the true distribution. As ABIC is reduced, the approximation of the model to the 'truth' is improved (Akaike 1977).

By evaluating the integral in eq. (20), we obtain

$$\begin{aligned} &L\left(\sigma^2_{\rm geod}, \alpha^2, \beta^2, \gamma^2|\mathbf{d}\right) \\ &= c\left(2\pi\sigma^2_{\rm geod}\right)^{-N/2} |\alpha^2\mathbf{G_s} + \beta^2\mathbf{G_t}|^{1/2} |\mathbf{E}(\gamma^2)|^{-1/2} |\mathbf{H}^T\mathbf{E}(\gamma^2)^{-1}\mathbf{H} \\ &\quad + \alpha^2\mathbf{G_s} + \beta^2\mathbf{G_t}|^{-1/2} \exp\left(-\frac{1}{2\sigma^2_{\rm geod}} s(\mathbf{a}^*)\right) + C, \end{aligned} \quad (21)$$

given that, after Yabuki & Matsu'ura (1992), we can express $s(\mathbf{a})$ in terms of $s(\mathbf{a}^*)$:

$$s(\mathbf{a}) = s(\mathbf{a}^*) + (\mathbf{a} - \mathbf{a}^*)^{\rm T}(\mathbf{H}^{\rm T}\mathbf{E}(\gamma^2)^{-1}\mathbf{H} + \alpha^2\mathbf{G_s} + \beta^2\mathbf{G_t})(\mathbf{a} - \mathbf{a}^*). \quad (22)$$

If ABIC is to be minimized, then by definition, the marginal likelihood L must be maximized, that is, we need to find the values of the hyperparameters such that

$$\frac{\partial L}{\partial \sigma^2_{\rm geod}} = \frac{\partial L}{\partial \alpha^2} = \frac{\partial L}{\partial \beta^2} = \frac{\partial L}{\partial \gamma^2} = 0. \quad (23)$$

The first condition allows us to analytically obtain an estimate for the geodetic data variance, $\sigma^2_{\rm geod}$, in terms of the other hyperparameters:

$$\sigma^2_{\rm geod} = s(\mathbf{a}^*)/N. \quad (24)$$

Substituting this quantity, with eq. (21) into eq. (19), the following expression for ABIC is obtained:

$$\begin{aligned} \mathrm{ABIC}(\alpha^2, \beta^2, \gamma^2) = {} & N \log s(\mathbf{a}^*) - \log|\alpha^2\mathbf{G_s} + \beta^2\mathbf{G_t}| \\ & + \log|\mathbf{H}^T\mathbf{E}(\gamma^2)^{-1}\mathbf{H} + \alpha^2\mathbf{G_s} + \beta^2\mathbf{G_t}| \\ & + \log|\mathbf{E}(\gamma^2)| + \kappa. \end{aligned} \quad (25)$$

Here, κ is a constant term independent of the hyperparameters α^2, β^2 and γ^2. Given eq. (6), the quantity $\log|\mathbf{E}(\gamma^2)|$ in this specific case is given by

$$\log|\mathbf{E}(\gamma^2)| = \log|\mathbf{E}_{\rm geod}| + \log|\mathbf{E_{seis}}| + N_{\rm seis}\log\gamma^2, \quad (26)$$

with the posterior covariance $\mathbf{C}$ of the model parameters $\mathbf{a}$, which can be used as a measure of uncertainty, given by

$$\mathbf{C} = \sigma^2_{\rm geod}(\mathbf{H}^{\rm T}\mathbf{E}(\gamma^2)^{-1}\mathbf{H} + \alpha^2\mathbf{G_s} + \beta^2\mathbf{G_t})^{-1}. \quad (27)$$

Since we are only interested in the hyperparameter values that correspond to the minimum of ABIC, and not the absolute value of ABIC itself, we solve for a minimum in ABIC, neglecting the constant term κ. This solution is obtained numerically by a parameter search over values of α^2, β^2 and γ^2, systematically focusing the search around the vicinity of the ABIC minimum.

3 APPLICATION TO THE 1997 MANYI, TIBET, EARTHQUAKE

The 1997 November 8 Manyi earthquake was a large ($M_{\rm w} \sim 7.6$) left-lateral strike-slip event, which occurred in a remote region of northern Tibet (Fig. 2). No near-field seismological (i.e. strong motion) data exist for this earthquake, and to our knowledge no field information has been collected on the surface rupture; our only means of obtaining information on the earthquake are 'remote' methods—in this case, from satellite imagery, space geodesy and teleseismic data. Given the clear, coherent InSAR data available (e.g. Peltzer *et al.* 1999; Funning *et al.* 2007), and that the size of the event lends itself well to broad-band seismic modelling (e.g. Velasco *et al.* 2000), the Manyi earthquake is an excellent test case for our joint-inversion methodology.

As well as allowing an evaluation of our method, a joint inversion may also allow us to reconcile some of the differences between published InSAR and seismic source models for the Manyi earthquake. The body wave inversion model of Velasco *et al.* (2000) identified five subevents for the earthquake, all located within 25 km of the hypocentre. Conversely, our InSAR distributed slip model (Funning *et al.* 2007) shows that slip is much more widely distributed, with at least one area of elevated slip located ~80 km west of the region of maximum slip. We hope to evaluate, for instance, if certain areas of slip can be moved in space and time without significant degradation to the seismic and/or InSAR data fits.

3.1 The fault model in space and time

In order to efficiently compute data kernels, and for reasons of parsimony, we define a simple fault model for our test case. Using the average strike of the aftershock zone and mapped fault trace (258°), along with the results of our detailed InSAR modelling (Funning *et al.* 2007), a single, vertical fault plane is defined, extending for 180 km along-strike and 18 km downdip, approximately the dimensions of the rupture area obtained previously. In order to image heterogeneous slip, we divide the fault area into 6 × 6 km patches; hence $I = 30$, $J = 3$ and $M_{\rm geod} = 90$ as defined (Section 2.1). The rake on each patch is fixed at −5° on the basis of our earlier analysis of the InSAR data (Funning 2005; Funning *et al.* 2007).

To enable time-varying slip on the fault, we define in addition six time steps in which slip can take place on each patch ($K = 6$). To increase the efficiency of the calculations, we assume that the first time step, and therefore slip, commences on each patch when a maximum speed rupture front of constant velocity emanating from the earthquake hypocentre reaches an individual patch. Slip need not occur at the first time step (thus, lower rupture velocities than that of the maximum speed rupture front are allowed), but must conclude at the final time step. In our inversion scheme, each time step represents a 4-s interval with an isosceles-triangular source–time function with a rise time of 2 s. The interval between successive time steps is 2 s. These values were selected following a preliminary analysis of the teleseismic waveform data. Further details on the selection of rupture velocities and time steps will be given later.

3.2 Inversion of InSAR data

We use pairs of ERS-2 images bracketing the Manyi earthquake, three frames of data from each of three tracks (Fig. 2 and Table 1), to produce coseismic interferograms (Fig. 3). The data are exactly the same as those used in an earlier study (Funning *et al.* 2007), in which a detailed description of data processing and reduction is given. As before, we employ curvature-based quadtree sampling (e.g. Simons *et al.* 2002) to reduce numbers of datapoints to 1500–2000 per track. The data are further reduced by removing points located within 3000 m of the segmented fault trace obtained by comparison of

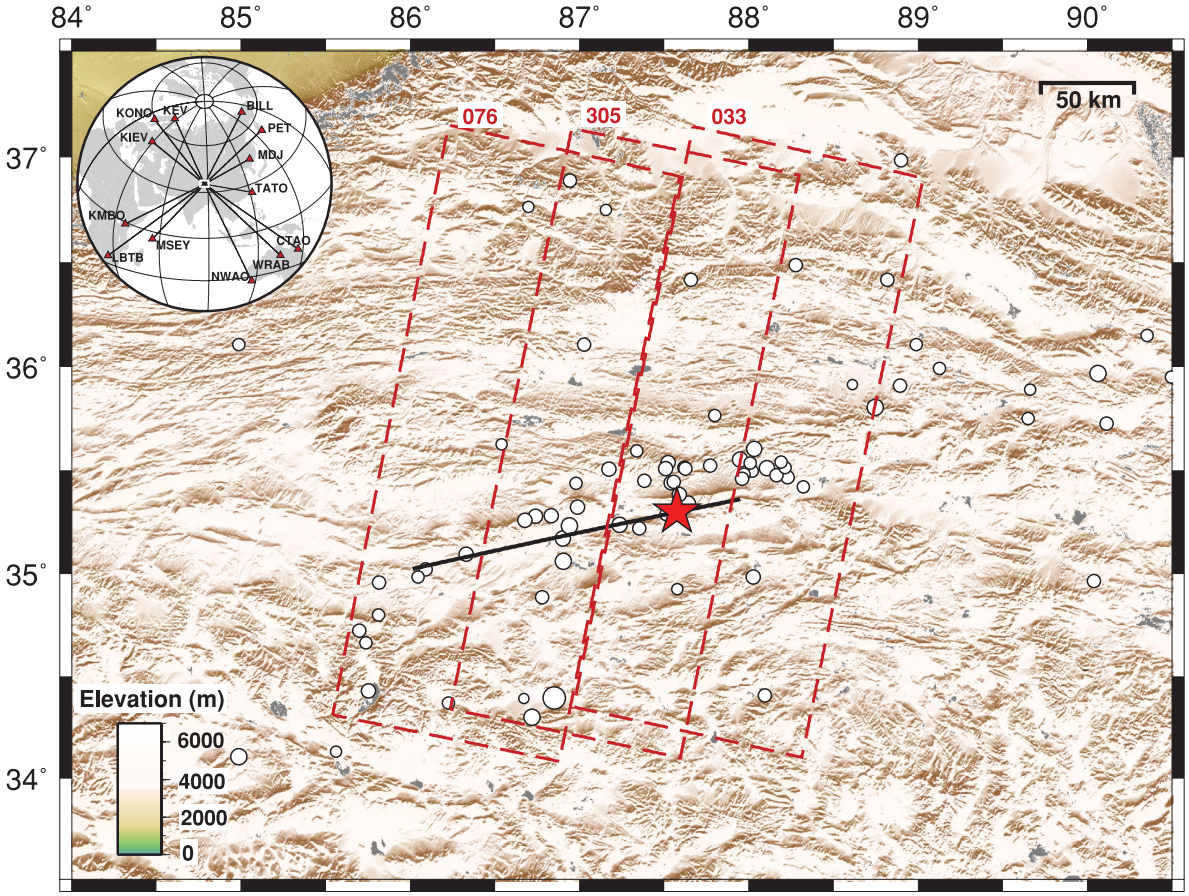

Figure 2. Location map for the Manyi earthquake epicentral region. The distribution of aftershocks (white circles), whose locations relative to the main shock (red star; from Steck *et al.* 2001) were estimated by joint hypocentre determination methods (Dewey 1971, ; David Robinson, personal communication, 2008), define a linear zone approximately 150–200 km long. This zone is covered by data from three ERS satellite tracks (areas of coverage defined by red dashed lines). Surface elevation data shown are from the Shuttle Radar Topography Mission (Farr & Kobrick 2000). Inset map shows location of the study region within Asia, and azimuths, distances and great circle paths to the 13 seismic broad-band stations used in this study.

Table 1. Descending track ERS-SAR data used in this study. In all cases, data are from frames 2871–2907 on the tracks mentioned.

	Date 1	Sensor 1	Date 2	Sensor 2	$B_\perp/m^a$	$\lvert h_a\rvert/m^b$	$\Delta t_{\rm post}/{\rm d}^c$
Track 076	1997 March 16	ERS-2	1997 November 16	ERS-2	65	154	8
Track 305	1997 August 19	ERS-2	1997 December 02	ERS-2	4	2560	24
Track 033	1997 May 22	ERS-2	1997 December 18	ERS-2	−38	263	40

[a]Perpendicular baseline at image centre.
[b]Altitude of ambiguity at image centre.
[c]Duration of post-seismic period covered.

optical imagery, digital topography, interferometric coherence and radar image azimuth offsets (Funning *et al.* 2007); this is necessary in order to prevent simplified fault geometries from violating near-fault displacement polarities. The kernels $\mathbf{H}_{\rm geod}$ are calculated for these locations by considering each patch as a dislocation within an elastic half-space (Lamé elastic parameters $\lambda = \mu = 30$ GPa; Okada 1985). We add to the kernel terms that solve for the best-fitting static shift and tilt for each InSAR data set, to mitigate zero-level ambiguities and errors in orbital data. In this inversion, and all others shown in this study, we use the Fast Non-Negative Least-Squares algorithm (Bro & De Jong 1997) when solving for best-fitting model parameters, to prevent reversals in fault slip. In light of the work of Fukahata & Wright (2008), who demonstrated the importance of using full covariance matrices, rather than assuming independent data, we estimate data covariance $\mathbf{E}_{\rm geod}$ by radially averaging the autocorrelation function of an undeforming portion of each interferogram (e.g. Wright *et al.* 2003; Funning *et al.* 2005; Parsons *et al.* 2006). The covariance versus distance relationship thus obtained is approximately exponential in form, and can be used to estimate the expected covariance between datapoint pairs based upon their separation distance.

In a geodetic data-only inversion, only the relative importance of data and spatial smoothing must be considered, and thus ABIC becomes a function of a single hyperparameter, α^2, as shown by Yabuki & Matsu'ura (1992). In order to quantify the behaviour of ABIC, and find the slip model corresponding to the ABIC minimum, we evaluate ABIC over a wide range of values of α^2 spanning over 20 orders of magnitude. Our results are plotted in Fig. 4. At low values of α^2 there is an approximately log-linear relationship between ABIC and α^2, flattening out into a curve, as the ABIC

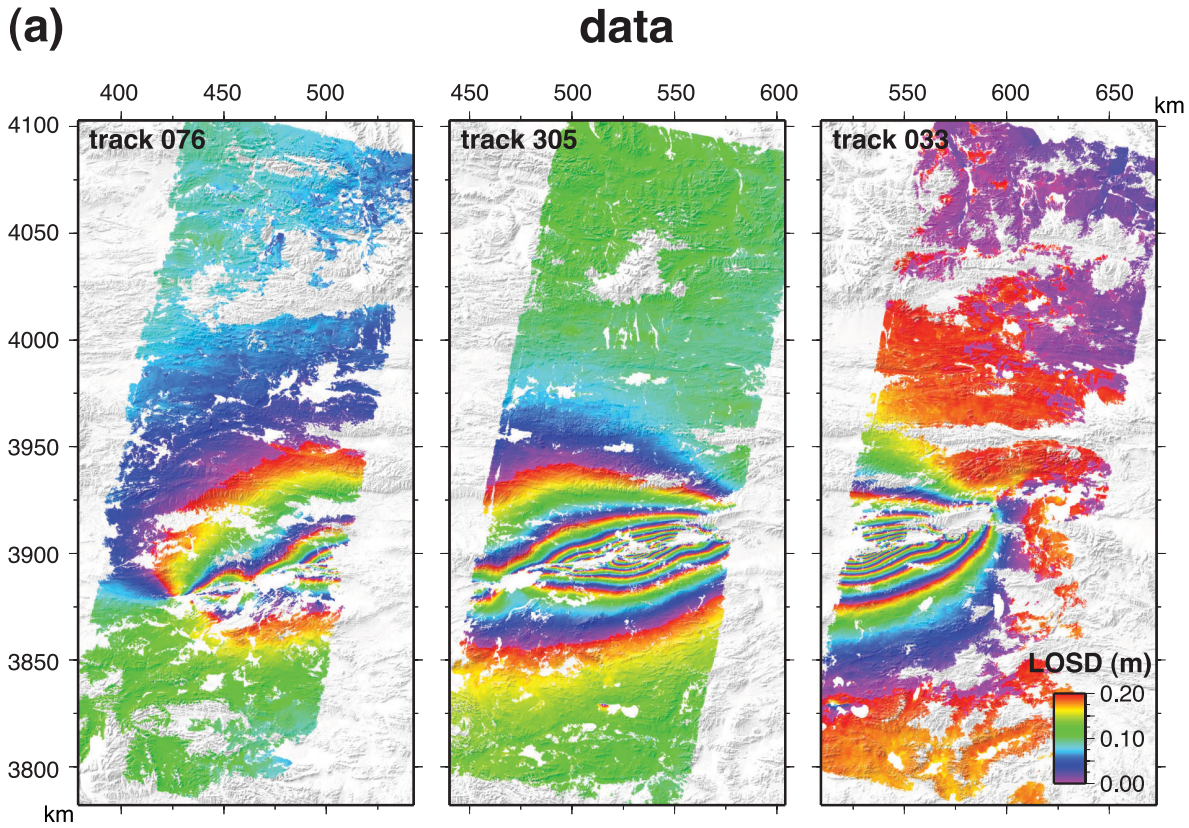

Figure 3. Observed, modelled and residual interferograms of the Manyi earthquake. (a) Observed interferogram data, unwrapped, converted to line-of-sight displacement (LOSD) of the ground and rewrapped at 0.2 m intervals. They are overlaid upon the Shuttle Radar Topography Mission digital elevation model used in the data processing, illuminated from an incidence angle of 45° from the northeast. The peak offset in LOS displacement is 2.1 m in both the central (305) and eastern (033) tracks, equivalent to a horizontal offset of ∼7 m if pure strike-slip deformation is assumed. Coordinates are given in UTM km, zone 45. (b) Synthetic model and residual (model subtracted from data) interferograms for each inversion (InSAR data only, seismic waveform data only and joint InSAR and seismic data). Note the substantial difference in the number of residual fringes between the InSAR and joint-inversion models (which fit well) and the seismic inversion model (which fits poorly).

minimum is approached at $\alpha^2 = 0.35$. As α^2 increases above the value corresponding to the ABIC minimum, ABIC rises sharply, as modelled slip becomes oversmoothed and fits the data less well, before flattening out into a plateau at values above $\sim 10^5$, reflecting the zero slip fault boundary conditions in our methodology (Section 2.2), which drive the fault model towards zero slip, and therefore a constant large misfit, at high α^2.

The slip model corresponding to the ABIC minimum, along with formal uncertainties in slip, is plotted in Fig. 5. Several concentrations of slip can be identified in the model, with peak slip of 7.2 m occurring at a distance of 132–144 km along-strike (here and in all subsequent discussions, along-strike distances are measured from the western termination of the fault), a secondary peak (of up to 6.7 m of slip) between 108 and 126 km along-strike and minor peaks (of 3.2 and 2.4 m of slip) at along-strike distances of 162–168 and 30–42 km (respectively). 92 per cent of moment release occurred within the top 12 km of the fault plane, with peak slip values occurring mostly within the upper 6 km. The slip pattern is broadly consistent with that obtained from detailed modelling of the same data (Funning *et al.* 2007), despite the many simplifications to the fault model and the fixed rake used in this case. The formal 1σ uncertainties in slip increase monotonically with depth, as may be expected, with typical values of 10–15 cm in the upper 6 km of the fault increasing to 30–35 cm at depths of 12–18 km; however, this level of uncertainty is an order of magnitude lower than the modelled slip values in the upper 12 km of the fault where most slip occurs, and so we are confident that the slip values we recover are significant.

Synthetic and residual interferograms for this model are plotted in Fig. 3. The model deformation pattern (rewrapped at 20 cm intervals) appears visually to be a good match for the data; residual fringes are confined only to the nearest near-field (within ∼15 km from the fault) on all three tracks. The fit of the deformation pattern in the far-field suggests that the model captures the bulk deformation field well; the local differences between data and model in the near-field can be ascribed to the crudeness of the assumed fault geometry and coarseness of the fault discretization that render the model unable to resolve fine details of the deformation pattern. The total seismic moment for the InSAR-only inversion is $(1.90 \pm 0.02) \times 10^{20}$ Nm (quoted formal uncertainty is at the 1σ level). We compare the moment estimates obtained by various studies with our models in Table 2. Here, the moment estimate of our InSAR-only inversion is comparable with existing seismological estimates (e.g. Velasco *et al.* 2000) and InSAR-derived models based on crude, uniform-slip fault discretizations (Funning *et al.* 2007). The moment estimate is somewhat lower than that we previously obtained through detailed modelling where slip and rake are permitted to vary spatially (2.84×10^{20} Nm; Funning *et al.* 2007); there are several potential reasons for such a discrepancy. First, the difference could be a result of the simplified fault geometry and rake used in this study, compared to a detailed reversing-dip, variable rake model. In our previous study, models constructed using a

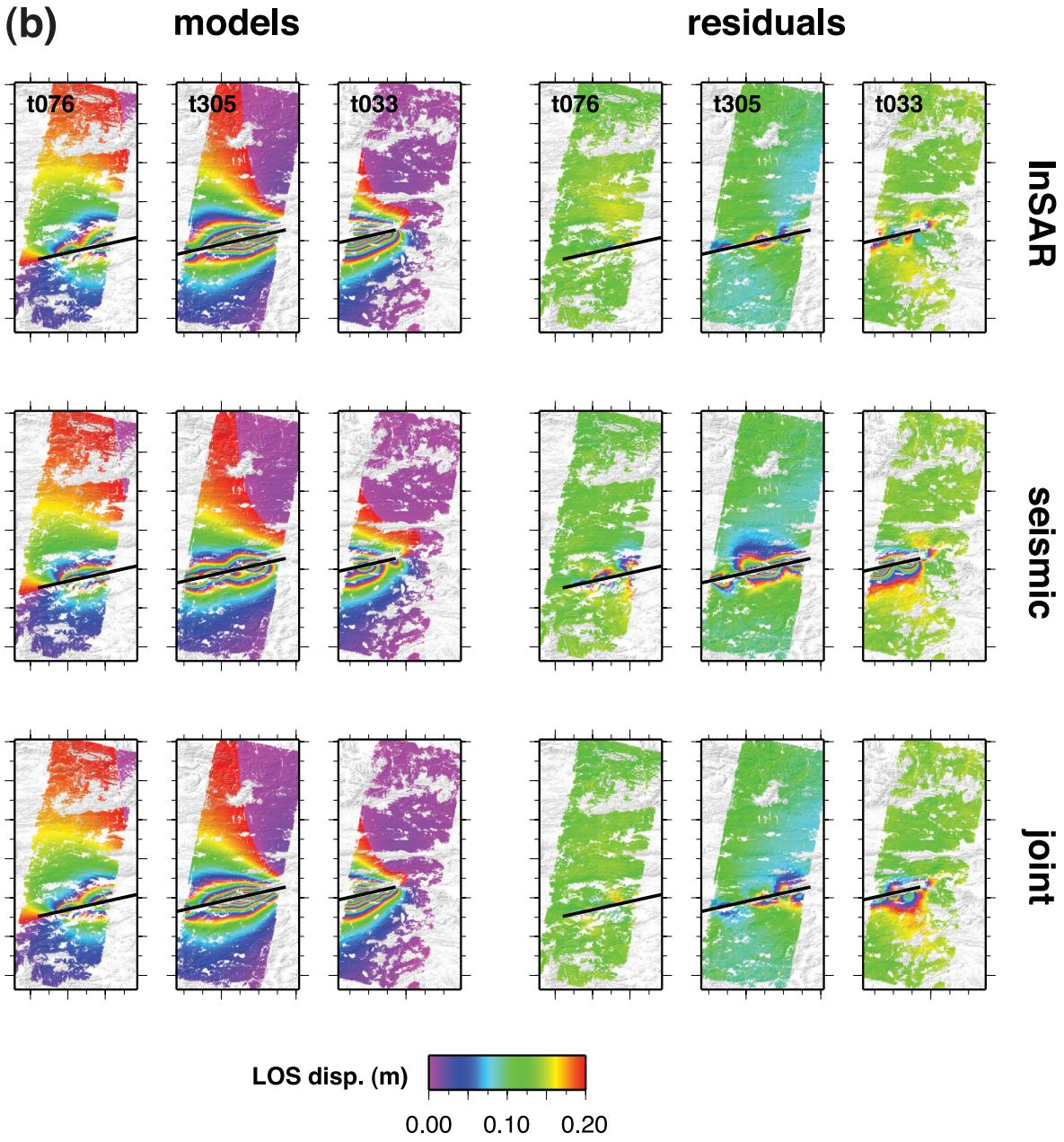

Figure 3. (*Continued.*)

reversing-dip fault geometry had higher moments than those with a single, near vertical dip, and models where the rake was allowed to vary had higher moments than those where the rake was fixed (Table 2). Second, it is possible that some of the deficit in moment may represent unmodelled large shallow slip in this simplified model, due to the exclusion of datapoints within 3 km of the surface fault trace. Third, the earlier InSAR study used a different method to estimate the smoothing parameter (a quantity analogous to the α^2 hyperparameter in this case)—the so-called 'L-curve' method, whereby a smoothing parameter value is selected as a compromise between improving data misfit and reducing solution roughness. It is possible that this different method may have resulted in a bias to the seismic moment, for example, by extending slip into areas where it may not have occurred through oversmoothing (Yabuki & Matsu'ura 1992).

3.3 Inversion of broad-band teleseismic waveform data

For our seismic data inversion we take vertical component data from 13 Global Seismic Network stations with an approximately even azimuthal distribution (Fig. 2) and which are at teleseismic distances from the source. The data are deconvolved for instrument response, bandpass filtered between 0.01 and 1 Hz, normalized by peak amplitude, and subsampled at 0.5 s intervals. From these data, and given the location of our fault model in space, and a Jeffreys–Bullen velocity model (Jeffreys & Bullen 1940), we locate the earthquake hypocentre on the fault plane, at a distance of 147 km along-strike from the western end of the fault and a depth of 15 km. Seismic data kernels are computed using the method of Kikuchi & Kanamori (1991) assuming overlapping isosceles-triangular source–time functions with a 4 s duration and a 2 s repeat

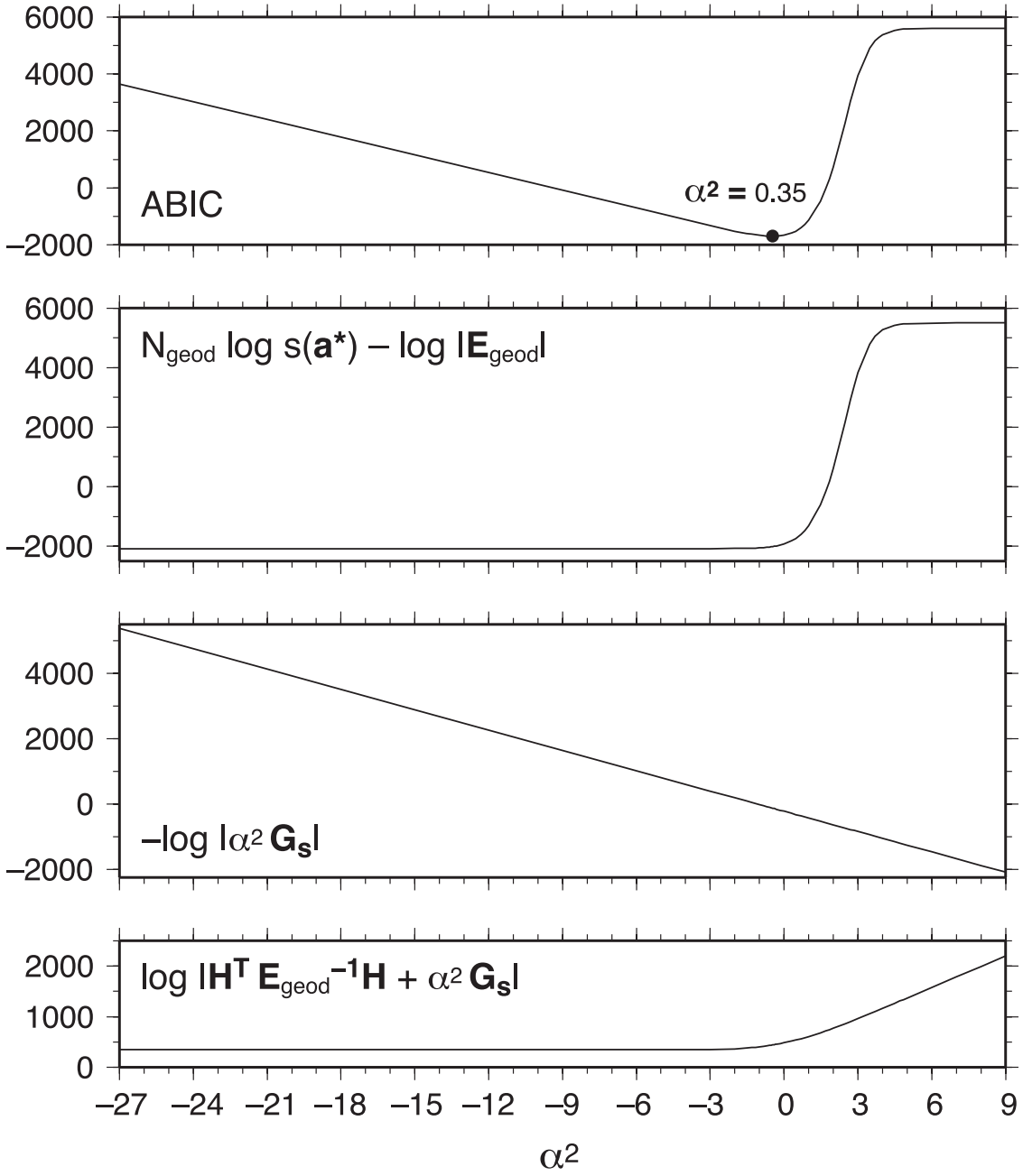

Figure 4. Form and components of ABIC for the InSAR-only inversion. Top panel: ABIC shows a well defined minimum at $\alpha^2 = 0.35$, representing the optimal objective weighting of the spatial smoothing. Below panel: Components of ABIC that sum to give the above curve. Further details given in the main text.

interval, for point sources located at the centre of each fault patch. A maximum rupture velocity of 3.2 km s^{-1} is used, selected by running inversion tests over a range of potential velocities to find a minimum variance (e.g. Yagi *et al.* 2004). We find that the complexity of the event can mostly be captured with six time steps; adding more time steps does not improve the data fit, but significantly increases the number of model parameters, and hence computation time. Again, we use a Jeffreys–Bullen velocity model (Jeffreys & Bullen 1940) in these computations. The covariance matrix $\mathbf{E}_{\mathbf{seis}}$ is assumed to be diagonal, with terms corresponding to the variances of the noise on each individual seismic trace.

A seismic data-only inversion is controlled by two hyperparameters, α^2 and β^2, controlling the relative importance of spatial and temporal smoothing, respectively. We search over a range of both values, as before, to find the ABIC minimum, located at $\alpha^2 = 0.03$ and $\beta^2 = 6.5$ (Fig. 6a). The pattern of ABIC in two dimensions is similar to that for the 1-D case, in that at low values of either α^2 or β^2, ABIC tends towards high values, while at high values of both α^2 and β^2, ABIC tends to plateau at a value significantly higher than its minimum value. In this set-up, the minimum in ABIC is less sensitive to changes in α^2 than β^2 with a wide range of values of α^2—spanning approximately 3 orders of magnitude—giving values of ABIC close to that of the minimum.

The moment rate function and modelled slip associated with the ABIC minimum model are plotted in Figs 6(b) and (c), respectively. The evolution of slip on the fault shown is interpolated at 5 s

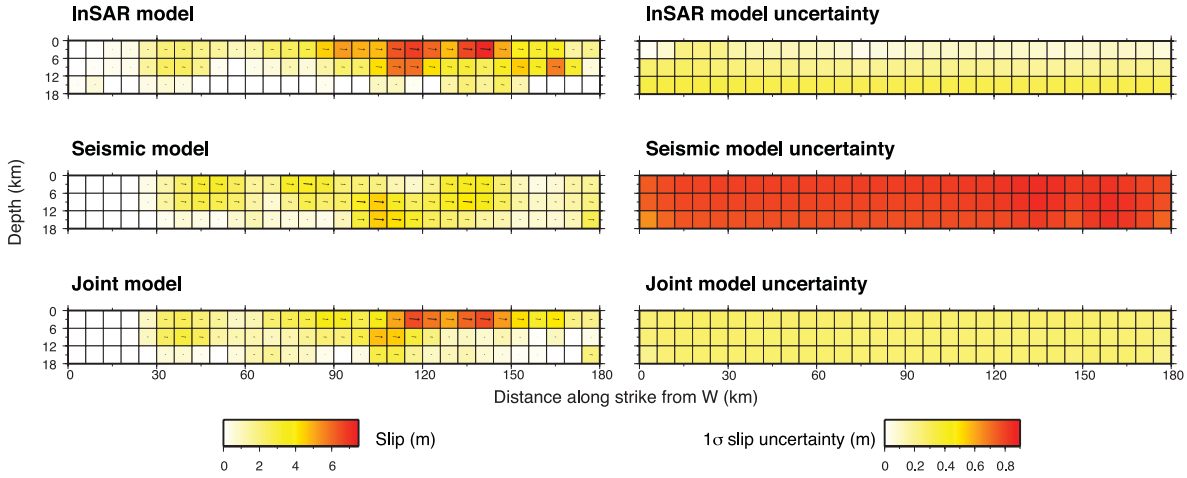

Figure 5. Comparison of total slip and uncertainty in total slip for the three inversion models. In general, the seismic model shows more heterogeneous and lower values of slip than the InSAR model, which has peaks in slip concentrated in the upper 12 km of the fault. The joint-inversion result broadly follows the pattern of the InSAR model, but with peak slip confined to the upper 6 km of the fault and reduced slip at its eastern end.

Table 2. Catalogued and published seismic solutions for the mechanism of the Manyi earthquake.

	Lat.	Long.	Depth km^{-1}	Strike	Dip	Rake	$M_0/10^{20}$ Nm
Seismic models							
Harvard CMT	35.33[a]	86.96[a]	16.4[a]	79	69	2	2.23
NEIC	35.069	87.325	38.0	70	89	2	1.40
Velasco *et al.* (2000)	35.246[b]	87.351[b]	12[b]	255	76	−5	1.80
This study, seismic only	35.302[b]	87.596[b]	15[b]	258	90	−5	1.63 ± 0.08
InSAR models							
Funning *et al.* (2007) single fault	35.258[c]	87.358[c]	10.1[d]	260	85.2	−3	1.68
Funning *et al.* (2007) uniform slip	Variable[e]	Variable[e]	12.6[d]	Variable[e]	86.5	−11.2	2.18
Funning *et al.* (2007) uniform slip	Variable[e]	Variable[e]	11.7/14.2[d,f]	Variable[e]	−85.5/84.3[f,g]	2.6/−9.3[f]	2.63
Funning *et al.* (2007) variable slip	Variable[e]	Variable[e]	20[d]	Variable[e]	−85.5/84.3[f,g]	2.9/−9.3[f]	2.52
Funning *et al.* (2007) variable slip	Variable[e]	Variable[e]	20[d]	Variable[e]	−85.5/84.3[f,g]	Variable	2.84
This study, InSAR only	35.196[c]	86.978[c]	18[d]	258	90	−5	1.90 ± 0.02
This study, joint inversion	35.196[c]	86.978[c]	18[d]	258	90	−5	1.73 ± 0.01

[a]Centroid location.
[b]Location of hypocentre.
[c]Location of midpoint of the model fault trace.
[d]Bottom depth of fault (all InSAR models rupture to the surface).
[e]Fault is divided into 11 subfaults with variable strike, geometry given in Funning *et al.* (2007).
[f]Two values here indicate a change of parameters along strike.
[g]Indicates reversal of dip along strike; negative dips indicate dip to the south.

intervals from the model of slip over multiple time steps at each model subfault, given the initiation of slip at the arrival of the maximum rupture velocity wave front at the centre of that subfault. Slip, initiating at the hypocentre, propagates bilaterally at speeds close to the maximum rupture velocity (3.0–3.2 km s^{-1}) for the first 20 s, at which point the eastward rupture reaches the eastern end of the fault. This earlier period of the earthquake is marked by a broad double peak in the moment rate function. The westward rupture continues subsequently for a further 28 s, with a concentration of slip at around 30–54 km along strike marked as a secondary peak in the moment rate function at around 37 s. The spatial pattern of total slip is significantly more heterogeneous, rough and uncertain than that obtained from the InSAR-only inversion (Fig. 5). Whereas the InSAR data inversion showed that the highest slip was concentrated on the upper 12 km of the eastern half of the fault, the highest slip in the seismic-only model is divided into five discrete areas of the fault, located at both the shallowest and deepest portions of the fault, peaking at ∼5 m, with 1σ uncertainties of around 90 cm across the whole fault plane. The M_0 of (1.63 ± 0.08) × 10^{20} Nm reflects the lower estimated slip, and the greater uncertainty in the slip of individual fault patches.

The ABIC minimum model provides a reasonable fit to the waveform data (Fig. 7) particularly in the earlier portion of the record. Later portions of the seismograms from some of the stations, such as LBTB and KMBO, show poor fits, which may be due to unmodelled fault complexity at the western end of the fault, or (in the case of LBTB and KMBO) due to some heterogeneity on the ray paths to the two stations which are close in azimuth. Nevertheless, for the majority of waveforms, the majority of the recorded signal is accounted for. If synthetic interferograms are produced on the basis of the modelled total slip, however, a very poor fit is obtained, with near-fault line-of-sight displacement residuals of the order of 1 m in the areas where the seismic slip inversion predicts low surface slip on the eastern half of the fault (Fig. 3). The misfit of this model, represented by a weighted residual sum of squares statistic, is a factor of 4 greater than that estimated for the InSAR-only

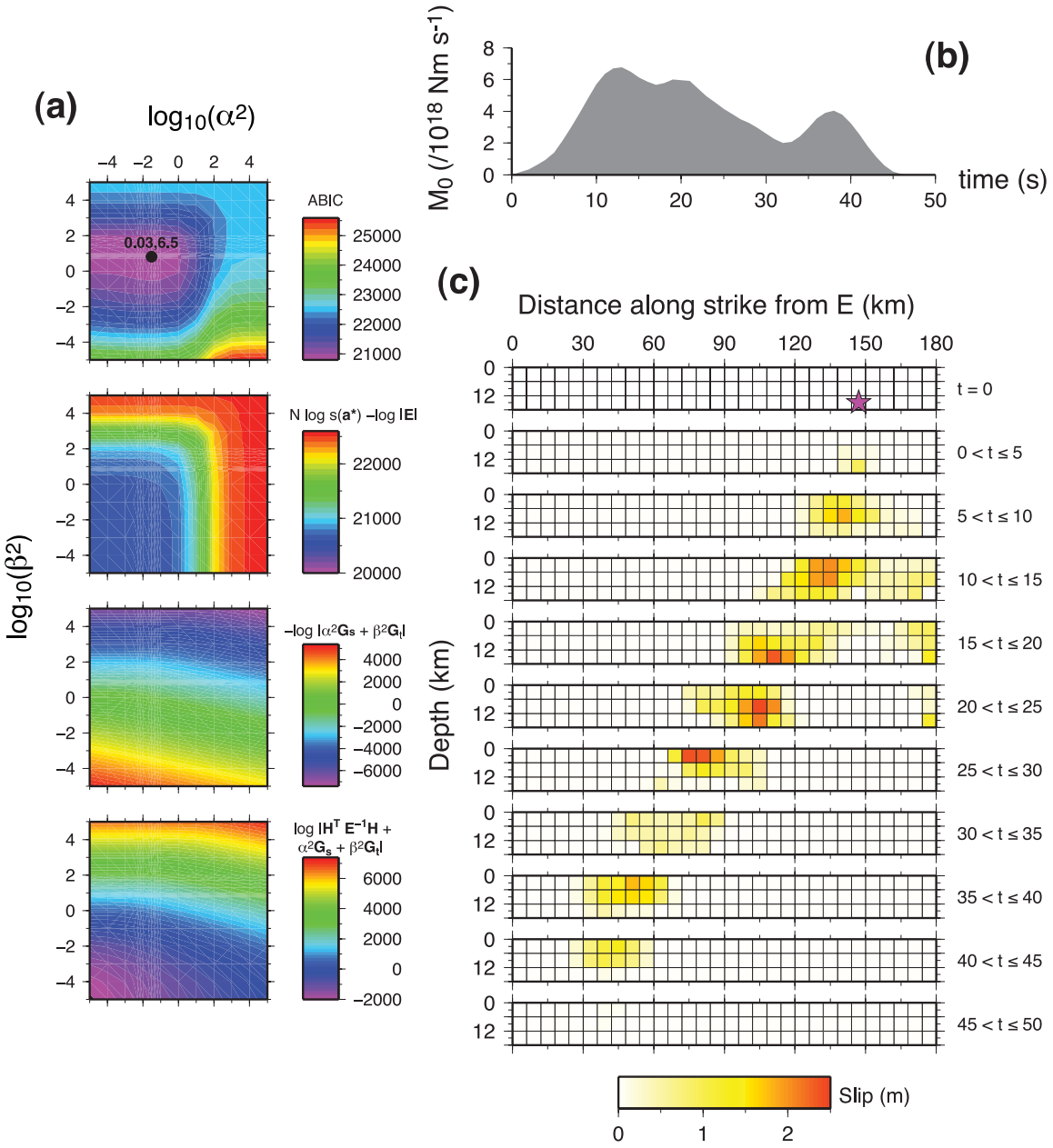

Figure 6. ABIC, source–time function and slip evolution for the seismic-only inversion. (a) Form of ABIC and constituent components. The ABIC minimum is found at $\alpha^2 = 0.03$, $\beta^2 = 6.5$. Details on the 2-D form of ABIC are given in the main text. (b) Moment–rate (source–time) function for the earthquake rupture, comprising of a broad first double peak (0–32 s) followed by a smaller, secondary peak (32–48 s). (c) Evolution of slip in the seismic-only model, shown in 5-s time-slices. Magenta star indicates location of the earthquake hypocentre. Slip propagates bilaterally for the first 20 s, and westward for a further 28 s.

inversion (Table 3). The failure of the model to agree with the independent InSAR data set could have two potential explanations—first, that our teleseismic data are not sufficient to accurately constrain the spatial slip pattern in this case, or secondly, that the two data sets are simply not compatible (for instance, the InSAR data could be significantly affected by additional, non-coseismic signals).

3.4 The joint inversion

In order to resolve whether the difference in fault slip models is due to an incompatibility in the input data sets, we run a full joint inversion including both the InSAR and teleseismic waveform data, and identical kernel functions to those utilized in the single data set inversions. Here our inversion is controlled by all three of the hyperparameters given in our derivation (Section 2), with α^2 and β^2 reflecting spatial and temporal smoothing weights as in the seismic inversion mentioned above, and the third hyperparameter γ^2 reflecting the relative weighting of the two data sets. Larger values of γ^2 are consistent with greater influence of the InSAR data set, and vice versa. The distribution of ABIC with respect to the hyperparameters is evaluated, as before, by a parameter search, which is densified in the region of the global minimum, in order to find the ABIC minimum.

The distribution of ABIC is plotted in Fig. 8(a), in terms of a series of 'slices' through hyperparameter space at different values

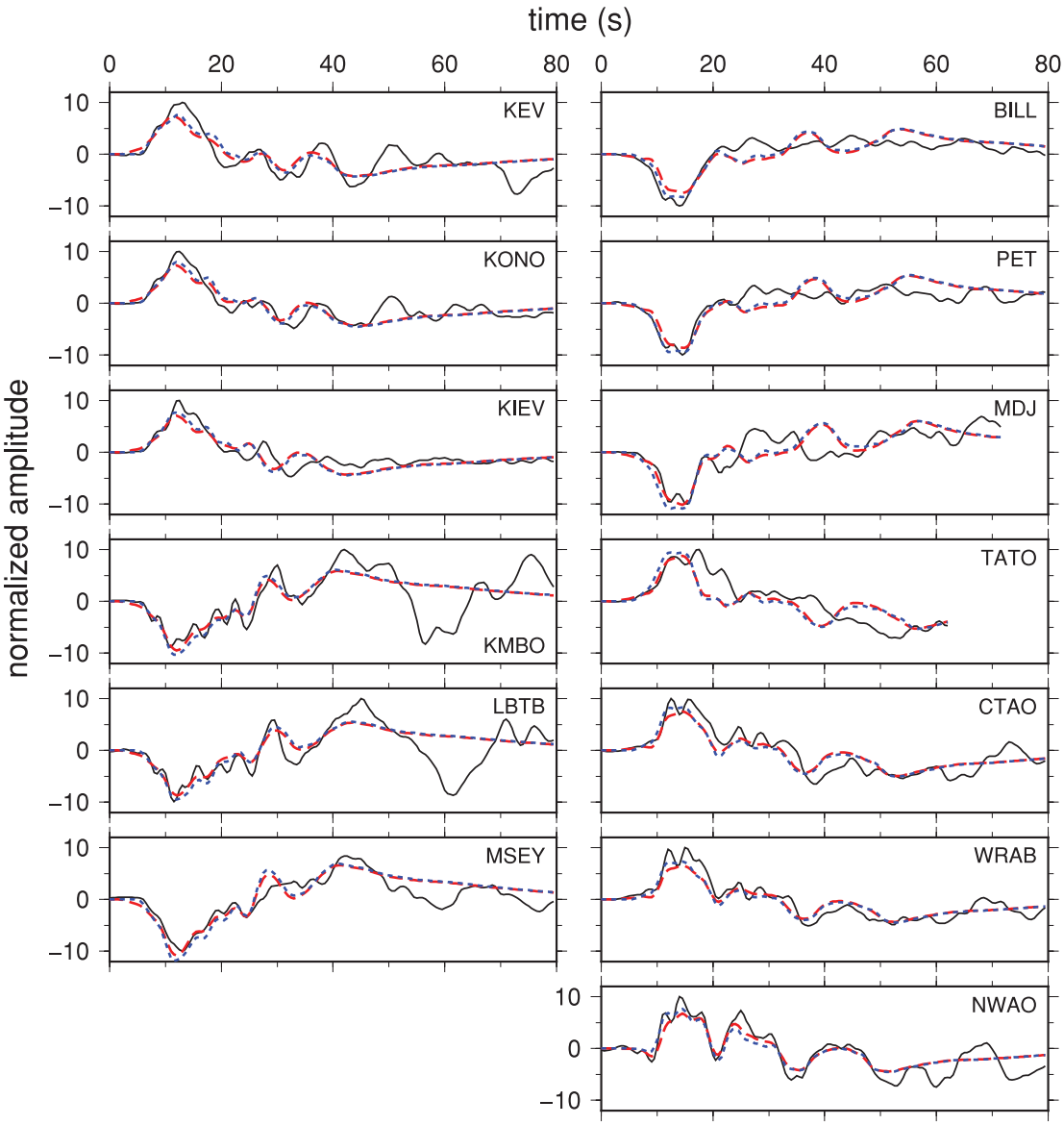

Figure 7. Fit of the seismic-only and joint inversions to the waveform data. The recorded waveforms are represented by black solid lines, synthetic waveforms from the seismic-only model by red dashed lines and synthetic waveforms from the joint-inversion model by blue dotted lines. The fit is reasonable for most stations, particularly in the early portion of the traces; differences in fit between the seismic-only and joint inversions are generally very small. Waveforms are ordered azimuthally by station (see Fig. 2 inset).

Table 3. Calculated model misfits to the InSAR and seismic input data sets.

Model	InSAR misfit	Seismic misfit
InSAR-only	32469[a] (0.0012)[b]	n/a
Seismic-only	132118[a] (0.0643)[b]	3153[a] (0.3602)[b]
Joint	38089[a] (0.0087)[b]	11254[a] (0.3657)[b]

[a]Weighted residual sum of squares between observed and modelled values, e.g. $(\mathbf{d}_{\mathrm{geod}} - \mathbf{H}_{\mathrm{geod}}\mathbf{a}_{\mathbf{total}})^{\mathrm{T}} \mathbf{E}_{\mathrm{geod}}^{-1}(\mathbf{d}_{\mathrm{geod}} - \mathbf{H}_{\mathrm{geod}}\mathbf{a}_{\mathbf{total}})$.
[b]Normalized misfit, e.g. $(\mathbf{d}_{\mathbf{seis}} - \mathbf{H}_{\mathrm{seis}}\mathbf{a})^{\mathrm{T}} (\mathbf{d}_{\mathbf{seis}} - \mathbf{H}_{\mathrm{seis}}\mathbf{a})/(\mathbf{d}_{\mathbf{seis}}{}^{\mathrm{T}}\mathbf{d}_{\mathbf{seis}})$.

of γ^2. Each 'slice' has associated with it a local minimum in ABIC, corresponding to the optimal model for that level of relative weighting between the InSAR and waveform data sets. At low values of γ^2 (<0.01), the surface of the ABIC function is quite flat overall, suggesting that neither temporal nor (especially) spatial smoothing provides a relatively strong constraint on the optimal model for inversions where teleseismic waveform data are dominant. As γ^2 increases, the response of ABIC to the smoothing constraints becomes more apparent, with plateauing of ABIC at high values of β^2 as well as high values of α^2. At larger values of γ^2 (>1), the value at which ABIC plateaus increases dramatically. At intermediate values of γ^2 (in the range 0.01–0.1), a prominent global minimum is developed, and the optimum values of the hyperparameters can therefore be determined ($\alpha^2 = 0.9$, $\beta^2 = 35$, $\gamma^2 = 0.025$).

The source–time function and spatial pattern of slip for the model associated with the ABIC minimum are shown in Figs 8(b) and (c), respectively. The overall shape of the moment rate function is very similar to that of the seismic-only model, differing only in the strength of the second peak of the initial double peak in moment release, which is larger in the joint-inversion model; this similarity suggests that the seismic data impose a strong constraint

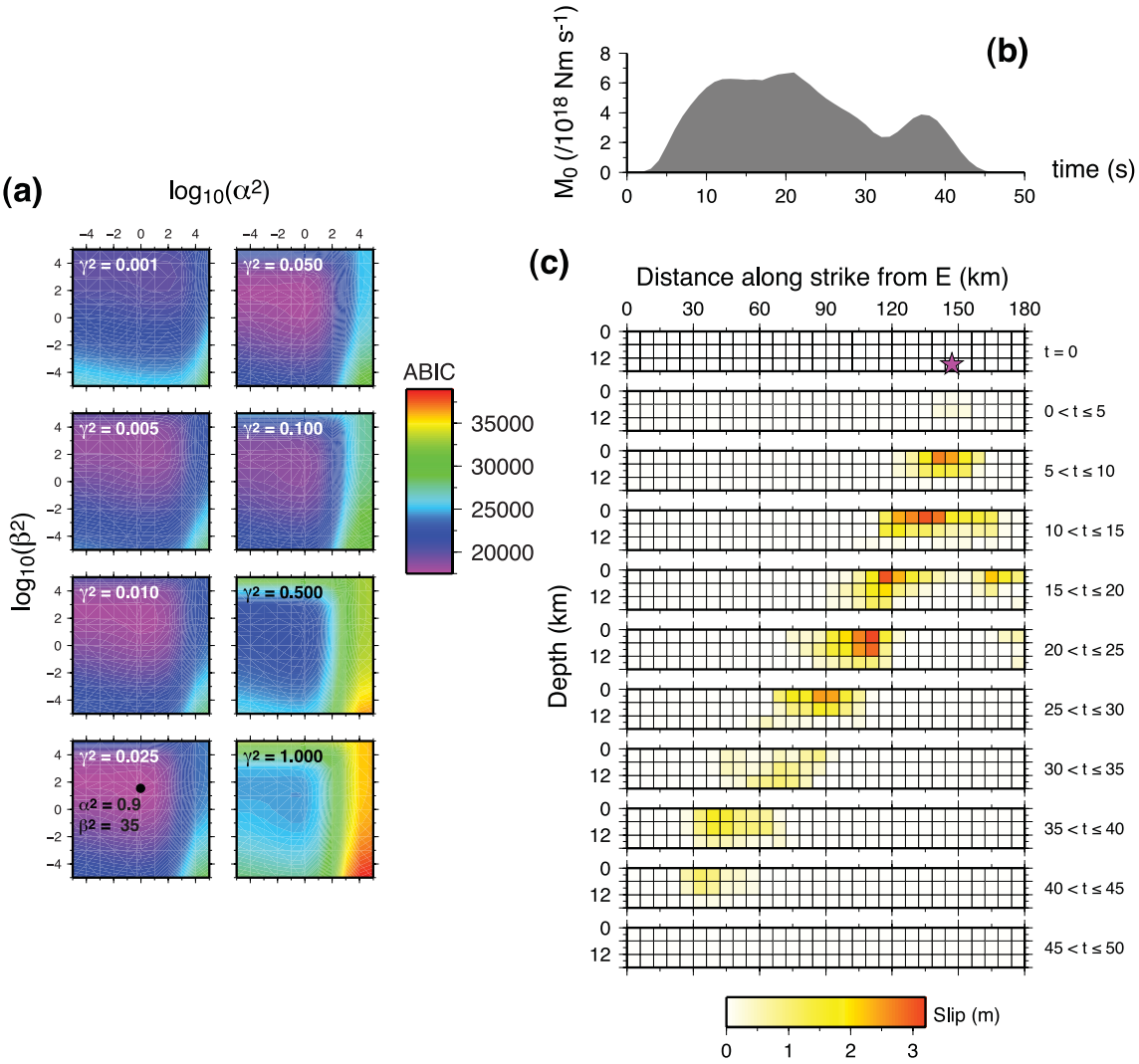

Figure 8. ABIC, source–time function and slip evolution for the joint inversion. (a) Distribution of ABIC in three-hyperparameter space. The ABIC minimum is found at $\alpha^2 = 0.9$, $\beta^2 = 35$, $\gamma^2 = 0.025$. (b) Moment–rate (source–time) function for the earthquake rupture, comprising of a broad first double peak (0–32 s) followed by a smaller, secondary peak (32–45 s). (c) Evolution of slip in the joint-inversion model, shown in 5-s intervals. Magenta star indicates location of the earthquake hypocentre. Slip propagates bilaterally for the first 20 s, and westwards for a further 25 s. Note the change in colour scale compared with the seismic inversion (Fig. 6c); here the peak slip in each time interval is typically 25–30 per cent larger.

on the rate of moment release in both models. Otherwise, as may be expected for such a similar moment rate pattern, the bulk behaviour of the models is very similar (e.g. the location and timing of the later concentration of slip at towards the western end of the fault rupture). The predominant difference between the spatiotemporal slip patterns of the seismic-only and joint-inversion models is that the peak slip in the joint model is both larger (peaking at around 7 m) and shallower (confined to the upper 6 km of the fault); these features can be best seen in the total slip patterns (Fig. 5), which suggest that the joint-inversion solution may have a closer affinity to the InSAR-only inversion model in terms of the spatial pattern of slip. The values of slip on each patch are typically much greater than the uncertainties in those values, which are of the order of 30 cm, substantially better than the uncertainties in the seismic-only model, and only marginally worse than those of the InSAR-only model. The seismic moment obtained $(1.73 \pm 0.01) \times 10^{20}$ Nm, is intermediate to the two single data set inversions, showing perhaps the influence of the seismic data in restricting deeper slip when compared with the InSAR-only case.

The fit to data for the joint-inversion model is good, although degraded when compared with the InSAR-only model (Table 3). Residual interferograms (Fig. 3) show that the fit to data for both the InSAR-only and joint models is comparable for the western InSAR track (076), but degraded on the central (305) and eastern (033) tracks. This is most likely the result of the restriction on deeper slip placed on the central and eastern portions of the model fault by the seismic data, and suggests that the deeper slip favoured by the InSAR-only model, that reduces the amplitude of the residual fringes on the corresponding tracks, may have a post-seismic origin. In contrast, the fit of the joint model is almost indistinguishable, both qualitatively and quantitatively from the seismic-only model fit when comparing waveforms (Fig. 7 and Table 3). It appears that the two data sets are indeed compatible, and that a model can be obtained that honours both; specifically, the teleseismic waveform data provide only a weak constraint on the spatial pattern of slip such that the major patches of slip can be rearranged, mostly by moving them to shallower depths, without strongly affecting the misfit.

4 DISCUSSION AND CONCLUSIONS

As we have demonstrated, a joint inversion between InSAR data and teleseismic waveform data holds a number of advantages over the cases where each data set is inverted alone. The fine spatial resolution of InSAR provides strong constraints on the spatial pattern of fault slip, particularly at shallow depths, but models based on InSAR data alone contain no information about the temporal evolution of fault slip. The detailed temporal variations captured by teleseismic waveform data, on the other hand, provide strong constraints on the rate of moment release with time, but have a much weaker control on the spatial locations of slipping areas, which in the case of the Manyi earthquake manifests itself in an erroneously rough slip distribution that violates the near-field InSAR displacements. Our joint inverse model of the Manyi event combines the advantageous characteristics of both data sets, showing a slip pattern that is strongly influenced by, and consistent with, the InSAR data, and a moment release history that is consistent with the teleseismic data.

Here we explore some implications of our inversion results, and issues raised by our methodology.

4.1 How should values of the hyperparameters be interpreted?

The hyperparameters α^2 and β^2 control the relative weights of the different prior smoothing constraints (spatial smoothing and temporal smoothing, respectively) with respect to the data. Their optimal values, as obtained in our joint-inversion scheme, are a function of the explanatory power of the data (i.e. how well they constrain the modelled slip). If the data were more effective at constraining the spatial slip pattern, for instance, we would expect α^2 to be smaller, when compared with less effective data, as the prior information on slip smoothness would be less important in the inversion. Scaling the kernel matrix, **H**, on the other hand, changes the 'balance' between the data and the prior constraints as influences on the inversion, and will cause a change in the optimal hyperparameter values. For example, if the values in the kernel matrix were scaled by a factor of 10, by changing the units of data from metres to decimetres, we would expect the optimal values of α^2 and β^2 to be scaled by a factor of 10^2.

It is not appropriate to directly compare the values of α^2 between the individual data set inversions (Figs 4 and 6), as they do not use the same kernels or data. On the other hand, we can compare the values of α^2 between the InSAR-only inversion and the joint inversion (Figs 4 and 8), and between the seismic-only inversion and the joint inversion (Figs 6 and 8)—although in the latter case, to account for the optimal relative weighting of the seismic data in the joint inversion, the hyperparameter values for the seismic inversion should be multiplied by 40 (i.e. $1/\gamma^2$). According to these comparisons, the spatial smoothing constraint of the joint inversion (0.9) is stronger than that of the InSAR-only inversion (0.35), but weaker than that of the seismic-only inversion ($0.03 \times 40 = 1.2$), implying that the InSAR data require weaker smoothing than the seismic data due to the strength of the constraint they provide on the spatial pattern of fault slip.

The hyperparameter γ^2, controls the relative weight of the two data sets with respect to each other. In keeping with the formulation presented earlier, high values of γ^2 result in a joint-inverse model dominated by geodetic data, and vice versa. To illustrate the effect of varying the relative weights of the data sets, we show slip inversion results, along with values of ABIC and smoothing hyperparameter values, for a range of different values of γ^2 in Fig. S1. A continuum of slip models can be seen between the InSAR- and seismic-dominated end-members, that can be used to assess how strongly individual slip features are required by the different data sets.

4.2 Are differences in moment estimates between models due to post-seismic deformation in the InSAR data?

The remote location of the Manyi earthquake meant that there were no post-seismic GPS data collected; all our knowledge of post-seismic deformation due to the earthquake comes from InSAR data. Post-seismic interferograms spanning the period from 8 d to 3 yr after the earthquake show a deformation signal that has the same polarity and a maximum displacement in a similar location along-strike as the coseismic deformation (Ryder *et al.* 2007). This is consistent with, if not definitive evidence of, post-seismic afterslip on an expanded portion of the earthquake-bearing fault. Ryder *et al.* (2007) estimate that if the observed deformation were all due to afterslip, up to 0.7 m of afterslip could have occurred on this expanded fault, mostly at depths greater than 15 km, with the total moment released equivalent to around 20 per cent of the coseismic moment. Given that the interferograms we use in this study cover the earliest portion of the post-seismic period, it is likely that we have captured some portion of post-seismic signal in our InSAR data that may have biased the models using them as a constraint, mapping that additional deformation into the estimated coseismic slip distribution. An additional complication is that the different interferograms cover different amounts of the post-seismic period, between 8 and 40 d (Table 1).

We can attempt to place bounds on the amount of (i.e. moment due to) post-seismic deformation in the model by extrapolating the earliest post-seismic afterslip rate obtained by Ryder *et al.* (2007). If we assume that an average of 24 d of deformation is preserved in the data (the mean post-seismic duration of the three interferograms), and use the expanded fault geometry and average early post-seismic slip rate obtained by Ryder *et al.*, we obtain a total moment release of 0.04×10^{20} Nm, which is significantly less than the difference in moment between the joint-inversion and InSAR-only models, even at the 95 per cent confidence level, of $(0.17 \pm 0.04) \times 10^{20}$ Nm. Some of this difference may be a result of the length of post-seismic interval chosen; however, given the approximately linear deformation rate obtained by Ryder *et al.* over the first few months of the post-seismic observation period, even an expansion of the interval considered to 40 d, the longest post-seismic interval covered by our data, would not account for more than half of the discrepancy between models. Of course, other explanations for, or contributions to, the discrepancy in moment are possible. There could be problems with the simplified fault geometry used in this study, for instance, which would likely affect the InSAR inversion more than the inversions including the seismic data, given the lower sensitivity of the latter data set to fault geometry. The assumption of fixed rake could also affect the moment estimate from InSAR data, given that rake strongly affects the relative proportions of horizontal and vertical surface displacement, and the much greater sensitivity of InSAR data to vertical displacements. Another potential explanation could be that the limited azimuthal coverage of the seismic data is insufficient to capture the full moment release of the earthquake. In order to more definitively assess the reasons for the differences in moment estimates between the different inversions, a more complete examination of each of these effects would be necessary, which is beyond the scope of this study.

It is at least plausible, then, that a portion of unmodelled rapid post-seismic deformation, occurring over the first 8 d following the earthquake, could be present in the InSAR data, and thus the InSAR-only model. This is corroborated by the location of the largest InSAR residuals for the joint-inversion model being in a similar location to the early post-seismic displacements observed by Ryder *et al.*—in the eastern half of track 305 and the western half of track 033, within 30 km of the fault trace (Fig. 3)—consistent with the deformation signals having a similar origin. Such early afterslip would be similar in timing, if not in amplitude, to the rapid post-seismic afterslip observed following the 2004 Parkfield, California, earthquake, where deformation approximately equal in moment to that of the main shock occurred over the first 3 d following the earthquake (Johanson *et al.* 2006; Johnson *et al.* 2006).

4.3 Application of the method to variable fault geometries

The case study presented here of the Manyi earthquake is intended as a demonstration of concept, specifically of how our ABIC joint-inversion scheme can estimate the optimal relative weighting of two different data sets. We used a fixed fault geometry as a means of reducing the number of model parameters under consideration, and thus the computation time. However, this is not a necessary constraint—geometric parameters such as fault dip or strike can be incorporated into this inversion scheme, increasing its flexibility, albeit at the expense of additional computation time, as we shall explain.

The value of ABIC for a given set of hyperparameter values is strongly affected by the value of the misfit parameter $s(\mathbf{a}^*)$ (eqs 17 and 25), quantifying the fit of the model predictions to the data and their compatibility with the prior information. It is conceivable that by varying the fault geometry (e.g. by changing the strike or dip of the fault) a more compatible model could be found, resulting in a smaller value of $s(\mathbf{a}^*)$, and therefore of ABIC. In this case, the parameters controlling fault geometry would act effectively as additional hyperparameters in the inversion, and could be included in the parameter search. Such an approach was taken by Fukahata & Wright (2008) in their InSAR study of the slip distribution of the 1995 Dinar, Turkey, earthquake. In that study, in addition to a hyperparameter controlling the spatial roughness of the slip distribution (analogous to α^2 in this study), fault dip was explicitly treated as a second hyperparameter. The minimum-ABIC solution thus provided both the optimal level of roughness and the optimal fault dip appropriate for the data and fault dimensions used.

The practicality of such an approach is controlled in large part by the computation time required to calculate the necessary kernel matrices $\mathbf{H}_{\mathbf{geod}}$ and/or $\mathbf{H}_{\mathbf{seis}}$, which will change as the fault geometry changes. In the geodetic-only case considered by Fukahata and Wright, the kernels were calculated using the analytical solutions for a rectangular dislocation in an elastic half-space (e.g. Okada 1985), a calculation that is very rapid. On the other hand, the calculation of the seismic kernels $\mathbf{H}_{\mathbf{seis}}$ used in this study is not analytical; rather, it is a numerical calculation that requires several orders of magnitude more computation time. Changes to the fault geometry would require recalculating the kernels for each fault patch as well as the re-estimation of the hypocentre location on the fault plane, and, potentially, of the maximum rupture velocity (see Section 3.3 for more details). In addition to the need to recalculate the kernels for changes in fault geometry, the addition of more hyperparameters to the parameter search will also result in an exponential increase in the number of computations of ABIC. These issues may be mitigated in part by parallelization of the computations and/or the use of more efficient methods for searching parameter space; we intend to explore these possibilities in future work.

ACKNOWLEDGEMENTS

This work was supported by the Natural Environment Research Council (NERC) through the Centre for the Observation and Modelling of Earthquakes and Tectonics (COMET), as well as through a research studentship to GJF. GJF also acknowledges support from a Hertford College Senior Scholarship, and a Lindemann Trust Postdoctoral Fellowship. Thanks to David Robinson for providing aftershock relocations and Isabelle Ryder for supplying the details of her post-seismic model results; Tim Wright is thanked for useful discussions. We are grateful to Takeshi Sagiya and one anonymous reviewer for providing constructive comments that have helped us to improve the manuscript. All ERS synthetic aperture radar data are copyrighted by the European Space Agency, and were acquired under project AOE-621. Some of the figures were made using the public domain Generic Mapping Tools (Wessel & Smith 1998).

REFERENCES

Akaike, H., 1977. On entropy maximization principle, in *Application of Statistics,* pp. 27–41, ed. Krishnaiah, P.R., North-Holland.

Akaike, H., 1980. Likelihood and the Bayes procedure, in *Bayesian Statistics,* pp. 143–166, eds Barnardo, J.M., DeGroot, M.H., Lindley, D.V. & Smith, A.F.M., Valencia University Press.

Asano, K., Iwata, T. & Irikura, K., 2005. Estimation of source rupture process and strong ground motion simulation of the 2002 Denali, Alaska earthquake, *Bull. seism. Soc. Am.,* **95,** 1701–1715.

Bayes, T., 1763. An essay towards solving a problem in the doctrine of chances, *Phil. Trans. R. Soc. Lond. A,* **53,** 370–418.

Bro, R. & De Jong, S., 1997. A fast non-negativity-constrained least squares algorithm, *J. Chemometr.,* **11,** 392–401.

Clévédé, E., Bouin, M.-P., Bukchin, B., Mostinskiy, A. & Patau, G., 2004. New constraints on the rupture process of the 1999 August 17 Izmit earthquake deduced from estimates of stress glut rate moments, *Geophys. J. Int.,* **159,** 931–942.

Delouis, B., Giardini, D., Lundgren, P. & Salichon, J., 2002. Joint inversion of InSAR, teleseismic and strong motion data for the spatial and temporal distribution of earthquake slip: application to the 1999 Izmit mainshock, *Bull. seism. Soc. Am.,* **92**(1), 278–299.

Dewey, J.W., 1971. Seismicity studies with the method of joint hypocenter determination, *PhD thesis,* University of California, Berkeley, USA.

Farr, T. & Kobrick, M., 2000. Shuttle Radar Topography Mission produces a wealth of data, *EOS, Trans. Am. geophys. Un.,* **81,** 583–585.

Fukahata, Y. & Wright, T.J., 2008. A non-linear geodetic data inversion using ABIC for slip distribution on a fault with an unknown dip angle, *Geophys. J. Int.,* **173,** 353–364.

Fukahata, Y., Yagi, Y. & Matsu'ura, M., 2003. Waveform inversion for ABIC for seismic source processes with two sorts of prior constraints: comparison between proper and improper formulations, *Geophys. Res. Lett.,* **30**(6), doi:10.1029/2002GL16293.

Fukahata, Y., Nishitani, A. & Matsu'ura, M., 2004. Geodetic data inversion using ABIC to estimate slip history during one earthquake cycle with viscoelastic slip-response functions, *Geophys. J. Int.,* **156,** 140–153.

Funning, G.J., 2005. Source parameters of large shallow earthquakes in the Alpine-Himalayan belt from InSAR and waveform modelling, *PhD thesis,* University of Oxford.

Funning, G.J., Parsons, B., Wright, T.J., Jackson, J.A. & Fielding, E.J., 2005. Surface displacements and source parameters of the 2003 Bam (Iran) earthquake from Envisat advanced synthetic aperture radar imagery, *J. geophys. Res.,* **110**(B09406), doi:10.1029/2004JB003338.

Funning, G.J., Parsons, B. & Wright, T.J., 2007. Fault slip in the 1997 Manyi, Tibet earthquake from linear elastic modelling of InSAR displacements, *Geophys. J. Int.,* **169,** 1009–1027.

Hernandez, B., Cotton, F. & Campillo, M., 1999. Contribution of radar interferometry to a two-step inversion of the kinematic process of the 1992 Landers earthquake, *J. geophys. Res.,* **104,** 13 083–13 099.

Jeffreys, H. & Bullen, K.E., 1940. *Seismological Tables,* British Association for the Advancement of Science.

Johanson, I.A., Fielding, E.J., Rolandone, F. & Bürgmann, R., 2006. Coseismic and postseismic slip of the 2004 Parkfield earthquake from space-geodetic data, *Bull. seism. Soc. Am.,* **96,** S269–S282.

Johnson, K.M., Bürgmann, R. & Larson, K., 2006. Frictional properties on the San Andreas fault near Parkfield, California inferred from models of afterslip following the 2004 earthquake, *Bull. seism. Soc. Am.,* **96**(4B), S321–S338.

Kikuchi, M. & Kanamori, H., 1991. Inversion of complex body waves III, *Bull. seism. Soc. Am.,* **81,** 2335–2350.

Kim, A. & Dreger, D., 2008. Rupture process of the 2004 Parkfield earthquake from near-fault seismic waveform and geodetic records, *J. geophys. Res.,* **113,** B07308, doi:10.1029/2007JB005115.

Okada, Y., 1985. Surface deformation due to shear and tensile faults in a half-space, *Bull. seism. Soc. Am.,* **75**(4), 1135–1154.

Parsons, B., Wright, T.J., Rowe, P., Andrews, J., Jackson, J.A., Walker, R., Khatib, M.M. & Talebian, M., 2006. The 1994 Sefiadbeh (eastern Iran) earthquakes revisited: new evidence from satellite radar interferometry and carbonate dating about the growth of an active fold above a blind thrust fault, *Geophys. J. Int.,* **164,** 202–217.

Peltzer, G., Crampé, F. & King, G., 1999. Evidence of nonlinear elasticity of the crust from the M_w 7.6 Manyi (Tibet) earthquake, *Science,* **286,** 272–276.

Rolandone, F., Dreger, D., Murray, M. & Bürgmann, R., 2006. Coseismic slip distribution of their 2003 m_w 6.6 San Simeon earthquake, California, determined from GPS measurements and seismic waveform data, *Geophys. Res. Lett.,* **33,** L16315, doi:10.1029/2006GL027079.

Ryder, I., Parsons, B., Wright, T.J. & Funning, G.J., 2007. Post-seismic motion following the 1997 Manyi (Tibet) earthquake: InSAR observations and modelling, *Geophys. J. Int.,* **169,** 1009–1027.

Salichon, J., Delouis, B., Lundgren, P., Giardini, D., Constantini, M. & Rosen, P., 2003. Joint inversion of broadband teleseismic and interferometric synthetic aperture radar (InSAR) data for the slip history of the $m_w = 7.7$ Nazca ridge (Peru) earthquake of 12 November 1996, *J. geophys. Res.,* **108,** doi:10.1029/2001JB000913.

Simons, M., Fialko, Y. & Rivera, L., 2002. Coseismic deformation from the 1999 M_w 7.1 Hector Mine, California earthquake as inferred from InSAR and GPS observations, *Bull. seism. Soc. Am.,* **92,** 1390–1402.

Steck, L.K., Velasco, A.A., Cogbill, A.H. & Patton, H.J., 2001. Improving regional seismic event location in China, *Pure appl. Geophys.,* **36,** 192–205.

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

Vallée, M. & Bouchon, M., 2004. Imaging coseismic rupture in far field by slip patches, *Geophys. J. Int.,* **156,** 615–630.

Velasco, A.A., Ammon, C.J. & Beck, S.L., 2000. Broadband source modeling of the November 8, 1997, Tibet ($m_w = 7.5$) earthquake and its tectonic implications, *J. geophys. Res.,* **105,** 28 065–28 080.

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

Wen, Y.-Y. & Ma, K.-F., 2010. Fault geometry and distribution of asperities of the 1997 Manyi, China ($M_w = 7.5$), earthquake: Integrated analysis from seismological and InSAR data, *Geophys. Res. Lett.,* **37,** doi:10.1029/2009GL041976.

Wessel, P. & Smith, W.H.F., 1998. New, improved version of generic mapping tools released, *EOS, Trans. Am. geophys. Un.,* **79,** 579.

Wright, T.J., Lu, Z. & Wicks, C., 2003. Source model for the M_w 6.7, 23 October 2002, Nenana Mountain Earthquake (Alaska) from InSAR, *Geophys. Res. Lett.,* **30**(18), doi:10.1029/2003GL018014.

Wu, C.J., Takeo, M. & Ide, S., 2001. Source process of the Chi-Chi earthquake: a joint inversion of strong motion data and global positioning system data with a mulitfault model, *Bull. seism. Soc. Am.,* **91,** 1128–1143.

Yabuki, T. & Matsu'ura, M., 1992. Geodetic data inversion using a Bayesian information criterion for spatial distribution of fault slip, *Geophys. J. Int.,* **109,** 363–375.

Yagi, Y. & Fukahata, Y., 2008. Importance of covariance components in inversion analyses of densely sampled observed data: an application to waveform data inversion for seismic source processes, *Geophys. J. Int.,* **175,** 215–221.

Yagi, Y., Mikumo, T., Pacheco, J. & Reyes, G., 2004. Source rupture process of the Tecoman, Colima, Mexico earthquake of 22 January, 2003, determined by joint inversion of teleseismic body-wave and near-source data, *Bull. seism. Soc. Am.,* **94,** 1795–1807.

SUPPORTING INFORMATION

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

Figure S1. The relationship between the relative weighting parameter γ^2, the other hyperparameters and the fault slip distribution (http://gji.oxfordjournals.org/lookup/suppl/doi:10.1093/gji/ggt406/-/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.