Simultaneous Bayesian Estimation of Non-Planar Fault Geometry and Spatially-Variable Slip

Rishabh Dutta[1], Sigurjón Jónsson[1] and Hannes Vasyura-Bathke[1,2]

[1]King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.
[2]Now at University of Potsdam, Potsdam, Germany.

Key Points:

- Non-planar fault geometries are parameterized using a set of polynomials allowing for along-strike and down-dip variations of the fault geometry.
- The non-planar fault geometrical parameters are estimated simultaneously with spatially-variable slip from geodetic data using Bayesian inference.
- Estimated fault slip asperities and their locations are found to be biased when simple planar faults are assumed in presence of non-planar fault geometries.

Corresponding author: Rishabh Dutta, `rishabh.dutta@kaust.edu.sa`

Abstract

Large earthquakes are usually modeled with simple planar fault surfaces or a combination of several planar fault segments. However, in general, earthquakes occur on faults that are non-planar and exhibit significant geometrical variations in both the along-strike and down-dip directions at all spatial scales. Mapping of surface fault ruptures and high-resolution geodetic observations are increasingly revealing complex fault geometries near the surface and accurate locations of aftershocks often indicate geometrical complexities at depth. With better geodetic data and observations of fault ruptures, more details of complex fault geometries can be estimated resulting in more realistic fault models of large earthquakes. To address this topic, we here parametrize non-planar fault geometries with a set of polynomial parameters that allow for both along-strike and down-dip variations in the fault geometry. Our methodology uses Bayesian inference to estimate the non-planar fault parameters from geodetic data, yielding an ensemble of plausible models that characterize the uncertainties of the non-planar fault geometry and the fault slip. The method is demonstrated using synthetic tests considering checkerboard fault-slip patterns on non-planar fault surfaces with spatially-variable dip and strike angles both in the down-dip and in the along-strike directions. The results show that fault-slip estimations can be biased when a simple planar fault geometry is assumed in presence of significant non-planar geometrical variations. Our method can help to model earthquake fault sources in a more realistic way and may be extended to include multiple non-planar fault segments or other geometrical fault complexities.

1 Introduction

With increasing availability and improving spatial and temporal resolution of geodetic data, more details of earthquake fault geometries and slip can be determined. Detailed estimates of fault model parameters are beneficial for better understanding of the earthquake mechanics at the different fault systems in the world. However, for the same earthquake, notably dissimilar coseismic fault-slip models have been produced by different authors depending on their diverse estimation methods and modeling assumptions, e.g. regarding the fault geometry, elastic layering, smoothing parameters, etc. (Lay, 2018; Mai & Thingbaijam, 2014; Razafindrakoto, Mai, Genton, Zhang, & Thingbaijam, 2015). Bayesian inference of earthquake sources allows for constraining the posterior probability distributions of the different fault model parameters and can also take uncertain knowledge of the elastic layering, elastic parameters, etc. into account through model covariances or *a priori* constraints (Duputel, Rivera, Fukahata, & Kanamori, 2012; Dutta, Jónsson, Wang, & Vasyura-Bathke, 2018; Fukuda & Johnson, 2008, 2010; Matsu'ura, Noda, & Fukahata, 2007; Minson, Simons, & Beck, 2013; Yagi & Fukahata, 2008). Despite this flexibility, most studies to date have approximated the earthquake source either as a single planar fault or a combination of several planar fault segments. Furthermore, when more complex fault geometries are used, they are usually estimated a priori and have not been varied in the Bayesian estimation because parameterizing complex non-planar geometries can result in hundreds of additional model parameters to be estimated together with a dramatic increase of Green's function calculations.

While most earthquake faults are considered planar in coseismic fault slip modeling, multiple lines of evidence show that faults are typically more complex and can, for example, consist of en echelon segments, have bends, be curved, or warped at different spatial scales (e.g., Duman, Emre, Dogan, & Ozalp, 2005; Klinger, 2010; Maerten, Resor, Pollard, & Maerten, 2005; Martel, 1999; Wesnousky, 1988). The growth of non-planar faults can be explained by either linkage of pre-existing, discontinuous, non-coplanar structures (Bürgmann & Pollard, 1994a; Cruikshank & Aydin, 1994; Segall & Pollard, 1980, 1983), during propagation as non-coplanar shear fractures (Cox & Scholz, 1988a, 1988b; Vermilye & Scholz, 1998), or due to heterogeneous mechanical conditions along a fault (Martel, 1999). Elastic analyses show that faults remain planar only if the stress drop across the fault is uniform, it is surrounded by homogeneous, and isotropic rock, and the far-field stress is uniform (Bürgmann, Pollard, & Martel, 1994b; Martel, 1999). Otherwise, the fault grows as a non-planar fault as the rupture propagates. Exhumed faults and mapped surface earthquake ruptures (e.g., San Andreas, Zirkuh, Chi-Chi, Wenchuan fault ruptures, etc.) show approximately self-similar or self-affine fractal characteristics that can continue to scales as large as tens to hundreds of kilometers, indicating that in general faults are non-planar (Candela et al., 2012; Power & Tullis, 1995).

While earthquake fault complexities at the surface can be directly observed from fault maps and surface displacement discontinuities in geodetic data (Wesnousky, 2008), fault complexities at depth are indicated by aftershocks and other earthquake locations (Dutta et al., 2018; Improta et al., 2019; Kaven & Pollard, 2013). As the availability of high-resolution geodetic data from GNSS, InSAR and optical images increases, more and more details of non-planar earthquake fault geometries can be constrained, not only at the surface but also at depth. Many studies have used multiple planar fault segments to represent geometrical fault complexities, with the fault-strike and -dip angles of each segment estimated from the geodetic data considering uniform slip (e.g., Jónsson, Zebker, Segall, & Amelung, 2002; Reilinger et al., 2000; Shen et al., 2009; Sudhaus & Jónsson, 2011). However, when planar fault segments are used to describe curved or warped surface fault ruptures or change in fault shape at depth, it usually results in unphysical gaps and/or intersections of fault segments leading to slip singularities at those geometric irregularities. In an attempt to address this problem, Maerten et al. (2005) used triangular dislocation elements in an elastic half-space to construct a non-planar fault model for the 1999 Hector Mine earthquake and showed a 32% fit improvement to observed geodetic data compared to when using planar fault segments. In addition, the 1995 Kozani-Grevena earthquake (Resor, Pollard, Wright, & Beroza, 2005) and the 2003 M_W6.8 Chengkung earthquake (Hsu, Yu, & Chen, 2009) were explained by faults with curvilinear tiplines that were constructed based on aftershock locations. Furthermore, non-planar geometries have been used for several other earthquakes leading to better fit to geodetic data, e.g., for the 2003 M_W6.8 Zemmouri (Belabbès, Wicks, Çakir, & Meghraoui, 2009), 1994-2004 Al Hoceima-Morroco (Akoglu et al., 2006), 2010 M_W6.9 Yushu (Jiang et al., 2013), 2008 M_W7.1 Yutian (Furuya & Yasuda, 2011), 2008 M_W 6.9 Iwate-Miyagi (Abe, Furuya, & Takada, 2013), and 2008 M_W6.4 Balochistan earthquakes (Usman & Furuya, 2015).

For all the studies mentioned above and for most other estimation studies using non-planar earthquake fault geometries, the complex fault geometry was determined *a priori* before fault slip was estimated. The selection of the fault geometry has usually been based on other sources of information, e.g. aftershock locations, geological maps, mapped surface ruptures, seismic reflection/refraction profiles, borehole data and/or slab models for subduction-zone earthquakes. However, as there is a trade-off between the choice of fault geometry and the amount of fault slip estimated (Ragon, Sladen, & Simons, 2018; Razafindrakoto et al., 2015), the estimated slip would likely be biased when the fault geometry (whether planar or non-planar) is fixed and different possible geometries neglected. While varying planar geometries has often been carried out in earlier estimation studies (Elliott et al., 2016; Fukahata & Wright, 2008), varying non-planar geometries has rarely been attempted. Bathke, Nikkhoo, Holohan, and Walter (2015) varied non-planar geometry of the caldera ring-fault at Tendurek volcano (Turkey) to better explain ring-like InSAR displacements. Also, Wan, Shen, Burgmann, Sun, and Wang (2017) varied the listric geometry of the Beichuan fault segment constrained using geodetic observations of the 2008 M_W 7.9 Wenchuan earthquake.

In this paper, we introduce a method for simultaneously estimating complex non-planar earthquake fault geometry and spatially-variable fault slip. We parametrize the non-planar fault geometry with a set of polynomial parameters that allow for fault curvature both along the fault strike as well as in the down-dip direction. Using Bayesian inference, an ensemble of fault model parameters corresponding to a posterior distribu-

revision1/rev_figs/Figure1.jpg

Figure 1: Examples of non-planar fault geometry: (a) A fault with horizontal curvilinear tiplines and a uniform dip angle (close to vertical). Inset shows the trace of the top edge of the fault. (b) A fault with straight horizontal tiplines but curved in the down-dip direction. Inset shows the profile of the curved vertical tipline. (c) A fault with both curved horizontal and vertical tiplines such that the top-edge strike is generally different from that of the bottom edge. Insets show (ii) the top and (iii) the bottom edge of the fault. The X, Y and Z axes can have arbitrary units of length.

tion is estimated that accords to the data likelihood and a priori constraints. Such an ensemble of fault geometrical parameters yields both fault-location uncertainties and fault-dip angle uncertainties along both the strike and down-dip directions of the fault. We then demonstrate this simultaneous estimation of non-planar geometry and slip distribution using three synthetic test cases, and compare them with results when planar geometries are used.

2 Model Parametrization and the Forward Model

We parametrize the complex non-planar fault geometry with only a few polynomial parameters (less than 10) to vary the non-planar fault geometrical structure either along the strike or the down-dip direction or both with various degrees of freedom. Since we consider a non-planar finite fault in our parametrization, the top and bottom edges of the fault are termed horizontal tiplines and other edges are termed vertical tiplines in this paper. The two horizontal tiplines each lie at a particular z-plane of the 3D Cartesian coordinate system in which our finite fault is defined. Three examples of complex fault geometries are shown in Fig. 1. The first example is a fault with curvilinear horizontal tiplines and a constant fault-dip angle (close to vertical), whereas the second example is a fault with straight horizontal but curved vertical tiplines. The third example is a fault that includes both curved horizontal tiplines and a curvature in the down-dip direction. We later demonstrate how to simultaneously determine these fault geometries along with spatially-variable slip from geodetic data using Bayesian inference.

The fault model parameters in our problem, denoted by $\boldsymbol{\theta}$ in the M-dimensional model space $\mathbb{M}$, are a combination of fault geometrical parameters $\boldsymbol{m}$, fault-slip parameters $\boldsymbol{s}$, and hyperparameters σ_1^2, σ_2^2, $\cdots$,σ_p^2, such that : $\boldsymbol{\theta} = [\boldsymbol{m}\ \boldsymbol{s}\ \sigma_1^2\ \sigma_2^2\ \cdots\ \sigma_p^2]^T$. The geometrical model parameters $\boldsymbol{m}$ constrain the 3D fault geometry with the fault surface comprising of both curved horizontal and vertical tiplines, while the slip model parameters $\boldsymbol{s}$ are the slip values on this non-planar fault surface. The hyperparameters (σ_i^2) control the weights of direct/indirect priors with respect to the data likelihood (see Section 3.2). Our problem utilizes geodetic data (e.g., GNSS and/or InSAR) denoted as data vector $\boldsymbol{d}$ in the N-dimensional data space $\mathbb{N}$ that relates to the model parameter vectors $\boldsymbol{m}$ and $\boldsymbol{s}$ through: $\boldsymbol{d} = G(\boldsymbol{m}, \boldsymbol{s}) + \epsilon$, where $G(\boldsymbol{m}, \boldsymbol{s}) = G'(\boldsymbol{m})\boldsymbol{s}$ is the predicted data and ϵ is the error vector. The finite non-planar fault constrained by our parametrization is discretized using a fixed number of triangular dislocation elements (TDEs) so that the fault is meshed without any gaps or overlaps that can arise when using rectangular elements for a non-planar surface. This finite fault is placed within a homogeneous and isotropic elastic half-space and the predicted data $G'(\boldsymbol{m})\boldsymbol{s}$ is computed using an analytic solution (Meade, 2007) with $\boldsymbol{s}$ describing the slip on the TDEs.

The different types of non-planar fault geometries shown in Fig. 1 can be parametrized by the technique described below. However, the steps to follow to construct the non-planar

revision1/rev_figs/Figure2.jpg

Figure 2: Construction of a non-planar fault surfaces using multiple geometrical parameters. Continued on the next page ...

Figure 2: (a) The green line shows the top edge of the fault at depth ${}^{1}z$ and dots denote its regular discretization with coordinates of three points A, B and P shown. The inset shows strike ψ_i calculated at i^{th} point F as the average of the angles ∠MEF ($\psi^{'}_{i-1}$) and ∠NFG ($\psi^{'}_{i}$). (b) Green lines show the down-dip polynomials passing through the discretized top-edge of the fault. Points A, ^{2}A, ..., ^{q}A are the discretizations of the polynomial passing through point A on the top edge. The inset shows the points F and ${}^{q'}$F on the line F ${}^{q'}$F in the $z = {}^{1}z$ plane passing through the ith point F on the top edge, where the curve EFG is a portion of the top edge, also shown in the inset of (a). The x- and y-coordinates of the polynomial in the down-dip direction passing through point F (ith discretization on top edge) are obtained by discretizing the line F ${}^{q'}$F which strikes at angle (ψ_i−90°) in the plane $z = {}^{1}z$. The z-coordinates are determined by Eq. (1) using the geometric parameters D_1 and D_2. (c) Gray lines show the down-dip curves after adjusting them such that the fault's lower edge strikes differently from the top edge. Light green and dark green lines show the previous and adjusted fault-bottom edges, respectively. The inset shows a part of the fault's bottom edge changing from curve ^{q}A ^{q}Q ^{q}F to curve ^{q}A ^{q*}Q ^{q}F depending on the geometric parameters S_1 and S_2 (see Eq. A.1). (d) The resulting fault geometry discretized using triangular dislocation elements (TDEs). Note: The x- and y- coordinates in the cartesian system are denoted as double-indexed variable (for e.g., ${}^{1}x_2$), where the top-left index denotes the discretization in the down-dip direction and the bottom-right index denotes that in the along-strike direction.

fault geometry and the number of parameters may differ for different types of non-planar faults depending on the complexity desired. Fig. 2 shows the parametrization steps for a fault geometry with curved horizontal and vertical tiplines and they are as follows:

(i) We first determine the top edge of the fault surface (curve AP in Fig. 2a) and discretize this top edge at regular intervals (points A, B, $\cdots$, P in Fig. 2a). We often have good *a priori* information about the fault trace, e.g., from geological maps, mapped surface ruptures, coseismic interferograms or image offsets. Such a fault trace can be considered as a single linear segment or a set of several connected linear fault segments with different strike angles. In case of a buried fault, the top edge of the fault can be parametrized using a 2^{nd} or 3^{rd} degree polynomial. The top edge (curve AP in Fig. 2a) is then discretized with equidistant points using piece-wise linear segments.

(ii) To allow for curvature in the down-dip direction, we introduce two parameters D_1 and D_2 (or three depending on the level of complexity desired in the down-dip direction) to define polynomials that pass through the discretized points of the top edge (for e.g., curve A ^{q}A in Fig. 2b where point A is on the top edge). Each of these polynomials follow the equations:

$$\begin{aligned} z &= D_2(x^2 - ({}^{1}x_i)^2) + D_1(x - {}^{1}x_i) + {}^{1}z; \\ y &= \tan(\psi_i)(x - {}^{1}x_i) + {}^{1}y_i, \end{aligned} \tag{1}$$

where D_1 and D_2 are the two parameters that are the same for all these down-dip polynomials (curve A ^{q}A, B ^{q}B, etc.). The terms x, y and z are the x-, y- and z- coordi-

revision1/rev_figs/Figure3.jpg

Figure 3: Effects of the geometrical parameters in varying the complex non-planar fault geometry. (a) Fault vertical cross-sections showing how parameter D_1 controls the down-dip slope of the down-dip polynomial. Inset shows the fault in 3D with different D_1 values. (b) Parameter D_2 controls the curvature of the fault in the down-dip direction. Inset shows the fault in 3D with different D_2 values. (c) Insets (i)-(iii) show S-shaped and D-shaped faults in 3D with different S_1 values. Inset (iv) shows the corresponding color-coded fault-bottom tiplines. (d) Insets (i)-(iii) show the 3D faults with change in intensity of the along-strike curvature for different values of S_2. Inset (iv) shows the corresponding color-coded fault-bottom tiplines. The X, Y and Z axes can have arbitrary units of length.

nates of these polynomials, respectively, which pass through the ith point of the top edge $({}^1x_i, {}^1y_i, {}^1z)$. ψ_i is the azimuth of the ith polynomial (see Fig. 2a).

(iii) The polynomials are truncated at a desired depth and they are discretized at regular distance intervals in the down-dip direction. At this point, the curvature of the fault's lower edge in the fault-strike direction is the same as that of the top edge of the fault.

(iv) To change the curvature of the fault's lower edge in the along-strike direction between points ^{q}A and ^{q}F (in Fig. 2c), we introduce two more parameters S_1 and S_2 defining a polynomial that has a different curvature compared to the corresponding top edge and that passes through these points (i.e., solid curve ^{q}A ^{q}F vs. dashed curve ^{q}A ^{q}F in Fig. 2c inset). The equations of these polynomials are explained in Appendix A. Similarly, as an example, the curvature between points ^{q}F and ^{q}P at the fault's lower edge is also changed using two more parameters S_3 and S_4 (Fig. 2c).

(v) The curvature difference between the lower and top edges of the fault is then linearly propagated in the up-dip direction for the discretized points (Fig. 2d).

This procedure results in a finite non-planar fault surface with curved horizontal (along-strike direction) and vertical (down-dip direction) tiplines. The number of fault geometrical parameters used may however depend on the desired level in curvature complexity of the fault. Fig. 3 shows how the two parameters in the down-dip direction (D_1 and D_2) and two parameters along strike direction (S_1 and S_2) can vary the curvature of the fault in those two directions. Parameter D_1 changes the slope of the fault in the down-dip direction and parameter D_2 effectively changes the curvature, as expected from Eq. 1. Thus, as shown in Fig. 3a-b, these parameters can be varied to generate a fault geometry with a constant dip angle or with varying dip angles. The parametrization is not limited to generating fault geometries with increasing dip angle with depth, but it can also be changed to include listric or arbitrarily dipping faults. Fig. 3c-d shows the curvature of the fault in the along-strike direction changing with varying parameters S_1 and S_2. The curvature is S-shaped when the parameter S_1 is between 0 and 2, while it is D-shaped curved either outward or curved inward when the value of this parameter is greater than 2 or less than 0, respectively. The parameter S_2 either lessens or amplifies the along-strike curvature, with the curvature increasing with higher absolute value of S_2. These two parameters S_1 and S_2 defined between control points at certain constant depth can be used to map a fault with varying dip angles along strike. Multiple pairs of such control points at different depths (including top edge) can also be used resulting in a complex fault with different curvature at different depths.

After the complex non-planar fault is constructed and discretized, we use TDEs to tie the discretized points (see Fig. 2d) to avoid gaps or nonphysical crossings (or overlays). These TDEs however tend to have different sizes (i.e., different length of sides and hence different surface area) with respect to each other when the fault geometrical parameters are varied. Due to the use of polynomial parameters, the fault curvature is usually smooth both in the along-strike and down-dip directions except close to the control points.

3 Bayesian Inference

In this study, we use a stochastic (probabilistic) approach to estimate the model parameters instead of a deterministic approach that yields a single best set of estimated model parameters. Due to the non-uniqueness of the optimization problem and uncertainties in data (and/or uncertainties in parametrization or modeling scheme), the deterministic result may not be robust and is potentially inaccurate. After being introduced in the 1980s in geophysics (Tarantola & Valette, 1982), the stochastic approach has been used in many fault model parameter estimations with various flavors (Dettmer, Benavente, Cummins, & Sambridge, 2014; Duputel, Agram, Simons, Minson, & Beck, 2014; Dutta et al., 2018; Fukuda & Johnson, 2008, 2010; Matsu'ura et al., 2007; Minson et al., 2014; Monelli, Mai, Jónsson, & Giardini, 2009; Sudhaus & Jónsson, 2009; Yagi & Fukahata, 2011). Here, we use Bayesian inference, where the posterior probability distribution of the model parameters for given data can be determined using two sources of information on these parameters, namely, a priori information and a physical relation between the data and the parameters (Tarantola, 2005). The multidimensional posterior probability density function (PDF) defined on the model parameter space, referred to as the posterior density, can be estimated approximately by sampling using various Monte Carlo methods. In addition, estimating the 1D/2D marginal densities of the model parameters from the posterior density can be useful in characterizing the features and uncertainties of each of those model parameters. The posterior density of the model parameters $p(\boldsymbol{\theta}|\boldsymbol{d})$ is given as (Tarantola, 2005):

$$p(\boldsymbol{\theta}|\boldsymbol{d}) \propto p(\boldsymbol{\theta}) \cdot \mathcal{L}(\boldsymbol{\theta}), \tag{2}$$

where $p(\boldsymbol{\theta})$ represents the prior density of the model parameters and $\mathcal{L}(\boldsymbol{\theta})$ is the likelihood function.

3.1 Likelihood Function

The likelihood function acts as a goodness of fit of the model parameters with respect to the data, including information about the uncertainties of the data measurement process as well as uncertainties in the modeling scheme (or its parametrization) or of other parameters (for e.g., Earth structure, etc.) that are not varied in the problem (Minson et al., 2013). The likelihood function can be given as (Duputel et al., 2014; Tarantola, 2005):

$$\mathcal{L}(\boldsymbol{\theta}) = \int_{\mathbb{D}_a} \rho_D(\boldsymbol{d}|\boldsymbol{d}_a)\, \mathcal{C}(\boldsymbol{d}_a|\boldsymbol{\theta})\, d\boldsymbol{d}_a, \tag{3}$$

where $\boldsymbol{d}_a$ is the actual true displacement that we are interested in during the measurement process. However, the measurements are affected by errors and the data $\boldsymbol{d}$ is only a single realization of a stochastic data vector representing uncertain measurements. $\rho_D(\boldsymbol{d}|\boldsymbol{d}_a)$ is then the density of the data $\boldsymbol{d}$ conditioned on $\boldsymbol{d}_a$ and $\mathcal{C}(\boldsymbol{d}_a|\boldsymbol{\theta})$ is the probability density for $\boldsymbol{d}_a$ conditional on $\boldsymbol{\theta}$. In our problem, we assume that the observed data $\boldsymbol{d}$ has Gaussian distributed measurement errors with zero mean. The variability in the data is described by the covariance matrix $\sigma_1^2\Sigma_d$, where the hyperparameter σ_1^2 is a scaling factor. This Gaussian probability density $\rho_D(\boldsymbol{d}|\boldsymbol{d}_a)$ is then given as:

$$\rho_D(\boldsymbol{d}|\boldsymbol{d}_a) = (2\pi\sigma_1^2)^{-N/2}|\Sigma_d|^{-1/2} \exp\left[-\frac{1}{2\sigma_1^2}(\boldsymbol{d}-\boldsymbol{d}_a)^T\Sigma_d^{-1}(\boldsymbol{d}-\boldsymbol{d}_a)\right]. \tag{4}$$

The conditional density $\mathcal{C}(\boldsymbol{d}_a|\boldsymbol{\theta})$ represents the correlation between the model parameters where uncertainties related to Earth structure (layer thickness, elastic parameters), or the problem scheme can be introduced (Fukuda & Johnson, 2008; Yagi & Fukahata, 2008). Generally, when such model uncertainties are neglected, the conditional probability becomes: $\mathcal{C}(\boldsymbol{d}_a|\boldsymbol{\theta}) = \delta(\boldsymbol{d}_a - G(\boldsymbol{m}, \boldsymbol{s}))$, which we use in this study. However, high error correlation resulting from modeling uncertainties are often present and ignoring them can result in underestimated model uncertainties or bias in parameters estimated (Dettmer, Dosso, & Holland, 2007). The correlated errors can be minimized by downsampling the data such that the distance between data points exceeds the correlation length of the original data. However, for many fault parameter estimation problems, the data are already downsampled or limited. Uncertainties in the earth structure (Duputel et al., 2014; Yagi & Fukahata, 2008) or fault geometry (Ragon et al., 2018) have been quantified to estimate the covariance of the model prediction errors Σ_p, which was incorporated in the conditional probability $\mathcal{C}(\boldsymbol{d}_a|\boldsymbol{\theta})$ as following:

$$\mathcal{C}(\boldsymbol{d}_a|\boldsymbol{\theta}) = (2\pi\sigma_2^2)^{-N/2}|\Sigma_p|^{-1/2} \exp\left[-\frac{1}{2\sigma_2^2}\Big(\boldsymbol{d}_a - G(\boldsymbol{m},\boldsymbol{s})\Big)^T \Sigma_p^{-1}\Big(\boldsymbol{d}_a - G(\boldsymbol{m},\boldsymbol{s})\Big)\right], \quad (5)$$

where the covariance matrix Σ_p is scaled by the hyperparameter σ_2^2.

The likelihood function is thus obtained by combining Eqs. (3), (4) and (5) as:

$$\mathcal{L}(\boldsymbol{\theta}) = \eta(\sigma_1^2, \sigma_2^2) \times \exp\left[-\frac{1}{2}\Big(\boldsymbol{d} - G(\boldsymbol{m},\boldsymbol{s})\Big)^T \Sigma_\psi^{-1}\Big(\boldsymbol{d} - G(\boldsymbol{m},\boldsymbol{s})\Big)\right], \quad (6)$$

where Σ_ψ is the full covariance matrix, and $\eta(\sigma_1^2, \sigma_2^2)$ is the normalizing factor defined as (Duputel et al., 2014; Tarantola, 2005):

$$\begin{aligned} \Sigma_\psi &= \sigma_1^2\Sigma_d + \sigma_2^2\Sigma_p, \\ \eta(\sigma_1^2, \sigma_2^2) &= (2\pi\sigma_1^2\sigma_2^2)^{-N/2}\,|\Sigma_d|^{-1/2}\,|\Sigma_p|^{-1/2}\,\left|(\sigma_1^2\Sigma_d)^{-1} + (\sigma_2^2\Sigma_p)^{-1}\right|^{-1/2}. \end{aligned} \quad (7)$$

In our study, ignoring the model prediction errors results in the full covariance matrix defined as: $\Sigma_\psi = \sigma_1^2\Sigma_d$, and the normalizing factor $\eta(\sigma_1^2) = (2\pi\sigma_1^2)^{-N/2}|\Sigma_d|^{-1/2}$.

3.2 Direct or Indirect Priors

The Bayesian approach allows the inclusion of any prior information on the model parameters defined by a prior probability density, which can restrict the model solution space. Direct prior information restricts the model parameters firmly within a permissible range (Matsu'ura et al., 2007). On the other hand, indirect priors regulate the structure of the stochastic model based on some physical consideration implemented in the problem (e.g., regularization, spatial slip smoothness, etc.). In some past fault model parameter estimation studies, various direct a priori constraints on the fault model parameters have been used (Dutta et al., 2018; Hashimoto, Noda, Sagiya, & Matsu'ura, 2009; Jackson, 1979; W. Xu, Dutta, & Jónsson, 2015), e.g., moment tensor solutions or locations of aftershocks to constrain the fault geometry, or the mainshock moment magnitude to constrain the total slip magnitude on the fault, etc. Here, we use slip smoothness as an indirect prior constraint to reduce the roughness of the slip distribution, similar to Fukuda and Johnson (2008, 2010). For this, the slip smoothness prior $p(\boldsymbol{s}|\sigma_3^2)$, where $\boldsymbol{s}$ is a vector containing the slip at M_L TDEs, restricts rough slip changes between adjacent TDEs (Maerten et al., 2005). This prior consists of slip model roughness $({\sigma_3}^{-2}||L\boldsymbol{s}||_2)$, where σ_3^2 is a hyperparameter that scales the smoothness constraint and L is the discrete second-order finite-difference operator (Laplacian). The slip smoothness prior can be formalized as:

$$p(\boldsymbol{s}|{\sigma_3}^2) = (2\pi{\sigma_3}^2)^{-M_L/2}|L^TL|^{1/2} \times \exp\left[-\frac{1}{2{\sigma_3}^2}\Big(L\boldsymbol{s}\Big)^T\Big(L\boldsymbol{s}\Big)\right]. \quad (8)$$

Instead of the general finite-difference formulation, we use the following approximation of the discrete Laplacian operator (∇^2) due to the use of TDEs (Maerten et al 2005):

$$\nabla^2 s_i = \frac{2}{M_i} \sum_{j=1}^{3} \frac{s_j - s_i}{h_{ij}}, \tag{9}$$

where for the i^{th} TDE with adjacent elements j (where, $j = 1, 2, 3$), s_i represents slip value, and h_{ij} represents distance between centroids of the i^{th} and j^{th} elements, and $M_i = \sum_{j=1}^{3} h_{ij}$. We obtain the sparse-matrix smoothing operator L after superposing the above relation for all the TDEs of the fault.

3.3 Posterior Distribution

The posterior density $p(\boldsymbol{\theta}|\boldsymbol{d})$ (Eq. 2) is obtained by combining Eqs. (2), (6), and (8). Considering positivity constraints on the slip parameters or constraining them within certain bounds, the posterior distribution can be given as:

$$p(\boldsymbol{\theta}|\boldsymbol{d}) \propto \begin{cases} p(\boldsymbol{m}) \cdot p(\boldsymbol{s}|{\sigma_3}^2) \cdot p(\sigma_3^2) \cdot p(\sigma_1^2) \cdot \mathcal{L}(\boldsymbol{\theta}), & \text{if } \boldsymbol{s}_\alpha \leq \boldsymbol{s} \leq \boldsymbol{s}_\beta \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where the probabilities $p(\sigma_1^2)$ and $p(\sigma_3^2)$ are given such that $\sigma_1^2 \sim \mathcal{LU}[\alpha_1 \; \beta_1]$ and $\sigma_3^2 \sim \mathcal{LU}[\alpha_3 \; \beta_3]$ are log-uniformly distributed with α_i and β_i corresponding to lower and upper limits of the logarithmic scale, which are chosen subjectively. Probability $p(\boldsymbol{m})$ is given such that $\boldsymbol{m} \sim \mathcal{U}[\boldsymbol{m}_\alpha \; \boldsymbol{m}_\beta]$ is uniformly distributed between the bounds $\boldsymbol{m}_\alpha$ and $\boldsymbol{m}_\beta$ that are subjectively chosen. The terms $\boldsymbol{s}_\alpha$ and $\boldsymbol{s}_\beta$ are the lower and upper bounds of the slip parameters.

3.4 Sampling Technique

Multidimensional posterior probability densities, where the posterior outcome is estimated only point-wise through a numerical/analytic method, are typically sampled using Markov Chain Monte Carlo (MCMC) sampling techniques (Gelman et al., 2013; Gilks, Richardson, & Spiegelhalter, 1995). MCMC sampling of a target probability distribution consists of generating a reversible Markov chain, such that the resulting equilibrium distribution is similar to that of the target distribution. Most of these MCMC methods are not effective when the posterior probability densities are high-dimensional, multi-modal, very peaked, flat, etc.. Here, we use a variant of the CATMIP algorithm (Minson et al., 2013) that is based on the Transitional Markov Chain Monte Carlo algorithm of Ching and Chen (2007). This method belongs to the class of sequential particle filter methods (Chopin, 2002), which combines transitioning (tempering of simulated annealing) and resampling (replication and mutation of genetic algorithm) with MCMC sampling.

This method implements the idea proposed by Beck and Au (2002) to construct a series of intermediate probability densities that transitions from the prior probability density $p(\boldsymbol{\theta})$ to the target probability density $p(\boldsymbol{\theta}|\boldsymbol{d})$ by increasingly sampling the intermediate probability densities in the following way:

$$\begin{gathered} p_j(\boldsymbol{\theta}) \propto p(\boldsymbol{\theta}) \cdot p(\boldsymbol{d}|\boldsymbol{\theta})^{\gamma_j}, \\ j = 0, ..., J \quad and \quad 0 = \gamma_0 < \gamma_1 < \cdots < \gamma_J = 1 \end{gathered} \tag{11}$$

where j is the transition stage number and $p(\boldsymbol{d}|\boldsymbol{\theta})$ is the data likelihood. In cases when the geometry of the target probability density $p(\boldsymbol{\theta}|\boldsymbol{d})$ is dramatic (e.g., high-dimensional, strongly correlated, highly peaked, flat or multi-modal), it is not easy to sample it. Hence, small changes in geometry of consecutive intermediate probability densities, i.e. from $p_j(\boldsymbol{\theta})$ to $p_{j+1}(\boldsymbol{\theta})$, leads to efficiently obtaining the samples. These adaptive intermediate probability densities are chosen depending on the criteria that if the target probability density drastically varies from the prior probability density, there are more transition stages as compared to when the variation is low. The transitional stages (for example transition from stage j to stage $j+1$) are controlled by the coefficient γ_{j+1} corresponding to the next stage, which is chosen adaptively such that the coefficient of variation of $p(\boldsymbol{d}|\boldsymbol{\theta}_j)^{\gamma_{j+1}-\gamma_j}$ is equal to a chosen threshold (Beck & Zuev 2013). At the next stage, the samples from the intermediate probability density at the current stage are resampled according to the weights determined by the ratio of the corresponding data likelihoods at the next and current stages ($p(\boldsymbol{d}|\boldsymbol{\theta}_j)^{\gamma_{j+1}-\gamma_j}$). This causes unlikely models to be rejected in favor of more likely models making this technique more robust against the dimension of the target probability density. To make the samples from each stage distinct from each other, we employ MCMC sampling (adaptive Metropolis Hastings algorithm) technique (Haario, Saksman, & Tamminen, 2001) within each stage that samples the corresponding intermediate probability density. Appendix B summarizes the different steps in this sampling technique.

The resulting ensemble of model samples representing the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{d})$ can then be inferred by estimating the maximum a posteriori (MAP) model $\hat{\boldsymbol{\theta}}$, mean model $\overline{\boldsymbol{\theta}}$, or 1D/2D marginal probability densities. Theoretically, maximum a posteriori model $\hat{\boldsymbol{\theta}}$ estimate is the mode of the posterior distribution and mean model $\overline{\boldsymbol{\theta}}$ estimate is its mean, which can be obtained by:

$$\begin{aligned}\hat{\boldsymbol{\theta}} &= \arg\max_{\boldsymbol{\theta}} \;\; p(\boldsymbol{\theta}|\boldsymbol{d}) \\ \overline{\boldsymbol{\theta}} &= \int_{\mathbb{M}} \boldsymbol{\theta}' \, p(\boldsymbol{\theta}'|\boldsymbol{d}) \, d\boldsymbol{\theta}'\end{aligned} \tag{12}$$

The 1D marginal posterior probability density of a specific parameter can be obtained by integrating the posterior probability density over the entire model parameter space, except for the model parameter of interest:

$$M_i(\theta_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\boldsymbol{\theta}|\boldsymbol{d}) \prod_{j=1,\, j\neq i}^{N} d\theta_j . \tag{13}$$

4 Synthetic Tests

In this section, we test the method of simultaneously estimating non-planar fault geometry and spatially-variable slip as discussed in Section 2 for three synthetic fault slip models. First, synthetic data generated from these fault models were used to estimate slip distributions on pre-assumed planar and non-planar faults. Then, Bayesian analyses were used to estimate non-planar geometries simultaneously with spatially-variable slip and the results compared with those when planar geometries are assumed.

4.1 Normal faulting example

For the first synthetic test, we consider a fault-slip model consisting of a checkerboard-like slip pattern on a listric normal fault (Fig. 4a). The synthetic data for this normal fault were generated by firstly constructing the geometry such that the fault has the following features: (i) the top edge of the fault is at 1 km below the surface, (ii) the bottom edge of the fault is at 9 km, (iii) the down-dip curvature of the fault resembles that of a listric fault, and (iv) the fault's lower edge is curved in the along-strike direction (Fig. 4a). The fault dips steeply at Plane B compared to at Plane A (Fig. 4b). After determining its geometry, the fault was discretized using triangular dislocation elements and a normal-component slip imposed in a checkerboard pattern with maximum slip of 4 m and a minimum of 0 m (Fig. 4a). The slip was slightly spatially smoothed to avoid sharp changes from 0 to 4 m. The moment magnitude of the resulting listric normal faulting event is 7. The three components of the ground displacements as GNSS stations would yield due to the normal-slip on this fault were then calculated on a rectangular ground surface grid using the analytical TDE solution by Meade (2007). Finally, to make the synthetic data figures/Figure4.jpg

Figure 4: The non-planar fault model used in the first test of simultaneously estimating non-planar geometry and spatially-variable slip. (a) The listric normal fault geometry in 3D with a checkerboard slip distribution. The green line is the projection of the top edge of the fault on the surface. (b) Fault profiles at the vertical planes A and B shown in (a). (c) Quadtree subsampled synthetic surface displacements resulting from the normal faulting in (a), with arrows showing the horizontal displacements and colored squares the vertical displacements. The green and black lines are the surface projection of the top and other edges of the fault, respectively.

more realistic, Gaussian noise with standard deviation proportional to the displacement magnitude was added to the ground displacements. We used Quadtree sub-sampling of this dense surface displacements to reduce the number of data points (Fig. 4c), yielding about 100 observation locations.

4.2 Bias in Slip Estimates using Planar Faults

We first used these synthetic ground displacement observations to estimate spatially-variable slip for several different fault geometries, which we fixed before the estimation, using linear regularized non-negative least-squares (RNNLSQ) optimization (Altman & Gondzio, 1999). Figure 5 shows the resulting spatially-variable slip estimates using the following four different fault geometries: (a) the reference listric fault geometry curved in both down-dip and along-strike directions that was used to generate the synthetic data, (b) the reference fault geometry without any geometrical variations in the along-strike direction, (c) the planar fault geometry that is the best 3D fit to the reference fault geometry, and (d) the planar fault geometry that is the best 3D fit to the deeper parts (below 5 km) of the reference listric fault geometry. We use a weighted measure of variance reduction (Suppl. 1) to compare how well the fault-slip model explains the synthetic surface displacement observations.

Not surprisingly, the fault slip model with the reference fault geometry (Fig. 5a) can explain the synthetic dataset the best with about 99.6% variance reduction. However, planar faults are the most common assumption for fault slip estimation studies worldwide, and here the two planar fault slip models (Fig. 5c,d) have lower variance reductions of 95 % and 91.1%. Although these planar fault models appear to resolve both the shallow and deeper slip asperities with more than 90% variance reduction, the estimated slip distributions are quite different from the reference one. For example, slip on the shallow and deeper slip patches of the planar faults is over-estimated by 100% and 50%, respectively. In addition, there is also a bias in the depths at which the slip asperities are estimated. The fault slip model with the simplified reference non-planar fault geometry with no along-strike variations also exhibits all the four slip asperities, but the maxima slip values are over-estimated where the dip angle is different from the reference fault (Fig. 5b).

4.3 Simultaneous Bayesian Estimation

We now apply our proposed method of simultaneously estimating non-planar finite-fault geometry and spatially-variable slip. For this, we used the Bayesian inference described in Section 3 to estimate the variability of the fault geometrical and slip parameters. The geometrical parametrization we used is with 4 parameters (as explained in Section 2) to vary the non-planar finite-fault geometry with the top edge fixed. We as-

revision1/rev_figs/Figure10.jpg

Figure 5: Results of linear least-squares slip inversions (RNNLSQ) with different predefined fault geometries. (a) The reference fault geometry with variable dip angle, both down-dip and along-strike. (b) The reference fault again, but with varying dip angle only in down-dip direction. (c) A planar fault that best fits the reference fault geometry. (d) A planar fault that fits the shallower parts (below 5 km depth) of the reference fault geometry. (e-f) Profiles of the different fault geometries along cross-sections A and B are shown. The insets show the change in dip angle of the faults with depth along the corresponding planes.

figures/Figure5.jpg

Figure 6: Marginal prior probability densities of the model parameters. (a) Uniform prior probability densities of the geometrical parameters and hyperparameters with dashed magenta lines indicating the reference values. (b) Marginal prior probability densities of the slip parameters overlaid on the correlation of the corresponding slip patch and the slip patch outlined by red line. The inset shows an example of the bimodal prior slip probability density of the slip patch outlined by magenta line. (c) Fault geometry and slip distribution corresponding to the prior median fault geometrical and slip parameters.

sume the top edge of this fault is well constrained based on the discontinuity of the surface displacement due to the fault reaching close to the surface. In this example, we also fix the fault-bottom depth at 9 km below the surface. The two geometrical parameters S_1 and S_2 vary the curvature of the fault in the along-strike direction and parameters D_1 and D_2 vary the curvature of the fault in the down-dip direction. This non-planar geometry was discretized with 96 TDEs, i.e., larger patches than the one used to generate the synthetic surface displacement observations, and we only estimated the normal-component slip assuming there is no strike-slip. Thus the geometrical parameters (D_1, D_2, S_1 and S_2), the 96 slip parameters (for 96 TDEs) and the hyperparameters (σ_1^2 and σ_3^2) resulted in a total of 102 parameters to estimate.

As mentioned in section 3.2, Bayesian estimation allows the use of any a priori information about the fault geometrical or slip parameters along with the physical relation of the data with the model parameters. Despite this flexibility, we chose a uniform distribution as the prior probability for the geometrical parameters and log-uniform prior distribution for the hyperparameters (Fig. 6a). However, the prior probabilities of the slip parameters have to be set such that there are enough probable samples to start the sampling and we also want the slip values that are strictly positive. For this, Minson et al. (2013) used a Dirichlet distribution as the prior probability for each slip patch such that the total moment of the slip parameter ensemble followed a Gaussian distribution about a plausible event moment magnitude. Instead of using the moment magnitude based Dirichlet prior probability for the slip patches, we used the synthetic observations to determine this prior probability density. For this, the ensemble of non-planar fault geometries following the uniform (prior) probability as mentioned above was used to obtain an ensemble of slip distributions, such that each slip distribution corresponding to a sample non-planar fault geometry was estimated by using the observed synthetic data (in Fig. 4c) and linear RNNLSQ optimization (Altman & Gondzio, 1999). The ensemble of such slip distributions was then used as the prior probability of each of the slip figures/Figure6.jpg

Figure 7: A few samples of fault geometries and corresponding slip distributions of the prior distribution ensemble.

revision1/rev_figs/Figure7.jpg

Figure 8: Progression of the fault geometry and slip distribution during the estimation process showing the posterior median sample of the fault parameters at several transition stages of the SMC sampling.

revision1/rev_figs/geopar_evol.jpg

Figure 9: 1D marginal densities of the geometrical parameters and hyperparameters obtained at several different intermediate stages of the sampling process.

patches. Co-incidentally, the prior probabilities of the slip patches follow a bi-modal distribution where each mode represents a Dirichlet distribution (Fig. 6b). The resulting slip prior ensemble correlations of each slip patch with the rest of the slip patches are more than 0.7 (Fig. 6b). Although highly correlated, such prior probability for the slip patches ensures that the sampling technique initiates with probable samples and a total moment close to the final moment estimate. Otherwise, the data likelihoods cause the starting samples to be rejected in the first resampling stage resulting in failure of the sampling. Fig. (6c) shows that sample of the slip distribution and the fault geometry, which corresponds to prior median values of the fault geometrical and slip parameters. Fig. 7 shows some starting sample fault geometries and slip distributions belonging to the prior probability of the fault model parameters. These sample fault geometries show the extent to which the fault can be warped or twisted in both directions, along-strike and down-dip.

The multi-dimensional posterior probability density of the fault geometrical and slip parameters, which contains information of the indirect/direct priors and the data likelihood, can be sampled to obtain the most probable fault parameter values as described in section 3.4. Here we sampled the posterior probability density using the SMC technique with 20,000 Markov chains and a chain length of 150, a sampling procedure that took 28 stages and was robust such that the obtained samples converged to the posterior probability density (Fig. S2). During the intermediate stages, the prior probability density transitions itself to the posterior probability density with increasing contribution of the data likelihood (Figs. 8, 9, 10, S4, S5, S6, S7). Fig. 8 shows the median fault geometry and slip distribution at several of these transitional stages of the sampling. While the SMC sampling progressively explores through more probable estimates of the fault model parameters, during the earlier stages of the sampling it converges to probable estimates of hyperparameters (Fig. 9). At the intermediate stages, the sampling converges to constraints of the fault geometrical parameters (Fig. 9), and finally to constraints of the spatially-variable slip (Fig. 10).

revision1/rev_figs/slip_evol.jpg

Figure 10: 1D marginal densities of the four selected slip asperities (slip patch indices 3, 33, 51 and 81) at several different intermediate stages of the sampling process.

revision1/rev_figs/Figure8.pdf

Figure 11: Modeling results for the fault geometry. (a) 1D/2D marginal posterior probability densities of the fault geometrical parameters and the hyperparameters. (b) The posterior median sample of the fault geometry and slip distribution. (c) Comparison of the reference fault geometry with the posterior ensemble at Plane A, and (e) at Plane B. (d) Comparison of the fault-dip angle of the reference fault geometry with the posterior median sample at plane A, and (f) at Plane B.

The resulting 1D marginal probability density of the geometrical parameters are mostly skewed or multi-modal describing the highly non-linear relation of the data with these parameters (Fig. 11a). The 1D marginal probability densities of the hyperparameters on the other hand exhibit log-normal distributions. The two parameters controlling the curvature in the down-dip direction (D_1 and D_2) are positively correlated to each other. This correlation demonstrates that the steeper faults are more curved in the down-dip direction, and vice versa. The negative value of S_1 shows the listric fault is curved as D-shape (inward) in the along-strike direction. This parameter S_1 is negatively correlated to the parameter S_2 that determines the intensity of the curvature in the along-strike direction. However, the correlation coefficient for S_1 and S_2 is lower in magnitude for the more probable samples than that for the less probable samples showing a non-linear correlation of these two fault geometrical parameters (Fig. 11a).

The estimated fault geometry is listric like the reference fault (Fig. 11). The estimated fault-dip angle matches the reference dip angle better for deeper portions of the fault. While the estimated geometrical parameters do not agree with the reference values, the reference geometry lies within their 95% confidence intervals. The estimated fault geometry agrees within the 95% confidence interval with the reference fault geometry for depths greater than 6 km, while at shallower depths, it is steeper than the reference geometry. This can be explained as being due to the coarse fault discretization and the subsampling of the data used in this estimation (Supp. 3, Figs. S9 and S10). The uncertainty of the estimated fault geometry increases with depth, as expected (Fig. 11c,e). But, the higher uncertainty of the fault geometry at $\sim$3.5 km depth, compared to that at $\sim$7 km depth, can be attributed to the inaccurate location of the fault at shallower depths.

The estimated posterior median fault slip model has slip asperities that agree well with slip maxima of the reference model, recovering the checkerboard pattern successfully (Fig. 12a). The maxima slip values are approximately 4 m, with a few patches with lower or higher slip values likely due to the coarse discretization of the fault. Fig. 12a shows the median and the 1D marginal probability densities of the slip values at each of the TDEs. The 1D marginal probability densities of the slip values are mostly Gaussian or truncated Gaussian distributions (due to positivity or maximum cut-off constraints on slip values). The standard deviation is high ($>$1 m) for slip asperity patches at larger

revision1/rev_figs/Figure9.jpg

Figure 12: Modeling results for fault slip. (a) 1D marginal posterior probability densities of the fault slip overlaid on the corresponding posterior median slip of each fault patch. (b) An example probability density in more detail for one selected fault patch (from top-left). (c) The posterior standard deviation of the fault slip values. (d) The synthetic data, modeled predicted data and the data residuals corresponding to the posterior median of fault parameters. (e) Distribution of variance reduction obtained from the ensemble of fault model parameters.

revision1/rev_figs/Figure13.jpg

Figure 13: Modeling results for a thrust faulting case. (a) Reference slip model with checkerboard-like thrust slip pattern on a non-planar fault. (b) Posterior median sample of the estimated result. (c) The estimated non-planar fault geometry and its uncertainties compared to the reference geometry at plane A and (d) at plane B.

revision1/rev_figs/Figure14.jpg

Figure 14: Modeling results for a strike-slip faulting case. (a) Reference slip model with checkerboard-like strike-slip pattern on a non-planar fault. (b) Posterior median sample of the estimated result. (c) The estimated non-planar fault geometry and its uncertainties compared to the reference geometry at planes A and B.

depths, whereas at shallower depths it is about 0.5-0.9 m (Fig. 12c). For the patches with low slip values, the standard deviation at larger depth is about 0.5 m and at shallower depths it is about 0.2-0.3 m. The ensemble of the fault slip models has a variance reduction ranging from 98.6 % to 99.3 % (Fig. 12e), with a median value of 99.03 %. The spatially-variable slip is estimated robustly and is similar to the reference slip without any significant biases related to inaccurate fault geometry.

4.4 Thrust and strike-slip faulting examples

We tested the simultaneous Bayesian estimation of non-planar fault geometry and spatially-variable slip on two other cases that included thrust and strike-slip faulting. For each of these tests, fault slip models consisting of checkerboard-like slip patterns were considered (Figs. 13a, 14a, S11, S14). For the thrust-faulting reference model, the fault-dip angle varies in both the along-strike and down-dip directions such that the fault is shallower dipping at plane B compared to that at plane A (Fig. 13c,d). From these models, we calculated 3D surface displacements with added Gaussian noise, which we then used in the Bayesian estimation of the geometrical and slip parameters.

The modeling results show that in the thrust-faulting case the non-planar fault geometry is mostly recovered, with a down-dip curvature that agrees well with the reference model geometry for depths shallower than 20 km, but deviates slightly below that, where the fault-dip angle is steeper than in the reference model. Not surprisingly, the spatially-variable slip is better constrained at shallower depths on the fault than at greater depths, where it does not agree well with the reference model (Fig. 13b).

For the strike-slip faulting test case, the reference fault has variable fault-dip angle ranging from 85° to 70° in the along-strike direction, but constant fault-dip angle in the down-dip direction (Fig. 14a). The modeling results when compared with the reference fault model show that the estimated geometry at both the planes A and B are similar (Fig. 14c). The estimated fault-slip pattern resembles the reference fault slip patches, although the slip maxima patches of the posterior median sample shows higher slip by about 25%.

5 Discussion

Many fault-slip models are usually published after each well-recorded major earthquake, but the resulting slip models are often quite different from each other. The reason is related to a combination of factors, such as what datasets were used, how the earth structure and the fault geometry were parameterized, how the model parameter estimation was set up, and what optimization scheme was used (Razafindrakoto et al., 2015). Even in Bayesian estimations of fault-slip models that are used to address possible fault-model discrepancies, some aspects of the model are usually pre-assumed, e.g., the Earth structure (layer thickness, elastic parameters, etc.), the fault geometry, etc. Efforts of integrating uncertainties in the Earth structure (Duputel et al., 2014) or uncertainties in fault geometry (Ragon et al., 2018) as model prediction covariances in confluence with the data uncertainties have shown to reduce bias in fault slip estimations. In our work, we extend previous studies by parameterizing a non-planar fault geometry such that it can be estimated simultaneously with the slip distribution. With this geometry parameterization one can estimate spatially-variable fault-strike and -dip angles, both in the along-strike and down-dip directions, such that the fault surface is allowed to twist and warp to explain the observed data. This flexibility eliminates the need for *a priori* assumptions about the fault geometry and the Bayesian inference provides all the associated uncertainties of the estimated fault parameters.

Estimations of variable non-planar fault geometries have usually been avoided in earthquake fault estimation studies as they require calculating the Green's function for every perturbation in the fault geometry, making such flexibility computationally impractical in kinematic source model estimations. Instead, researchers have usually estimated or assumed simple planar fault geometries before determining spatio-temporal details of the fault slip. However, when the Green's function calculation is fast and robust (such as in static finite-fault estimations), it is possible to allow for local variations in geometry as well as estimating the slip. While this has been done in a few studies, uncertainties in fault position and dip angle in both down-dip and along-strike directions have rarely been considered. This is important to do as the simultaneous estimation of local variations in both fault geometry and slip can be far from robust. While simple pre-assumed planar or non-planar fault geometries can be good first-order approximations of the real fault geometries, the increasing availability, resolution, and quality of geodetic data (e.g., InSAR and image offsets) allow resolving of fault parameters such that the estimation of local fault geometry complexities are warranted.

As demonstrated by our results of the listric normal fault, estimated fault slip assuming planar faults may explain observed displacement data well (more than 90% variance reduction in our example), even when the true source fault is non-planar. However, the estimated spatial variations in slip can be under- or over-estimated of the different parts of the fault due to the planar approximation of the local fault location and dip angle. Using geodesy, fault slip variations are typically better resolved at the shallower fault depths than deeper (e.g., Simons, Fialko, & Rivera, 2002), but even shallow slip values can be strongly biased if the local fault location or dip angle is inaccurate (Ragon et al., 2018). Using the fault geometry parameterization presented here, where variable fault-dip angle and location can be estimated along with the slip, thus helps to eliminate this potential bias in slip estimations. Also, the non-planar fault geometry can be estimated solely based on the observed data or it can be constrained using prior physical information in the Bayesian framework. While it is known that the estimated geometry and slip are generally less well constrained at larger depths, Bayesian inference also helps to quantify these uncertainties and provides confidence levels for all the estimated fault parameters. The different synthetic test cases in our study demonstrate that our approach works well in estimating the overall non-planar fault geometry simultaneously with the spatially-variable slip to 20 to 30 km depths for earthquakes with different slip mechanisms and moment magnitude ranging between 7 and 8. However, the quality of the estimation result can vary depending on the style of faulting and the depth of the source fault. In addition, lack of data coverage, e.g., in case of an event occurring offshore, extensive water bodies, or InSAR decorrelation, would influence the estimation results.

Estimating complex non-planar fault geometries has rarely been attempted before due to the high computation cost of each forward calculation of the model (i.e., the computation of $G(\boldsymbol{m}, \boldsymbol{s})$ in our study), making the MCMC sampling impractical. However, in our study for the listric normal faulting test scenario, the average computation time of $G(\boldsymbol{m}, \boldsymbol{s})$ was only about 0.3 seconds. This means that for the 28 stages of the SMC sampling and 20000 Markov chains of length 150, the total computation time was about 7000 CPU core hours. This computation time scales up with more data observations, finer fault discretizations, and more sampling stages. When studying large earthquakes (M_W7.5 and larger), more data points and a larger number of TDEs would typically be needed than we used in this study. Also, more sampling stages would be needed for convergence due to the increased complexity in the fault geometry and slip. For example, it took 49 stages with 10000 Markov chains of length 150 for the finer discretizations of 384 TDEs (i.e., 4 times of the normal faulting case) to be resolved from the same dataset in Supp. 3. The resulting computational time was about 17150 CPU core hours (i.e., 2.5 times of the normal faulting case). The high computation cost can be kept within acceptable bounds by parallelization of CPU clusters (Minson et al., 2013) or GPU architectures (Lee, Yau, Giles, Doucet, & Holmes, 2010). In addition, as the fault geometry parameterization has only a few parameters, the Green's functions ($G'(\boldsymbol{m})$ in Section 2) can be pre-computed and stored (Heimann et al., 2019; Vasyura-Bathke et al., 2020). The pre-computed Green's function databases could then be used during the Bayesian inference, leading to drastically reduced computation costs.

In cases when there is scarce surface displacement data, using non-planar geometries may seem like an over-parameterization of the problem. Such over-parameterizations can lead to unrealistic models that fit noise features in the data or models with unrealistically high uncertainties, whereas using overly simple models with fewer parameters (under-parametrization) can result in biased solutions as mentioned above. However, the model parametrization scheme for the fault model estimation problem (e.g., the number of control points pairs to constrain the fault-top/bottom edge curvature, or the coarseness of the fault-slip discretization, etc.) can be either selected *ad hoc* or based on some model selection metric. The different parameterization schemes can be statistically compared by calculating the Bayesian evidence (i.e., the denominator in the Bayes' theorem), which quantitatively embodies Occams' razor (Knuth, Habeck, Malakar, Mubeen, & Placek, 2015; Madigan & Raftery, 1994; Von der Linden, Dose, & Von Toussaint, 2014). The evidence, which is an integral over the entire parameter space of the product of the prior and likelihood, can be estimated directly from posterior sampling (Berkhof, Van Mechelen, & Gelman, 2003; Chib & Jeliazkov, 2001; Chib & Ramamurthy, 2010) or other numerical techniques (Knuth et al., 2015). Apart from the measure of Bayesian evidence, different measures, e.g., Akaike and Bayesian information criteria (Burnham & Anderson, 2004), Akaike's Bayesian information criterion (Funning, Fukahata, Yagi, & Parsons, 2014; Yabuki & Matsu'Ura, 1992), deviance information criterion (Kowsari, Halldorsson, Hrafnkelsson, & Jonsson, 2019), etc., provide a faster alternative to balance the model parametrization accuracy against complexity. These model selection techniques can thus help determining the appropriate complexity of the model given the prior information and the quantity, coverage, and quality of the available data.

The complex non-planar fault in our synthetic tests is modeled within an isotropic and homogeneous elastic half-space. More realistic earth models that consider depth-dependent elastic parameters typically show deeper slip centroid estimates and more slip at depth compared to solutions that use homogenous medium (Hearn & Bürgmann, 2005) and there can be considerable differences in how well the models fit observed data (Wang & Fialko, 2018). The effect of uncertain depth-dependent elastic parameters can be included in slip estimations through model prediction error covariances (Duputel et al., 2014), which requires determination of sensitivity kernels of how model predictions change with elastic parameter modifications. While not used in our synthetic tests here, similar sensitivity kernels of the model prediction could be included in determining Σ_p (Eq. 5) during the simultaneous Bayesian estimation of fault geometry and spatially-variable slip.

A major question in the earthquake source estimation community has been what causes the shallow slip deficit seen in many fault-slip estimation solutions for major earthquakes (Fialko, Sandwell, Simons, & Rosen, 2005). The shallow slip deficit is the reduction (usually more than 10%) in inferred fault slip at shallower depths compared to slip at intermediate depths in the crust. Various explanations for this apparent shallow slip deficit have been proposed, such that low initial tectonic stress in the low-rigidity shallow crust (Rybicki, 1992; Rybicki & Yamashita, 1998), bulk inelastic yielding of the near-fault host rocks in the shallow crust (Fialko et al., 2005), shallow velocity-strengthening fault friction leading to shallow post-seismic afterslip and interseismic creep (Kaneko & Fialko, 2011), and slip estimation bias due to lack of near-field data (X. Xu et al., 2016). Here we show that the depth of inferred slip can be biased when planar fault geometries are used in slip estimations for source faults that are in reality non-planar. In the case related to the normal listric fault presented in our study, the estimated slip asperities are deeper and there is less estimated shallow fault slip than in the reference non-planar fault slip model, resulting in a clear shallow slip deficit (Fig. 5). While this type of fault-slip biases does not resolve the slip-deficit question, it offers yet another possibility for the apparent shallow slip deficit and might help explaining some cases of shallow slip deficit.

The non-planar fault parameterization introduced here can be useful when studying earthquakes occurring in subduction-zones, on listric normal faults, on faults with varying dip angle along strike, and in other cases of non-planar faulting. In addition, this parameterization can be extended to multiple fault branches, with the geometry of each fault branch described by its own set of polynomial parameters. Estimating complex non-planar geometries does not only eliminate fault slip biases in many cases, but it can also have consequences for studies that are based on biased results. For example, Barrientos and Ward (1990) used a planar fault to estimate the fault slip for the 1960 ($M_W 9.5$) Chile megathrust earthquake and reported that the fault slip occurred on isolated fault slip patches at 80-110 km depth. They then suggested that the isolated fault areas left unruptured had experienced postseismic aseismic slip. However, Moreno, Bolte, Klotz, and Melnick (2009) showed that no such isolated slip patches result when using a more realistic non-planar fault geometry, demonstrating that the isolated slip patches were merely an artifact of using a planar fault. Using non-planar faults instead of planar faults can also lead to more realistic near-fault ground motion calculations (Passone & Mai, 2017), help the understanding of the physics of fault ruptures (Aochi, Fukuyama, & Matsu'ura, 2000), and hence improve seismic hazard assessments (Aochi & Fukuyama, 2002).

6 Conclusions

We have introduced a method to parametrize non-planar earthquake fault geometries using a few polynomial parameters, which can be estimated simultaneously with spatially-variable fault slip from geodetic data using Bayesian inference. The non-planar fault surfaces are discretized with triangular dislocation elements and the surface is re-meshed each time the geometrical fault parameters are updated in the estimation process. The Bayesian inference allows the incorporation of prior information about the fault surface, such as from mapping of surface fault ruptures or from aftershock locations, or about the smoothness of the fault slip. It also provides the full posterior probability distribution of the estimated geometrical and fault slip parameters, yielding information on how well these parameters are constrained by the data and how they are correlated to one another.

We demonstrate the applicability of the method by using three synthetic tests of normal, thrust, and strike-slip fault models, all with variable fault-dip and -strike angles and with a checkerboard-like fault slip distribution. While the resulting ensemble of estimated geometrical parameters exhibits multi-modal and skewed distributions with strong correlation between parameters, the complex non-planar fault geometry and the main slip asperities are mostly well resolved. Our results also show that when planar fault geometries are assumed in presence of non-planar faulting, significant fault slip estimation biases can result with strong over- or under-estimation of fault slip asperities as well as incorrect determination of the locations of these asperities.

A Model Parametrization

In Section 2, parameters S_1 and S_2 control the curvature between points ${}^q\mathrm{A}$ and ${}^q\mathrm{F}$, while parameters S_3 and S_4 control it between ${}^q\mathrm{F}$ and ${}^q\mathrm{P}$ at the bottom of the fault (Fig. 2c). In the example in Section 2, these along-strike parameters define two polynomials, between ${}^qA\,{}^qF$ and ${}^qF\,{}^qP$ (Fig. 2). For S_1 and S_2 and the corresponding control points ${}^q\mathrm{A}$ and ${}^q\mathrm{F}$, we first consider the Z-plane in which these control points lie, i.e., the $z = {}^qz$ plane. The x- and y-coordinates (2D Cartesian coordinates) of the control points in this Z-plane are $({}^qx_1, {}^qy_1)$ and $({}^qx_i, {}^qy_i)$, respectively. The 2D Cartesian coordinate system at this Z-plane is then transformed using an isotropic scaling factor K and rotation angle ω:

$$K = \frac{2}{\sqrt{({}^qx_i - {}^qx_1)^2 + ({}^qy_i - {}^qy_1)^2}} \quad and \quad \omega = \tan^{-1}\left(\frac{{}^qy_i - {}^qy_1}{{}^qx_i - {}^qx_1}\right). \tag{A.1}$$

The transformation matrix for the x-y coordinates can then be given as:

$$\mathcal{A} = K \cdot \begin{bmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{bmatrix}. \tag{A.2}$$

In the new 2D Cartesian coordinate system, the coordinates of points ${}^q\mathrm{A}$ and ${}^q\mathrm{F}$ are (0,0) and (2,0), respectively. For different pairs of control points at different depths, the scaling factor K and rotation angle ω change according to the coordinates of these control points. However, the transformed coordinates of the control points are still (0,0) and (2,0). The curve ${}^qA\,{}^qF$ is then obtained from the polynomial: $y' = S_2 \cdot x'(x' - 2)(x' - S_1)$, where x' and y' are the coordinates of the polynomial in the transformed coordinate system. The coordinates of the new polynomial is then transformed back to the original coordinate system using the transformation matrix $\mathcal{A}^{-1}$. In this original coordinate system, the new polynomial is tied to the rest of the discretized fault. For this, it is discretized at its intersection with the vertical projection of down-dip polynomials passing through the discretized top edge (curve Q ^{q}Q in Fig. 2c). Then the difference in the distance between the original curve and the modified curve (i.e., distance between points ^{q}Q and ^{q*}Q in Fig. 2c inset) is then decreased linearly at nearby depths to generate a gradual change in the along-strike curvature with depth. The same procedure is followed for different sets of control points (that might be located also on the top horizontal tipline) and polynomials, which can be generated depending on the corresponding along-strike parameters.

B SMC Sampling

The posterior probability density $p(\boldsymbol{\theta}|\boldsymbol{d})$ in our study (Eq. 9) is sampled using Sequential Monte Carlo sampling (Sec. 3.4). This sampling technique can be summarized in the following steps:

(i) Set $j = 0$ and coefficient $\gamma_0 = 0$. Generate K samples of geometrical parameters $\boldsymbol{m}_j = \{\boldsymbol{m}_{j,1}, \cdots, \boldsymbol{m}_{j,K}\}$ and hyperparameters $(\sigma_i^2)_j = \{(\sigma_i^2)_{j,1}, \cdots, (\sigma_i^2)_{j,K}\}$ from a uniform prior probability density $p_0^{'}(\{\boldsymbol{m}, \sigma_i^2\})$ and estimate K sets of spatially-variable slip solutions $\boldsymbol{s}_j = \{\boldsymbol{s}_{j,1}, \cdots, \boldsymbol{s}_{j,K}\}$ on the corresponding K samples of non-planar fault geometries using the synthetic data and linear regularized non-negative least-squares optimization (RNNLSQ) at stage $j = 0$. Set the ensemble of samples $\boldsymbol{\theta}_j = \{\boldsymbol{\theta}_{j,1}, \cdots, \boldsymbol{\theta}_{j,K}\}$ for stage j, such that the kth element $\boldsymbol{\theta}_{j,k} = \{\boldsymbol{m}_{j,k}, \boldsymbol{s}_{j,k}, (\sigma_i^2)_{j,k}\}$.

(ii) Set $j = j + 1$ and choose γ_{j+1} such that the coefficient of variation of $\boldsymbol{w}^T$ is equal to a threshold value, where $\boldsymbol{w}^T = \{w_1, \cdots, w_K\}$ is a weight vector given as:

$$w_k\left(\boldsymbol{\theta}_{j,k}\right) = \frac{p_{j+1}(\boldsymbol{\theta})}{p_j(\boldsymbol{\theta})} = \frac{p(\boldsymbol{\theta}_{j,k})\ p(\boldsymbol{d}|\boldsymbol{\theta}_{j,k})^{\gamma_{j+1}}}{p(\boldsymbol{\theta}_{j,k})\ p(\boldsymbol{d}|\boldsymbol{\theta}_{j,k})^{\gamma_j}} = p(\boldsymbol{d}|\boldsymbol{\theta}_{j,k})^{\gamma_{j+1}-\gamma_j} \tag{B.1}$$

(iii) Resample the samples obtained at the previous stage j, i.e., $(\boldsymbol{\theta}_j)$ using the probabilities p_j to obtain resampled samples $\boldsymbol{\Theta}_j$, where

$$p_{j,k} = \frac{w_k(\boldsymbol{\theta}_{j,k})}{\sum_{l=1}^{K} w(\boldsymbol{\theta}_{j,l})} \tag{B.2}$$

(iv) Evaluate the weighted sample covariance with $\boldsymbol{\theta}_j$ and p_j using the following relations:

$$\begin{aligned} \bar{\boldsymbol{\theta}}_j &= \sum_{k=1}^{K} p_{j,k}\boldsymbol{\theta}_{j,k} \\ C_j &= \sum_{k=1}^{K}(\boldsymbol{\theta}_{j,k} - \bar{\boldsymbol{\theta}}_j)(\boldsymbol{\theta}_{j,k} - \bar{\boldsymbol{\theta}}_j)^T p_{j,k} \end{aligned} \tag{B.3}$$

(v) Use resampled samples $\boldsymbol{\Theta}_j$ as seeds for generating N_{steps} samples of the intermediate probability density using Metropolis Hastings algorithm with a Gaussian proposal density that has covariance $\delta^2 C_j$. The covariance is adapted within a Markov chain and is controlled by coefficient δ, where $\delta = a + bR$, and R is the acceptance rate. The parameters, a and b are chosen according to the dimension of the problem. In our case, they were empirically chosen as $a = \frac{1}{90}$ and $b = \frac{89}{90}$.

(vi) Collect the final sample from each of the K Markov chains and assign them as samples $(\boldsymbol{\theta}_{j+1})$ for stage $j + 1$.

(vii) Repeat the steps (ii) to (vi) until $\gamma_{j+1} \geq 1$.

Acknowledgments

This article benefited from constructive comments by Associate editor Andrew Hooper, and an anonymous reviewer. We thank Birgir Hrafnkelsson (Univ. Iceland) for commenting on a part of the manuscript. The research reported upon in this publication was supported by King Abdullah University of Science and Technology (KAUST) under award number BAS/1/1353-01-01. The simulations were run in Ibex (formerly called Noor/SM-C/Dragon) high performance computing clusters in KAUST. The data and Matlab codes for the calculations and visualizations are available from OSF public repository (`https://osf.io/7vdm9/`). Python codes for the SMC sampling are available from Zenodo and Github repositories (Dutta, 2020, `https://github.com/rishabhdutta/SMC-python`).

References

Abe, T., Furuya, M., & Takada, Y. (2013). Nonplanar fault source modeling of the 2008 Mw 6.9 Iwate-Miyagi inland earthquake in northeast Japan. *Bull. Seismol. Soc. Am.*, *103*(1), 507–518.

Akoglu, A. M., Cakir, Z., Meghraoui, M., Belabbes, S., El Alami, S. O., Ergintav, S., & Akyüz, H. S. (2006). The 1994-2004 Al Hoceima (Morocco) earthquake sequence: Conjugate fault ruptures deduced from InSAR. *Earth Planet. Sci. Lett.*, *252*(3–4), 467–480.

Altman, A., & Gondzio, J. (1999). Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. *Optim. Method. Softw.*, *11*(1–4), 275–302.

Aochi, H., & Fukuyama, E. (2002). Three-dimensional nonplanar simulation of the 1992 Landers earthquake. *J. Geophys. Res.- Sol. Ea.*, *107*(B2), ESE–4.

Aochi, H., Fukuyama, E., & Matsu'ura, M. (2000). Spontaneous rupture propagation on a non-planar fault in 3-D elastic medium. *Pure Appl. Geophys.*, *157*(11-12), 2003–2027.

Barrientos, S. E., & Ward, S. N. (1990). The 1960 Chile earthquake: inversion for slip distribution from surface deformation. *Geophys. J. Int.*, *103*(3), 589–598.

Bathke, H., Nikkhoo, M., Holohan, E., & Walter, T. R. (2015). Insights into the 3D architecture of an active caldera ring-fault at Tendürek volcano through modeling of geodetic data. *Earth Planet. Sc. Lett.*, *422*, 157–168.

Beck, J., & Au, S.-K. (2002). Bayesian updating of structural models and reliability using Markov Chain Monte Carlo simulation. *J. Eng. Mech.*, *128*(4), 380–391.

Belabbès, S., Wicks, C., Çakir, Z., & Meghraoui, M. (2009). Rupture parameters of the 2003 Zemmouri (Mw6.8), Algeria, earthquake from joint inversion of interferometric synthetic aperture radar, coastal uplift, and GPS. *J. Geophys. Res.- Sol. Ea.*, *114*(B03406).

Berkhof, J., Van Mechelen, I., & Gelman, A. (2003). A bayesian approach to the selection and testing of mixture models. *Stat. Sinica*, *13*(2), 423–442.

Bürgmann, R., & Pollard, D. D. (1994a). Strain accommodation about strike-slip fault discontinuities in granitic rock under brittle-to-ductile conditions. *J. Struct. Geol.*, *16*(12), 1655–1674.

Bürgmann, R., Pollard, D. D., & Martel, S. (1994b). Slip distributions on faults: effects of stress gradients, inelastic deformation, heterogeneous host-rock stiffness, and fault interaction. *J. Struct. Geol.*, *16*(12), 1675–1690.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.*, *33*(2), 261–304.

Candela, T., Renard, F., Klinger, Y., Mair, K., Schmittbuhl, J., & Brodsky, E. E. (2012). Roughness of fault surfaces over nine decades of length scales. *J. Geophys. Res.- Sol. Ea.*, *117*(B8).

Chib, S., & Jeliazkov, I. (2001). Marginal likelihood from the Metropolis-Hastings output. *J. Am. Stat. Assoc.*, *96*(453), 270–281.

Chib, S., & Ramamurthy, S. (2010). Tailored randomized block MCMC methods with application to DSGE models. *J. Econom.*, *155*(1), 19–38.

Ching, J., & Chen, Y.-C. (2007). Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. *J. Eng. Mech.*, *133*(7), 816–832.

Chopin, N. (2002). A sequential particle filter method for static models. *Biometrika*, *89*(3), 539–552.

Cox, S. J. D., & Scholz, C. H. (1988a). Rupture initiation in shear fracture of rocks: An experimental study. *J. Geophys. Res.- Sol. Ea.*, *93*(B4), 3307–3320.

Cox, S. J. D., & Scholz, C. H. (1988b). On the formation and growth of faults: an experimental study. *J. Struct. Geol.*, *10*(4), 413–430.

Cruikshank, K. M., & Aydin, A. (1994). Role of fracture localization in arch formation, Arches National Park, Utah. *Geol. Soc. Am. Bull.*, *106*(7), 879–891.

Dettmer, J., Benavente, R., Cummins, P. R., & Sambridge, M. (2014). Trans-dimensional finite-fault inversion. *Geophys. J. Int.*, *199*(2), 735–751.

Dettmer, J., Dosso, S. E., & Holland, C. W. (2007). Uncertainty estimation in seismo-acoustic reflection travel time inversion. *J. Acoust. Soc. Am.*, *122*(1), 161–176.

Duman, T. Y., Emre, O., Dogan, A., & Ozalp, S. (2005). Step-over and bend structures along the 1999 Duzce earthquake surface rupture, North Anatolian fault, Turkey. *Bull. Seismol. Soc. Am.*, *95*(4), 1250–1262.

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.

Duputel, Z., Rivera, L., Fukahata, Y., & Kanamori, H. (2012). Uncertainty estimations for seismic source inversions. *Geophys. J. Int.*, *190*(2), 1243–1256.

Dutta, R. (2020, June). *Python codes for Sequential Monte Carlo sampling technique.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.3908931` doi: 10.5281/zenodo.3908931

Dutta, R., Jónsson, S., Wang, T., & Vasyura-Bathke, H. (2018). Bayesian estimation of source parameters and associated Coulomb failure stress changes for the 2005 Fukuoka (Japan) Earthquake. *Geophys. J. Int.*, *213*(1), 261–277.

Elliott, J., Jolivet, R., González, P., Avouac, J.-P., Hollingsworth, J., Searle, M., & Stevens, V. (2016). Himalayan megathrust geometry and relation to topography revealed by the Gorkha earthquake. *Nat. Geosci.*, *9*(2), 174.

Fialko, Y., Sandwell, D., Simons, M., & Rosen, P. (2005). Three-dimensional deformation caused by the Bam, Iran, earthquake and the origin of shallow slip deficit. *Nature*, *435*(7040), 295.

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*(2), 353–364.

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

Fukuda, J., & Johnson, K. M. (2010). Mixed linear-non-linear inversion of crustal deformation data: Bayesian inference of model, weighting and regularization parameters. *Geophys. J. Int.*, *181*(3), 1441–1458.

Funning, G. J., Fukahata, Y., Yagi, Y., & Parsons, B. (2014). A method for the joint inversion of geodetic and seismic waveform data using ABIC: application to the 1997 Manyi, Tibet, earthquake. *Geophys. J. Int.*, *196*(3), 1564–1579.

Furuya, M., & Yasuda, T. (2011). The 2008 Yutian normal faulting earthquake (Mw 7.1), NW Tibet: Non-planar fault modeling and implications for the Karakax Fault. *Tectonophysics*, *511*(3-4), 125–133.

Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis.* Chapman and Hall/CRC.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. (1995). *Markov chain Monte Carlo in practice.* Chapman and Hall/CRC.

Haario, H., Saksman, E., & Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, *7*(2), 223–242.

Hashimoto, C., Noda, A., Sagiya, T., & Matsu'ura, M. (2009). Interplate seismogenic zones along the Kuril-Japan trench inferred from GPS data inversion.

Nat. Geosci., *2*(2), 141.
Hearn, E. H., & Bürgmann, R. (2005). The effect of elastic layering on inversions of GPS data for coseismic slip and resulting stress changes: Strike-slip earthquakes. *Bull. seism. Soc. Am.*, *95*(5), 1637–1653.
Heimann, S., Vasyura-Bathke, H., Sudhaus, H., Isken, M. P., Kriegerowski, M., Steinberg, A., & Dahm, T. (2019). A Python framework for efficient use of pre-computed Green's functions in seismological and other physical forward and inverse source problems. *Solid Earth*, *10*(6), 1921–1935.
Hsu, Y.-J., Yu, S.-B., & Chen, H.-Y. (2009). Coseismic and postseismic deformation associated with the 2003 Chengkung, Taiwan, earthquake. *Geophys. J. Int.*, *176*(2), 420–430.
Improta, L., Latorre, D., Margheriti, L., Nardi, A., Marchetti, A., Lombardi, A. M., … others (2019). Multi-segment rupture of the 2016 Amatrice-Visso-Norcia seismic sequence (central Italy) constrained by the first high-quality catalog of Early Aftershocks. *Sci. Rep.*, *9*(1), 6921.
Jackson, D. D. (1979). The use of a priori data to resolve non-uniqueness in linear inversion. *Geophys. J. Roy. Astr. S.*, *57*(1), 137–157.
Jiang, G., Xu, C., Wen, Y., Liu, Y., Yin, Z., & Wang, J. (2013). Inversion for coseismic slip distribution of the 2010 M w 6.9 Yushu Earthquake from InSAR data using angular dislocations. *Geophys. J. Int.*, *194*(2), 1011–1022.
Jónsson, S., Zebker, H., Segall, P., & Amelung, F. (2002). Fault slip distribution of the 1999 M w 7.1 Hector Mine, California, earthquake, estimated from satellite radar and GPS measurements. *Bull. seism. Soc. Am.*, *92*(4), 1377–1389.
Kaneko, Y., & Fialko, Y. (2011). Shallow slip deficit due to large strike-slip earthquakes in dynamic rupture simulations with elasto-plastic off-fault response. *Geophys. J. Int.*, *186*(3), 1389–1403.
Kaven, J. O., & Pollard, D. D. (2013). Geometry of crustal faults: Identification from seismicity and implications for slip and stress transfer models. *J. Geophys. Res.- Sol. Ea.*, *118*(9), 5058–5070.
Klinger, Y. (2010). Relation between continental strike-slip earthquake segmentation and thickness of the crust. *J. Geophys. Res.- Sol. Ea.*, *115*(B7).
Knuth, K. H., Habeck, M., Malakar, N. K., Mubeen, A. M., & Placek, B. (2015). Bayesian evidence and model selection. *Digital Signal Process.*, *47*, 50–67.
Kowsari, M., Halldorsson, B., Hrafnkelsson, B., & Jonsson, S. (2019). Selection of earthquake ground motion models using the deviance information criterion. *Soil Dyn. Earthq. Eng.*, *117*, 288–299.
Lay, T. (2018). A review of the rupture characteristics of the 2011 Tohoku-Oki Mw 9.1 earthquake. *Tectonophysics*, *733*, 4–36.
Lee, A., Yau, C., Giles, M. B., Doucet, A., & Holmes, C. C. (2010). On the utility of graphics cards to perform massively parallel simulation of advanced Monte Carlo methods. *J. Comput. Graph. Stat.*, *19*(4), 769–789.
Madigan, D., & Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *J. Am. Stat. Assoc.*, *89*(428), 1535–1546.
Maerten, F., Resor, P., Pollard, D., & Maerten, L. (2005). Inverting for slip on three-dimensional fault surfaces using angular dislocations. *Bull. seism. Soc. Am.*, *95*(5), 1654–1665.
Mai, P. M., & Thingbaijam, K. K. S. (2014). SRCMOD: An online database of finite-fault rupture models. *Seismol. Res. Lett.*, *85*(6), 1348–1357.
Martel, S. J. (1999). Mechanical controls on fault geometry. *J. Struct. Geol.*, *21*(6), 585–596.
Matsu'ura, M., Noda, A., & Fukahata, Y. (2007). Geodetic data inversion based on Bayesian formulation with direct and indirect prior information. *Geophys. J. Int.*, *171*(3), 1342–1351.
Meade, B. J. (2007). Algorithms for the calculation of exact displacements, strains,

and stresses for triangular dislocation elements in a uniform elastic half space. *Comput. Geosci.*, *33*(8), 1064–1075.

Minson, S. E., Simons, M., Beck, J., Ortega, F., Jiang, J., Owen, S., … Sladen, A. (2014). Bayesian inversion for finite fault earthquake source models-II: the 2011 great Tohoku-oki, Japan earthquake. *Geophys. J. Int.*, *198*(2), 922–940.

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.

Monelli, D., Mai, P., Jónsson, S., & Giardini, D. (2009). Bayesian imaging of the 2000 Western Tottori (Japan) earthquake through fitting of strong motion and GPS data. *Geophys. J. Int.*, *176*(1), 135–150.

Moreno, M. S., Bolte, J., Klotz, J., & Melnick, D. (2009). Impact of megathrust geometry on inversion of coseismic slip from geodetic data: Application to the 1960 Chile earthquake. *Geophys. Res. Lett.*, *36*(16).

Passone, L., & Mai, P. M. (2017). Kinematic earthquake ground-motion simulations on listric normal faults. *Bull. Seism. Soc. Am.*, *107*(6), 2980–2993.

Power, W., & Tullis, T. (1995). Review of the fractal character of natural fault surfaces with implications for friction and the evolution of fault zones. In *Fractals in the Earth Sciences* (pp. 89–105). Springer.

Ragon, T., Sladen, A., & Simons, M. (2018). Accounting for uncertain fault geometry in earthquake source inversions-I: theory and simplified application. *Geophys. J. Int.*, *214*(2), 1174–1190.

Razafindrakoto, H. N. T., Mai, P. M., Genton, M. G., Zhang, L., & Thingbaijam, K. K. S. (2015). Quantifying variability in earthquake rupture models using multidimensional scaling: application to the 2011 Tohoku earthquake. *Geophys. J. Int.*, *202*(1), 17–40.

Reilinger, R., Ergintav, S., Bürgmann, R., McClusky, S., Lenk, O., Barka, A., … others (2000). Coseismic and postseismic fault slip for the 17 August 1999, M= 7.5, Izmit, Turkey earthquake. *Science*, *289*(5484), 1519–1524.

Resor, P., Pollard, D., Wright, T., & Beroza, G. (2005). Integrating high-precision aftershock locations and geodetic observations to model coseismic deformation associated with the 1995 Kozani-Grevena earthquake, Greece. *J. Geophys. Res.- Sol. Ea.*, *110*(B9).

Rybicki, K. R. (1992). Strike-slip faulting in the presence of low-rigidity inhomogeneities. *Bull. Seism. Soc. Am.*, *82*(5), 2170–2190.

Rybicki, K. R., & Yamashita, T. (1998). Faulting in vertically inhomogeneous media and its geophysical implications. *Geophys. Res. Lett.*, *25*(15), 2893–2896.

Segall, P., & Pollard, D. D. (1980). Mechanics of discontinuous faults. *J. Geophys. Res.- Sol. Ea.*, *85*(B8), 4337–4350.

Segall, P., & Pollard, D. D. (1983). Nucleation and growth of strike slip faults in granite. *J. Geophys. Res.- Sol. Ea.*, *88*(B1), 555–568.

Shen, Z.-K., Sun, J., Zhang, P., Wan, Y., Wang, M., Bürgmann, R., … Wang, Q. (2009). Slip maxima at fault junctions and rupturing of barriers during the 2008 Wenchuan earthquake. *Nat. Geosci.*, *2*(10), 718.

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*(4), 1390–1402.

Sudhaus, H., & Jónsson, S. (2009). Improved source modelling through combined use of InSAR and GPS under consideration of correlated data errors: application to the June 2000 Kleifarvatn earthquake, Iceland. *Geophys. J. Int.*, *176*(2), 389–404.

Sudhaus, H., & Jónsson, S. (2011). Source model for the 1997 Zirkuh earthquake (MW= 7.2) in Iran derived from JERS and ERS InSAR observations. *Geophys. J. Int.*, *185*(2), 676–692.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter esti-*

mation (Vol. 89). siam.

Tarantola, A., & Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Rev. Geophys.*, *20*(2), 219–232.

Usman, M., & Furuya, M. (2015). Complex faulting in the Quetta Syntaxis: fault source modeling of the October 28, 2008 earthquake sequence in Baluchistan, Pakistan, based on ALOS/PALSAR InSAR data. *Earth Planets Space*, *67*(1), 142.

Vasyura-Bathke, H., Dettmer, J., Steinberg, A., Heimann, S., Isken, M. P., Zielke, O., … Jónsson, S. (2020, 01). The Bayesian Earthquake Analysis Tool. *Seismol. Res. Lett.*. doi: 10.1785/0220190075

Vermilye, J. M., & Scholz, C. H. (1998). The process zone: A microstructural view of fault growth. *J. Geophys. Res.- Sol. Ea.*, *103*(B6), 12223–12237.

Von der Linden, W., Dose, V., & Von Toussaint, U. (2014). *Bayesian probability theory: applications in the physical sciences.* Cambridge University Press.

Wan, Y., Shen, Z.-K., Burgmann, R., Sun, J., & Wang, M. (2017). Fault geometry and slip distribution of the 2008 Mw 7.9 Wenchuan, China earthquake, inferred from GPS and InSAR measurements. *Geophys. J. Int.*, *208*(2), 748–766. doi: 10.1093/gji/ggw421

Wang, K., & Fialko, Y. (2018). Observations and modeling of coseismic and postseismic deformation due to the 2015 Mw 7.8 Gorkha (Nepal) earthquake. *J. Geophys. Res.- Sol. Ea.*, *123*(1), 761–779.

Wesnousky, S. G. (1988). Seismological and structural evolution of strike-slip faults. *Nature*, *335*(6188), 340.

Wesnousky, S. G. (2008). Displacement and geometrical characteristics of earthquake surface ruptures: Issues and implications for seismic-hazard analysis and the process of earthquake rupture. *Bull. Seismol. Soc. Am.*, *98*(4), 1609–1632.

Xu, W., Dutta, R., & Jónsson, S. (2015). Identifying active faults by improving earthquake locations with InSAR data and Bayesian estimation: the 2004 Tabuk (Saudi Arabia) earthquake sequence. *Bull. Seismol. Soc. Am.*, *105*(2A), 765–775.

Xu, X., Tong, X., Sandwell, D. T., Milliner, C. W., Dolan, J. F., Hollingsworth, J., … Ayoub, F. (2016). Refining the shallow slip deficit. *Geophys. J. Int.*, *204*(3), 1867–1886.

Yabuki, T., & Matsu'Ura, M. (1992). Geodetic data inversion using a Bayesian information criterion for spatial distribution of fault slip. *Geophys. J. Int.*, *109*(2), 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*(1), 215–221.

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.

Figure1.

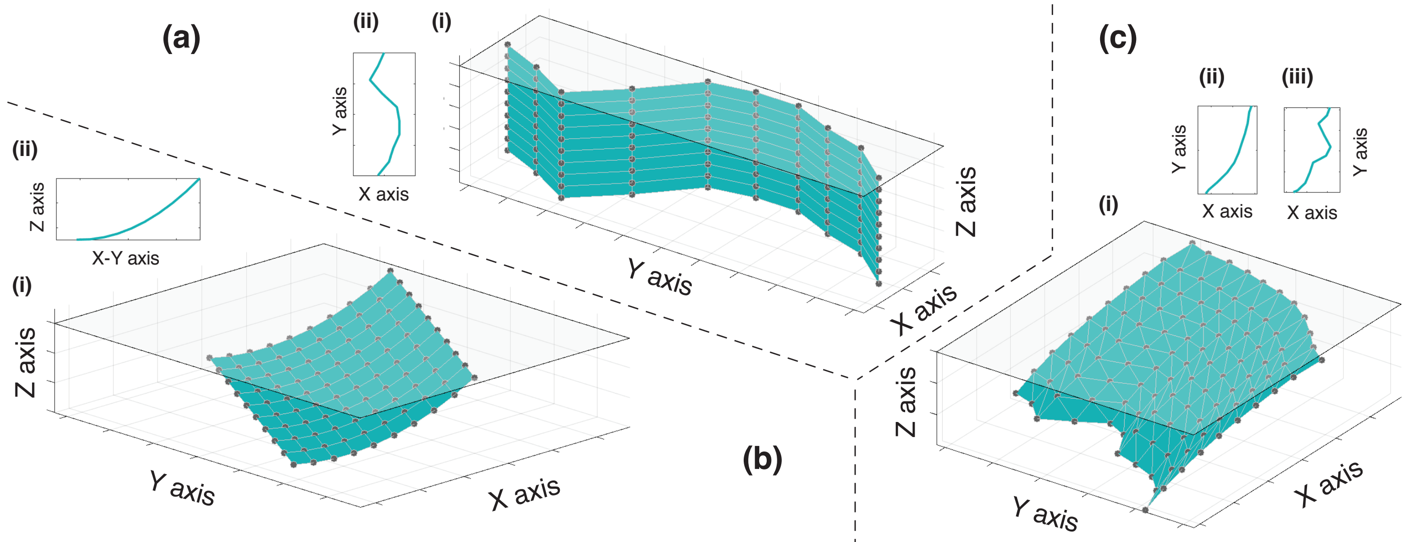

(a)
(ii)
Z axis
X-Y axis
(i)
Z axis
Y axis
X axis
(b)
(ii)
Y axis
X axis
(i)
Z axis
Y axis
X axis
(c)
(ii)
Y axis
X axis
(iii)
Y axis
X axis
(i)
Z axis
Y axis
X axis

Figure2.

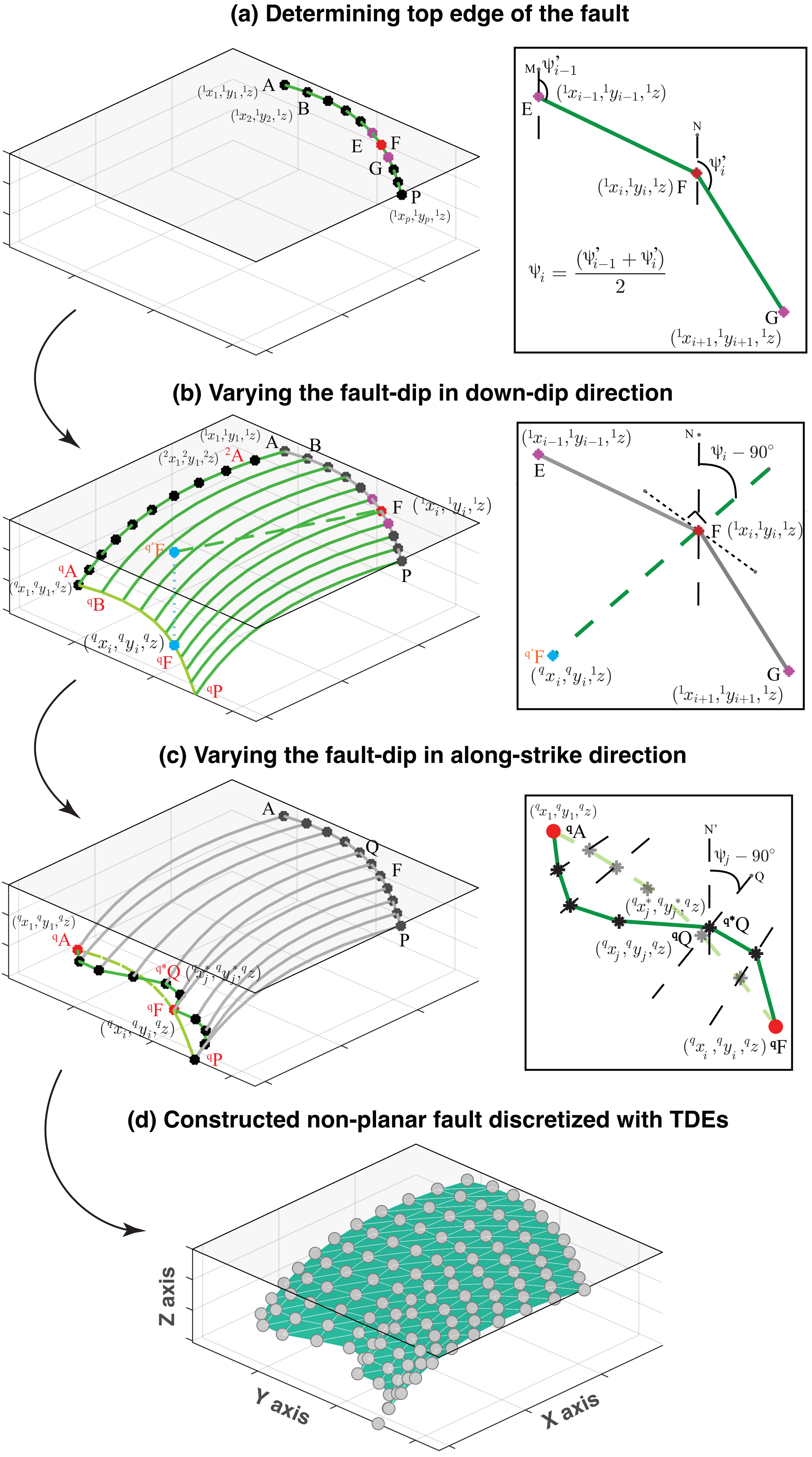
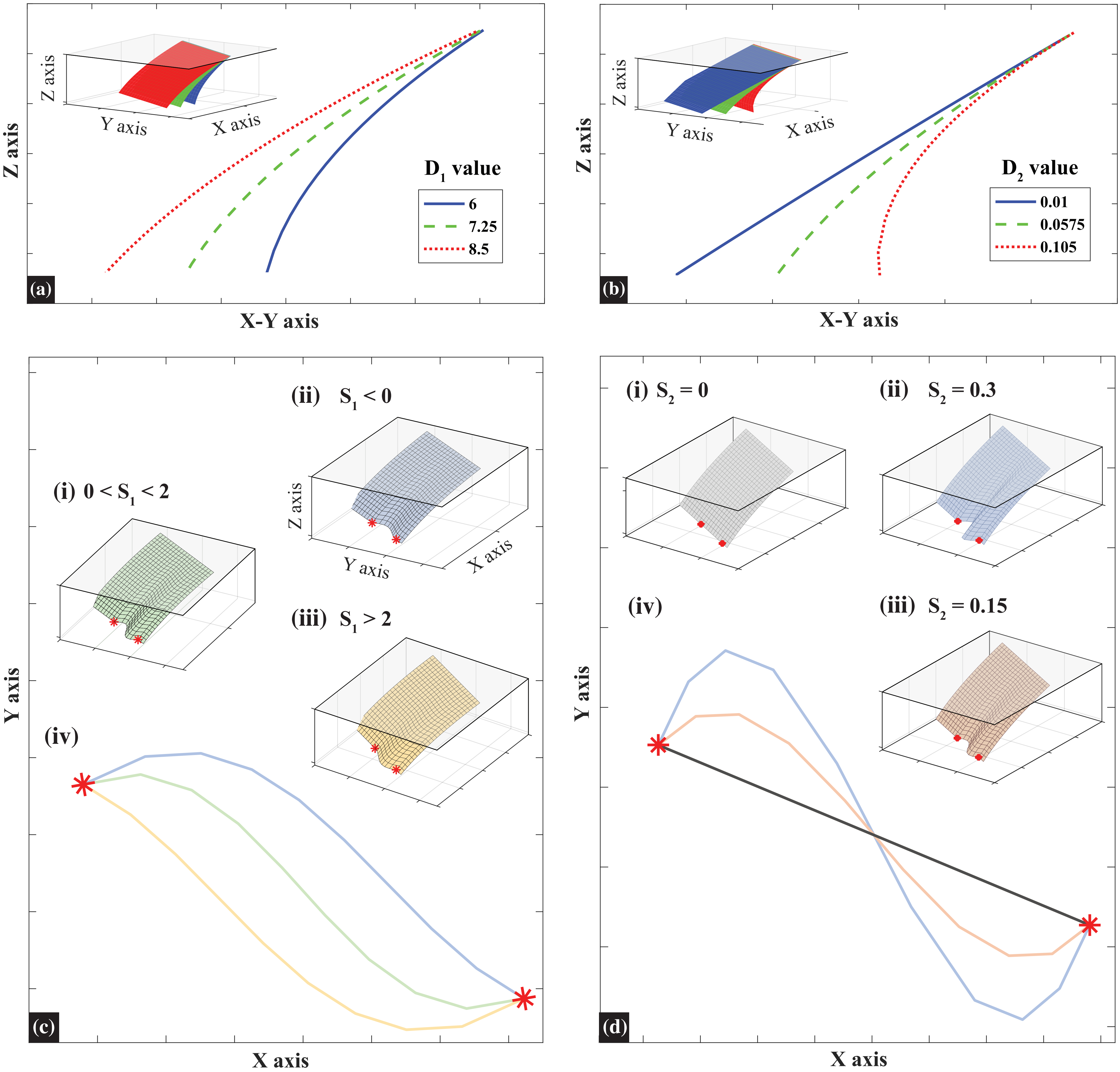

(a) Determining top edge of the fault
A
B
E
F
G
P
M
N
$\psi_i = \frac{(\psi'_{i-1} + \psi'_i)}{2}$
(b) Varying the fault-dip in down-dip direction
$\psi_i - 90°$
(c) Varying the fault-dip in along-strike direction
Q
N'
$\psi_j - 90°$
(d) Constructed non-planar fault discretized with TDEs
Z axis
Y axis
X axis

Figure3.



(a)
Z axis
Y axis
X axis
Z axis
D1 value
6
7.25
8.5
X-Y axis
(b)
Z axis
Y axis
X axis
Z axis
D2 value
0.01
0.0575
0.105
X-Y axis
(c)
(i) 0 < S1 < 2
(ii) S1 < 0
Z axis
Y axis
X axis
(iii) S1 > 2
(iv)
Y axis
X axis
(d)
(i) S2 = 0
(ii) S2 = 0.3
(iii) S2 = 0.15
(iv)
Y axis
X axis

Figure4.

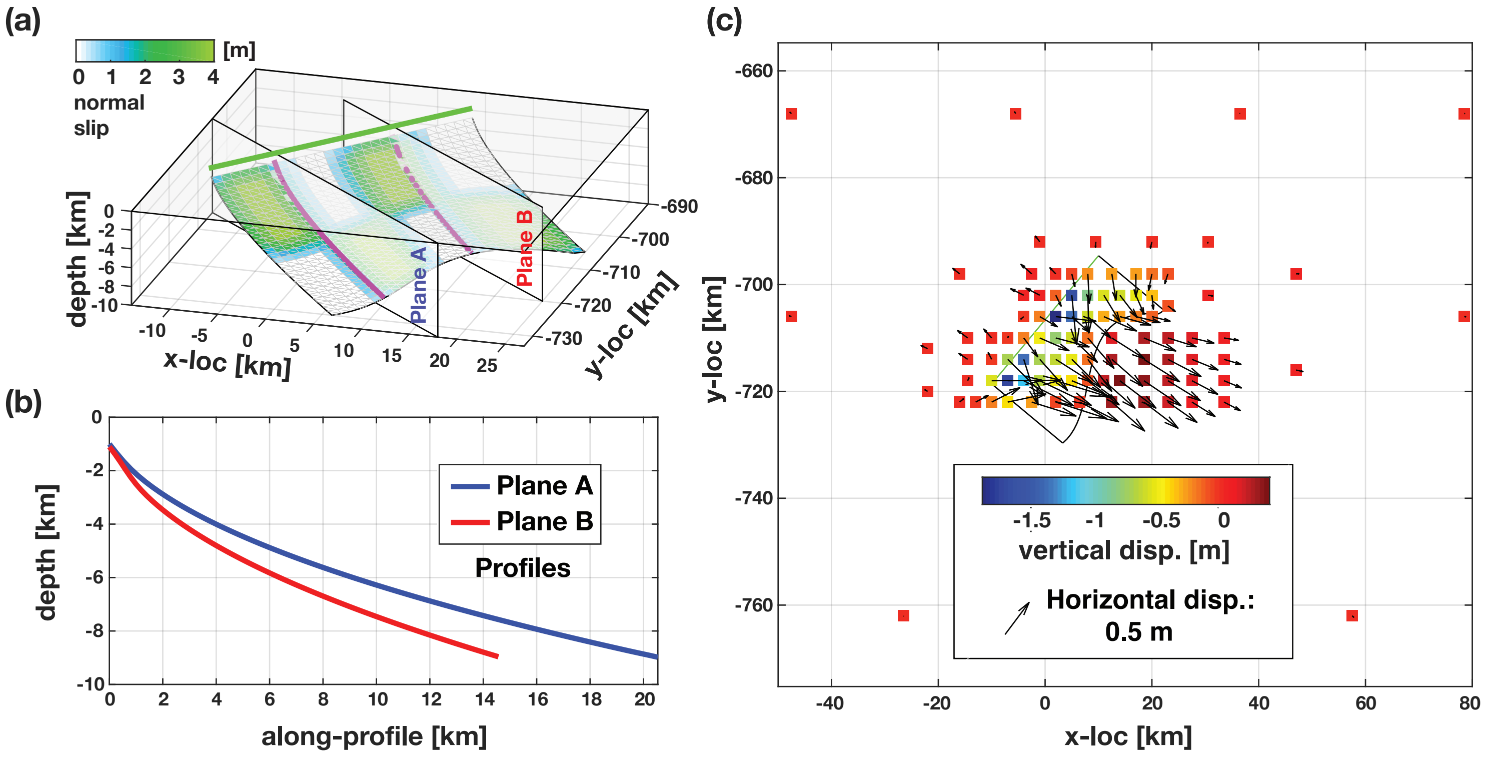
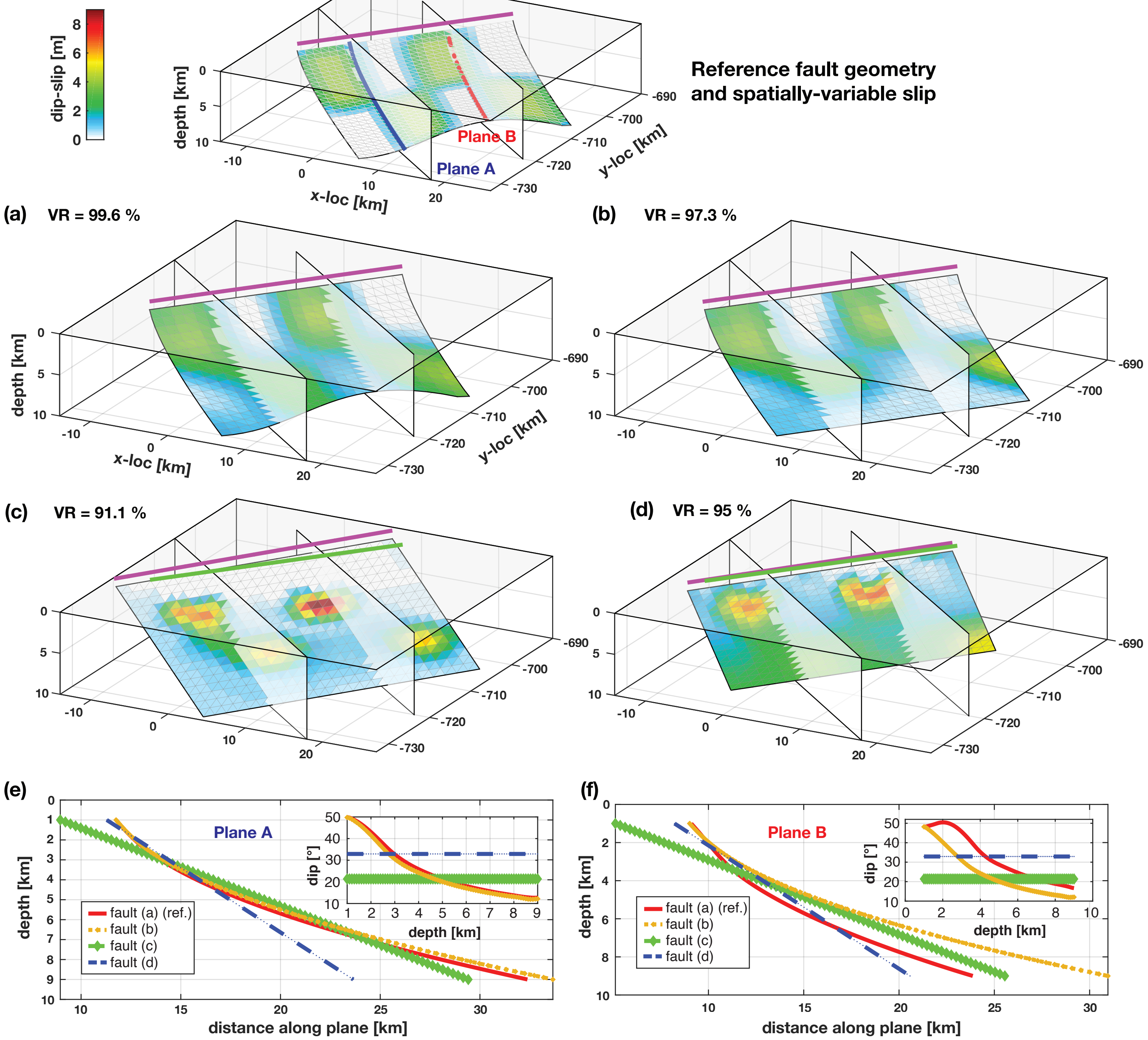

(a)
[m]
0 1 2 3 4
normal
slip
depth [km]
0
-2
-4
-6
-8
-10
-10
-5
0
5
10
15
20
25
x-loc [km]
-690
-700
-710
-720
-730
y-loc [km]
Plane A
Plane B
(b)
depth [km]
along-profile [km]
Plane A
Plane B
Profiles
(c)
y-loc [km]
-660
-680
-700
-720
-740
-760
x-loc [km]
-40
-20
0
20
40
60
80
-1.5 -1 -0.5 0
vertical disp. [m]
Horizontal disp.:
0.5 m

Figure5.



Reference fault geometry and spatially-variable slip
dip-slip [m]
Plane A
Plane B
x-loc [km]
y-loc [km]
depth [km]
(a) VR = 99.6 %
(b) VR = 97.3 %
(c) VR = 91.1 %
(d) VR = 95 %
(e)
(f)
fault (a) (ref.)
fault (b)
fault (c)
fault (d)
dip [°]
depth [km]
distance along plane [km]

Figure6.

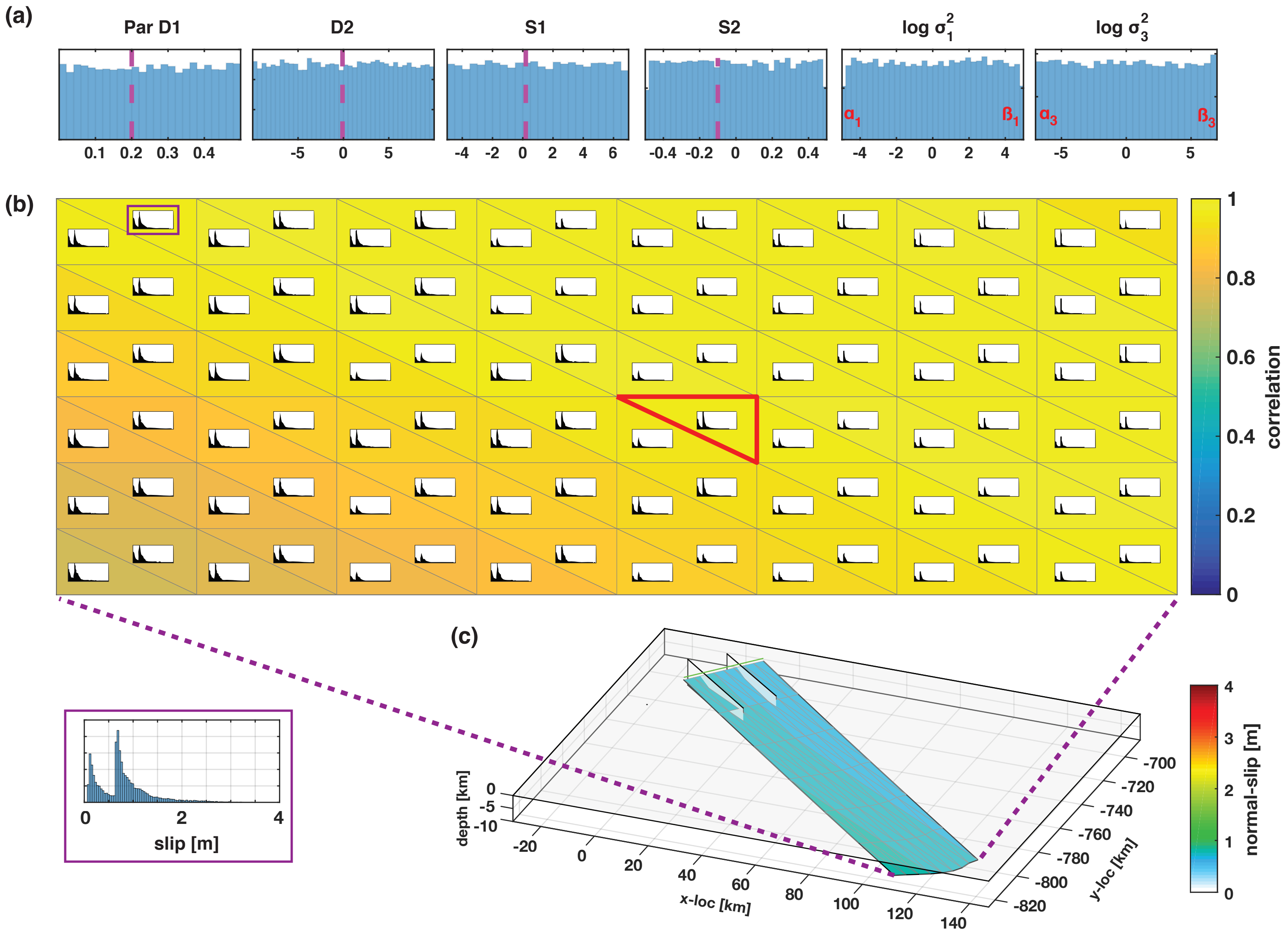
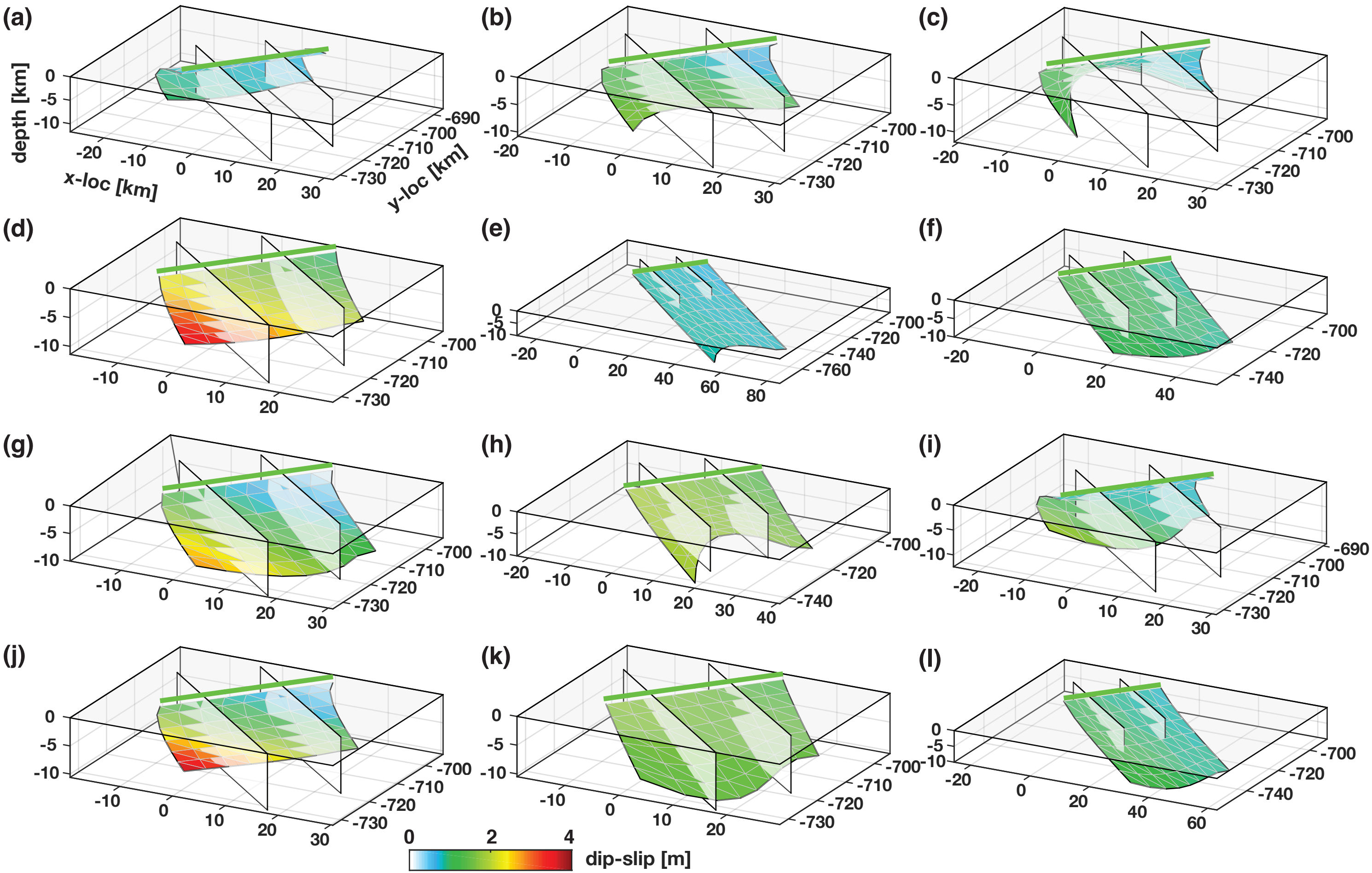
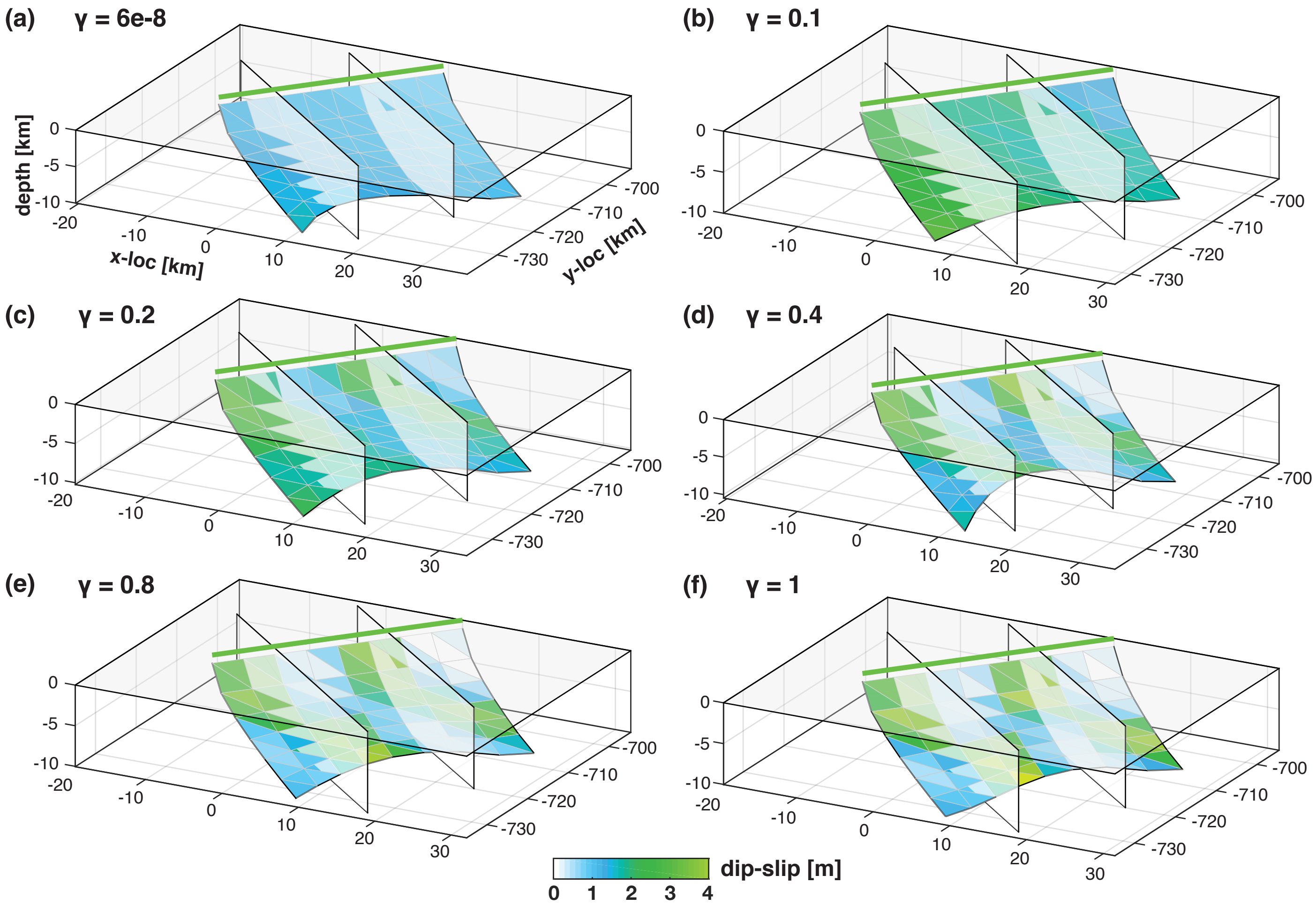

(a)
Par D1
D2
S1
S2
log σ_1^2
log σ_3^2
α_1
β_1
α_3
β_3
0.1 0.2 0.3 0.4
-5 0 5
-4 -2 0 2 4 6
-0.4 -0.2 0 0.2 0.4
-4 -2 0 2 4
-5 0 5
(b)
correlation
0 0.2 0.4 0.6 0.8 1
(c)
slip [m]
0 2 4
depth [km]
0 -5 -10
x-loc [km]
-20 0 20 40 60 80 100 120 140
y-loc [km]
-700 -720 -740 -760 -780 -800 -820
normal-slip [m]
0 1 2 3 4

Figure7.



(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
depth [km]
x-loc [km]
y-loc [km]
dip-slip [m]
0
2
4

Figure8.



(a) γ = 6e-8
(b) γ = 0.1
(c) γ = 0.2
(d) γ = 0.4
(e) γ = 0.8
(f) γ = 1
depth [km]
x-loc [km]
y-loc [km]
dip-slip [m]
0
1
2
3
4

Figure9.

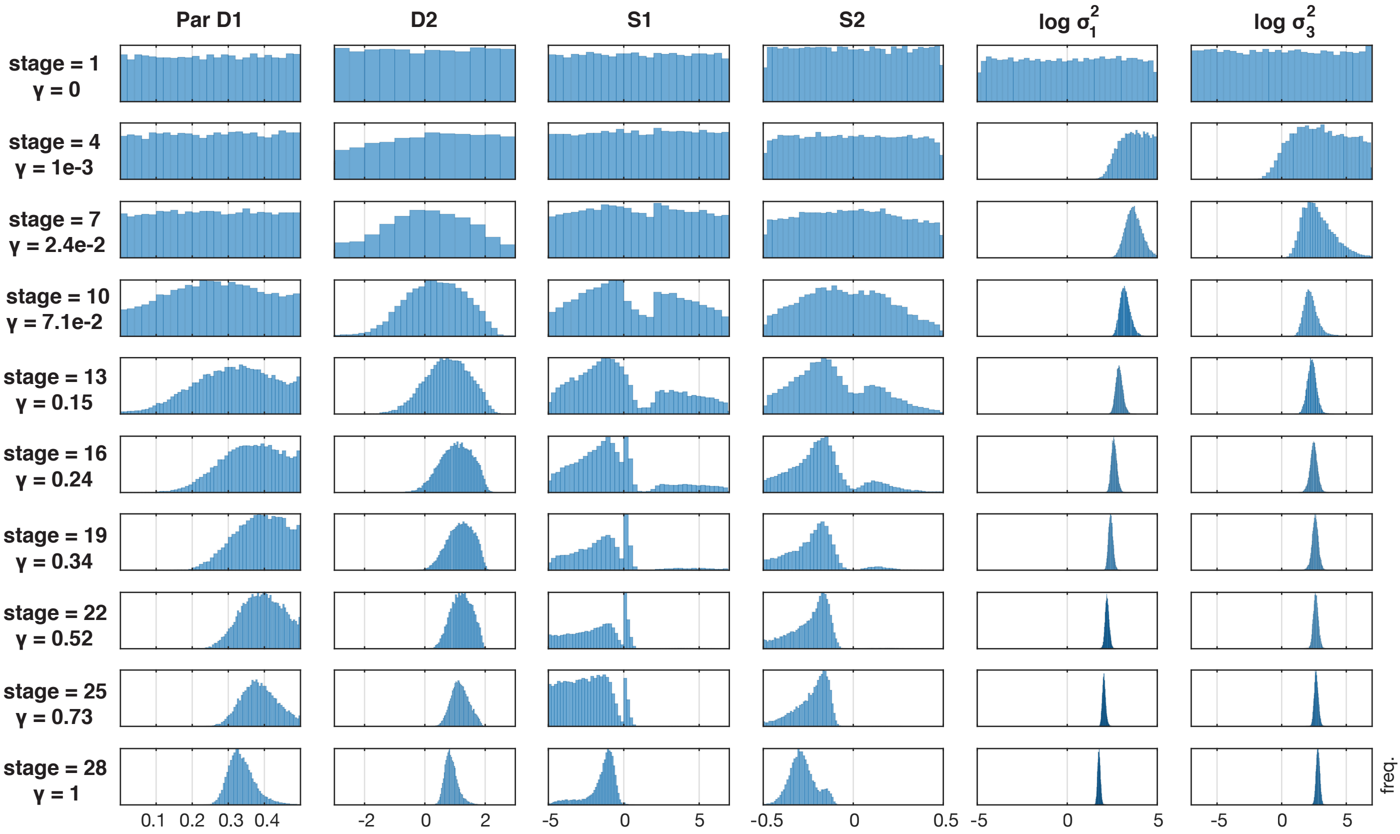
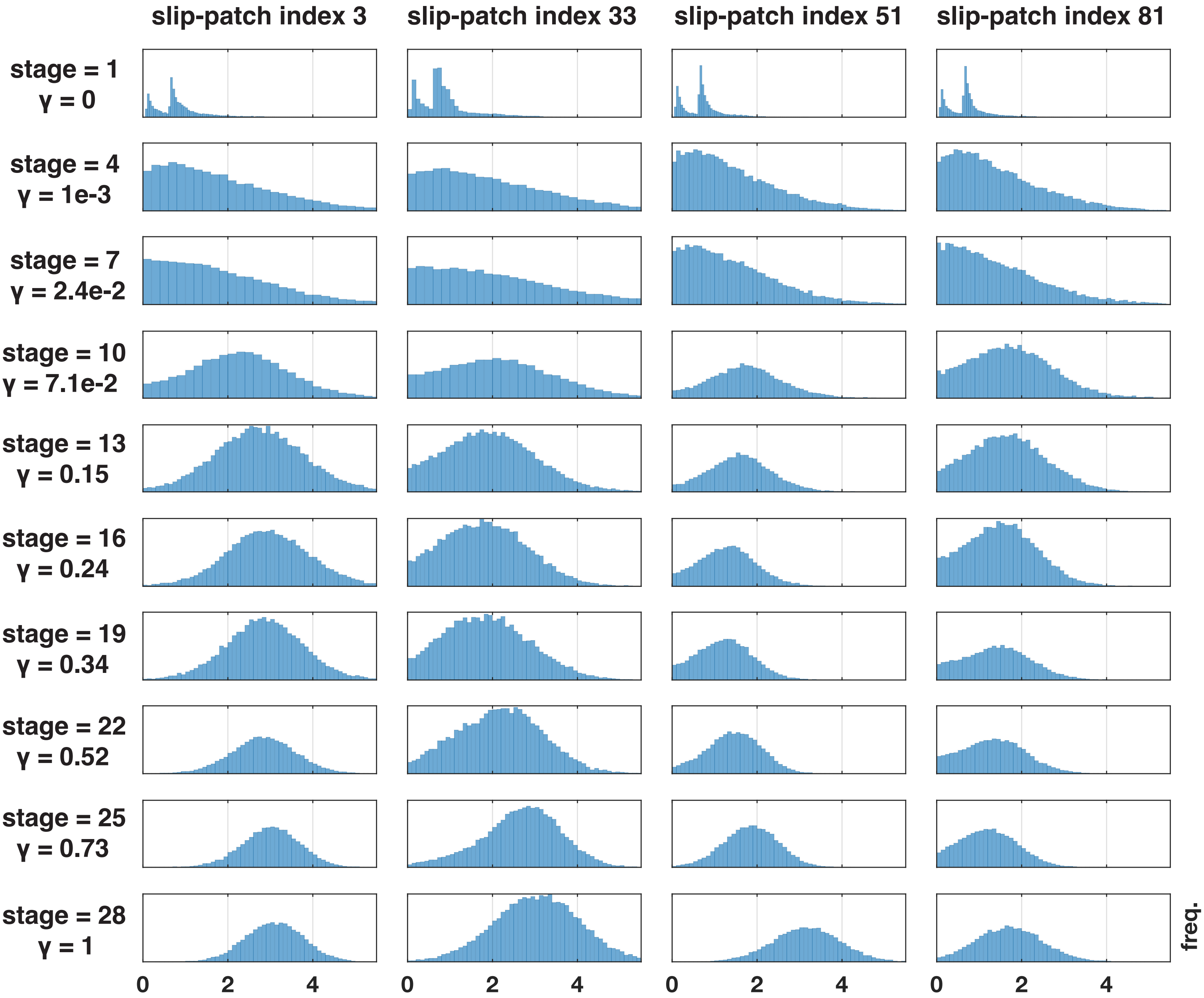
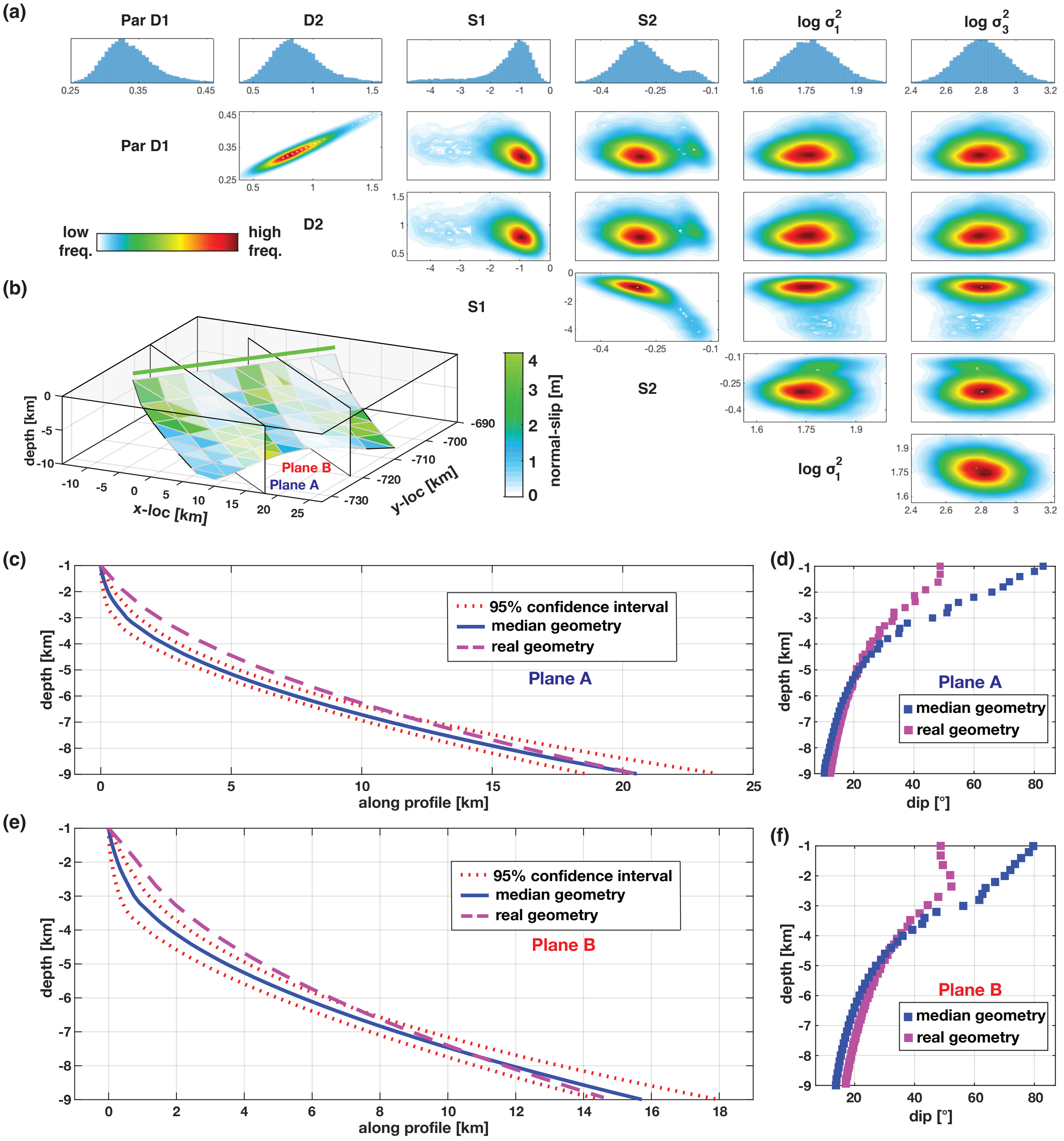

Par D1
D2
S1
S2
log σ_1^2
log σ_3^2
stage = 1
γ = 0
stage = 4
γ = 1e-3
stage = 7
γ = 2.4e-2
stage = 10
γ = 7.1e-2
stage = 13
γ = 0.15
stage = 16
γ = 0.24
stage = 19
γ = 0.34
stage = 22
γ = 0.52
stage = 25
γ = 0.73
stage = 28
γ = 1
freq.
0.1 0.2 0.3 0.4
-2 0 2
-5 0 5
-0.5 0 0.5
-5 0 5
-5 0 5

Figure10.



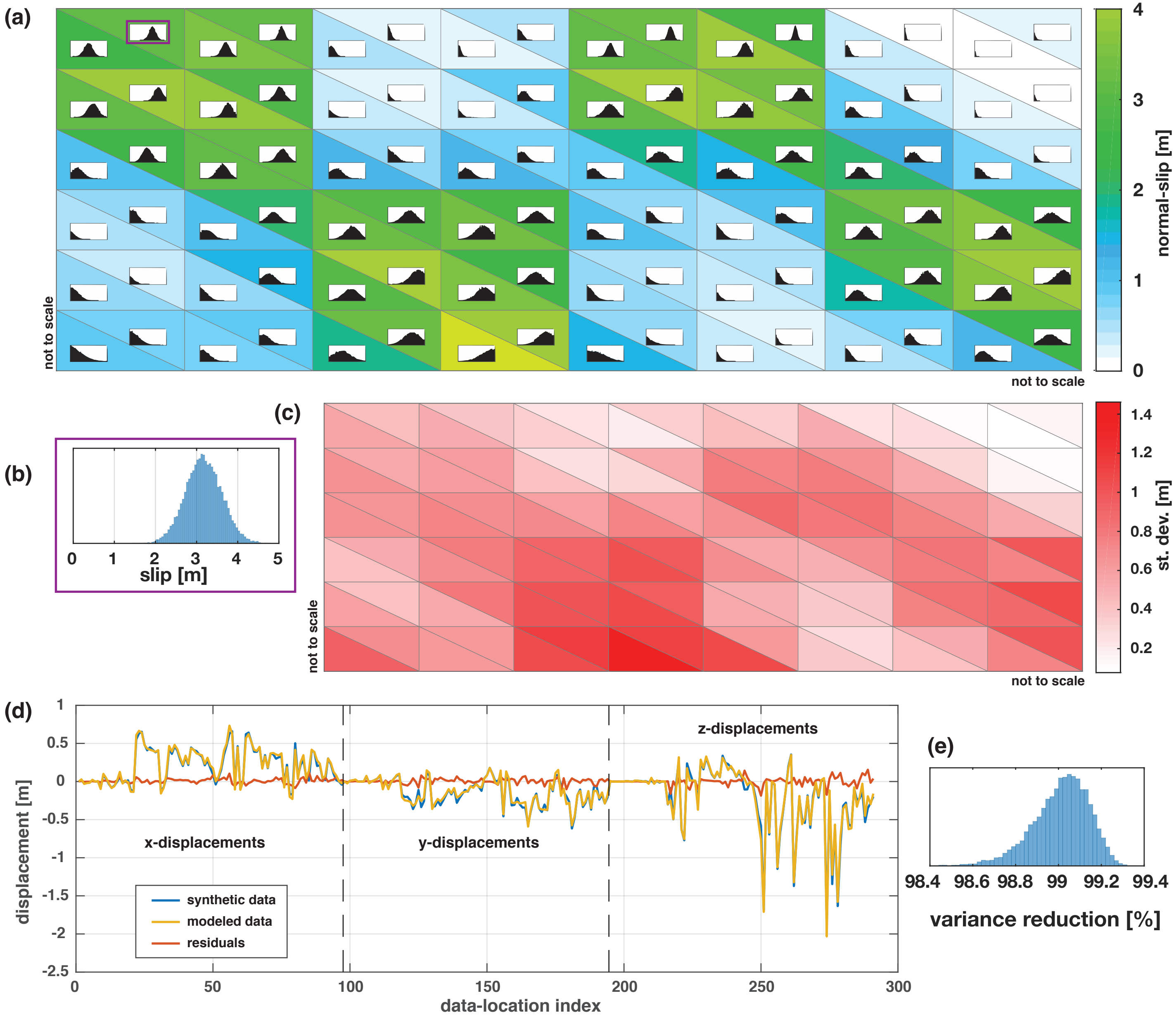
slip-patch index 3
slip-patch index 33
slip-patch index 51
slip-patch index 81
stage = 1
γ = 0
stage = 4
γ = 1e-3
stage = 7
γ = 2.4e-2
stage = 10
γ = 7.1e-2
stage = 13
γ = 0.15
stage = 16
γ = 0.24
stage = 19
γ = 0.34
stage = 22
γ = 0.52
stage = 25
γ = 0.73
stage = 28
γ = 1
0
2
4
freq.

Figure11.



(a)
Par D1
D2
S1
S2
log σ_1^2
log σ_3^2
low freq.
high freq.
(b)
depth [km]
x-loc [km]
y-loc [km]
Plane B
Plane A
normal-slip [m]
(c)
95% confidence interval
median geometry
real geometry
Plane A
depth [km]
along profile [km]
(d)
Plane A
median geometry
real geometry
depth [km]
dip [°]
(e)
95% confidence interval
median geometry
real geometry
Plane B
depth [km]
along profile [km]
(f)
Plane B
median geometry
real geometry
depth [km]
dip [°]

Figure12.



(a)
not to scale
not to scale
normal-slip [m]
0
1
2
3
4
(b)
0 1 2 3 4 5
slip [m]
(c)
not to scale
not to scale
st. dev. [m]
0.2
0.4
0.6
0.8
1
1.2
1.4
(d)
displacement [m]
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
x-displacements
y-displacements
z-displacements
synthetic data
modeled data
residuals
0 50 100 150 200 250 300
data-location index
(e)
98.4 98.6 98.8 99 99.2 99.4
variance reduction [%]

Figure13.

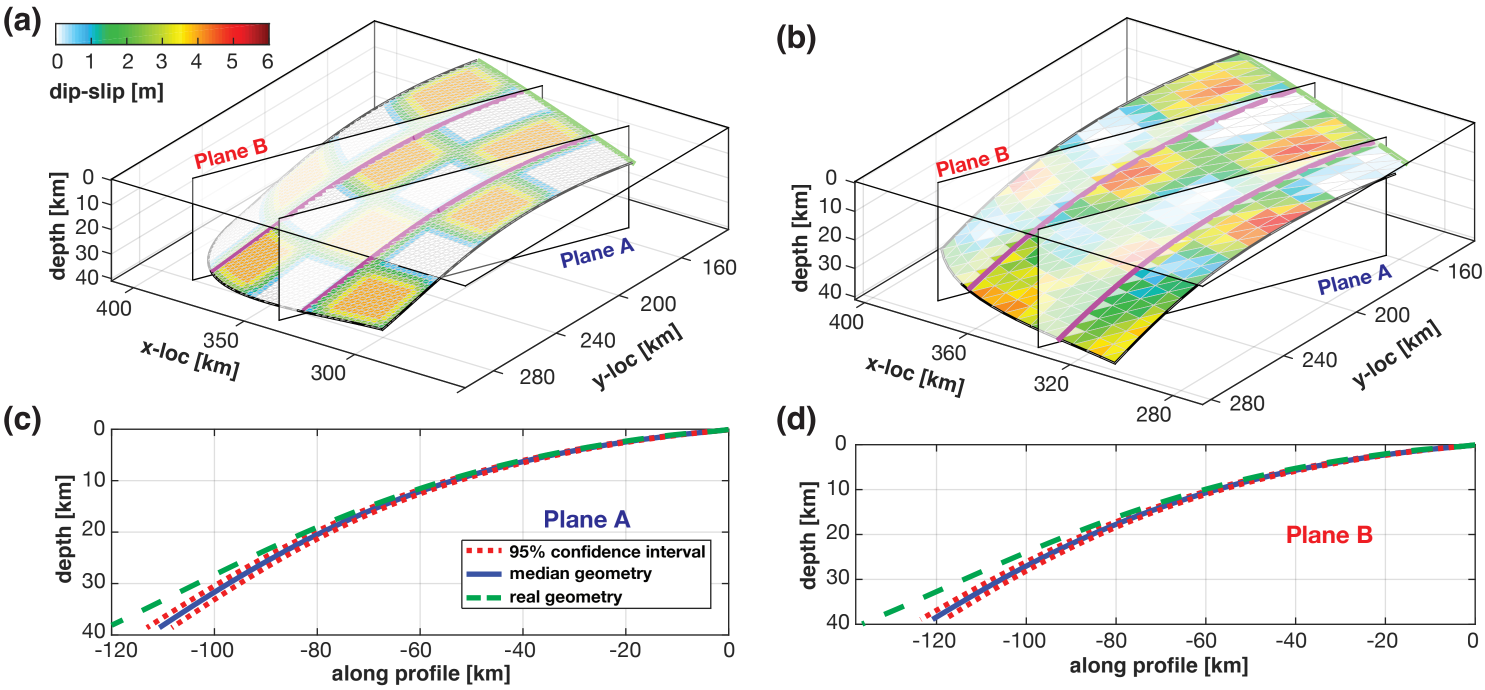
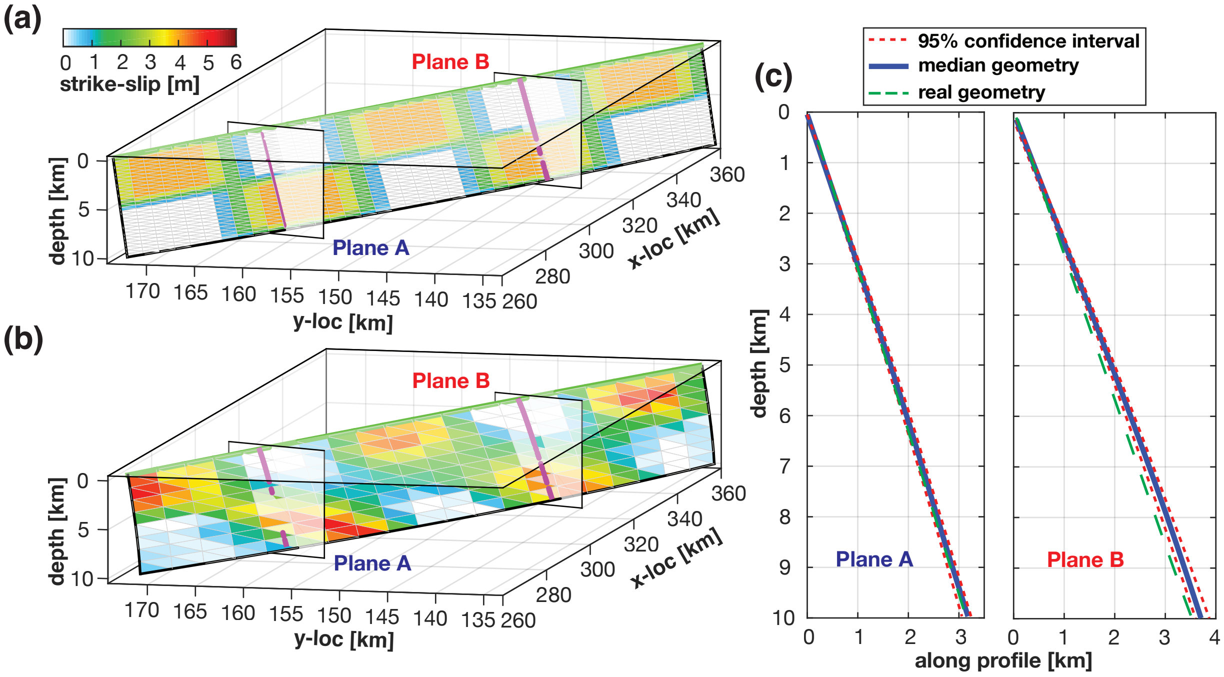

(a)
dip-slip [m]
0 1 2 3 4 5 6
Plane B
Plane A
depth [km]
x-loc [km]
y-loc [km]
(b)
Plane B
Plane A
depth [km]
x-loc [km]
y-loc [km]
(c)
Plane A
95% confidence interval
median geometry
real geometry
depth [km]
along profile [km]
(d)
Plane B
depth [km]
along profile [km]

Figure14.


(a)
0 1 2 3 4 5 6
strike-slip [m]
Plane B
Plane A
depth [km]
0
5
10
170 165 160 155 150 145 140 135
y-loc [km]
260 280 300 320 340 360
x-loc [km]
(b)
Plane B
Plane A
depth [km]
0
5
10
170 165 160 155 150 145 140 135
y-loc [km]
260 280 300 320 340 360
x-loc [km]
(c)
95% confidence interval
median geometry
real geometry
depth [km]
0 1 2 3 4 5 6 7 8 9 10
Plane A
Plane B
0 1 2 3
0 1 2 3 4
along profile [km]