An EPIC Tikhonov Regularization: Application to Quasi-Static Fault Slip Inversion

F. Ortega-culaciati, M. Simons, J. Ruiz, L. Rivera, N. Díaz-salazar

JGR Solid Earth

RESEARCH ARTICLE
10.1029/2020JB021141

An EPIC Tikhonov Regularization: Application to Quasi-Static Fault Slip Inversion

F. Ortega-Culaciati[1], M. Simons[2], J. Ruiz[1], L. Rivera[3], and N. Díaz-Salazar[1]

[1]Departamento de Geofísica, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile, [2]Seismological Laboratory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA, [3]Université de Strasbourg, CNRS, ITES UMR 7063, Strasbourg, France

Key Points:

- Common approaches to spatial regularization are susceptible to unstable biases in estimates of subsurface fault slip
- We introduce a novel spatially variable regularization scheme called Equal Posterior Information Condition (EPIC) Tikhonov
- Use of EPIC Tikhonov results in robust estimates of fault slip, in terms of stability, bias, and ease of uncertainty analysis

Supporting Information:

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

Correspondence to:

F. Ortega-Culaciati,
ortega.francisco@uchile.cl

Citation:

Ortega-Culaciati, F., Simons, M., Ruiz, J., Rivera, L., & Díaz-Salazar, N. (2021). An EPIC Tikhonov regularization: Application to quasi-static fault slip inversion. *Journal of Geophysical Research: Solid Earth*, *126*, e2020JB021141. https://doi.org/10.1029/2020JB021141

Received 8 OCT 2020
Accepted 2 JUN 2021

Abstract Imaging subsurface fault slip from surface observations is essential to improving our understanding of the physics of earthquakes and tsunamis. As the estimation of subsurface fault slip is inherently ill-posed, common inversion methods usually require a regularization term to counteract instabilities. Such regularization introduces biases in inferred slip estimates. Here, we discuss the effects that prior information, implied by a given regularization scheme, can have on fault slip estimates. We propose a novel Equal Posterior Information Condition (EPIC)—based Tikhonov regularization that generalizes the concept of prior information. The EPIC determines variances of prior information based on a chosen form of the structure of the posterior covariance matrix. In the context of subduction zone earthquakes, use of the EPIC counterbalances the spatial heterogeneity of observational constraints on fault slip, improving stability, quality and interpretability of fault slip estimates. We investigate the efficiency of the EPIC in the context of various synthetic fault slip distributions. We also demonstrate the methodology by inferring co-seismic slip from geodetic data for the 2011 (M_w 9.0) Tohoku-Oki, Japan, earthquake, obtaining robust slip estimates that are similar to those inferred using a unregularized fully Bayesian approach.

Plain Language Summary Inferring the distribution of relative motion across earthquake faults in Earth's subsurface using observations of motion at the surface is an essential component of improving our knowledge of the physics of earthquakes and tsunamis. However, surface observations only provide limited resolution and thus the estimation problem is non-unique and may be unstable in the presence of noise. Therefore, prior information or regularization is commonly used to control the solutions. We analyze the influence that different regularization schemes have on inferred images of fault motion. In addition, we propose a novel Equal Posterior Information Condition (EPIC) -based Tikhonov regularization, that selects a spatially variable strength of regularization to obtain images with spatially uniform precision. In synthetic tests, this new approach shows an improvement of the quality, interpretability and stability of fault motion estimates in the presence of different noise realizations. We investigate the quality of the EPIC-based regularization on simulated earthquake scenarios and demonstrate the methodology by inferring images of fault slip for the March 11, 2011, (M_w 9.0) great Tohoku-Oki, Japan, earthquake.

1. Introduction

Fault slip during all the stages of the seismic cycle can produce measurable deformation at the surface of the Earth. We are interested in the inverse problem of inferring subsurface fault slip given surface observations. For the quasi-static problem in a linear elastic media, the relationship between slip $\mathbf{m}$ on a discretized fault and crustal surface displacements $\mathbf{d}$ can be written as $\mathbf{d} = \mathbf{Gm}$, where the Green's functions $\mathbf{G}$ represents the response of the medium due to unitary slip on each element of a discretized fault surface (e.g., Segall & Harris, 1987).

Given a fine enough fault discretization, the prediction of any model inference is insensitive (within data uncertainties) to the high frequency component of the slip distribution on the fault surface, leading to an ill-posed discrete linear inverse problem (e.g., Hansen, 1998). Moreover, the ability to infer fault slip, typically constrained by onland observations in the subduction earthquake example, is variable across the subsurface fault, with the least constrained portions of the fault located further away from the observations (e.g.,

Lohman, 2004; Loveless & Meade, 2011; Ortega-Culaciati, 2013; Pritchard et al., 2002; Simons et al., 2002). Thus, any errors on either the measurements or the prediction of the approximate physical model, can map into spurious features in the slip distribution, leading to unstable and biased estimates.

There are two end-member approaches that deal with the inherent instability of the ill-posed inverse problem: An unregularized, fully Bayesian approach (e.g., Minson et al., 2013; Tarantola, 2005), and a more expedient but biased optimization approach using some form of regularized cost function minimization (e.g., Segall & Harris, 1987). In the Bayesian approach, the solution of the inverse problem is the posterior joint probability density function (pdf) of model parameters, which describes the likelihood of model parameters constrained by prior information and observations. In the optimization approach, the estimated model corresponds to the model that maximizes the posterior joint pdf of the regularized problem, and therefore depends on the type and amount of regularization imposed (e.g., Tarantola, 2005). Both approaches can be described using a Bayesian formulation of the inverse problem, being different in the specific assumptions on data and model parameter uncertainties and in the methodology used to solve the inverse problem (sampling or optimization methods).

An important advantage of the Bayesian approach over the classic optimization techniques is the ability to limit oneself to physically justifiable prior information (e.g., we do not expect to find normal slip on a subduction megathrust earthquake) without imposing other prior behavior (e.g., smoothness) beyond that imposed by the model parameterization itself (Minson et al., 2013). However, to avoid over-fitting, this approach requires a careful description of both observational and model prediction uncertainties (Duputel et al., 2014).

Optimization approaches use some form of regularization, not necessarily based on the physics of the problem, to avoid numerical instabilities and over-fitting of errors (in both the observations and forward model prediction). The optimization approach inherently requires one to determine the amount of regularization applied through the choice of model parameterization (e.g., Barnhart & Lohman, 2010; Pritchard et al., 2002) or regularization parameter—sometimes posed as a penalty of the L_1 or L_2 norm of the regularization or as a correlation length (e.g., Evans & Meade, 2012; Radiguet et al., 2011; Segall & Harris, 1987; Tarantola & Valette, 1982). The range of values of the regularization parameter defines a family of solutions of the inverse problem. Typically, selection of the best regularization parameter is performed by using methods that search for a balance between misfit and regularization norms, such as an L-curve (e.g., Hansen & O'Leary, 1993) or U-curve (e.g., Krawczyk-Stańdo & Rudnicki, 2007; Wang et al., 2018); that search for a balance between over-fitting and over-smoothing based on data and model resolution, such as jRi (Barnhart & Lohman, 2010); that minimize the mean squared error (MSE) of estimated model parameters (e.g., Cai et al., 2004; Wang & Gu, 2020); or by using some criteria based on a balance between how well observations are explained by model predictions and to what level the observations are independently resolved, such as Cross-Validation (e.g., Aster et al., 2013; Craven & Wahba, 1979; Wahba, 1990), Akaike Bayesian Information Criterion (e.g., Akaike, 1980), or some other form of model class selection (e.g., Muto & Beck, 2008).

Implementation of the unregularized Bayesian approach can be computationally intensive, potentially requiring the evaluation of a large ensemble of forward models (Duputel et al., 2014; Minson, 2010; Minson et al., 2013). Thus, while we generally prefer to advance Bayesian methodologies that rely only on physically justifiable prior information, we continue to pursue implementations of the least squares optimization approach that can be used to quickly solve large slip inversion problems using minimal computational resources. The optimization approach becomes a practical choice at times when rapid solutions are necessary, or for problems with a very large number of model parameters that defy sampling methods.

2. Methodology

Including a regularization term in an inverse problem involves introducing prior information or beliefs on what model parameters should be. For instance, smoothing constraints tend to discard rough slip models as candidate solutions of the inverse problem. As particular examples, several studies inferring interseismic coupling in the Japan Trench megathrust from onshore GNSS data (e.g., Hashimoto et al., 2009; Loveless & Meade, 2010, 2011; Mazzotti et al., 2000; Suwa et al., 2006) obtain coupling models that differ substantially due to different choices of prior assumptions or regularization methods. Thus, one must evaluate how the

chosen regularization scheme biases the solution and how well a particular approach deals with the inherent instabilities of the ill-posed inverse problem.

We frame our discussion in the context of the 2011 (M_w 9.0) Tohoku-Oki, Japan, earthquake (e.g., Simons et al., 2011, see Text S4). Here, we analyze slip estimates constrained by three component displacement data generated from various synthetic slip models evaluated at the locations of 738 GNSS stations of the Japanese GEONET network located in northern Honshu, Japan (Minson et al., 2014; Sagiya et al., 2000). We choose this scenario due to the particularly large distance between the onland observations and the shallowest portions of the megathrust fault in this region, which results in a large spatial variability of the resolving power that the observations have on inferred fault slip (Loveless & Meade, 2011). Here, we analyze the stability and biases of inferred slip estimates when using prior information implied by commonly used regularization schemes as well as for a novel regularization scheme, we call the Equal Posterior Information Condition (EPIC) Tikhonov regularization.

In order to probe the behavior of all the regularization schemes analyzed in this work, we design various synthetic slip models, subsequently using them to generate noisy synthetic datasets. From the work of Lévêque et al. (1993) for seismic tomography inversions, we understand that the quality of the assessment on the recovery of the "True slip model" depends on the shape and size of the synthetic model used for such purpose. Thus, we perform our analysis for a set of synthetic checkerboard slip models at different scales and a set of synthetic slip models resembling various earthquake scenarios with an elliptical shape, with moment magnitudes M_w 8.0, M_w 8.5, M_w 9.0, with spatial extent based on scaling laws (e.g., Stirling et al., 2013), and located at different positions on the subduction megathrust. We construct a 3D triangulated fault surface based on the geometry of the subduction megathrust from Simons et al. (2011). We compute synthetic crustal displacements using a layered elastic medium (Herrmann, 2013). The elastic structure is taken from Simons et al. (2011), which is based on the 3D tomography from NIED (Japan National Research Institute of Earth Science and Disaster Prevention) and the work of Takahashi et al. (2004). We consider synthetic data with three realizations of random noise (Noise 1, 2, and 3) from the same unbiased normal distribution. The standard deviation of the random noise is set to 5 mm for the horizontal components and 10 mm for the vertical component of the synthetic GNSS data. For each regularization scheme, we infer slip using the generalized cross validation (GCV) technique to select the optimum regularization parameter by minimizing the GCV loss function (Craven & Wahba, 1979). We refer the reader to Text S4 for further details.

We demonstrate the vie of EPIC Tikhonov regularization in a real example, by estimating fault slip during the 2011 (M_w 9.0) Tohoku-Oki, Japan, earthquake as constrained by co-seismic offsets measured by the GEONET network. When we consider the actual Tohoku-Oki earthquake observations, we also include positivity constraints on fault slip. We parameterize slip in two directions, one parallel to the slip vector defined by the Tohoku-Oki earthquake focal mechanism (rake parallel), and a perpendicular one (see Minson et al., 2013, Text S4). However, for the synthetic tests, we only allow slip in the rake parallel direction, in order to avoid needing to interpret complexity that arises from correlations between slip in the two orthogonal directions.

We acknowledge that the defined forward model is an approximation of the true physics underlying the slip inversion problem, for instance, due to uncertain fault geometry and elastic structure. Furthermore, techniques have been proposed for dealing with such epistemic uncertainties (e.g., Duputel et al., 2014; Minson et al., 2013; Ragon et al., 2018; Yagi & Fukahata, 2011). The inclusion of epistemic uncertainties increases the variances of the misfit between each observation and its corresponding forward model prediction, improving the robustness of slip estimates, as it avoids overfitting of errors (e.g., Duputel et al., 2014, 2015; Minson et al., 2013, 2014; Yagi & Fukahata, 2011). At the same time, epistemic uncertainties contributes to the covariances (correlations) between the misfit of different modeled data (Duputel et al., 2014; Ragon et al., 2018; Yagi & Fukahata, 2011). Such covariances are often neglected in slip inversions. In our work, when using real Tohoku-Oki earthquake data, we include epistemic uncertainties using the approach of Minson et al. (2014) (see Text S4). While we prefer approaches that include model prediction correlations (e.g., Duputel et al., 2014), we use the uncorrelated model prediction uncertainties of Minson et al. (2014) for consistency, as we compare our results with those of Minson et al. (2014). For our comparative analysis

of regularization schemes, we do not include epistemic uncertainties, as we constrain slip models with synthetic datasets, thus considering an exact forward model.

2.1. On the Estimation of Quasi-Static Fault Slip

Following Tarantola and Valette (1982), the general inverse problem can be formulated using a Bayesian framework as

$$f^{post}(\mathbf{m}) = \frac{1}{\nu}\mathcal{L}(\mathbf{m} \mid \mathbf{d}) f^{prior}(\mathbf{m}) \tag{1}$$

where $\mathbf{m}$ is the model parameter vector representing slip on the fault, $f^{prior}(\mathbf{m})$ is the pdf describing prior information on model parameters, $\mathcal{L}(\mathbf{m} \mid \mathbf{d})$ is the likelihood of the predicted data (surface displacements) given a set of model parameters $\mathbf{m}$ and observed data $\mathbf{d}$, $f^{post}(\mathbf{m})$ is the posterior pdf on model parameters, and ν is a normalization constant. As much as possible, we follow the notation of Tarantola (2005) throughout our development.

Typically, the uncertainties on both observations and model predictions are assumed to be unbiased, following a multivariate normal distribution with zero mean and covariance matrices $\mathbf{C_d}$ and $\mathbf{C_p}$, respectively (e.g., Duputel et al., 2014; Minson et al., 2013; Tarantola, 2005). Thus, given the linear relationship between data and model parameters, $\mathbf{d = Gm}$, the likelihood function can be written as,

$$\mathcal{L}(\mathbf{m} \mid \mathbf{d^{obs}}) = const. \cdot \exp\left(-\frac{1}{2}\left(\mathbf{Gm} - \mathbf{d^{obs}}\right)^{\top} \mathbf{C}_{\chi}^{-1}\left(\mathbf{Gm} - \mathbf{d^{obs}}\right)\right) \tag{2}$$

where $\mathbf{d^{obs}}$ is a vector containing the observations, $\mathbf{Gm}$ the model prediction, $\mathbf{C}_{\chi} = \mathbf{C_d} + \mathbf{C_p}$ is the misfit covariance matrix, and *const*. is an unknown normalization constant (e.g., Tarantola, 2005). We note that the assumption that model prediction ($\mathbf{C_p}$) is unbiased and follows a Gaussian distribution is clearly an approximation (e.g., Duputel et al., 2014).

Among the infinite ways to prescribe prior information on model parameters, $\mathbf{m}$, one approach is to rely on the beliefs that we have in some *ad hoc* and uncertain quantity, $\mathbf{h}$, related to the model parameters, $\mathbf{m}$, through some mathematical relationship, $\mathbf{h} = \mathcal{F}(\mathbf{m})$. For instance, the quantity $\mathbf{h}$ may represent a spatial derivative of slip values (e.g., slip roughness $\mathcal{F}(\mathbf{m}) = \nabla^2\mathbf{m}$), or actual values of slip (i.e., $\mathcal{F}(\mathbf{m}) = \mathbf{m}$). In the context of a linear inversion problem, we are interested in the case of a linear relationship $\mathbf{h = Hm}$. Since $\mathbf{h}$ is an uncertain quantity, we consider a stochastic model $\mathbf{h} = \mathbf{h^o} + \epsilon_{\mathbf{h}}$, where $\mathbf{h^o}$ is some prescribed deterministic *a priori* value (i.e., a particular realization of $\mathbf{h}$) and the stochastic variable $\epsilon_{\mathbf{h}}$ describes the uncertainties on the determination of $\mathbf{h^o}$. A wide range of prior pdf's can be used to describe $\epsilon_{\mathbf{h}}$. As a practical choice, uncertainties on $\mathbf{h}$ can be modeled as an unbiased normal distribution with covariance matrix $\mathbf{C_h}$, thus prior information on $\mathbf{h}$ is represented by a normal distribution $\mathcal{N}(\mathbf{h^o}, \mathbf{C_h})$, with pdf

$$\rho_h(\mathbf{h}) = const. \cdot \exp\left(-\frac{1}{2}\left(\mathbf{h} - \mathbf{h^o}\right)^{\top} \mathbf{C_h^{-1}}\left(\mathbf{h} - \mathbf{h^o}\right)\right) \tag{3}$$

where *const*. is a normalization constant. We then construct the prior information on model parameters as the likelihood that the model parameters are consistent with the prior value of the quantity $\mathbf{h = Hm}$,

$$f^{prior}(\mathbf{m}) = \rho_h\left(\mathbf{h(m)}\right) = const. \cdot \exp\left(-\frac{1}{2}\left(\mathbf{Hm} - \mathbf{h^o}\right)^{\top} \mathbf{C_h^{-1}}\left(\mathbf{Hm} - \mathbf{h^o}\right)\right) \tag{4}$$

where *const*. is a normalization constant. In this sense, regularization or prior information is not treated differently from observational constraints. The only key difference lies in that observations typically come from experimental measurements of physical phenomena being studied, while prior information or regu-

larization represents a belief on how model parameters should behave. Also, following Tarantola (2005), prior information must be obtained independently of observational constraints.

Despite the assumptions of Gaussian uncertainties for observations, a linear forward model, and linear prior information, the inverse problem in Equation 1 defines an infinite ensemble of posterior models. From such an ensemble, we are interested in a particular model that maximizes the posterior pdf $f^{post}(\mathbf{m})$, which is also the solution of the general linear least squares problem,

$$\min_{\mathbf{m}} \quad \underbrace{(\mathbf{Gm}-\mathbf{d}^{\mathbf{obs}})^{\top}\mathbf{C}_{\chi}^{-1}(\mathbf{Gm}-\mathbf{d}^{\mathbf{obs}})}_{\phi_E(\mathbf{m})} + \underbrace{(\mathbf{Hm}-\mathbf{h}^{\mathbf{o}})^{\top}\mathbf{C}_{\mathbf{h}}^{-1}(\mathbf{Hm}-\mathbf{h}^{\mathbf{o}})}_{\phi_R(\mathbf{m})} \tag{5}$$

where the cost function is composed by $\phi_E(\mathbf{m})$, a measure of goodness of a particular model to predict the observations, and $\phi_R(\mathbf{m})$ a measure of goodness of such model to be consistent with the chosen prior information implied by the regularization of the inverse problem. Equation 5 can be written in a more familiar form,

$$\min_{\mathbf{m}} \quad \underbrace{\| \mathbf{W}_{\chi}(\mathbf{Gm}-\mathbf{d}^{\mathbf{obs}}) \|_2^2}_{\phi_E(\mathbf{m})} + \underbrace{\| \mathbf{W}_{\mathbf{h}}(\mathbf{Hm}-\mathbf{h}^{\mathbf{o}}) \|_2^2}_{\phi_R(\mathbf{m})} \tag{6}$$

where the misfit and regularization weights follow $\mathbf{W}_{\chi}^{\top}\mathbf{W}_{\chi} = \mathbf{C}_{\chi}^{-1}$ and $\mathbf{W}_{\mathbf{h}}^{\top}\mathbf{W}_{\mathbf{h}} = \mathbf{C}_{\mathbf{h}}^{-1}$.

The solution of the general linear least squares problem (Equation 5) follows a normal distribution with posterior mean (and maximum likelihood) model

$$\tilde{\mathbf{m}} = \left(\mathbf{G}^{\top}\mathbf{C}_{\chi}^{-1}\mathbf{G} + \mathbf{H}^{\top}\mathbf{C}_{\mathbf{h}}^{-1}\mathbf{H}\right)^{-1}\left(\mathbf{G}^{\top}\mathbf{C}_{\chi}^{-1}\mathbf{d}^{\mathbf{obs}} + \mathbf{H}^{\top}\mathbf{C}_{\mathbf{h}}^{-1}\mathbf{h}^{\mathbf{o}}\right) \tag{7}$$

and posterior covariance matrix

$$\tilde{\mathbf{C}}_{\mathbf{m}} = \left(\mathbf{G}^{\top}\mathbf{C}_{\chi}^{-1}\mathbf{G} + \mathbf{H}^{\top}\mathbf{C}_{\mathbf{h}}^{-1}\mathbf{H}\right)^{-1} = \left(\mathbf{P} + \mathbf{H}^{\top}\mathbf{C}_{\mathbf{h}}^{-1}\mathbf{H}\right)^{-1} \tag{8}$$

(e.g., see Tarantola, 2005), where $\mathbf{P} = \mathbf{G}^{\top}\mathbf{C}_{\chi}^{-1}\mathbf{G}$ is the precision matrix of the unregularized problem, or equivalently, the mapping of the misfit covariance matrix ($\mathbf{C}_{\chi}$) into the model parameter space (see Text S1).

We quantify uncertainty in the inferred fault slip by analyzing the variances and spatial correlations of slip estimates directly from the posterior covariance matrix $\tilde{\mathbf{C}}_{\mathbf{m}}$. Posterior standard deviations provide information on the expected amplitude of slip perturbations given $\mathbf{C_d}$ and $\mathbf{C_p}$, and can be used to define confidence intervals on the value of estimated model parameters. Off diagonal elements of $\tilde{\mathbf{C}}_{\mathbf{m}}$ provide information on the pairwise covariance of estimated model parameters. To facilitate the interpretation of the covariance, we consider the Pearson correlation coefficient,

$$\varrho_{ij} = \frac{\left[\tilde{\mathbf{C}}_{\mathbf{m}}\right]_{ij}}{\sqrt{\left[\tilde{\mathbf{C}}_{\mathbf{m}}\right]_{ii}\left[\tilde{\mathbf{C}}_{\mathbf{m}}\right]_{jj}}} \tag{9}$$

which has extreme values of -1 and 1 for fully negatively and positively correlated model parameters, respectively. A null Pearson correlation indicates no correlation or linear dependency between model parameters. In that sense, a set of estimated model parameters that are positively or negatively correlated will tend to be perturbed from the estimated (mean) model, in the same direction due to uncertainties in observations, model prediction and prior information, with perturbation amplitudes controlled by their posterior marginal standard deviations.

When model parameters are defined from a spatial discretization of a continuous scalar field (e.g., a component of slip on a fault plane), a measure of correlation between model parameters can be used to estimate a

spatial correlation length (e.g., Mai & Beroza, 2002; Tarantola & Valette, 1982). The form of spatial correlation used in this work is given by the exponentially decaying function

$$\varrho_{ij} = e^{-\left(\frac{d_{ij}}{\lambda_i}\right)^2} \tag{10}$$

where d_{ij} is the distance between the discretized elements of i-th and j-th model parameter and λ_i is the spatial correlation length for the i-th discretized element (e.g., Radiguet et al., 2011; Tarantola, 2005; see Text S5). We interpret the calculated correlation lengths as spatially averaged uncertainties, in the sense that at any fault element, estimated slip is correlated to some degree with the ones at neighboring fault elements. Thus, λ_i serves as a proxy for a measure of spatial resolution of slip estimates. In this regard, we characterize model resolution based on posterior uncertainties, following Tarantola and Valette (1982).

2.2. Common Regularization Schemes

Commonly used prior information consist on prescribing some, typically null, values for slip or to search for spatially smooth slip distributions (e.g., Harris & Segall, 1987; Radiguet et al., 2011; Segall & Harris, 1987). For example, Radiguet et al. (2011) uses the approach given by Tarantola and Valette (1982), in which they impose a spatial correlation length on slip, as well as a bias toward a prior slip model $\mathbf{m^o}$. The implementation of the correlation between model parameters is achieved by designing a prior covariance matrix $\mathbf{C_m}$ that models prior uncertainties on $\mathbf{m^o}$ as correlated in space, in which correlation decays with the distance between discretized elements of the fault surface. In terms of Equation 4, this form of regularization is equivalent to setting $\mathbf{H = I}$ (the identity matrix), $\mathbf{h^o = m^o}$ (typically arbitrarily set as $\mathbf{m^o = 0}$), and covariance matrix $\mathbf{C_h = C_m}(\lambda)$, where

$$\left[\mathbf{C_m}(\lambda)\right]_{ij} = \sigma_m^2 e^{-\frac{d(i,j)}{\lambda}}. \tag{11}$$

Here, $d(i, j)$ is the distance between subfaults i and j, σ_m^2 is the prior variance on slip values and λ is a correlation length, used as a regularization parameter. Note that in this case, correlation decays exponentially with distance, but the choice of such a functional form is arbitrary (e.g., see Radiguet et al., 2011; Tarantola, 2005). Here, variances of prior information are all equal to $\sigma_h^2 = \sigma_m^2$. This operator imposes a correlation that is stronger for model parameters that are more proximal in physical space, i.e., a smoothing constraint on slip. This operator also imposes a condition in which model parameters are constrained to be close to their prior value $\mathbf{m^o}$, thus biasing model estimates to such values. For brevity, we refer to this regularization scheme as $C_{\mathbf{m}}^{\lambda}$. Note that λ_i and λ from Equations 10 and 11 have a different meaning. While λ_i represents a vector of posterior correlation lengths on slip estimates, λ is a scalar quantity representing a prescribed prior correlation length on slip.

Another common form of prior information on slip is given by Tikhonov regularization (Tikhonov, 1963), in which the spatial derivatives of a given order of the model parameters (e.g., slip) are biased toward a null value (e.g., Harris & Segall, 1987; Segall & Harris, 1987). The inverse problem using Tikhonov regularization is solved through the damped least squares method,

$$\min_{\mathbf{m}} \| \mathbf{W}_\chi(\mathbf{Gm} - \mathbf{d^{obs}}) \|_2^2 + \varepsilon^2 \| \mathbf{Hm} \|_2^2. \tag{12}$$

where ε^2, known as the regularization or damping parameter, controls the trade-off between the observational and regularization terms, which represent the misfit of physical, $\mathbf{W}_\chi\mathbf{Gm} = \mathbf{W}_\chi\mathbf{d^{obs}}$, and regularization, $\mathbf{Hm = 0}$, equations, respectively. The prior information implied by Tikhonov regularization can be represented in terms of Equation 4 by prescribing $\mathbf{h^o = 0}$, and a choice of $\mathbf{H}$ proportional to a finite difference approximation of the zeroeth order ($\mathbf{H = I}$), first order ($\mathbf{H} = \nabla$) or second order ($\mathbf{H} = \nabla^2$) spatial differential operators. Zeroeth order Tikhonov (T0) regularization biases model estimates to null values ($\mathbf{m^o = 0}$) without imposing any smoothing constraint. First (T1) and second (T2) order Tikhonov bias model estimates to smooth models with null first and second order spatial derivatives, respectively, without bias-

ing model estimates toward a specific prior slip value. In addition, we set $\mathbf{W_h} = \varepsilon\mathbf{I}$ so problems (6) and (12) are equivalent, meaning that all the regularization equations in (6) have the same weight or importance ε. In terms of the normal distribution used to describe prior information on $\mathbf{h}$, the prior covariance matrix implied by Tikhonov regularization becomes $\mathbf{C_h} = \varepsilon^{-2}\mathbf{I}$. Such definition of prior information considers that all components of $\mathbf{h}$ have independent and identically distributed uncertainties, with variances that are all equal to the reciprocal of the regularization parameter ($\sigma_h^2 = \varepsilon^{-2}$).

All the aforementioned regularization schemes impose prior information with the same variance for all the regularized quantities. For instance, T2 regularization imposes the same amount of slip smoothing for all fault patches regardless of their locations. As observations typically do not constrain or resolve slip homogeneously along the fault surface (e.g., Loveless & Meade, 2011; ,Pritchard et al., 2002; Simons et al., 2002 see Text S2), we are motivated to develop a spatially variable regularization scheme.

Spatially variable regularization schemes have been proposed either by changing the size of the fault elements in order to obtain a quasi-uniform resolving power, requiring little additional regularization (Barnhart & Lohman, 2010; Pritchard et al., 2002; Simons et al., 2002, see Text S2) or by the use of second order Tikhonov regularization with a spatially variable strength (Lohman, 2004; Ortega-Culaciati, 2013; Wang et al., 2019, see Text S2).

The first approach imposes a correlation length by setting constant slip over variably sized fault patches, but depends on additional prior information (e.g., direction of slip) and on the specific algorithm used to determine the fault discretization. This approach also leads to model discretizations with large regions of uniform slip behavior and sharp discontinuities, potentially limiting the usability of the inferred slip distribution. As an aside, one could reduce this problem by modifying the approach to include smoothly varying basis functions instead of piece-wise constant (e.g., Yabuki & Matsu'ura, 1992), but we do not explore such an approach here.

For the second approach, Lohman (2004) defines an iterative method which relies on applying a spatially variable penalty parameter using a low-rank approximation of the model resolution matrix $\mathbf{R}$ (Backus & Gilbert, 1968) of the regularized inverse problem. However, the model resolution matrix $\mathbf{R}$ intrinsically assumes that observations can be related to a true model as $\mathbf{d^{obs}} = \mathbf{Gm^{true}}$, neglecting uncertainties in both observations and the forward model. Wang et al. (2019) proposes the use of a least squares inversion with a method to define spatially variable weighting to the Laplacian second order smoothing constraint based on slip estimates. However, while such regularization scheme achieves larger slip values—compared with the unweighted Laplacian—that are closer to true values, it defines an iterative nonlinear approach in which prior information standard deviations (reciprocal of Laplacian weights) are updated based on the data that constrains the slip model. As another option, Ortega-Culaciati (2013) defines a sensitivity modulated second order Tikhonov (ST2) regularization, that empirically modifies variances of prior information implied by T2 regularization. Here, the variances of prior information on the second spatial derivative of slip at the i-th fault patch, are set proportional to the sensitivity $s_i = \dfrac{P_{ii}}{\max\limits_k P_{kk}}$, where $\mathbf{P} = \mathbf{G}^\top\mathbf{C}_\chi^{-1}\mathbf{G}$ is the precision matrix of the unregularized inverse problem (see Text S2). Figure 1a shows the square root of the sensitivity of fault slip for the model setup defined for the synthetic tests analyzed in this study. In terms of Equations 4 and 5, prior information implied by ST2 can be represented by prescribing $\mathbf{h^o} = \mathbf{0}$, $\mathbf{H} = \nabla^\mathbf{2}$ and $\mathbf{C_h} = \varepsilon^{-2}\mathbf{S}$, where $S_{ij} = \delta_{ij}s_i$ is the sensitivity matrix, and produces a spatially variable bias toward smooth slip models. Here, ε^{-2} has the same meaning as in standard Tikhonov regularization.

2.3. An EPIC Tikhonov Regularization

We aim to design a spatially variable regularization scheme that accounts for the heterogeneous influence of the observations on slip estimates. Here, we use an arbitrarily fine discretized fault surface (only limited by computer resources) and adopt a regularization that allows for a smoothly varying correlation length of the inferred values of slip, without using any prior information about the actual values of slip but just on their derivatives.

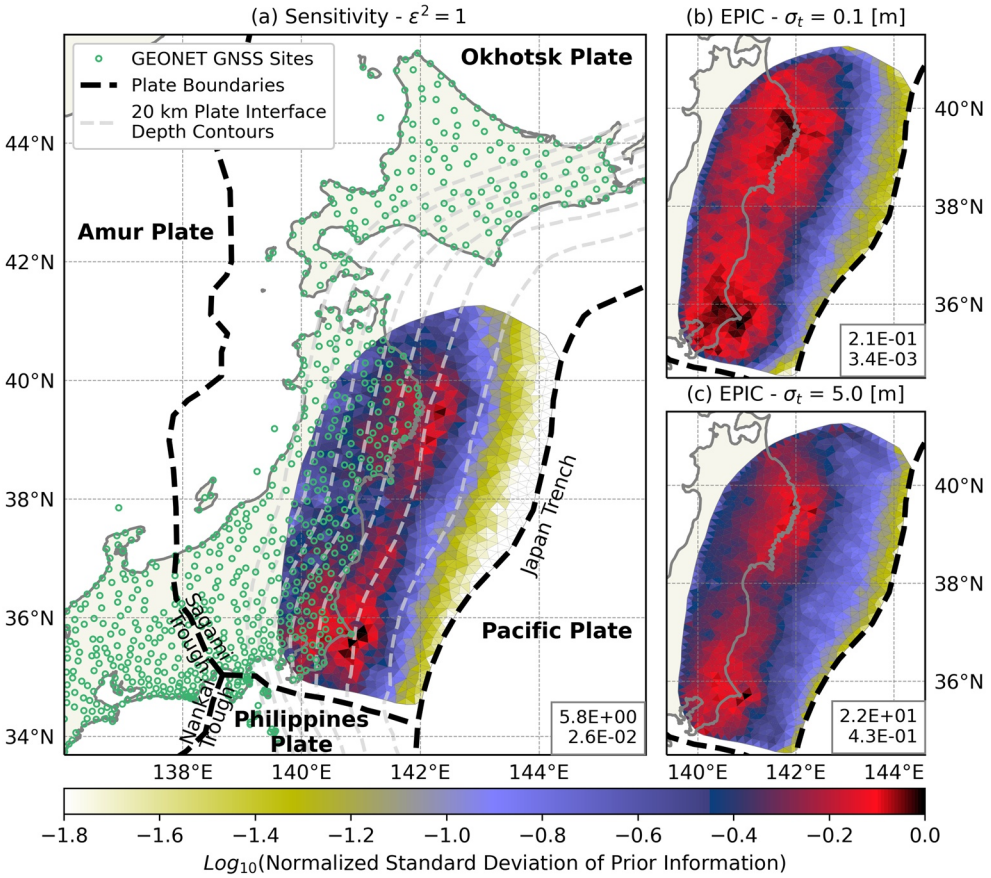

Figure 1. Comparison of standard deviations of prior information, that is, diagonal elements of $\sqrt{\mathbf{C_h}}$, for (a) Sensitivity (ST2) regularization with $\varepsilon^2 = 1$ (Ortega-Culaciati, 2013), and for EPIC (ET2) regularization (this work) considering two cases of target posterior standard deviations (σ_t) for slip estimates, (b and c). For each case, the color bar shows standard deviations normalized by their maximum value. Lower-right insets show the unnormalized minimum and maximum standard deviations of prior information in units of $\mathbf{h} = \nabla^2\mathbf{m}$. Panel (a) also displays major tectonic features within the study area and location of GEONET GNSS sites used in this work.

We propose to define an unknown diagonal covariance matrix $\mathbf{C_h}$ for prior information. We then ask what would be the values of such variances that in turn produce the same amount of information gain for every model parameter. Thus, we define an EPIC on model parameters, $\mathbf{m}$, where information gain accounts for both observations and prior knowledge, with respect to a homogeneous state of information (see Tarantola, 2005, Section 1.2.4). The state of homogeneous information is used to represent lack of knowledge about the slip distribution. Here, we do not use prior information as the reference state of information. Instead, we use prior information as an auxiliary source of knowledge or beliefs on specific characteristics of model parameters, and explore the impact of equalizing the amount of information gained on slip. As demonstrated later, equalizing the information gain results in uniformly distributed variances for estimated slip across the fault. Thus, uncertainties in observations or model predictions, will induce spurious slip that will have no preferred location at the fault surface, uniformly stabilizing the solution of the inverse problem.

The concept of differential entropy is used as a proxy to quantify the information gained on slip during the inversion process, with respect to the state of homogeneous information (see Text S3, Cover & Thomas, 1991). From Equations 7 and 8, the posterior marginal distribution for the i-th model parameter (m_i) is a normal distribution $\mathcal{N}(\tilde{m}_i, \tilde{\sigma}^2_{m_i})$, where $\tilde{m}_i$ is the i-th estimated model parameter value and $\tilde{\sigma}^2_{m_i}$ is the i-th diagonal element of $\tilde{\mathbf{C}}_\mathbf{m}$. The information gained on slip parameter m_i from data and prior information, can be quantified by the differential entropy $h(m_i)$ of the posterior marginal distribution for m_i as

$$h(m_i) = \frac{1}{2}\ln(2\pi e\tilde{\sigma}^2_{m_i}) \tag{13}$$

which depends only on the posterior marginal variance $\tilde{\sigma}^2_{m_i}$ (see Text S3). Therefore, EPIC is equivalent to imposing that all posterior marginal variances are the same, which can be formulated as:

$$\left[\tilde{\mathbf{C}}_{\mathbf{m}}\right]_{ii} = \left[\left(\mathbf{G}^\top\mathbf{C}_{\mathbf{x}}^{-1}\mathbf{G} + \mathbf{H}^\top\mathbf{C}_{\mathbf{h}}^{-1}\mathbf{H}\right)^{-1}\right]_{ii} = \sigma_t^2 \tag{14}$$

where the target posterior variance σ_t^2 of the estimated (slip) model is used as a regularization parameter. As is the case for higher order Tikhonov regularization, $\mathbf{H}$ already induces local correlations of slip. We define $\mathbf{C_h}$ as a diagonal matrix, whose terms control the strength of the induced correlations. Then, for a given value of σ_t, we compute $\mathbf{C_h}$ by solving the nonlinear system of Equation 14 using a nonlinear least squares method based on a Trust Region Reflective optimization algorithm (Coleman & Li, 1996). We emphasize that, as in most regularization schemes mentioned in Section 2.2, we define $\mathbf{C_h}$ in a parametric way, prescribing prior information on $\mathbf{h}$ with unknown covariance matrix $\mathbf{C_h}$. For EPIC Tikhonov, $\mathbf{C_h}$ is defined as an unknown diagonal matrix. Therefore, EPIC must be understood not as a condition defining prior information, but as complement of a classical model class selection method, selecting values of the diagonal of $\mathbf{C_h}$ that in turn produces a desired class of models with uniformly distributed posterior variances. We refer the reader to Text S3 for further details and implementation.

When using the EPIC for first or second order Tikhonov regularization, we refer to those as ET1 and ET2 regularization schemes, respectively. Figures 1b and 1c show the model setup defined for the synthetic tests analyzed in this study and the standard deviations of prior information ($\sqrt{\mathbf{C_h}}$) computed using ET2 regularization for two cases of target posterior variances σ_t^2 on slip. The EPIC induces prior information that compensates for the variability of the influence that observations have on slip estimates across the fault. Here, we note that the regions less constrained by the data correspond to the ones with lower sensitivity, or lower marginal variances of the data misfit projected onto model parameter space (see Text S1). Thus, model estimates at those regions are the most prone to produce overfitting, driven by observational and model prediction errors.

Although ST2, ET1, and ET2 define spatially variable regularization through variable variances of prior information (i.e., in $\mathbf{C_h}$, see Figure 1), how this variability is implemented arises as a key difference. On one hand, ST2 regularization defines $\mathbf{C_h}$ based on the sensitivity of fault slip and the shape of $\mathbf{C_h}$ does not depend on the amount of regularization applied (i.e., the value of ε^2) or the structure of the regularization operator $\mathbf{H}$. On the other hand, EPIC estimates $\mathbf{C_h}$ for each given value of the regularization parameter σ_t, accounting for the structure on $\mathbf{H}$ (Figure 1, Text S4). Moreover, ST2 uses only the diagonal elements of the precision matrix of the unregularized problem ($\mathbf{P}$), while EPIC uses the whole matrix $\mathbf{P}$, fully exploiting the properties of the physical problem being solved. Such difference is key to increasing stability and to reducing bias. Table 1 summarizes the regularization schemes commonly used for least squares linear slip inversion, as well as the regularization scheme proposed in this work.

3. Results for Inversions Using Synthetic Data

In our comparative analysis of the impact of the regularization schemes on slip estimates, we performed inversions for all combinations of regularization schemes, synthetic slip models and realizations of data noise analyzed in this work. In all cases, the regularization parameter value is determined using GCV. Figures 2 and 3 show a selection of the results for representative inversion cases, recovering synthetic checkerboard and elliptic slip distributions, respectively. In addition, Figure 2 shows prior information and posterior slip standard deviations along with posterior slip correlation lengths. Figure 3 shows a comparison of synthetic and estimated slip along a representative trench perpendicular profile, for the case of an elliptic slip distribution.

The regularization schemes explored in this work can be judged by the bias they induce on slip estimates and by how well they deal with the inherent instability of the inverse problem. We define instability as the

Table 1
Regularization Schemes for Least Squares Slip Inversion Analyzed in This Work

Notation	Examples	Prior information			Impact on
		$\mathbf{H}$	$\mathbf{C_h}$	$\mathbf{h^o}$	$\mathbf{m - m^o}$
$C_{\mathbf{m}}^{\lambda}$	Tarantola (2005), Radiguet et al. (2011)[a]	$\mathbf{I}$	$\mathbf{C_m}(\lambda)$	$\mathbf{m^o}$	Minimizes amplitude Induces spatial averaging
T0	Harris and Segall (1987)[a,b]	$\mathbf{I}$	$\varepsilon^{-2}\mathbf{I}$	$\mathbf{m^o}$	Minimizes amplitude
T1	No uses found[b,c]	∇	$\varepsilon^{-2}\mathbf{I}$	$\nabla\mathbf{m^o}$	Minimizes spatial gradient Induces spatial averaging
T2	Segall and Harris (1987)[b,d]	∇^2	$\varepsilon^{-2}\mathbf{I}$	$\nabla^2\mathbf{m^o}$	Minimizes roughness Induces spatial averaging
ST2	Ortega-Culaciati (2013)[d]	∇^2	$\varepsilon^{-2}\mathbf{S}$	$\nabla^2\mathbf{m^o}$	Minimizes roughness Induces spatial averaging
ET1	This work[c,e]	∇	$\mathbf{C_h}(\sigma_t^2)$	$\nabla\mathbf{m^o}$	Minimizes spatial gradient Induces spatial averaging
ET2	This work[d,e]	∇^2	$\mathbf{C_h}(\sigma_t^2)$	$\nabla^2\mathbf{m^o}$	Minimizes roughness Induces spatial averaging

[a]Imposes $\mathbf{m^o} = \mathbf{0}$. [b]Corresponds to Tikhonov regularization when $\mathbf{h^o} = \mathbf{0}$, that is, $\mathbf{m^o} = \mathbf{0}$, $\nabla\mathbf{m^o} = \mathbf{0}$ or $\nabla^2\mathbf{m^o} = \mathbf{0}$, respectively. [c]Imposes $\nabla\mathbf{m^o} = \mathbf{0}$. [d]Imposes $\nabla^2\mathbf{m^o} = \mathbf{0}$. [e]Uses EPIC to estimate $\mathbf{C_h}$.

extent to which changes in the different realizations of noise impact estimates of slip. Results show that those regularization schemes that introduce prior information by seeking slip models that are close to an arbitrarily prescribed prior model ($\mathbf{m^o} = \mathbf{0}$ in this case), that is, T0 and $C_{\mathbf{m}}^{\lambda}$, produce relatively stable slip estimates for all the synthetic study cases. Nevertheless, while the true slip distribution is relatively well recovered for portions of the fault with high sensitivity, the inferred distribution is highly biased toward *a priori* values of slip ($\mathbf{m^o} = \mathbf{0}$ in this case) in regions with lower sensitivity (see Figures 2, S6, and S7). Moreover, when elliptical slip models are analyzed, T0 and $C_{\mathbf{m}}^{\lambda}$ regularization schemes tend to shift the peak location of the estimated slip distribution to regions closer to the location of the observations. These schemes also recover slightly smaller moment magnitudes when the true slip distribution is located far from the data (see Figures 3, S13, S14, S20, S21, S27, and S28).

Higher order Tikhonov regularization (T1 and T2) results in stable slip recovery of checkerboard models at fault regions with high sensitivity (see Figures 1a and 2). A tendency to smoother slip and larger instabilities occurs at less constrained regions of the fault (low sensitivity), with instabilities being almost negligible for T1 and very noticeable for T2 regularization (see Figures 2, S8, S9, and S34). For near coast earthquake scenarios, T1 and T2 regularization recover the true model but with noticeable instabilities at the near-trench portions of the fault, including large back-slip that could be ameliorated using positivity constraints. For earthquake scenarios with fault slip near to the trench, T1 and T2 induce slip estimates that tend to be smeared away from observations, setting peak slip toward the trench (see Figures 3, S35–S37).

ST2 regularization excels at recovering stable slip estimates for all checkerboard cases of synthetic slip (Figures 2 and S10). For all earthquake scenarios with true slip located mainly offshore, the peak of recovered slip tends to be located near the coast, slip estimates located in regions of low sensitivity (near trench) are over-smoothed, and slip instabilities appear at fault regions with higher sensitivity (see Figures 3, S17, S24, and S31).

The EPIC approach results in stable slip estimates for all checkerboard slip scenarios, with a smoothing that varies in a similar manner as the fault slip sensitivity to the observations. A stronger bias occurs in regions with lower slip sensitivity, while the true model is well-recovered in regions with higher sensitivity (see Figures 1a, 2, S11 and S12). The same bias is observed when inverting for synthetic earthquake scenarios, whereby estimated slip tends to get more filtered, with smaller peak amplitude, when the synthetic true elliptic slip model is located closer to the trench, further away from observations (see Figures 3, S18, S19, S25, S26, S32, and S33). Estimates of slip using ET2 are generally more stable than ones using ET1.

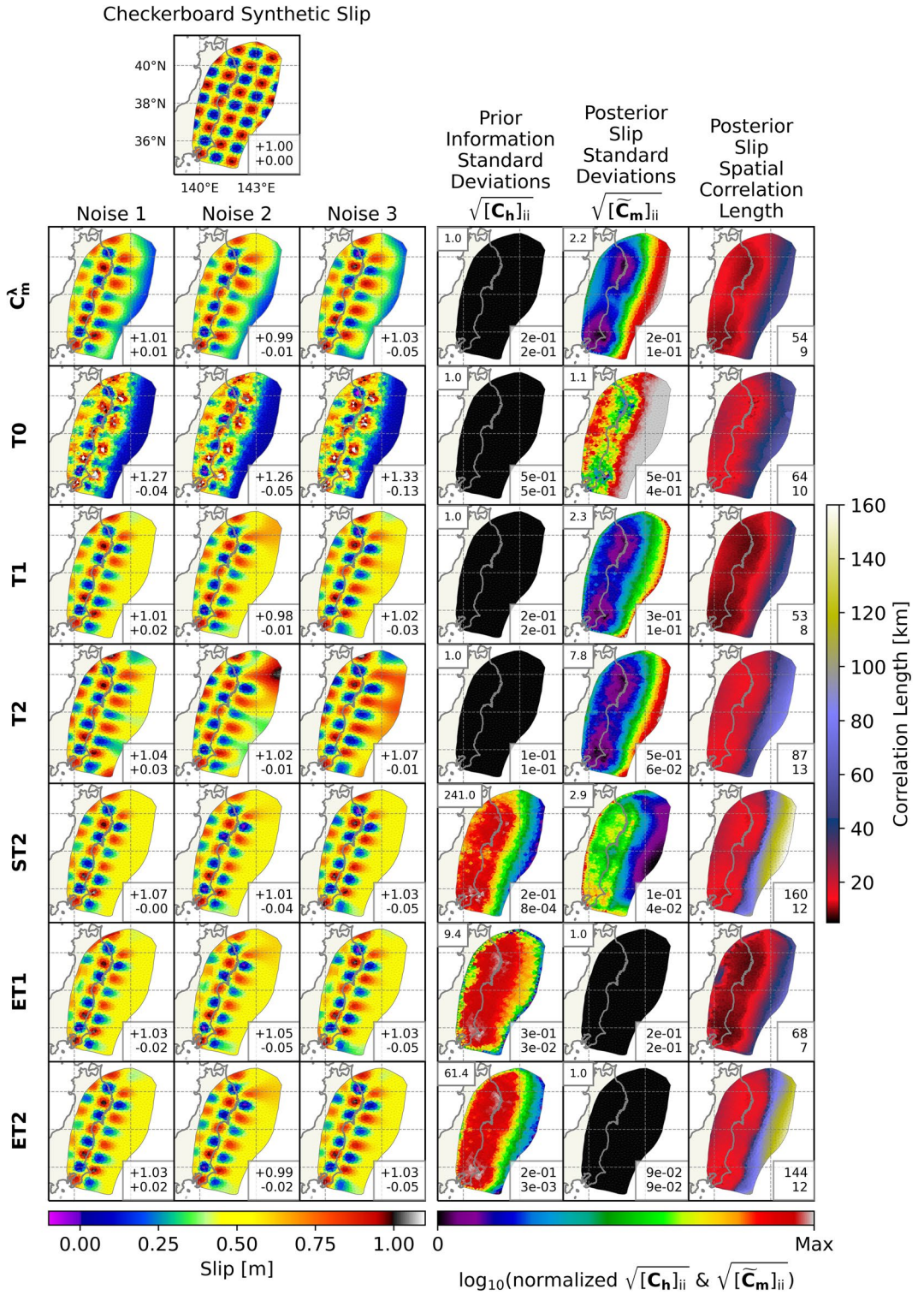

Figure 2. Slip estimates constrained by synthetic observations, generated from a prescribed true checkerboard (M_w 8.5) slip model, with added Noise 1, 2, and 3. Best model estimates are shown for inversion using regularization schemes in Table 1 and minimize the respective GCV cost function. Prior information and posterior slip standard deviations (normalized to their minimum value) and posterior slip correlation lengths are shown for each regularization scheme (for Noise 2 realization only). All recovered moment magnitudes equal M_w 8.5 ± 0.1. Bottom-right insets show maximum and minimum value of plotted field (without normalization), while top-left insets show the ratio of such values.

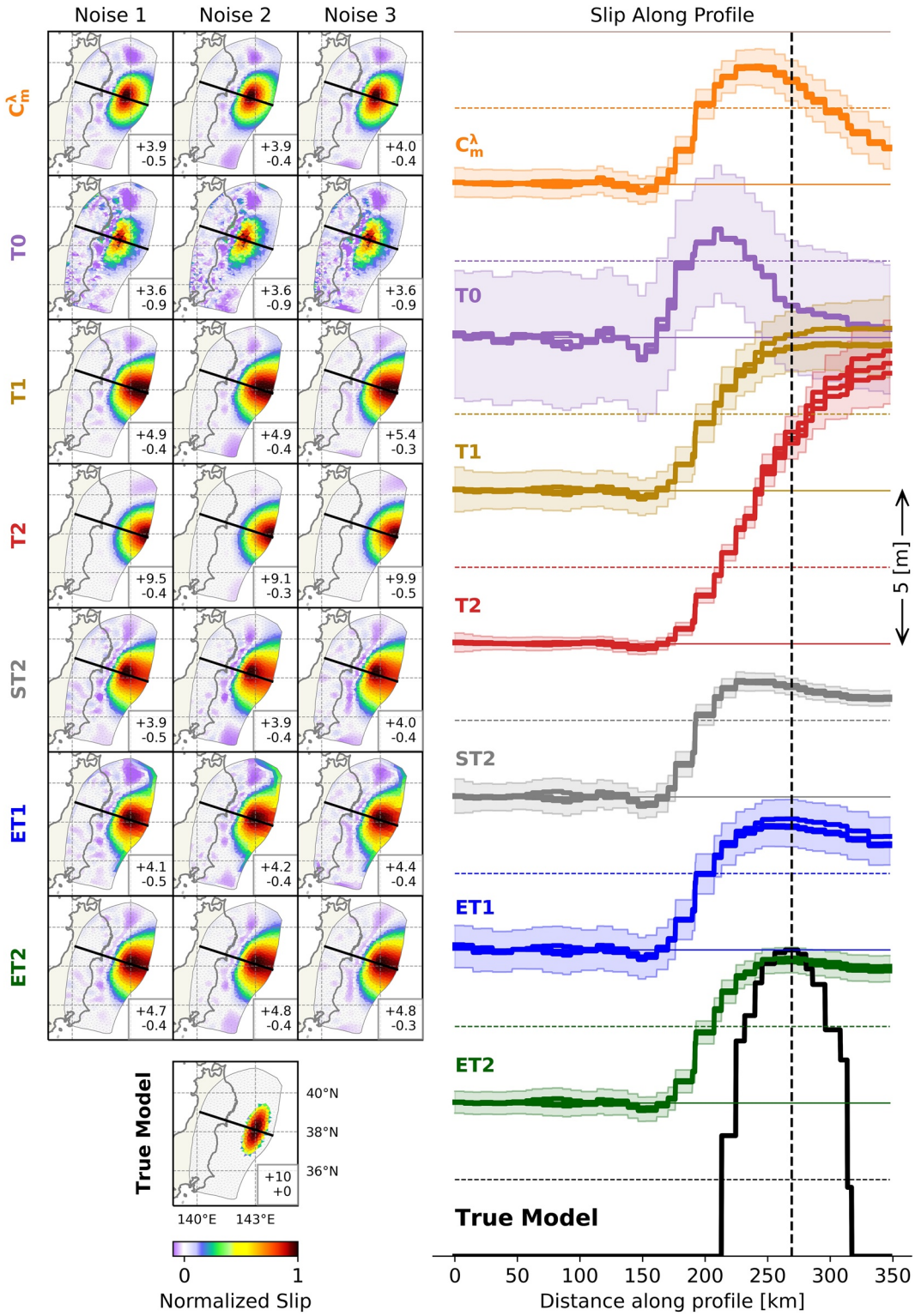

Figure 3. Slip estimates constrained by synthetic observations, generated from a prescribed true elliptical (M_w 8.5) slip model, with added Noise 1, 2, and 3. Best model estimates are shown for inversion using regularization schemes in Table 1 and minimize their respective GCV cost functions. Black line in slip maps indicate profile along with slip is plotted in the right panel. Profiles of slip for true and estimated models for all noise realizations are shown in the same scale. Solid horizontal lines indicate zero slip for the profile with the corresponding color. Shaded regions represent two standard deviations (95% confidence intervals) envelope for each three slip profiles set. All horizontal lines spacing corresponds to 2.5 m of slip. Vertical dashed line indicates location of maximum slip of the true synthetic distribution. Insets show maximum/minimum slip.

The impact of regularization schemes on slip estimates can also be analyzed in terms of the structure of the posterior covariance matrix of slip, $\tilde{\mathbf{C}}_{\mathbf{m}}$. Figure 2 shows posterior standard deviations and spatial correlation lengths on slip estimates for the checkerboard slip case with the Noise 2 realization. Similar values are obtained for different noise realizations and for the recovery of the elliptic synthetic slip shown in Figure 3. For classical regularization approaches, standard deviations of slip estimates vary throughout the fault, being larger in regions where slip instabilities are most noticeable or a larger bias to prior information is inferred. Instead, for EPIC Tikhonov regularization the standard deviations of inferred slip are by design uniformly distributed across the fault. For all regularization schemes, larger posterior slip correlation lengths are obtained for fault portions further away from the observations, presenting largest values at the trench, where slip estimates are less sensitive to the data.

For an exhaustive view of all our results, we direct the reader to the supporting information. For each regularization scheme and synthetic slip model, Figures S6–S33 show slip estimates as a function of the noise realization added to the synthetic data. Figures S34–S37 directly compare slip estimates as a function of the chosen regularization scheme. Figures S38–S51 show GCV curves for inversion results shown in Figures 2 and 3.

4. Discussion of Regularization Schemes Based on Synthetic Models

As expected, synthetic models are better recovered in regions of the fault that are closer to the observations (with higher sensitivity). Slip estimates are both more biased to prior information and more unstable for fault elements located further away from the onland observations (with lower sensitivity). The level of instability depends on the type and amount of regularization imposed.

In terms of the bias induced on model estimates, the analyzed regularization schemes can be characterized into two groups, ones that bias slip values toward a known prior slip model ($C_{\mathbf{m}}^{\lambda}$, T0), and ones that bias model estimates toward a smooth slip distribution, penalizing local spatial variations of slip ($C_{\mathbf{m}}^{\lambda}$, T1, T2, ST2, ET1, ET2). Justifying the choice of particular values for a prior slip model is impossible. Setting slip priors to null values throughout the fault is arbitrary. Instead, prescribing smooth slip as prior information is a practical way to partially take care of that problem. While still somewhat arbitrary, smoothing constraints can be at least weakly justified in the sense that we seek to only infer complexity required by the data and to ensure that unresolved complexity is appropriately averaged and stably inferred with minimal bias.

The stability of model estimates can be considered from the perspective of how observational noise produces perturbations on model estimates or by the changes in model estimates upon the choices of the "best" regularization parameter, given a model class selection method. From the first perspective, $C_{\mathbf{m}}^{\lambda}$ scheme tends to recover models in which slip instabilities are limited and minimally dependent on the noise realization. The model perturbations tend to be smoothed over wide regions due to the spatial correlation induced by the regularization scheme. The T0 scheme tends to induce larger slip instabilities that are noticeable dependent on the noise realization. Such instabilities are localized onto single fault slip patches due to lack of smoothing of the regularization scheme. For the recovery of elliptic synthetic slip, $C_{\mathbf{m}}^{\lambda}$ and T0 schemes tend to shift peak slip toward the coast as slip in the true model is located closer to the trench and the bias imposed toward a null slip model becomes more relevant. T1 and T2 regularization schemes show larger instabilities in regions that are less constrained by the observations, in which T2 allows for larger changes in slip compared to T1. Such instabilities are expressed as randomly distributed large slip variations near the trench when recovering checkerboard models, and with the largest slip located at the trench when recovering elliptical slip distributions (see Figures 2 and 3). ST2 regularization stabilizes slip estimates when recovering checkerboard models (see Figures 2 and S10). Nevertheless, for elliptical slip distributions, ST2 recovers over-smoothed slip at the regions with low sensitivity (near trench), shifting peak slip to a location closer to the observations. ST2 also shows large instabilities at the fault regions with high slip sensitivity (near coast), being highly dependent on the noise realization added to the data (see Figures 3, S17, S24, and S31). Thus, the empirical choice of prior information variances based on the sensitivity of fault slip is unreliable.

ET1 and ET2 regularization schemes excel in stabilizing the inverse problem solution for all analyzed synthetic slip cases. As a direct consequence of the EPIC, posterior slip variances are all the same, thus instabilities tend to be homogeneously distributed across the fault, with maximum amplitudes controlled by the choice of σ_t (Figures 2 and 3 and Text S1, S13). Figure 2 compares estimates of posterior marginal standard deviations and correlation lengths for all regularization operators analyzed in the study of synthetic slip cases. Traditional regularization schemes induce posterior standard deviations and correlation lengths that are both variable across the fault. While this in principle is not a problem, it becomes less practical as it is hard to separate the effects of amplitude and spatial distribution of slip estimates uncertainties. Instead, by design, EPIC Tikhonov regularization provides uniform standard deviation and varying correlation lengths on posterior information of slip, allowing a direct and effective analysis of model estimates spatial resolution in terms of their covariances or correlations. As the correlation length calculated for each slip fault discretization defines a region in which the estimated model instabilities are highly correlated, slip features smaller than the correlation length should not be inferred as they might be the result of overfitting. When comparing ET1 and ET2 with their nonEPIC counterpart (T1 and T2), we observe that imposing the EPIC results in larger correlation lengths in regions with smaller sensitivity (i.e., near the trench), where standard deviations of prior information are also smaller. Use of the EPIC Tikhonov regularization results in relatively stable and unbiased estimates of the slip averaged at scales corresponding to the correlation length and larger.

For all aforementioned regularization schemes, if one tries to avoid the instabilities by augmenting the amount of regularization from the optimum determined using GCV, for example, by increasing ε^2, λ or decreasing σ_t, one tends to over-smooth the solution at well-constrained portions of the fault (see Text S11). In contrast, if one lowers the value of the damping parameter in order to obtain a rougher solution at the well-constrained fault portions, slip becomes unstable near the trench. In comparison, EPIC Tikhonov produces model estimates that are the less sensitive to the choice of the regularization parameter value (see Text S11). This insensitivity becomes relevant as GCV or any other model class selection methods tend to select solutions of the inverse problem that are slightly over- or under-regularized (e.g., Krstajic et al., 2014).

Not surprisingly, for M_w 8.0 and M_w 8.5 synthetic elliptic slip scenarios, all regularization schemes fail to recover the true model average location and moment magnitude when synthetic slip is located near the trench (see Text S10, Figure S35—column 5, Figure S36—column 5). This behavior is not observed for M_w 9.0 synthetic slip scenarios, nor for the checkerboard ones. In all aforementioned cases, the best recovered model, in terms of location and magnitude M_w of slip, is achieved when using ET2. Therefore, slip at any fault portion needs to produce spatial variability in surface displacements with a high enough signal-to-noise ratio to allow recovery. If not, the regularization will set model estimates according to prior information, which in this case corresponds to either a null value of slip ($C_{\mathbf{m}}^{\lambda}$, T0), or null value of slip spatial derivatives (T1, T2, ST2, ET1, and ET2). For the regularization schemes that induce smoothing, actual values of slip will be related to the values of the nearest well constrained fault patches whose distance is of the order of the correlation length of the inferred fault slip.

5. Application to Tohoku-Oki (M_w9.0) 2011 Earthquake

5.1. Results and Discussion

Given the results of the tests with synthetic observations, we unsurprisingly favor ET2 regularization scheme as a smoothing prior. Figure 4 compares the behavior of EPIC Tikhonov and Tikhonov regularization using real observations from the Tohoku-Oki (M_w 9.0) 2011 earthquake. We consider onland crustal co-seismic offsets as seen by 738 GPS stations of the Japanese GEONET Network (Minson et al., 2014; Sagiya et al., 2000). Model prediction uncertainties are based on Minson et al. (2014) (see Text S4). We parameterize slip in two directions, one parallel to the slip vector defined by the Tohoku-Oki earthquake focal mechanism (rake-parallel direction), and a perpendicular one (rake-perpendicular direction) (see Minson et al., 2013, Text S4). Prior information is defined for the amplitude of rake-perpendicular slip ($\mathbf{h} = \mathbf{m}$), as we expect little to no slip in such direction. For the rake-parallel direction, prior information is defined for the second order spatial differential of rake-parallel slip ($\mathbf{h} = \nabla^2\mathbf{m}$). Hereafter, we focus on a comparison of ET2 and T2, as ET2 is an extension of the commonly adopted T2 regularization, and they both represent a

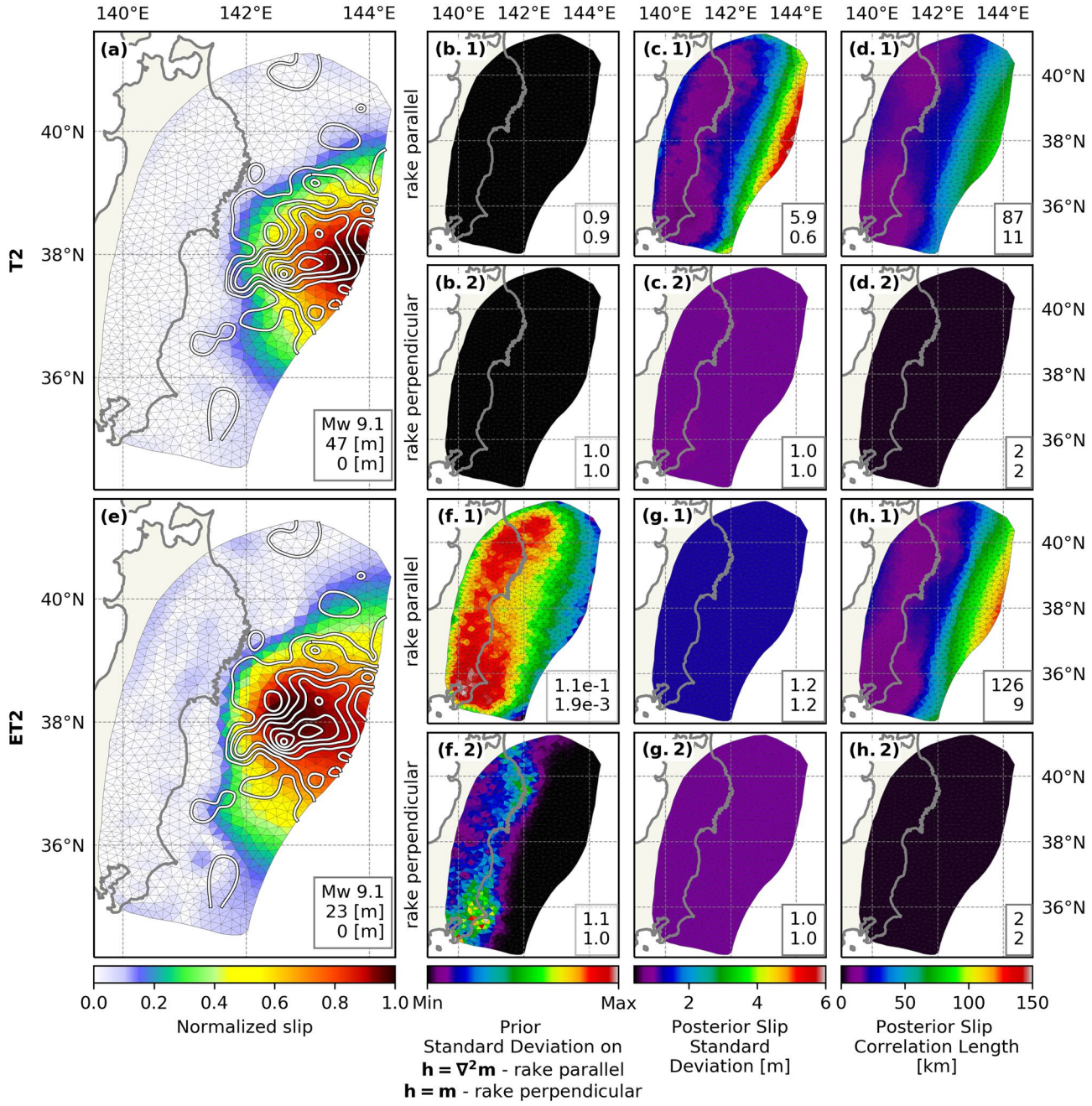

Figure 4. Comparison of 2011 (M_w 9.0) Tohoku-Oki earthquake slip. Panels (a and e) show slip estimates when using T2 and ET2 regularization, respectively. White lines are 10 m contours of co-seismic slip from Minson et al. (2014). Panels (b.1, b.2–f.1, f.2) show standard deviations of prior information, (c.1, c.2–g.1, g.2) show standard deviations of estimated slip and (d.1, d.2–h.1, h.2) show correlation lengths of slip estimates, all for rake parallel and perpendicular slip directions as well as T2 and ET2 regularization, respectively. Insets show minimum and maximum value of the plotted field as well as estimated moment magnitude for slip model plots.

smoothing prior. Panels (a and e) in Figure 4 show slip estimates minimizing GCV cost function, obtained using T2 and ET2 regularization schemes, respectively. Along with each slip model, we show standard deviations of prior and posterior information on model parameters as well as posterior correlation lengths, all for rake parallel and rake perpendicular slip components. In addition, in Text S12 (Figures S53–S61) we compare slip solutions, observations and model prediction, for nine selected values of the regularization parameter indicated in Figure S52, that adopt the same GCV cost function value for T2 and ET2 regularization, including under- and over-regularized solutions.

The best slip model obtained with ET2 closely resembles the slip model estimated using a fully Bayesian inversion method and a much more comprehensive data set including seismic, tsunami and ocean bottom geodetic offsets (Figure 4e, Minson et al., 2014). While the peak slip amplitude of ET2 is smaller due to strong smoothing, driven partially by the use of only onland GPS data, estimated regions with larger slip coincide. Instead, T2 regularization, as expected, tends to concentrate slip at the near trench portions of the subduction megathrust (Figure 4a). T2 regularization shows the smaller correlation lengths (compared to ET2) and the largest posterior variances of slip at the trench, where in turn, maximum slip is located. Inferences of maximum slip at the trench are found in several published models of the Tohoku-Oki earthquake (e.g., Bletery et al., 2014; Ide et al., 2011; Lay et al., 2011; Sun et al., 2014; Yue & Lay, 2011). As shown in the synthetic exercises, this kind of inference might be an artifact of the inversion if not using datasets capable of constraining shallow near trench slip, such as ocean bottom geodesy or tsunami (e.g., see sensitivity plots in Duputel et al., 2014, supporting information). In contrast, slip models constrained by ocean bottom geodetic and tsunami data along with GNSS or seismic data in a jointly manner, show limited slip at the trench (e.g., Koketsu et al., 2011; Minson et al., 2014; Simons et al., 2011; Yokota et al., 2011).

The aforementioned discrepancies in spatial location of estimated slip could be explained by the differences in the spatial distribution of standard deviations of prior and posterior information on model parameters. The main difference lies in that standard deviations of slip estimates are variable for T2 but uniform for ET2 across the fault (Figures 4c and 4g). As T2 shows larger slip variances at the trench, we expect larger model instabilities and bias in such region. Instead, the EPIC enforces that model perturbations upon observational and model prediction uncertainties have similar expected amplitudes and no preferred location on the fault. As a consequence of the EPIC, ET2 induces larger correlation lengths near the trench (fault regions less constrained by the data), thus averaging slip over a wider region compared with T2. At the same time, T2 and ET2 posterior slip correlation lengths are comparable near the coast, thus averaging slip over similarly sized regions for portions of the fault closer to the location of the observations.

5.2. On the Use of Positivity Constraints

The EPIC is defined for the regularized linear least squares problem, in which no positivity constraints are used. Positivity constraints are physically justifiable and are commonly used in slip inversions. Thus, we would like to explore the feasibility of using the EPIC for such inversions. However, if one would use positivity constraints on, for instance, rake parallel slip, formula 14 for the EPIC is no longer valid as the inverse problem becomes (weakly) nonlinear. In this case, the statistics of the solution of the inverse problem are different as posterior marginal pdf's could be better approximated by a log-normal pdf, and formula 8 for the posterior covariance matrix of model parameters is also no longer valid. Nevertheless, for large amplitude of slip estimates (very different from 0), the marginal pdf of the log-normal distribution representing parameters with positivity constraints could be locally approximated by a normal distribution. We use this simplification for expediency, but in the future it may be better to use a more elaborate approach to define the EPIC for the case in which positivity constraints are used. For such purpose, one could benefit from the approach of Nocquet (2018), in which nonnegativity constraints are introduced using a truncated multivariate normal prior distribution, deriving an efficient semi-analytical approach to compute posterior slip estimates and their marginal variances.

Making the aforementioned approximations, we tested ET2 regularization to invert for the Tohoku-Oki earthquake, using positivity constraints on rake parallel slip (see Figure 5). The slip models were selected with the L-curve criterion and have similar data misfit norm. In addition, due to the known caveats of such model class selection method (e.g., Fukuda & Johnson, 2008), in Text S13 we show slip models for several selected values of the regularization parameter across the L-curve. As a first observation, we see that the use of positivity constraints leads to more compact slip distributions. However, T2 regularization still induces peak slip at the trench. Instead, use of ET2 regularization with positivity constraints recovers slip distributions that are remarkably similar to smoothed versions of the unregularized co-seismic static slip distribution found by Minson et al. (2014). Particularly, ET2 and Minson et al. (2014) slip estimates infer the same location of maximum inferred slip (see slip envelopes in Figure 5). In contrast to T2 estimates, the spatial distribution of inferred slip using ET2 is very stable upon the choice of the regularization parameter value, varying mostly the degree of smoothness of the solution, but always representing a smoothed version of the

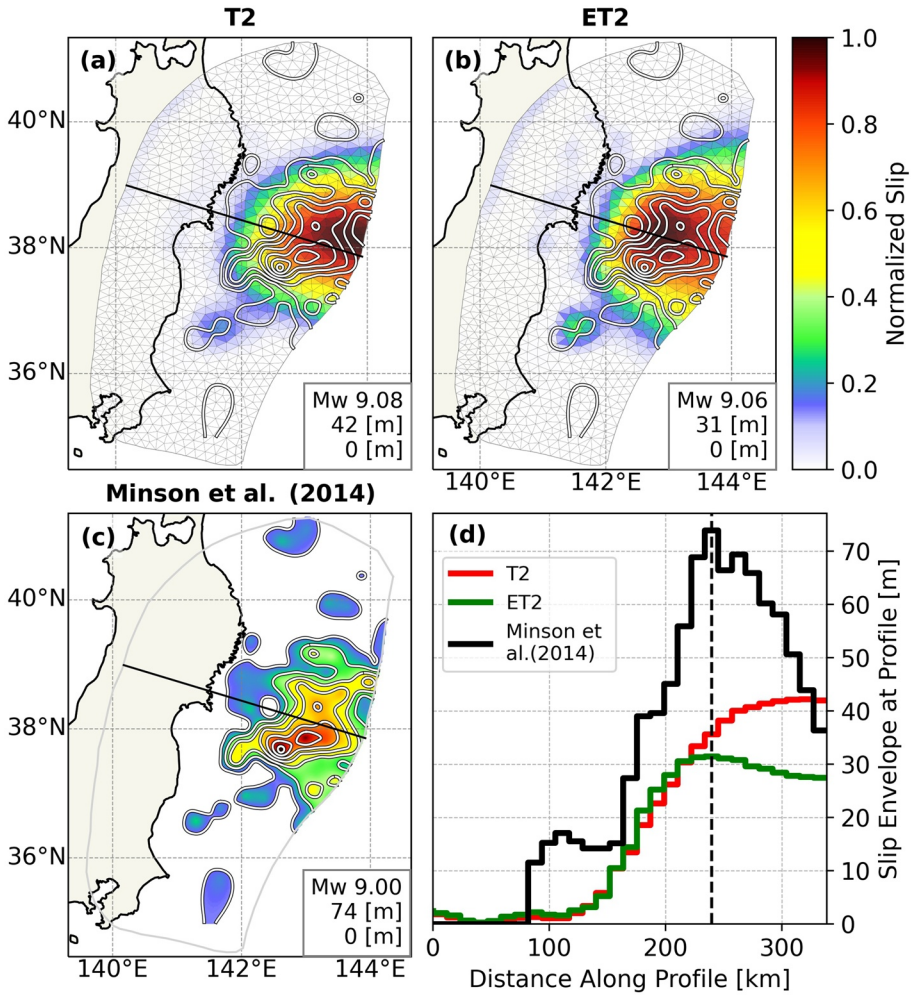

Figure 5. Slip distribution of the 2011 Tohoku-Oki (M_w 9.0) earthquake estimated using (a) T2 and (b) ET2 regularization schemes with positivity constraints on rake parallel slip. (c) Slip estimated by Minson et al. (2014). Insets show moment magnitude, maximum and minimum value of the slip distribution. Panel (d) shows envelope of each slip distribution projected onto the profile (black line on slip maps). Vertical dashed line indicates location of maximum slip envelope for Minson et al. (2014). See Text S13 for further detail.

reference slip distribution. Moment magnitudes, for all considered values of the regularization parameters range from M_w 9.0 to M_w 9.1 for T2 and ET2 regularization. T2 regularization results in near-trench slip values close to those inferred by Minson et al. (2014). Such similarity might be a coincidence as synthetic tests show that T2 regularization tends to concentrate largest slip at the trench.

6. Conclusions

Producing reliable estimates of the distributions of fault slip is a critical step to understand the physics controlling the behavior of faults throughout the seismic cycle. EPIC Tikhonov regularization produces robust slip estimates when considering stability, bias, and ease of uncertainty analysis of posterior model estimates. We also show an initial exploration of the use of EPIC Tikhonov regularization together with positivity constraints. Our preliminary tests are encouraging.

EPIC Tikhonov regularization is particularly useful for quasi-static slip inversions of slow motion at the fault, for example for estimating post-seismic slip, coupling (slip deficit) and slow slip events, in which only geodetic observations provide a spatially variable constraint on slip estimates. Sometimes, subduction megathrust slip models inferred using traditional least squares methodologies and constrained with far

onland data, present peak values of slip near the trench. Such behavior might be an artifact of the inversion scheme if slip is not well-constrained by appropriate observations. Based on our analyses, the proposed EPIC Tikhonov regularization seems to be free from such artifacts, as it sets maximum slip value close to its true location.

The examples in this work are focused on subduction zone environments because they are particularly challenging in terms of the spatially heterogeneous observations and their distance to fault elements. Nevertheless, the proposed methodology is equally applicable to other environments, due to depth dependence of sensitivity to slip and spatially heterogeneous surface observations.

Additionally, while in this work we use a fine fault mesh with quasi-uniformly sized elements, the methodology allows for spatially variably sized fault elements. Nevertheless, the largest fault elements must be small enough, as is required by our methodology. On one hand, having small fault elements minimizes the effects of stress concentrations resulting from using a piece-wise constant slip representation. Nevertheless, this problem could also be ameliorated using continuous basis functions representing slip. On the other hand, decreasing the size of fault patches allows for the selection of larger target posterior variances with feasible solutions for the EPIC. Therefore, the size of the fault mesh elements, will have an impact on the range of feasible target posterior variances (see Text S14).

Compared to other spatially variable regularization schemes, our approach is easily extendable to include seismic observations to perform kinematic slip inversions, as our approach achieves the spatially variable regularization using an arbitrary discretization of the considered fault geometry. Indeed, our approach can be applied to any ill-posed discrete linear inversion problem in which a second order Tikhonov smoothing constraint on finely discretized model parameters makes sense as prior information.

Data Availability Statement

Data sets for this research are included in this paper (and its supplementary information files): Minson et al. (2014). Python codes to solve the general linear least squares inversion problem using EPIC Tikhonov are available at https://doi.org/10.5281/zenodo.4922933, as well as in https://github.com/frortega/EPIC_LS.

Acknowledgments
F. Ortega-Culaciati acknowledges support from Proyecto Fondecyt 11140904 and 1181479, as well as from ANID PIA ACT 192169. J. Ruiz acknowledges support from Proyecto Fondecyt 1200679, and CONICYT PIA grant ACT172002. Powered@NLHPC: This research was partially supported by the supercomputing infrastructure of the NLHPC (ECM-02). We thank the editors and anonymous reviewers for providing insightful comments that helped to improve the clarity and quality of this manuscript.

References

Akaike, H. (1980). Likelihood and the Bayes procedure. *Bayesian Statistics*, *31*, 143–166. https://doi.org/10.1007/BF02888350

Aster, R. C., Borchers, B., & Thurber, C. H. (2013). *Parameter estimation and inverse problems*. Elsevier. https://doi.org/10.1016/C2009-0-61134-X

Backus, G., & Gilbert, F. (1968). The resolving power of gross earth data. *Geophysical Journal International*, *16*(2), 169–205. https://doi.org/10.1111/j.1365-246X.1968.tb00216.x

Barnhart, W. D., & Lohman, R. B. (2010). Automated fault model discretization for inversions for coseismic slip distributions. *Journal of Geophysical Research*, *115*(B10), B10419. https://doi.org/10.1029/2010JB007545

Bletery, Q., Sladen, A., Delouis, B., Vallée, M., Nocquet, J.-M., Rolland, L., & Jiang, J. (2014). A detailed source model for the Mw9.0 Tohoku-Oki earthquake reconciling geodesy, seismology, and tsunami records. *Journal of Geophysical Research*, *119*(10), 7636–7653. https://doi.org/10.1002/2014JB011261

Cai, J., Grafarend, E. W., & Schaffrin, B. (2004). The A-optimal regularization parameter in uniform Tykhonov-Phillips regularization—α-weighted BLE-. *International Association of Geodesy Symposia*, 309–324. https://doi.org/10.1007/978-3-662-10735-5_41

Coleman, T., & Li, Y. (1996). An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, *6*(2), 418–445. https://doi.org/10.1137/0806023

Cover, T., & Thomas, J. (1991). *Elements of information theory*. John Wiley & Sons, Inc. https://doi.org/10.1002/047174882X

Craven, P., & Wahba, G. (1979). Smoothing noisy data with spline functions—Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, *31*(4), 377–403. https://doi.org/10.1007/BF01404567

Duputel, Z., Agram, P., Simons, M., Minson, S., & Beck, J. (2014). Accounting for prediction uncertainty when inferring subsurface fault slip. *Geophysical Journal International*, *197*(1), 464–482. https://doi.org/10.1093/gji/ggt517

Duputel, Z., Jiang, J., Jolivet, R., Simons, M., Rivera, L., Ampuero, J.-P., et al. (2015). The Iquique earthquake sequence of April 2014: Bayesian modeling accounting for prediction uncertainty. *Geophysical Research Letters*, *42*(19), 7949–7957. https://doi.org/10.1002/2015GL065402

Evans, E. L., & Meade, B. J. (2012). Geodetic imaging of coseismic slip and postseismic afterslip: Sparsity promoting methods applied to the great Tohoku earthquake. *Geophysical Research Letters*, *39*(11), L11314. https://doi.org/10.1029/2012GL051990

Fukuda, J., & Johnson, K. M. (2008). A fully Bayesian inversion for spatial distribution of fault slip with objective smoothing. *Bulletin of the Seismological Society of America*, *98*(3), 1128–1146. https://doi.org/10.1785/0120070194

Hansen, P. C. (1998). *Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion*. Philadelphia: SIAM. https://doi.org/10.1137/1.9780898719697

Hansen, P. C., & O'Leary, D. P. (1993). The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing*, *14*(6), 1487–1503. https://doi.org/10.1137/0914086

Harris, R. A., & Segall, P. (1987). Detection of a locked zone at depth on the Parkfield, California, segment of the San Andreas fault. *Journal of Geophysical Research*, *92*(B8), 7945–7962. https://doi.org/10.1029/JB092iB08p07945

Hashimoto, C., Noda, A., Sagiya, T., & Matsu'ura, M. (2009). Interplate seismogenic zones along the Kuril–Japan trench inferred from GPS data inversion. *Nature Geosciences*, *2*(2), 141–144. https://doi.org/10.1038/ngeo421

Herrmann, R. B. (2013). Computer programs in seismology: An evolving tool for instruction and research. *Seismological Research Letters*, *84*(6), 1081–1088. https://doi.org/10.1785/0220110096

Ide, S., Baltay, A., & Beroza, G. C. (2011). Shallow dynamic overshoot and energetic deep rupture in the 2011 Mw 9.0 Tohoku-Oki earthquake. *Science*, *332*(6036), 1426–1429. https://doi.org/10.1126/science.1207020

Koketsu, K., Yokota, Y., Nishimura, N., Yagi, Y., Miyazaki, S. i., Satake, K., et al. (2011). A unified source model for the 2011 Tohoku earthquake. *Earth and Planetary Science Letters*, *310*(3), 480–487. https://doi.org/10.1016/j.epsl.2011.09.009

Krawczyk-Stańdo, D., & Rudnicki, M. (2007). Regularization parameter selection in discrete ill-posed problems—The use of the U-curve. *International Journal of Applied Mathematics and Computer Science*, *17*(2), 157–164. https://doi.org/10.2478/v10006-007-0014-3

Krstajic, D., Buturovic, L. J., Leahy, D. E., & Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics*, *6*(1), 1–15. https://doi.org/10.1186/1758-2946-6-10

Lay, T., Ammon, C. J., Kanamori, H., Xue, L., & Kim, M. J. (2011). Possible large near-trench slip during the 2011 M_w 9.0 off the Pacific coast of Tohoku earthquake. *Earth Planets and Space*, *63*(7), 687–692. https://doi.org/10.5047/eps.2011.05.033

Lévêque, J.-J., Rivera, L., & Wittlinger, G. (1993). On the use of the checker-board test to assess the resolution of tomographic inversions. *Geophysical Journal International*, *115*(1), 313–318. https://doi.org/10.1111/j.1365-246X.1993.tb05605.x

Lohman, R. B. (2004). *The inversion of geodetic data for earthquake parameters*. Doctoral dissertation, California Institute of Technology, Pasadena, CA, USA. https://doi.org/10.7907/9FYH-HD84

Loveless, J. P., & Meade, B. J. (2010). Geodetic imaging of plate motions, slip rates, and partitioning of deformation in Japan. *Journal of Geophysical Research*, *115*(B2), B02410. https://doi.org/10.1029/2008JB006248

Loveless, J. P., & Meade, B. J. (2011). Spatial correlation of interseismic coupling and coseismic rupture extent of the 2011 MW=9.0 Tohoku-Oki earthquake. *Geophysical Research Letters*, *38*(17), L17306. https://doi.org/10.1029/2011GL048561

Mai, P. M., & Beroza, G. C. (2002). A spatial random field model to characterize complexity in earthquake slip. *Journal of Geophysical Research*, *107*(B11), 10–11. https://doi.org/10.1029/2001jb000588

Mazzotti, S., Le Pichon, X., Henry, P., & Miyazaki, S. I. (2000). Full interseismic locking of the Nankai and Japan-west Kurile subduction zones: An analysis of uniform elastic strain accumulation in Japan constrained by permanent GPS. *Journal of Geophysical Research*, *105*(B6), 13159–13177. https://doi.org/10.1029/2000JB900060

Minson, S. (2010). *A Bayesian approach to earthquake source studies*. Doctoral dissertation, California Institute of Technology, Pasadena, CA, USA. https://doi.org/10.7907/3RT9-3215

Minson, S., Simons, M., & Beck, J. (2013). Bayesian inversion for finite fault earthquake source models. I—Theory and algorithm. *Geophysical Journal International*, *194*, 1–1726. https://doi.org/10.1093/gji/ggt180

Minson, S., Simons, M., Beck, J., Ortega, F., Jiang, J., Owen, S., et al. (2014). Bayesian inversion for finite fault earthquake source models—II: The 2011 great Tohoku-Oki, Japan earthquake. *Geophysical Journal International*, *198*(2), 922–940. https://doi.org/10.1093/gji/ggu170

Muto, M., & Beck, J. L. (2008). Bayesian updating and model class selection for hysteretic structural models using stochastic simulation. *Journal of Vibration and Control*, *14*(1–2), 7–34. https://doi.org/10.1177/1077546307079400

Nocquet, J.-M. (2018). Stochastic static fault slip inversion from geodetic data with non-negativity and bound constraints. *Geophysical Journal International*, *214*(1), 366–385. https://doi.org/10.1093/gji/ggy146

Ortega-Culaciati, F. H. (2013). *Aseismic deformation in subduction megathrusts: Central Andes and North-East Japan*. Doctoral dissertation, California Institute of Technology, Pasadena, CA, USA. https://doi.org/10.7907/BHDB-KH07

Pritchard, M. E., Simons, M., Rosen, P. A., Hensley, S., & Webb, F. H. (2002). Co-seismic slip from the 1995 July 30 Mw= 8.1 Antofagasta, Chile, earthquake as constrained by InSAR and GPS observations. *Geophysical Journal International*, *150*(2), 362–376. https://doi.org/10.1046/j.1365-246X.2002.01661.x

Radiguet, M., Cotton, F., Vergnolle, M., Campillo, M., Valette, B., Kostoglodov, V., & Cotte, N. (2011). Spatial and temporal evolution of a long term slow slip event: The 2006 Guerrero Slow Slip Event. *Geophysical Journal International*, *184*(2), 816–828. https://doi.org/10.1111/j.1365-246X.2010.04866.x

Ragon, T., Sladen, A., & Simons, M. (2018). Accounting for uncertain fault geometry in earthquake source inversions–I: Theory and simplified application. *Geophysical Journal International*, *214*(2), 1174–1190. https://doi.org/10.1093/gji/ggy187

Sagiya, T., Miyazaki, S., & Tada, T. (2000). Continuous GPS array and present-day crustal deformation of Japan. *Pure and Applied Geophysics*, *157*(11–12), 2303–2322. https://doi.org/10.1007/PL0002250710.1007/978-3-0348-7695-7_26

Segall, P., & Harris, R. (1987). Earthquake deformation cycle on the San Andreas Fault near Parkfield, California. *Journal of Geophysical Research*, *92*(B10), 10511–10525. https://doi.org/10.1029/JB092iB10p10511

Simons, M., Fialko, Y., & Rivera, L. (2002). Coseismic deformation from the 1999 Mw 7.1 Hector Mine, California, earthquake as inferred from InSAR and GPS observations. *Bulletin of the Seismological Society of America*, *92*(4), 1390–1402. https://doi.org/10.1785/0120000933

Simons, M., Minson, S., Sladen, A., Ortega, F., Jiang, J., Owen, S., et al. (2011). The 2011 magnitude 9.0 Tohoku-Oki earthquake: Mosaicking the megathrust from seconds to centuries. *Science*, *332*(6036), 1421–1425. https://doi.org/10.1126/science.1206731

Stirling, M., Goded, T., Berryman, K., & Litchfield, N. (2013). Selection of earthquake scaling relationships for seismic-hazard analysis. *Bulletin of the Seismological Society of America*, *103*(6), 2993–3011. https://doi.org/10.1785/0120130052

Sun, T., Wang, K., Iinuma, T., Hino, R., He, J., Fujimoto, H., et al. (2014). Prevalence of viscoelastic relaxation after the 2011 Tohoku-Oki earthquake. *Nature*, *514*(7520), 1–87. https://doi.org/10.1038/nature13778

Suwa, Y., Miura, S., Hasegawa, A., Sato, T., & Tachibana, K. (2006). Interplate coupling beneath NE Japan inferred from three-dimensional displacement field. *Journal of Geophysical Research*, *111*(B4), B04402. https://doi.org/10.1029/2004JB003203

Takahashi, N., Kodaira, S., Tsuru, T., Park, J.-O., Kaneda, Y., Suyehiro, K., et al. (2004). Seismic structure and seismogenesis off Sanriku region, northeastern Japan. *Geophysical Journal International*, *159*(1), 129–145. https://doi.org/10.1111/j.1365-246X.2004.02350.x

Tarantola, A. (2005). Inverse problem theory and methods for model parameter estimation. *Society for Industrial and Applied Mathematics*. https://doi.org/10.1137/1.9780898717921

Tarantola, A., & Valette, B. (1982). Inverse problems = quest for information. *Journal of Geophysics*, *50*(1), 159–170. https://geophysics-journal.com/article/28

Tikhonov, A. N. (1963). Resolution of ill-posed problems and the regularization method. *Doklady Akademii Nauk SSSR*, *151*, 501–504.

Wahba, G. (1990). Spline models for observational data. *Society for Industrial and Applied Mathematics*. https://doi.org/10.1137/1.9781611970128

Wang, L., & Gu, W. (2020). A-optimal design method to determine the regularization parameter of coseismic slip distribution inversion. *Geophysical Journal International*, *221*(1), 440–450. https://doi.org/10.1093/gji/ggaa030

Wang, L., Zhao, X., & Gao, H. (2018). A method for determining the regularization parameter and the relative weight ratio of the seismic slip distribution with multi-source data. *Journal of Geodynamics*, *118*, 1–10. https://doi.org/10.1016/j.jog.2018.04.005

Wang, L., Zhao, X., Xu, W., Xie, L., & Fang, N. (2019). Coseismic slip distribution inversion with unequal weighted Laplacian smoothness constraints. *Geophysical Journal International*, *218*(1), 145–162. https://doi.org/10.1093/gji/ggz125

Yabuki, T., & Matsu'ura, M. (1992). Geodetic data inversion using a Bayesian information criterion for spatial distribution of fault slip. *Geophysical Journal International*, *109*(2), 363–375. https://doi.org/10.1111/j.1365-246X.1992.tb00102.x

Yagi, Y., & Fukahata, Y. (2011). Introduction of uncertainty of Green's function into waveform inversion for seismic source processes. *Geophysical Journal International*, *186*(2), 711–720. https://doi.org/10.1111/j.1365-246X.2011.05043.x

Yokota, Y., Koketsu, K., Fujii, Y., Satake, K., Sakai, S., Shinohara, M., & Kanazawa, T. (2011). Joint inversion of strong motion, teleseismic, geodetic, and tsunami datasets for the rupture process of the 2011 Tohoku earthquake. *Geophysical Research Letters*, *38*(7), L00G21. https://doi.org/10.1029/2011GL050098

Yue, H., & Lay, T. (2011). Inversion of high-rate (1 sps) GPS data for rupture process of the 11 March 2011 Tohoku earthquake (Mw 9.1). *Geophysical Research Letters*, *38*(7), L00G09. https://doi.org/10.1029/2011GL048700