

## RESEARCH ARTICLE

10.1002/2017JB014300

## Bounding the moment deficit rate on crustal faults using geodetic data: Methods

Jeremy Maurer<sup>1</sup> , Paul Segall<sup>1</sup>, and Andrew M. Bradley<sup>2</sup><sup>1</sup>Department of Geophysics, Stanford University, Stanford, California, USA, <sup>2</sup>Center for Computing Research, Sandia National Laboratories, Albuquerque, New Mexico, USA

## Key Points:

- The geodetic moment deficit rate (MDR) measures earthquake potential without smoothness assumptions
- Two proposed methods based on constrained optimization provide reliable bounds on MDR; other methods can fail
- The estimated MDR at Parkfield excludes the observed 38 year recurrence time between 1966 and 2004 events

## Correspondence to:

J. Maurer,  
jlmaurer@stanford.edu

## Citation:

Maurer, J., P. Segall, and A. M. Bradley (2017), Bounding the moment deficit rate on crustal faults using geodetic data: Methods, *J. Geophys. Res. Solid Earth*, 122, doi:10.1002/2017JB014300.

Received 8 APR 2017

Accepted 23 JUL 2017

Accepted article online 27 JUL 2017

**Abstract** The geodetically derived interseismic moment deficit rate (MDR) provides a first-order constraint on earthquake potential and can play an important role in seismic hazard assessment, but quantifying uncertainty in MDR is a challenging problem that has not been fully addressed. We establish criteria for reliable MDR estimators, evaluate existing methods for determining the probability density of MDR, and propose and evaluate new methods. Geodetic measurements moderately far from the fault provide tighter constraints on MDR than those nearby. Previously used methods can fail catastrophically under predictable circumstances. The bootstrap method works well with strong data constraints on MDR, but can be strongly biased when network geometry is poor. We propose two new methods: the Constrained Optimization Bounding Estimator (COBE) assumes uniform priors on slip rate (from geologic information) and MDR, and can be shown through synthetic tests to be a useful, albeit conservative estimator; the Constrained Optimization Bounding Linear Estimator (COBLE) is the corresponding linear estimator with Gaussian priors rather than point-wise bounds on slip rates. COBE matches COBLE with strong data constraints on MDR. We compare results from COBE and COBLE to previously published results for the interseismic MDR at Parkfield, on the San Andreas Fault, and find similar results; thus, the apparent discrepancy between MDR and the total moment release (seismic and afterslip) in the 2004 Parkfield earthquake remains.

## 1. Introduction

Earthquake forecasts are based in part on the often brief historical record, paleoseismology and measurements of strain accumulation [e.g., *Field et al.*, 2014]. Geodetic data, including Global Navigation Satellite System (GNSS) and Interferometric Synthetic Aperture RADAR (InSAR) are used in various ways to estimate strain accumulation, often by either computing surface strain rates directly and extrapolating to depth [e.g., *Savage and Simpson*, 1997; *Ward*, 1998] or solving for the slip deficit rate (SDR) on modeled crustal faults. The SDR is the difference between the interseismic slip rate and the long-term or geologic slip rate on a given fault segment. These measurements can be used to estimate the *moment deficit rate* (MDR), which over time leads to a moment deficit that is released in earthquakes. The MDR and the SDR are first-order constraints on earthquake potential that can be estimated using geodetic measurements.

## 1.1. The Slip Deficit Rate for Earthquake Hazard

Estimating the SDR is typically done by making a set of kinematic assumptions to develop forward models and employing standard inverse techniques [e.g., *Segall and Harris*, 1986; *Murray and Langbein*, 2006; *Fialko*, 2006; *Meade and Loveless*, 2009; *Johnson and Fukuda*, 2010; *McCaffrey et al.*, 2013]. These models involve defining the fault system geometry and making assumptions about how the Earth's crust and faults behave over time. The majority of these models assume an elastic crust, and discretize faults with either rectangular [*Okada*, 1985] or triangular [*Maerten et al.*, 2005; *Meade*, 2007] elements (dislocations) in a half-space [e.g., *Segall and Harris*, 1986; *Johnson et al.*, 2001; *Jónsson et al.*, 2002; *McCaffrey*, 2002; *Fialko*, 2006; *Loveless and Meade*, 2011; *Xue et al.*, 2015]. Others have modeled the lower crust and/or upper mantle as viscoelastic [e.g., *Hashimoto et al.*, 2009; *Chuang and Johnson*, 2011; *Johnson*, 2013b]. SDR inversions have often been done within the context of elastic block models by estimating and subtracting deformation due to steady motion on block boundaries, then solving for back slip [e.g., *Murray et al.*, 2001; *Prescott et al.*, 2001; *Meade and Hager*, 2005; *Fukuda and Johnson*, 2008; *Minson et al.*, 2013; *Xue et al.*, 2015].

Even assuming fully elastic deformation and known fault geometry, the inverse problem of solving for back-slip rate on specified faults is underdetermined, requiring regularization to obtain a unique solution. Thus, any estimate of the SDR, and hence the associated earthquake potential, depends on the type of regularizing functional used (e.g., spatial smoothing) and the relative weighting of the data misfit versus the regularization measure. The continued development of InSAR for measuring interseismic deformation [e.g., *Jolivet et al.*, 2014; *Lindsey et al.*, 2014] will increase the number of measurements used for estimating SDR, but cannot overcome the fundamental nonuniqueness of estimating fault slip at depth from only surface measurements.

In sum, geodetic data alone cannot uniquely determine the spatially variable SDR; any part of the model space unconstrained by the data is constrained only by the regularization. Since earthquake potential is related at least in part to the accrued slip deficit, this can lead to serious consequences when the unconstrained part of the model contains significant slip deficit. For example, prior to the 2011 Tohoku-oki earthquake, back-slip inversions often assumed no slip deficit at the trench [e.g., *Hashimoto et al.*, 2009; *Nishimura et al.*, 2004; *Suwa et al.*, 2006]. Combined with spatial smoothing, this yielded slip-deficient zones that were considerably smaller than actually required by the data [e.g., *Loveless and Meade*, 2011] and were interpreted to preclude the possibility of a trench-rupturing earthquake. These assumptions were of course disproved by the occurrence of the  $M_w$ 9.0 2011 earthquake, with extremely large slip near the trench and a larger moment than had been thought possible.

### 1.2. Bayesian Approaches Include Uncertainty in the SDR

Several recent studies have approached this problem by estimating full probability distributions for SDR within a Bayesian framework, to determine the full suite of possible SDR models that fit the data. This approach does not require regularization, because a unique solution is not sought. Instead, a *prior* probability distribution must be proposed that specifies one's prior state of knowledge about interseismic slip rate. The output of these algorithms is the *posterior* distribution of SDR on each fault segment. This method can handle nonlinear models and non-Gaussian probability density functions (PDFs) by using sampling methods such as Markov Chain Monte Carlo (MCMC) methods to approximate the posterior PDFs on SDR [e.g., *Fukuda and Johnson*, 2008; *Minson et al.*, 2013; *Murray et al.*, 2014].

### 1.3. The Moment Deficit Rate for Hazard

In this study our principal concern is how geodetic data constrain seismic hazard. The MDR provides the single scalar measure most simply related to seismic potential of a fault segment. For this reason, and because the SDR itself cannot be uniquely determined from surface geodetic measurements, we focus here on the PDF of the MDR for a given fault system. MDR is proportional to the integral of the slip deficit rate over the fault (see equation (2)). When integrated over time the MDR yields an estimate of the moment deficit available to drive future earthquakes. The actual accrued moment may differ if some moment is released aseismically or previous earthquakes did not completely recover the available deficit. However, bounds on the estimated MDR serve as an estimate of the plausible range in earthquake potential based on geodetic observations and can be directly integrated into hazard assessments. The goal of this paper is to develop methods to place rigorous bounds on the interseismic MDR due to strain accumulation on a given fault segment.

#### 1.3.1. Previous Work

Numerous studies have estimated coseismic moment or interseismic MDR using geodetic data employing a variety of methods. Some studies have simply reported the moment from the best fitting or average SDR model [e.g., *Jónsson et al.*, 2002; *Murray and Langbein*, 2006; *Murray et al.*, 2014; *Schmidt et al.*, 2005; *Ryder and Burgmann*, 2008]. This, of course, gives no information on the uncertainty in the estimate of MDR.

A number of studies have used surface strain rate estimates to determine moment deficit by computing the largest principal strain rate within triangulated stations or on grid points and integrating these over the Earth's surface [*Ward*, 1994; *Savage and Simpson*, 1997; *Ward*, 1998], an approach based on *Kostrov* [1974]. This method does not require a fault model but assumes that the strain rate is depth independent and that the depth of seismic slip (here the *transition depth*) and elastic properties are known. Furthermore, these studies did not provide uncertainty in the moment deficit accumulation rate, and the results cannot be used to estimate moment deficit on specific faults. We do not consider these approaches in this study, because our focus is on assessing uncertainty in the MDR assigned to specific faults.

*Tong et al.* [2013] developed a novel approach to estimate MDR and applied it to the creeping section of the San Andreas Fault (SAF). Their method is based on assuming that the fault behaves as an infinitely long

strike-slip fault (antiplane strain), in which case they show that the MDR,  $\dot{M}_d$  can be estimated using the relation

$$\frac{\dot{M}_d}{L} = \lim_{W \rightarrow \infty} \frac{G\pi}{W} \int_{-W}^W xv(x)dx \quad (1)$$

where  $v(x)$  is the fault-parallel velocity at distance  $x$  perpendicular to the fault,  $L$  is the along-strike length of the fault, and  $G$  is the shear modulus. While this model is appealing from the perspective that no inversion for SDR is needed, many faults cannot be approximated as infinitely long, and in practice the integration width  $W$  may be limited due to the presence of other faults or coastlines. In fact, *Tong et al.* [2013] show a plot of MDR versus  $W$  for the creeping SAF that is not convergent for the range of distances employed.

One of the earliest attempts at bounding moment uncertainty was *Segall and Harris* [1986] and *Segall and Harris* [1987]. They varied the moment deficit systematically, by seeking SDR models as close as possible to either fully locked or rigid block motion and determined qualitatively when the predicted data no longer fit the observations or violated geologic bounds on slip rate. *Johnson et al.* [1994] estimated more rigorous bounds on coseismic moment during the 1992 Landers earthquake using a constrained optimization approach. Their algorithm, translated to the interseismic slip deficit problem, is to systematically sweep through values of MDR, from fully locked to fully creeping fault, and perform a constrained inversion for SDR fixed at each value of MDR. This gives the minimum data misfit as a function of MDR. Plotting moment versus misfit and setting a misfit threshold (*Johnson et al.* [1994] used L1 statistics) give bounds on the MDR. The challenge with this approach is to set the misfit threshold in a statistically meaningful manner, as discussed in detail in section 2.2. Others [*Murray and Segall*, 2002; *Maurer and Johnson*, 2014] used this constrained optimization method to estimate interseismic moment deficit on the Parkfield and creeping segments of the SAF, respectively.

*Johnson* [2013a] used a different approach to study the creeping section of the SAF, in which locked patches were assumed to be surrounded by patches creeping at constant resistive shear stress (i.e., no stress accumulation). By systematically varying the number of locked patches, *Johnson* [2013a] was able to determine the range of total locked fault area accumulating stress, an underlying assumption being that MDR associated with creeping areas is ultimately recovered aseismically.

Another approach to estimating bounds in interseismic MDR was given in *Murray and Segall* [2002], who used bootstrap resampling of the geodetic data to estimate a probability distribution on MDR. They found using simulations that the bootstrap method worked well at Parkfield for constraining the probability density function (PDF) for MDR. *Maurer and Johnson* [2014] also used the bootstrap to estimate uncertainty in the MDR on the creeping SAF.

### 1.3.2. Mathematical Definition of the MDR

The MDR is defined as the integral of the slip deficit rate over the area of the fault, multiplied by shear modulus,

$$\dot{M}_d = G \int_A (\dot{s}_\infty(\xi) - \dot{s}_{is}(\xi))d\xi = G \int_A \dot{s}_{bs}(\xi)d\xi = GA\dot{s}_{bs} \quad (2)$$

where  $\dot{s}_{is}$  is the interseismic slip rate,  $\dot{s}_\infty$  is the long-term or geologic slip rate,  $\dot{s}_{bs}$  is the so-called “back-slip” rate or slip deficit rate (SDR),  $\xi$  denotes position on the fault,  $G$  is shear modulus, and  $A$  is the area of the fault. The bar in the last equality indicates the average value over the fault. For the rest of this paper, we will consider the problem of estimating MDR in terms of back slip on the fault, so we drop the subscript for clarity.

For a fault discretized into  $N$  segments, MDR is given by

$$\dot{M}_d = G \sum_{i=1}^N \dot{s}_i a_i = GA\dot{s} \quad (3)$$

$\dot{s}$  is the weighted average of the SDR, where the weights are the fractional segment areas  $a_i$ , so that  $\sum_{i=1}^N a_i = A$ , the total fault area.

In this study we assume that fault geometry and long-term slip rates at all points on the seismogenic part of the fault are known. For the purposes of this study, we do not distinguish between geologic and geodetically inferred slip rates. While these do not always agree [e.g., *Chuang and Johnson*, 2011; *Evans et al.*, 2016], at least in part because geodetically inferred slip rates depend on model assumptions, we assume that any discrepancies have been resolved prior to estimation of the MDR. We do discuss, in section 3 on the application

to the Parkfield segment of the San Andreas Fault (SAF), how to incorporate uncertainty into the long-term slip rate. We further restrict our attention here to elastic models; in viscoelastic models the surface deformations are convolutions of past slip history, and slip rates are likely to vary with time throughout the earthquake cycle [Johnson and Segall, 2004; Chuang and Johnson, 2011; Hearn et al., 2013]. It is also critical to constrain the maximum possible *transition depth*: the maximum depth to which slip deficit may occur on the fault. This is not the same as the conventional *locking depth* in which the fault is fully locked at all shallower depths. The difference is that the degree of coupling can vary spatially above the transition depth.

The use of the *potency deficit rate* instead of MDR would be more appropriate in this analysis, because of the ambiguity in defining seismic moment for a fault bounding rocks with different elastic properties [e.g., Ampuero and Dahlen, 2005]. Our focus is uncertainty related to the data and Green's functions used to estimate moment, not the rheology of the Earth model [see also Diner and Özgün Konca, 2017]. For simplicity in this study we use the more customary MDR terminology.

#### 1.4. Goal of This Study

Our goal for this study is to develop methods for placing rigorous bounds on uncertainty in the MDR. We seek a method for deriving a distribution on MDR given data  $\mathbf{d}$ , denoted  $p(\hat{M}_d|\mathbf{d})$  that is

1. Convergent with mesh refinement: The solution should converge to a limit solution as the patch element size on the fault is decreased.
2. Consistent: As the data errors diminish and/or the model resolution improves, the estimate of the MDR should converge to the true MDR.
3. Independent of regularization.
4. Conservative: At a minimum, we require models that match available data equally well to have equal probability density.

In the following section, we evaluate methods previously used to derive the MDR estimate  $\hat{M}_d$  and present and evaluate new methods. We refer to  $\hat{M}_d$  and any quantity derived from it, such as a centered confidence interval (CI), as an estimator. We will show that previous methods can fail to provide reliable estimates of the PDF  $p(\hat{M}_d|\mathbf{d})$  under predictable circumstances, meaning that they fail to meet one or more of our four criteria above, while other methods do provide robust bounds on the MDR. Finally, we apply these methods to the Parkfield segment of the SAF to estimate the MDR on this fault and compare to previous estimates as an illustration of the method in practice.

## 2. Methods

### 2.1. The Bootstrap

The bootstrap estimator is a well-known method for nonparametric uncertainty quantification [e.g., Efron and Tibshirani, 1994] and has been successfully applied in previous studies to estimate the MDR [e.g., Murray and Segall, 2002; Maurer and Johnson, 2014]. No assumptions about the errors in the data need to be made, and it is straightforward to implement and interpret. Murray and Segall [2002] showed through a coverage test (explained further below) that the bootstrap confidence intervals matched expected intervals quite well, for the fault and geodetic network geometry at Parkfield. The bootstrap works by perturbing the data many times and each time computing the solution of interest.

To estimate MDR, the data are perturbed and used to solve for SDR on the fault. Then the corresponding MDR is computed using  $\hat{M}_d = \mathbf{a}^T \hat{\mathbf{s}}$ , where  $\hat{\mathbf{s}}$  is an  $N \times 1$  vector of back-slip rates (SDR) on the fault, and  $\mathbf{a}$  is the corresponding vector of patch areas (times a shear modulus to get units of moment rate).

Because we focus on fault-specific estimates of the MDR, all the methods we explore involve computing SDR models. As discussed in section 1, typical approaches to estimate SDR impose some form of regularization to obtain a unique solution, which we seek to avoid. We do employ bounds on slip rate, informed by geologic studies, as useful information to incorporate into SDR estimates.

The inverse problem for SDR including bounds on slip rate, with no other regularization, is given by

$$\begin{aligned} \min_{\hat{\mathbf{s}}} \Phi &\equiv \|\Sigma^{-1/2}(\mathbf{d} - \mathbf{G}\hat{\mathbf{s}})\|_2^2 \\ \text{s.t. } \mathbf{0} &< \hat{\mathbf{s}} < \hat{\mathbf{s}}_{\infty} \end{aligned} \quad (4)$$

where  $\mathbf{G}$  is the  $M \times N$  matrix of Green's functions, and  $\mathbf{d}$  is an  $M \times 1$  vector of data (such as GPS velocities).  $\Sigma$  is the data covariance matrix, and  $\Phi$  is the weighted residual misfit (squared) norm. Slip rates are constrained

to be between zero and the long-term slip rate  $\dot{s}_\infty$ , which can be a function of position on the fault ( $\xi$ ). Problem (4) may have a unique solution depending on whether the bounds constrain the part of the model unresolved by the data. Note that bound constraints on the back-slip rate means that this is a nonlinear estimator, and hence, one cannot, for example, compute degrees of freedom as the difference in dimension of the data and model spaces.

The bootstrap method simply replaces the original data in (4) with the bootstrapped (perturbed) data  $\mathbf{d}^*$  and uses the solution ( $\hat{\mathbf{s}}^*$ ) to compute a bootstrap estimate of the MDR:  $\hat{M}_d^* = \mathbf{a}^T \hat{\mathbf{s}}^*$ , where the star denotes the bootstrap estimate. This process is repeated many times resulting in a histogram of values for the MDR.

There are two possible techniques to perturb the data. One involves randomly sampling the data with replacement, which requires resampling the corresponding rows of  $\mathbf{G}$ . Resampling the data changes the network geometry, which for sparse networks can have a major impact on model resolution. The second approach is to resample the residuals from a best fitting model, adding the bootstrapped residuals to the original data at each iteration. Because this method does not change the network geometry, it is preferred for SDR inversions. However, both approaches have been used in the literature, and we show results using both methods.

In spite of previous success with the bootstrap method, there is reason for caution. Consider the case of resampling the data. At each iteration the rows of  $\mathbf{G}$  are resampled, corresponding to the data points chosen. In a linear unbounded inversion, the null space ( $\text{Ker}(\mathbf{G})$ ) is orthogonal to all the rows of  $\mathbf{G}$ :

$$\text{Ker}(\mathbf{G}) = \{ \mathbf{x} | \mathbf{g}_i^T \mathbf{x} = 0, i = 1, 2, \dots, M \} \quad (5)$$

where  $\mathbf{g}_i^T$  is the  $i$ th row of  $\mathbf{G}$ . All the bootstrap samples are from the orthogonal complement of the model null space, and hence do not provide *any* information about the model null space. The alternative method (resampling residuals) is equally problematic: it does not change the forward operator  $\mathbf{G}$  and therefore does nothing to sample from the model null space. This problem with the bootstrap has already been demonstrated by Page *et al.* [2009], and it implies that the bootstrap can fail to provide a reliable indicator of uncertainty in the MDR. We show some examples of this below.

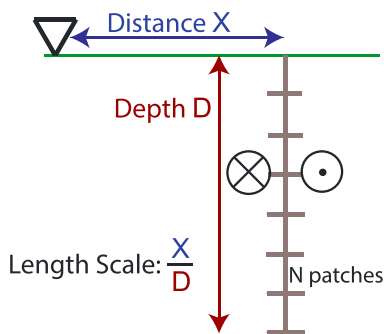
### 2.1.1. Method Validation

The analysis above indicates that the bootstrap could fail to provide reliable bounds on MDR under some conditions. We test for this possibility using simple synthetic examples. In general, we seek to determine whether a given method gives confidence bounds on MDR in the synthetic tests that satisfy the four criteria for a reliable estimator given in section 1.4. We take a Bayesian approach and interpret confidence bounds probabilistically, meaning that we assume some prior information about MDR. This differs from the traditional statistical view of confidence interval, which is not interpreted as a probability interval. Within a Bayesian context, we can relate a PDF to bounds on the MDR through the “credible interval” (CI). This is a nonunique designation: there are several possible ways to define credible intervals. For the purposes of this paper, we will define them as the inverse CDF corresponding to centered probability intervals; for example, the 95% CI is given by the values that correspond to 2.5% and 97.5% on the CDF.

We also use a “coverage test” as a performance metric for summarizing the results of many synthetic trials. The coverage test is a graph that plots the percentage of times a set of PDFs generated for a synthetic test include the true value for a given CI. A one-to-one line indicates perfect coverage: the percentage of distributions that includes the true value is equal to the confidence level. Coverage graphs that fall below the one-to-one line are “optimistic”; the distributions exclude the true value more often than the confidence level. Coverage plots that fall above the one-to-one line are conservative or “pessimistic”; they include the true value more than the specified level. In general, a coverage plot will only be optimal if the generating procedure for synthetic slip models is consistent with the estimator; for nonlinear estimators such a generating procedure may not exist. Therefore, we primarily use the coverage test as a metric to compare between different methods on the same test geometry; methods that are pessimistic (conservative) are preferred over methods that are optimistic.

### 2.1.2. Bootstrap Synthetic Test Results

We present simple synthetic test cases to illustrate the conditions where the bootstrap fails, in order to emphasize that the method should be used with caution. The first test is an antiplane fault (APF) with a strike-slip sense of motion and fixed transition depth  $D$  (identical to Figure 1 except with multiple stations). We use a set of 13 measurements distributed in two different configurations and two different noise levels (Figure 2). The stations are located at regular intervals, from 0.03 to 0.15 $D$  in the close case, and from 1.03 to 1.15 $D$  in the far case. Noise levels are such that the average signal-to-noise ratio (SNR) in the high-noise case is 2,



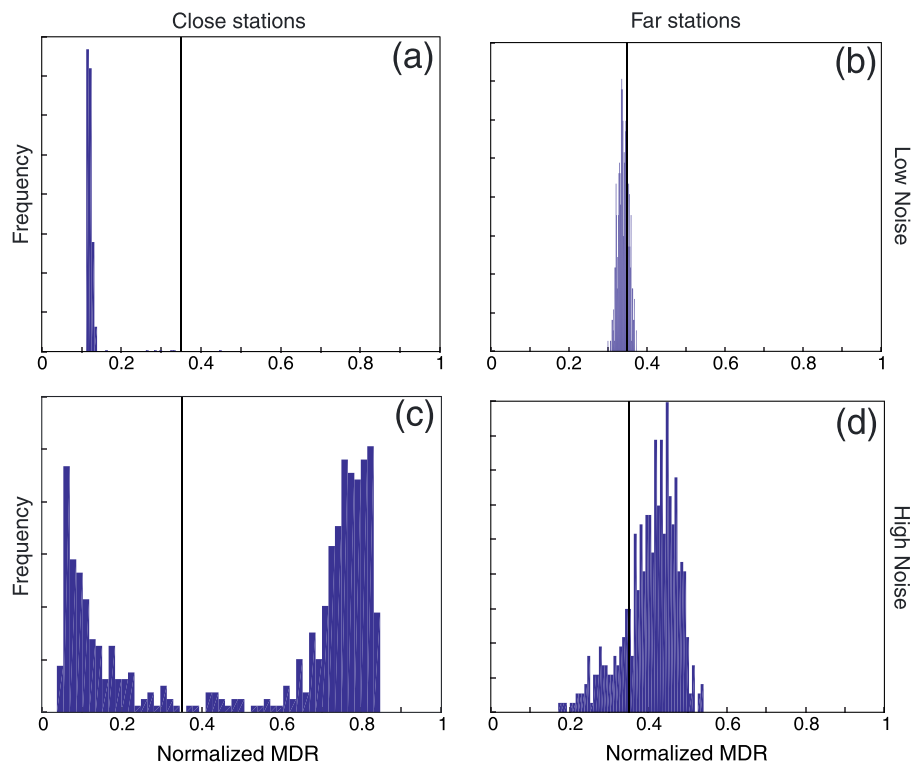
**Figure 1.** Antiplane strike-slip fault with depth  $D$  discretized into  $N$  segments, each of size  $D/N$ . A single GPS station at distance  $x$  from the fault leads to a length scale  $x/D$ . Slip on each patch can vary spatially and is bounded between 0 and 1.

random errors in the data, when the SNR is high but the estimate based on the original data is strongly biased, the resulting histogram for MDR can be catastrophically inaccurate (Figure 2a). When the SNR is low the large variance in the bootstrap estimates may include the true solution, but examination of the resulting histograms reveals that failure to sample from the model null space strongly distorts the estimated distribution of MDR (Figure 2c). In contrast, with far stations (Figures 2b and 2d) the bootstrap samples at least cover the true solution.

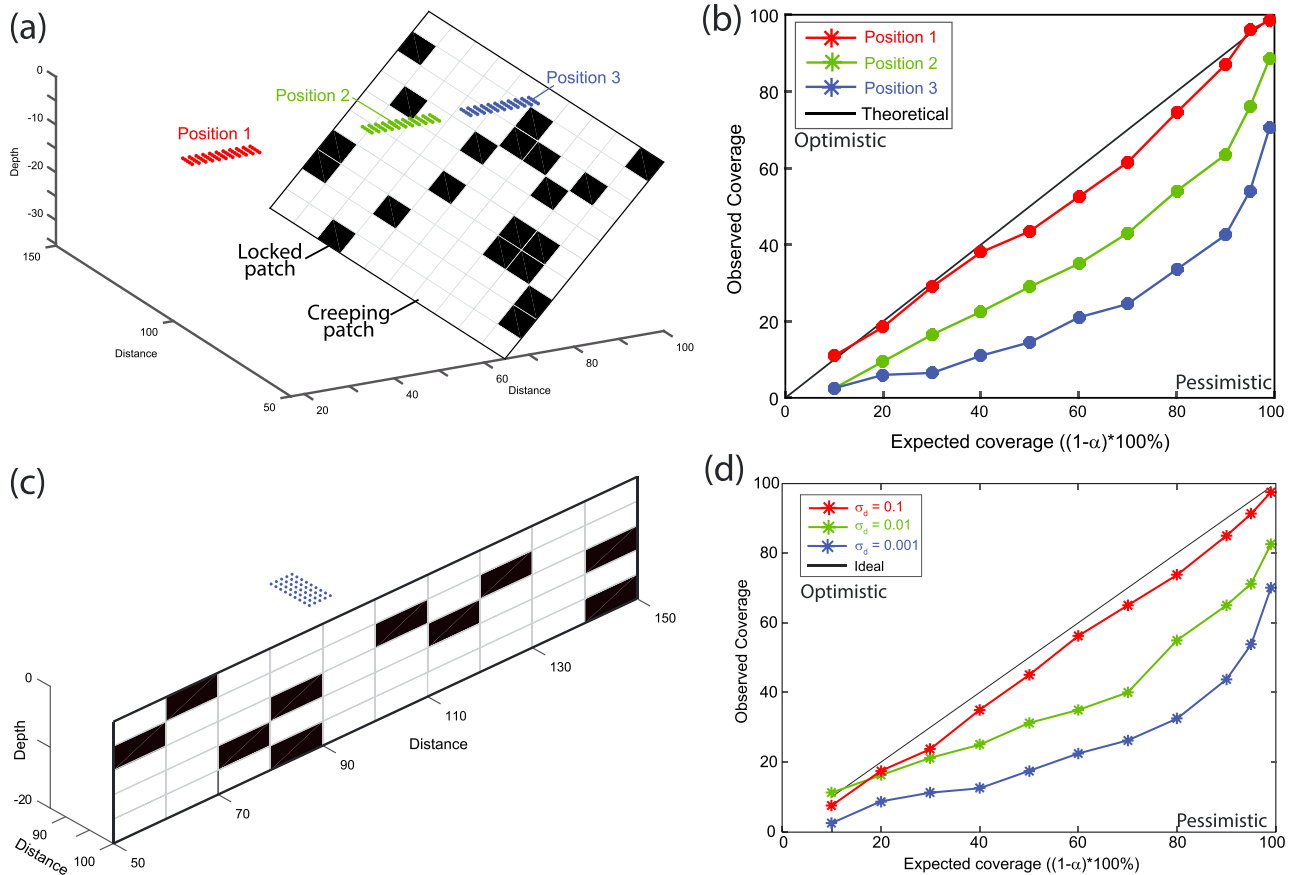
As additional examples, we consider two finite-fault models: a  $45^\circ$  dipping fault 50 km wide (Figures 3a and 3b) and a strike-slip fault 20 km wide (Figures 3c and 3d). Both faults are 100 km long. The dipping fault

and the average SNR in the low-noise case is 200. We use the residual bootstrap method with 500 bootstrap samples to estimate the posterior distribution on MDR. SDR on each fault patch in this simple test is bounded between 0 and 1. Since MDR is proportional to average slip deficit rate times shear modulus, the total possible MDR varies between zero and one times modulus, which we normalize to the unit interval. The true value in each case is  $\approx 0.35$ , or 35% of the total (35% coupled).

Bootstrap estimates can be strongly biased when the stations are close to the fault (Figure 2a). Since the bootstrap only samples



**Figure 2.** Example synthetic bootstrap results for the antiplane geometry of Figure 1, with 13 stations and 30 fault patches. ‘Normalized MDR’ is fraction of the total possible (fully locked). The true value is 0.35 in each case and shown by the vertical black lines. Columns are different station distances and rows are different noise levels. High noise: average SNR  $\approx 2$ , low noise: average SNR  $\approx 200$  (actual SNR varies with station distance). Close stations are regularly spaced between  $0.03$  and  $0.15D$ ; far set are between  $1.03$  and  $1.15D$ . (a) Low noise, close stations. (b) Low noise, far stations. (c) High noise, close stations. (d) High noise, far stations.



**Figure 3.** Three-dimensional bootstrap synthetic tests. (a) Dip-slip fault and station geometry. The fault is 100 by 50 km and dips at an angle of 45°. Three sets of 40 stations were used to estimate MDR using the residual resampling method. (b) Coverage test results for each set of stations in (a). (c) Strike-slip fault and station geometry. The fault is 100 × 20 km with a grid of 40 stations. (d) Coverage results for (c) using the data resampling method.  $\sigma_d$  is the standard deviation of the noise added to the synthetic data. The average amplitude of the signal is of order one, so the average SNR is approximately  $1/\sigma_d$ .

is discretized into 100 patches and the vertical fault into 50. We generate 200 random slip realizations, all at 20% total coupling, where each patch is either locked or creeping (those shown in Figures 3a and 3c) are only examples of each). Synthetic data are computed for each slip realization for sets of 40 stations; the location of this group of stations is varied to three different areas (as shown in Figure 3a), denoted by Positions 1, 2, and 3. We consider constant noise level ( $\sigma_d$ ) but varying station locations (Figures 3a and 3b) and constant location but varying noise level (Figures 3c and 3d). For each SDR realization, the data are resampled 400 times, resulting in 400 MDR estimates used to generate histograms. We use the residual bootstrap method for Figures 3a and 3b, and the data bootstrap method for Figures 3c and 3d. We use the empirical CDF of the MDR histogram to determine CIs and repeat this process for each of the 200 SDR models. The total set of 200 CIs are used to construct the coverage plots shown in Figure 3. To reiterate, the coverage plot shows on the y-axis the percentage of the 200 PDFs on MDR that include the true value for each confidence level on the x-axis. Curves falling below the one-to-one line are optimistic and exclude the true value too often; curves above the line are pessimistic (conservative) and include the true value more often than necessary.

The percentages of distributions that include the true value at the 95% CI in Figure 3b are 96% (Position 1), 76% (Position 2), and 54% (Position 3). For Figure 3d, percentages including the true value are 95% ( $\sigma_d = 0.1$ ), 71% ( $\sigma_d = 0.01$ ), and 54% ( $\sigma_d = 0.001$ ). For Figure 3d, the amplitude of the signals for these conditions is approximately one, with some variation between stations, so the corresponding signal-to-noise ratios (SNRs) are approximately 10, 100, and 1000, respectively.

Why is the bootstrap optimistic when there are no far-field stations, as observed in both examples? We previously noted that bootstrap resampling does not sample from the model null space. As an example, a single

station adjacent to the fault and with zero measurement error will obviously not constrain the SDR anywhere on the fault except close to the station. The bootstrap does not include the uncertainty from having no constraints elsewhere on the fault. If there are stations in the far field to constrain the MDR (Figure 3a), the bootstrap works well because stations in the far-field constrain the overall MDR. When no stations are far enough from the fault to constrain the total MDR, the bootstrap does not fully sample the allowed MDRs because it cannot sample the model null space (see also further discussion of this issue in section 2.5.2). For the strike-slip fault example, the coverage plot is better in the high-noise case (Figure 3d). Clearly, increasing noise in the data cannot improve the estimate of MDR. The source of the problem lies in the significant model null space for Figures 3c and 3d, leading to large biases in the maximum likelihood estimates (MLEs) for MDR. Analysis of the bootstrap PDFs compared to the distribution of MLEs for all 200 slip models shows that the bootstrap uncertainty is underestimated, and the effect is worse for the low-noise case, which is why the coverage plots are optimistic (see also Figure 2c). Put simply, in the low-noise case, the bias is large and the uncertainty estimated by the bootstrap is small. For the high-noise case, the bias is still large but the uncertainty is also large, so the bootstrap does better *on average*. For any realization (e.g., with actual data) the bootstrap distribution may be wildly inaccurate. Obviously, regularizing the problem would reduce, or eliminate, the model null space and change the bias in the bootstrap estimates. Our goal for this study is to consider estimates of the MDR apart from ad hoc regularization schemes; thus, we do not consider such approaches.

If the bootstrap is to be used to constrain MDR, synthetic tests with the fault and station geometry of interest and a variety of slip models should be performed to determine if the bootstrap will work for the network geometry in question, such as was done by *Murray and Segall* [2002]. In general it cannot be assumed a priori that the bootstrap method will correctly estimate the distribution of the MDR. When it fails the bootstrap can be wildly optimistic.

## 2.2. The *Johnson et al.* [1994] Approach

The approach of *Johnson et al.* [1994] to obtain bounds on the MDR is to modify the bounded problem (4), by adding an equality constraint on the MDR:

$$\begin{aligned} \min_{\mathbf{s}} \Phi_M(\mathbf{s}; \Sigma) &\equiv \|\Sigma^{-1/2}(\mathbf{d} - \mathbf{G}\mathbf{s})\|_2^2 \\ \text{s.t. } \mathbf{a}^T \mathbf{s} &= \dot{M}_{\text{test}} \\ \text{and } \mathbf{0} &< \mathbf{s} < \dot{\mathbf{s}}_{\infty} \end{aligned} \quad (6)$$

where  $\dot{M}_{\text{test}}$  is a fixed moment constraint. (Note that *Johnson et al.* [1994] minimized the L1 norm of the residuals, rather than the L2 norm as in (6).) Then, systematically sweep through the range of possible  $\dot{M}_{\text{test}}$  (from fully locked to fully creeping), to obtain a function of MDR,  $\Phi_M(\dot{M}_{\text{test}})$ . For a specified  $\dot{M}_{\text{test}}$ ,  $\Phi_M$  is the weighted L2 residual norm corresponding to the maximum likelihood solution for  $\mathbf{s}$  satisfying the constraints. The moment constraint provides for a unique solution to equation (6) for a given  $\dot{M}_{\text{test}}$ ; there may be an infinite number of SDR models that fit the data just slightly worse, but there are none that fit better. The curve  $\Phi_M(\dot{M}_{\text{test}})$  gives the *minimum misfit* as a function of MDR. For values of MDR with high misfit we can confidently assert that there is no SDR yielding the specified MDR that would fit the data better. This conclusion is independent of any assumed regularization of the SDR since none was applied.

Also note that while the bound constraints serve to reduce the uncertainty on MDR, they are not required in the formulation of equation (6), because the moment constraint is sufficient to provide a unique solution at each  $\dot{M}_{\text{test}}$ .

*Maurer and Johnson* [2014] applied this method to estimate the interseismic MDR on the creeping SAF, assuming that the misfits  $\Phi_M$  were drawn from a chi-square distribution. Bounds on the MDR were determined by finding the two points corresponding to the desired confidence level using  $\chi^2$  statistics.

This method fulfills three of our stated criteria from section 1.4 for an MDR estimator: it is convergent with mesh refinement, it is consistent, and models that fit the data equally well are equally likely. The problem is that if the nonlinear bound constraints on slip are active in (6), the residuals are not Gaussian, and  $\Phi_M$  will not be  $\chi^2$  distributed. Additionally, if the true data covariance and/or model errors are not properly characterized, the misfits may be systematic, non-Gaussian, and very large, so that no model fits the data acceptably well. It is often the case with SDR inversions that model errors dominate the misfit [e.g., *Duputel et al.*, 2014]. There is no obvious way to rigorously determine bounds on MDR under these circumstances, and thus the method can only be used qualitatively.



### 2.3. Estimating MDR From SDR Probability Distributions — “Naïve MCMC”

We noted in section 1 that many previous studies have developed Bayesian approaches for estimating SDR using various sampling algorithms such as Markov chain Monte Carlo (MCMC) methods (e.g., *Fukuda and Johnson* [2008] and *Minson et al.* [2013] use another sampling method).

A simple approach to obtaining a PDF on MDR is to compute the MDR for each SDR model obtained using a sampling method. The steps are as follows: first, generate posterior probability distributions on the SDR by sampling from

$$p(\dot{\mathbf{s}}|\mathbf{d}; \Sigma) \propto p(\mathbf{d}|\dot{\mathbf{s}}; \Sigma)p(\dot{\mathbf{s}}), \quad (7)$$

where  $p(\dot{\mathbf{s}}) \sim \mathcal{U}[0, \dot{s}_\infty]$  is the uniform prior distribution on SDR on each fault segment,  $p(\mathbf{d}|\dot{\mathbf{s}}; \Sigma)$  is the data likelihood function, which assuming normally distributed errors is proportional to  $\exp(-0.5(\mathbf{d} - \mathbf{G}\dot{\mathbf{s}})^T \Sigma^{-1}(\mathbf{d} - \mathbf{G}\dot{\mathbf{s}}))$ . The second step is to compute the MDR corresponding to each slip model using equation (3):  $\dot{M}_d = \mathbf{a}^T \dot{\mathbf{s}}$  for each SDR model  $\dot{\mathbf{s}}$  obtained via the sampling algorithm. This gives a histogram of MDR; confidence bounds can be estimated using this empirical distribution. We refer to this approach as “Naïve MCMC.”

#### 2.3.1. Testing Naïve MCMC

We used the simple APF geometry shown in Figure 1 to test the Naïve MCMC method. A single geodetic station is located a distance  $x/D = 0.55$ . Figure 4a shows several simulations with varying number of fault patches ( $N$ ). Note that the distribution on MDR becomes narrower and more biased toward the center of the bounds as  $N$  increases. The result may seem surprising: depending on  $N$ , the solution may completely exclude the true value of MDR, and for a given data set the solution is strongly mesh dependent.

#### 2.3.2. Why Does Naïve MCMC Fail?

Naïve MCMC seems like a very straightforward method, but consider what happens in the absence of data ( $p(\dot{\mathbf{s}}|\mathbf{d}; \Sigma) \propto 1$ ). Assume for simplicity that the patch area and prior bounds on SDR for every patch is the same; i.e.,  $a_i = GA/N$ , the total area divided by the number of patches. The prior covariance on SDR is assumed to be uncorrelated and designated by  $\sigma_s^2$ . Then the prior variance on MDR is given by

$$\sigma_M^2 = \text{Var} \left[ \sum_i a_i \dot{s}_i \right] = \sum_i a_i^2 \sigma_s^2 = \left( \frac{GA}{N} \right)^2 \sum_i \sigma_s^2 = \left( \frac{GA}{N} \right)^2 N \sigma_s^2 = G^2 \frac{A^2}{N} \sigma_s^2$$

i.e., the variances on MDR and SDR are not independent; they are linked by a factor that depends on the number of patch elements. This leads to mesh dependence in MDR if we fix the variance on SDR, as we have done by assigning bounds. Figure 4b shows the variance on MDR given a constant uniform prior on slip as a function of the number of patches on a fixed size fault. The theoretical variance of the MDR given no data are proportional to the *Bates Distribution*, the variance of the mean of  $N$  random uniformly distributed variables.

What is needed to eliminate mesh dependence? Consider the variance of MDR for two discretization schemes:  $\sigma_M^2 = N_1 a_1^2 \sigma_{s,1}^2$  and  $\sigma_M^2 = N_2 a_2^2 \sigma_{s,2}^2$ . Note that  $N_1 a_1 = N_2 a_2 = A$ , the total fault area. Then for mesh independence,

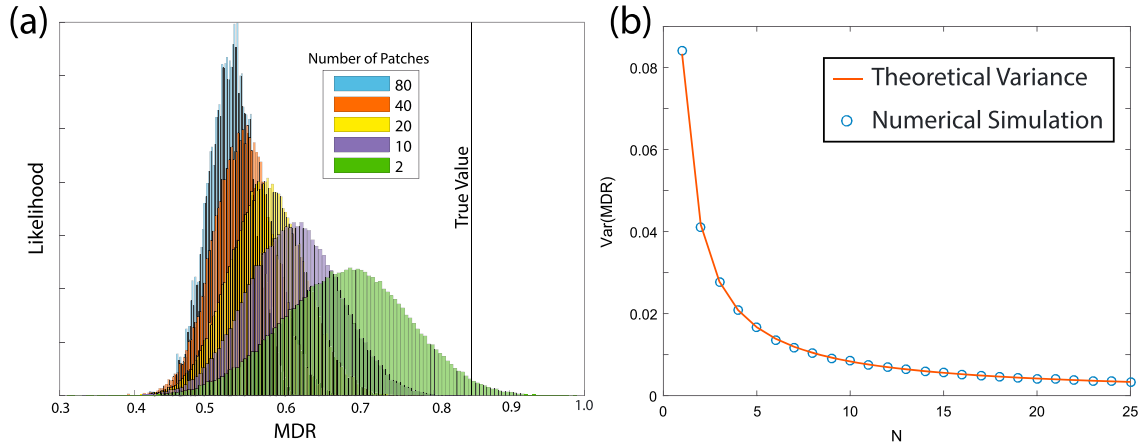
$$a_1 \sigma_{s,1}^2 = a_2 \sigma_{s,2}^2$$

implying that  $\frac{a_1}{a_2} \sigma_{s,1}^2 = \sigma_{s,2}^2$ . Thus, if, for example, mesh 2 is finer ( $a_2 < a_1$ ), then the prior variance on SDR in mesh 2 would need to be *larger*, by a factor of  $a_1/a_2$ , in order for the MDR variance to be mesh independent. This condition must be satisfied for mesh independence of any linear estimator, such as the Bayesian approach employed in this section.

#### 2.3.3. A Linear Bayesian Estimator — BEGS

In the Naïve MCMC method, we began by fixing bounds on slip rate, so the prior variance on MDR varies with  $N$ , leading to the mesh dependence. Consider instead fixing the prior variance on MDR and computing the corresponding variance on SDR; this solves the mesh-dependence problem, although one loses the point-wise bounds on SDR. In this section, we explore the use of Gaussian priors on slip rate with a variance that scales inversely with patch area; the use of Gaussian priors allows for an analytical solution [e.g., *Tarantola*, 2005]. The prior distribution on MDR is then given by

$$\dot{M}_d \sim \mathcal{N}(\mu_M, \sigma_M^2) = \mathcal{N}(\mathbf{a}^T \mu_s, \mathbf{a}^T \mathbf{S} \mathbf{a}) \quad (8)$$



**Figure 4.** (a) Synthetic tests of Naïve MCMC using the antiplane fault geometry (Figure 1) for different numbers of patches  $N$ .  $x/D = 0.55$  for this simulation. (b) Theoretical and simulated variance on MDR given no data.

where  $\mu_M$  and  $\sigma_M^2$  are the prior mean and variance for MDR and  $\mathbf{s} \sim \mathcal{N}(\mu_{\mathbf{s}}, \mathbf{S})$  is the (normal) prior distribution on slip. This gives  $\sigma_M^2 = \mathbf{a}^T \mathbf{S} \mathbf{a}$ . To illustrate, suppose that  $\mathbf{S}$  is diagonal. To determine  $(\sigma_{\mathbf{s}})_i^2$ , the  $i$ th diagonal entry of  $\mathbf{S}$ , note that

$$\sigma_M^2 = \mathbf{a}^T \mathbf{S} \mathbf{a} = \sum_i (\sigma_{\mathbf{s}})_i^2 a_i^2$$

For mesh independence, set

$$(\sigma_{\mathbf{s}})_i^2 = \frac{\sigma_M^2}{(\mathbf{e}^T \mathbf{a}) a_i}. \quad (9)$$

where  $\mathbf{e}$  is the vector of ones and  $\mathbf{e}^T \mathbf{a} = GA$ . For mesh independence, the prior variance on patch SDR is inversely related to the patch area.

Denote the posterior PDF on MDR

$$\hat{M}_d \sim \mathcal{N}(\hat{\mu}_M, \hat{\sigma}_M^2). \quad (10)$$

We show in Appendix A that

$$\begin{aligned} \hat{\mu}_M &= \mathbf{a}^T \mathbf{A}^{-1} \mathbf{c} \\ \hat{\sigma}_M^2 &= \mathbf{a}^T \mathbf{A}^{-1} \mathbf{a}, \end{aligned}$$

where  $\mathbf{A} \equiv \mathbf{S}^{-1} + \mathbf{G}^T \Sigma^{-1} \mathbf{G}$ , and  $\mathbf{c} \equiv \mathbf{G}^T \Sigma^{-1} \mathbf{d} + \mathbf{S}^{-1} \mu_{\mathbf{s}}$ . This result can be used to find the probability distribution on MDR given data, for Gaussian priors on SDR and MDR.

Correctly scaling the variance on SDR with mesh refinement eliminates mesh dependence. In addition, since we have specified all input probability distributions in terms of their mean and variance, the solution (10) is also the maximum entropy solution for MDR for Gaussian priors [e.g., Cover and Thomas, 2012, p. 254]. We refer to this method as the “Bayesian Estimator with Gaussian priors and Scaled variance” (BEGS).

BEGS solves the problem of mesh dependence, but does not make use of point-wise bounds on slip from geologic information. When bounds are known, the variance of the posterior distribution obtained using equation (10) is larger (i.e., less informative) than optimal. A simple example is a fault with prior variance equal to that of a uniform distribution between zero and  $\dot{M}_{\max}$  (fully locked):

$$\sigma_M^2 = \frac{(\dot{M}_d^{(\max)})^2}{12}$$

Suppose we now observe that MDR on half the fault is zero, but the other half of the fault is unobserved. The true posterior variance of MDR now is

$$\frac{(\dot{M}_d^{(\max)}/2)^2}{12} = \frac{1}{4} \frac{(\dot{M}_d^{(\max)})^2}{12} = \frac{1}{4} \sigma_M^2$$

i.e., a factor of 4 smaller than the original variance. However, for the Gaussian prior as assumed by (10), the posterior variance is

$$\sum_{i=1}^{N/2} (\sigma_{\dot{\mathbf{s}}})^2 a_i^2 = \frac{1}{GA} \sum_{i=1}^{N/2} a_i^2 \frac{\sigma_M^2}{a_i} = \frac{\sigma_M^2}{GA} \sum_{i=1}^{N/2} a_i = \frac{1}{2} \sigma_M^2$$

i.e., half the prior variance. Thus, the Gaussian prior is less informative than what is possible with bound constraints. To use bound constraints on both SDR and MDR and avoid mesh dependence requires a nonlinear estimator, which we consider in the next section.

#### 2.4. Constrained Optimization Bounding Estimator (COBE)

The Naïve MCMC estimator with fixed bounds on slip fails, because it is mesh dependent and biased with increasing number of patches. BEGS resolves this problem by specifying that variance on SDR scales with the number of patches but does not make use of point-wise bounds on SDR. The bootstrap and *Johnson et al.* [1994] estimators do not suffer from mesh dependence, but the bootstrap does not sample the model null space and the *Johnson et al.* [1994] method provides only qualitative bounds. We now consider a new approach to address all of these issues.

##### 2.4.1. Estimator Definition

Write the maximum likelihood function corresponding to equation (6) (minimizing misfit subject to bounds and a moment constraint):

$$\begin{aligned} p(\dot{M}_{\text{test}} | \mathbf{d}; \Sigma) &\equiv \max_{\dot{\mathbf{s}}} \text{lik}(\mathbf{d} | \dot{\mathbf{s}}; \Sigma) \\ \text{s.t.} \quad \mathbf{a}^T \dot{\mathbf{s}} &= \dot{M}_{\text{test}} \\ \text{and} \quad \mathbf{0} &\leq \dot{\mathbf{s}} \leq \dot{\mathbf{s}}_{\infty} \end{aligned} \quad (11)$$

or equivalently

$$\begin{aligned} p(\dot{M}_{\text{test}} | \mathbf{d}; \Sigma) &\equiv \max_{\dot{\mathbf{s}}} \text{lik}(\mathbf{d} | \dot{\mathbf{s}}; \Sigma) p_U(\dot{\mathbf{s}}; \mathbf{0}, \dot{\mathbf{s}}_{\infty}) \\ \text{s.t.} \quad \mathbf{a}^T \dot{\mathbf{s}} &= \dot{M}_{\text{test}} \end{aligned} \quad (12)$$

where  $p_U(\cdot; \mathbf{0}, \dot{\mathbf{s}}_{\infty})$  is the uniform PDF over  $[0, \dot{\mathbf{s}}_{\infty}]^N$ , the prior distribution on SDR. As often assumed, the data likelihood is Gaussian and parameterized by the data covariance matrix  $\Sigma$ . The solution to this problem is identical to that of the *Johnson et al.* [1994] method (equation (6)), because maximizing the likelihood is equivalent to minimizing the misfit subject to the constraints given. Note that when the data likelihood is uninformative (i.e.,  $\text{lik}(\mathbf{d} | \dot{\mathbf{s}}; \Sigma) = \text{constant}$  for all  $\dot{M}_{\text{test}}$ ), the likelihood function on MDR is uniform between the bounds.

With bounds on SDR, equation (12) is a nonlinear estimator just as with equation (6). We will show in section 2.4.3 that substituting a Gaussian prior for the uniform prior on SDR leads to a linear problem with a solution that is identical to the BEGS solution, the optimal solution for the linear problem. This provides some justification for the form of equation (12) in the uniform prior case.

The solution to equation (12) with uniform priors on slip is also the solution to equation (6), transformed by an exponential:

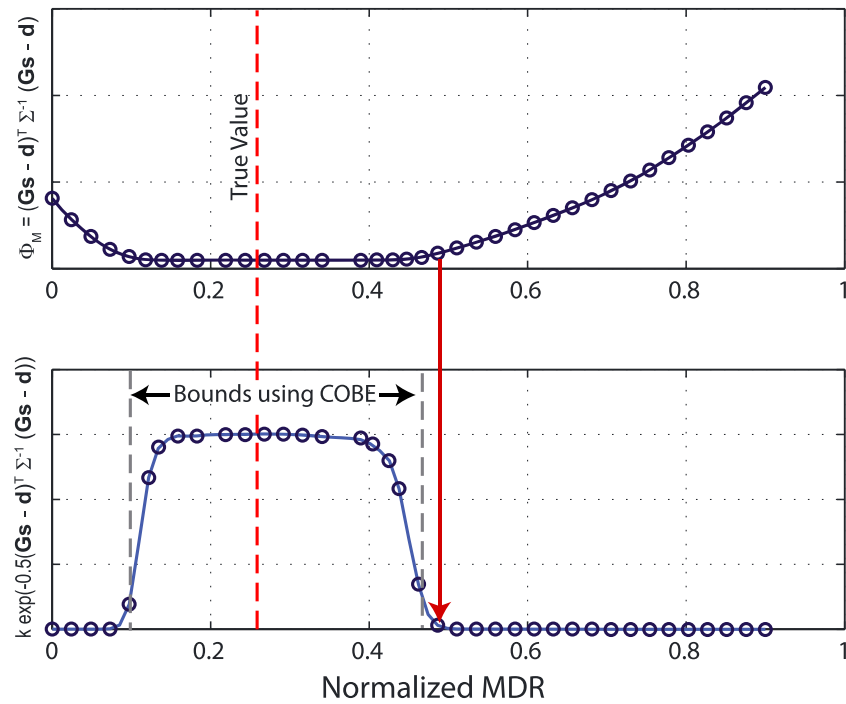
$$p(\dot{M}_d | \mathbf{d}; \Sigma) \propto e^{-\frac{1}{2} \Phi_M}, \quad (13)$$

where  $\Phi_M$  is the misfit for each MDR from equation (6). Figure 5 compares the original *Johnson et al.* [1994] method with the likelihood maximization method. Each point on the lower curve is the transformed misfit from the upper curve. Confidence bounds can be obtained, for example, using the cumulative distribution function (CDF) of the normalized PDF.

To summarize, for a suite of test values of MDR, solve the *Johnson et al.* [1994] system (equation (6)) subject to the bound constraints on slip and the moment constraint. Then use equation (13) and normalize to transform the set of misfits  $\Phi_M$  to the function  $p(\dot{M}_d | \mathbf{d}; \Sigma)$  (equation (12)). We refer to this method as the Constrained Optimization Bounding Estimator (COBE).

##### 2.4.2. Synthetic Tests of COBE

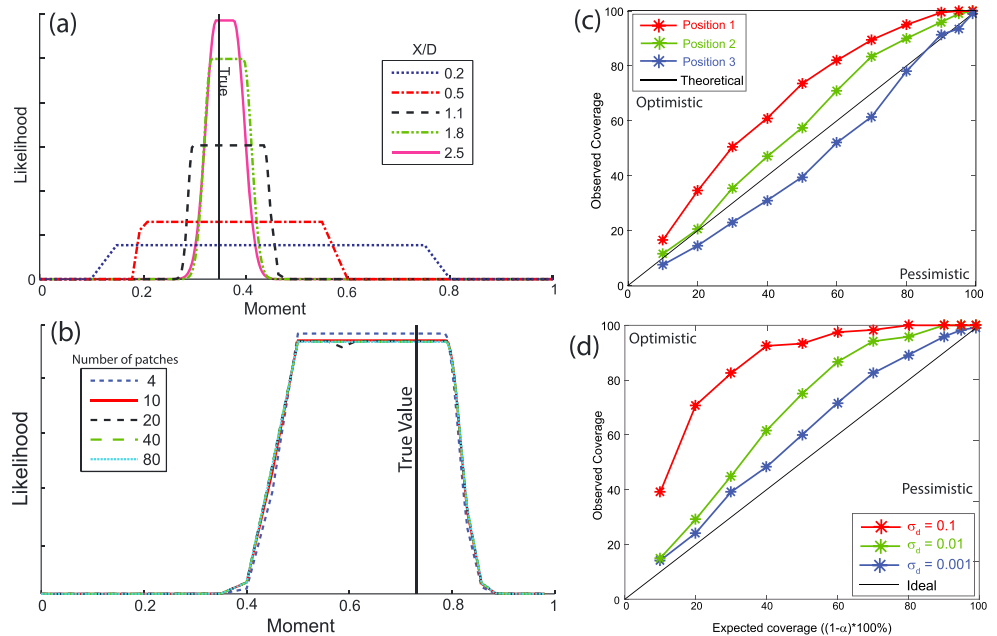
We tested COBE using a number of synthetic test geometries and fault configurations. Using the antiplane fault of Figure 1 and a single geodetic measurement, Figure 6a shows the performance for changing station distance  $x/D$ , but keeping the SNR roughly constant. The MDR estimate is more uncertain for stations closer to the fault, consistent with our earlier discussion, and different values of MDR that fit the data equally well are equally probable. Figure 6b shows results for several fault patch sizes, illustrating that COBE is mesh



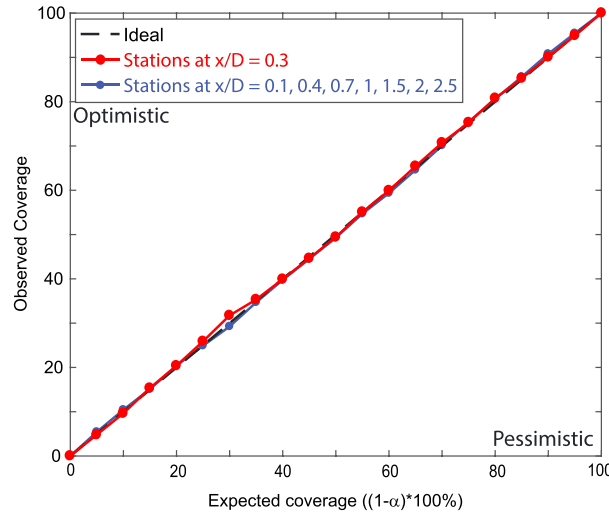
**Figure 5.** Example synthetic results using the *Johnson et al.* [1994] method and comparing to COBE. (Top plot)  $\Phi_M(\hat{M}_{test})$ . (Bottom plot)  $p(\hat{M}_d | \mathbf{d}; \Sigma) \propto \exp(-0.5\Phi_M(\hat{M}_{test}))$ .

independent, in the sense that the PDF converges as the mesh is refined. The COBE method is able to estimate MDR independent of fault element size while retaining point-wise bounds on SDR, unlike linear methods.

In addition to these single-station, antiplane fault tests, we also compare the performance of COBE to the bootstrap using the 3-D discretized fault geometries from Figure 3 with the same sets of 40 stations; results are



**Figure 6.** Synthetic test results for COBE. (a) Changing station distance for the fault geometry in Figure 1. SNR is approximately constant (i.e., the noise level decreases with distance from the fault) and is  $\approx 75$ . (b) Mesh refinement test. (c-d) Coverage results using finite faults shown in Figure 3a,c. (c) Changing station locations. (d) Changing noise level.



**Figure 7.** Coverage plots for COBLE/BEGS for a well-observed fault (many stations, including far field) and a poorly observed fault (single close station). The percentage of distributions that include the true value at the 95% CI are 94.9% (single station) and 95.4% (several stations).

shown in Figures 6c and 6d. COBE is more conservative than the bootstrap in all cases and does not exclude the true value at high confidence levels as the bootstrap does when MDR is poorly resolved. The bootstrap failed when the data did not strongly constrain MDR, because it cannot sample from the model null space. In contrast, when the data do provide strong constraints on MDR, COBE tends to be overly conservative while the bootstrap coverage plots are close to ideal. COBE searches the space of SDR models permitted by the data for all possible values of MDR, which includes the model null space, and uses the best fitting slip model to compute the likelihood at a given MDR.

To return to our simple thought experiment where half of the fault is known to have zero MDR, the variance given by COBE is exactly one fourth of the prior variance as we would like, because the COBE PDF is a

boxcar over the range  $[0, \dot{M}_{\max}/2]$ . COBE provides a mesh-independent estimate of MDR that can give uniform probability when bounds are known and thus makes use of point-wise bounds on the SDR from geologic information. Thus, it satisfies our criteria while also using as much information as possible.

Based on the synthetic tests and the comparison to the *Johnson et al.* [1994] estimator, we find that the COBE PDF satisfies the minimum criteria given in section 1.4: it is convergent with mesh refinement, it is consistent and conservative, and models that fit the data equally well are equally likely.

In the next section, we consider a related problem obtained by substituting Gaussian priors on slip and MDR in equation (12). This problem has an analytical solution, and we show that it is identical to the BEGS estimator, thus is the optimal solution for MDR for Gaussian priors.

### 2.4.3. The COB Linear Estimator (COBLE)

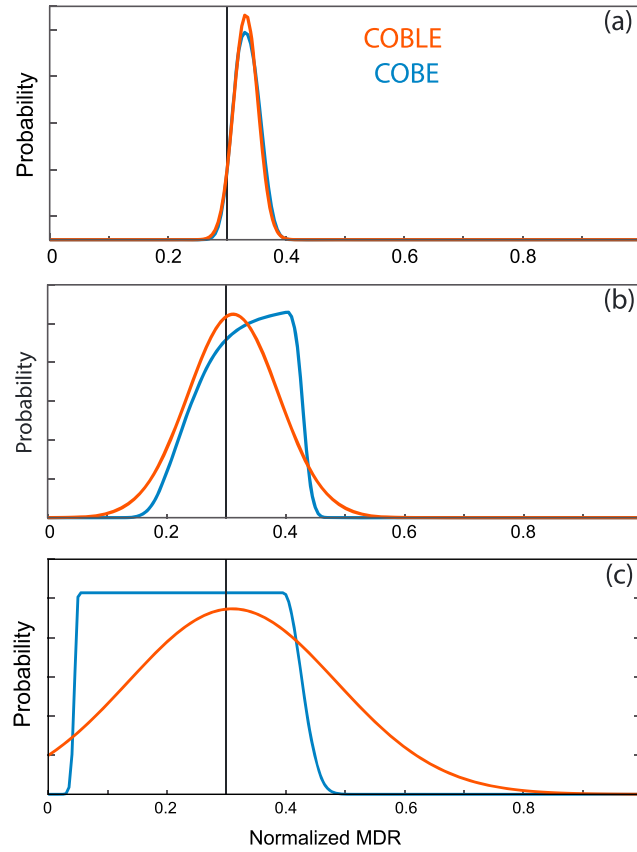
A simplification of COBE yields an estimator having a closed-form solution. Consider modifying (12) by changing the uniform prior on slip to a Gaussian prior:

$$\begin{aligned} \text{lik}(\dot{M}_d | \mathbf{d}) &= \max_{\dot{\mathbf{s}}} \text{lik}(\mathbf{d} | \dot{\mathbf{s}}) p_{\mathcal{N}}(\dot{\mathbf{s}}; \mu_{\dot{\mathbf{s}}}, \mathbf{S}) \\ \text{s.t. } \mathbf{a}^T \dot{\mathbf{s}} &= \dot{M}_{\text{test}} \end{aligned} \quad (14)$$

where now the prior  $p_{\mathcal{N}}(\cdot; \mu_{\dot{\mathbf{s}}}, \mathbf{S})$  is Gaussian, with prior mean ( $\mu_{\dot{\mathbf{s}}}$ ) and covariance ( $\mathbf{S}$ ), as for the linear Bayes estimator (BEGS) developed in section 2.3.3. Equation (14) differs from (12) only in the prior on SDR, which is Gaussian instead of uniform. We have already shown (section 2.3.3 and Appendix A) that BEGS is the solution for the linear problem. We prove in Appendix B that  $\text{lik}(\dot{M}_d | \mathbf{d})$  in equation (14) is exactly the BEGS solution (10) and that therefore, it is also the maximum entropy solution for MDR under Gaussian prior assumptions. Because it is a linear estimator and has the form of COBE except for the prior on slip rate, we refer to this method as the COB Linear Estimator, or COBLE. Figure 7 shows that COBLE has optimal coverage; we discuss further details of the coverage test for the linear estimator in Appendix C.

As noted, COBLE and COBE are identical except for the priors on SDR; therefore, (1) their PDFs should be similar when the MDR is well constrained (the data likelihood function is identical in both cases), and (2) differences between the two distributions (COBE versus COBLE) when the fault is not well observed are due to the difference in priors. Figure 8 compares three cases using the antiplane fault of Figure 1.

In the limit that the data strongly constrain the MDR, COBLE returns a Gaussian with mean and variance consistent with those of the data, while COBE returns essentially the same Gaussian but truncated to the bounds and normalized to integrate to 1. This situation is observed in Figure 8a: the moment is well constrained



**Figure 8.** COBE and COBLE for the Figure 1 geometry using (a) 15 stations up to  $3.5D$  from the fault, (b) 5 stations up to  $0.5D$  distance, and (c) one station at  $0.3D$ . Vertical black line is the true value. For each of these tests, average SNR  $\approx 10$  and the number of fault patches  $N = 20$ .

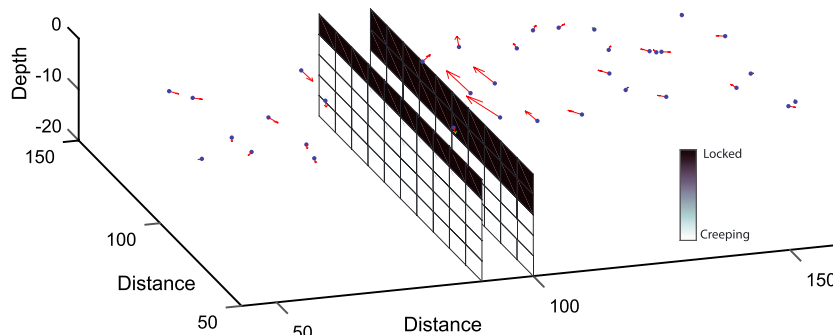
entire space of possible MDRs. COBE then varies both segments simultaneously, with two constraints in the optimization problem:

$$p(\hat{M}_1, \hat{M}_2 | \mathbf{d}; \Sigma) = \max_{\hat{\mathbf{s}}} \text{lik}(\mathbf{d} | \hat{\mathbf{s}}; \Sigma) p_{\mathcal{L}}(\hat{\mathbf{s}}; \mathbf{0}, \hat{\Sigma}_{\infty}) \quad (15)$$

$$\text{s.t. } [\mathbf{a}_1^T, \mathbf{0}^T] \hat{\mathbf{s}} = \hat{M}_1 \text{ and } [\mathbf{0}^T, \mathbf{a}_2^T] \hat{\mathbf{s}} = \hat{M}_2$$

where  $\hat{M}_1$  and  $\hat{M}_2$  are the moment constraints on segments 1 and 2, respectively.

Figure 9 shows an example with two parallel, closely spaced faults, both locked from the surface to a depth of 8 km (Fault 1) and 4 km (Fault 2) and creeping below to a depth of 20 km. The 35 stations randomly placed are used to solve for MDR using (15).

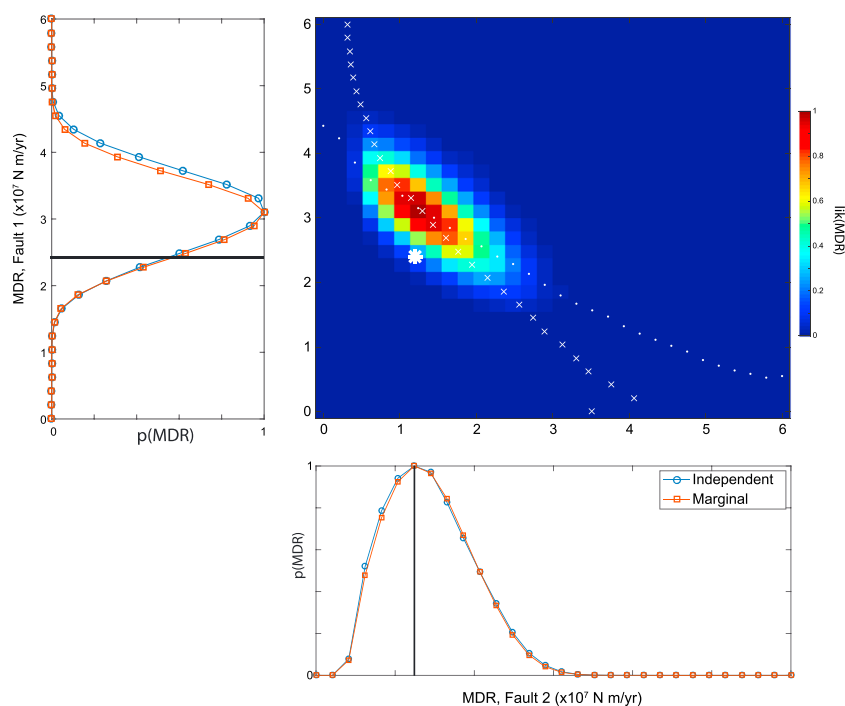


**Figure 9.** Example 2-fault geometry. Each fault is fully locked from the surface to a specified depth. Stations are randomly placed in this example.

by 15 stations at a range of distances including far-field stations, and COBLE and COBE are nearly identical. For Figure 8b, there are five stations all within half a transition depth of the fault, and the MDR is not as well constrained; some slight differences are now observed in the two PDFs due to their differing priors. Finally, in Figure 8c, there is only a single station located at  $0.3$  transition depths from the fault, the MDR is poorly resolved, and the differences between the two methods are clear. In the limit that no data is observed, COBE returns a uniform distribution between the bounds on MDR, while COBLE returns a Gaussian with the prior mean and variance.

#### 2.4.4. COBE and COBLE for Multiple Faults

The discussion so far has focused on a single fault, but most practical situations involve multiple faults or fault segments. It will be of interest to bound moment on segments individually, as well as determine correlations between segments. Consider a two-fault example: denote the patch area vector for fault 1 as  $\mathbf{a}_1$ , and similarly  $\mathbf{a}_2$  is the patch area vector for fault 2. The solution is to constrain MDR on both fault segments simultaneously and compute  $\Phi_M$  for the



**Figure 10.** COBE for two-fault geometry of Figure 9. Color plot shows the likelihood (normalized to have unit maximum). White star is the true MDR. Lines of white dots and crosses show how moment on one fault varies when the moment on only the other fault is fixed. White crosses are for Fault 1 fixed, white dots are for Fault 2 fixed. These are also shown in the marginal plots, denoted “Independent,” in comparison with marginal distributions obtained by integrating the full 2-D distribution.

Figure 10 shows the full 2-D distribution on MDR from (equation (15)) with MDR on each fault constrained separately (color plot and red lines in the marginal plots), and MDR constrained on one fault only (blue lines in the marginal plots). The white star in the color plot and vertical black lines in the marginal plots show the true value. The white crosses and dots on the color plot show the values of MDR from constraining one fault only, matching the corresponding blue lines in the marginal plots. The expected correlation between the MDR estimates for the two faults is clearly visible. The marginal and independent PDFs are equal within numerical error, as shown. Using the 2-D PDF, it is possible to find a confidence region that, for example, contains 95% of the PDF.

The two-fault example illustrates that the COBE method can be used to find correlations between faults, as well as the marginal PDFs on each segment. Implementing this for many faults would be a high dimensional grid search to obtain the high dimensional PDF on MDR for several faults. While this could be computationally challenging, each constrained inversion is completely independent and therefore the algorithm is trivially parallelizable and could scale to large systems. Furthermore, the COBLE estimator could be used as an initial step to determine the general region of high probability. Sampling from the COBLE PDF, evaluating COBE for each sample, and interpolating the result using a kernel method, for example, would give an efficient estimate of the COBE PDF.

### 2.5. Discussion and Conclusions on Methodology

We have analyzed previous methods used to estimate bounds on MDR and found that these are not fully satisfactory:

1. The bootstrap can fail when the data do not reasonably well constrain the moment.
2. The output of the *Johnson et al.* [1994] method is difficult to interpret because the bounds on MDR cannot be defined when unmodeled errors are present, which is the case with most real data, and additionally because bounds on slip-rate cause residuals to be nonnormally distributed.
3. Integrating SDR distributions obtained by sampling methods such as MCMC implicitly assumes that the prior variance on SDR is fixed; the result is mesh dependence and inaccurate solutions for MDR.

Regarding methods that do give satisfactory performance,

1. Correctly scaling the variance on SDR as a function of prior variance on MDR and patch size solves the problem of mesh dependence for the Bayesian approach and leads to the COBLE/BEGS method. Assuming Gaussian priors instead of uniform gives an analytical solution, a perfect coverage plot, and the maximum entropy solution for MDR assuming Gaussian prior distributions. However, this method does not make use of our knowledge of geologic bounds on SDR.
2. COBE can be shown to empirically satisfy our criteria for a reliable estimator using synthetic tests. COBE and COBLE differ only in their imposed priors on MDR and SDR, so when the data strongly constrain MDR, COBE is also the COBLE solution. When the data do not provide strong constraints on MDR, COBE provides conservative bounds.

### 2.5.1. Issues With the COBE Method

Parameters that strongly affect the solution obtained with COBE are the data covariance matrix and the bounds on SDR. We show some examples of how the solution changes with changing bounds on slip rate in the application section on Parkfield. Since bounds on slip-rate are always imperfectly known, one solution is to construct a suite of distributions using COBE with samples drawn from the distribution of long-term slip rates  $\dot{s}_\infty$ , which can then be integrated to give the final distribution on MDR. The estimator, as currently implemented, does not account for viscoelastic effects that could lead to a time-dependent SDR through the earthquake cycle [Johnson, 2013b].

A second issue is imperfect knowledge of the data covariance matrix  $\Sigma$ , which directly influences the likelihood, equation (6). The benefit of nonparametric methods such as the bootstrap is that the distribution of observational errors need not be specified in advance. Underestimating data uncertainty, or excluding uncertainty arising from the Earth model or transition depth all have significant impact on the estimated MDR. A detailed analysis of modeling errors is beyond the scope of this paper [see, e.g., Duputel et al., 2014]. Ideally, one could obtain an estimate of the data uncertainty independently of any geophysical model [e.g., Dmitrieva et al., 2015].

Our focus in this study has been to derive bounds on MDR with as few nondata-driven constraints as possible, resulting in the broadest possible PDFs that fit the data subject to the specified fault geometry and imposed long-term slip rate. There are clearly SDR models that are physically implausible, but nevertheless fit data acceptably well. A practical approach to deal with this important issue might be to present results constrained only by data, such as COBE, in addition to results including increasingly restrictive physical constraints. We leave it to future work to implement more realistic physical constraints that also honor the data, although note the work of Johnson [2013a].

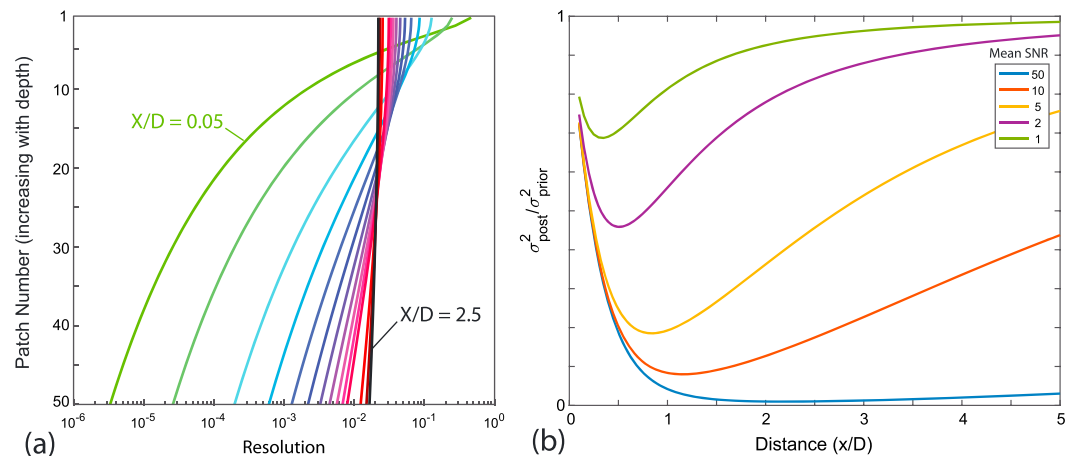
There is also uncertainty in how much of the present-day MDR maps to future seismic moment release as earthquakes. Other factors, such as postseismic moment release, creep, fraction of the moment deficit released in the last earthquake, and off-fault deformation need to be considered for a total moment budget analysis.

### 2.5.2. Optimal Geodetic Station Geometry for Estimating Moment

When do geodetic data provide strong constraints on the MDR? This question was investigated by Ziv et al. [2013] for the SAF near Parkfield, CA. That study showed (1) the GNSS stations around Parkfield are geometrically suboptimal for estimating the MDR, and (2) examples of random networks that have better resolution properties in all cases have stations located farther away from the fault than the Parkfield network they considered. Finally, (3) they showed that a single station can resolve the total MDR if that station is not too close to the fault.

Similar to their analysis, we use the APF geometry of Figure 1 with  $N = 50$  patches and a suite of station distances to test the sensitivity of geodetic measurements to MDR for various distances to the fault (Figure 11). We compute the singular value decomposition (SVD) of the Green's function matrix  $\mathbf{G} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  and the resolution matrix  $\mathbf{R} = \mathbf{V}_1\mathbf{V}_1^T$ , where  $\mathbf{V}_1$  is the singular vector spanning the row space of  $\mathbf{G}$  (the dimension of the model null space  $N - 1 = 49$ ). (Note that the SVD only makes sense in the context of linear vector spaces, which do not exist when bound constraints on SDR are active.) The diagonal of  $\mathbf{R}$  is plotted with patch depth in Figure 11a, which gives an indication of how well each slip element is resolved. For  $x/D$  small the diagonal element of  $\mathbf{R}$  for the top patch is close to one, but for deeper elements  $\text{diag}(\mathbf{R})$  is very small; conversely, for  $x/D$  large the diagonal elements are small but nearly constant. This is because when the station is far from the fault, every patch contributes to the data, while for a close station the contribution to the data from deep slip is





**Figure 11.** (a) Diagonal of the resolution matrix for the fault geometry in Figure 1, plotted as a function of patch number (increasing with depth). Lines are colored by distance  $x/D$ , which varies between 0.05 and 2.5 as shown. (b) Examples of the ratio of posterior to prior variance for MDR as a function of distance for a fully locked dislocation model estimated using COBLE/BEGS. Noise level is fixed, so actual SNR changes with distance.

very small. Since MDR is proportional to average SDR, a station close to the fault provides very little constraint on SDR at depth, and consequently the uncertainty in MDR is high regardless of the SNR of the data.

We also show the ratio of the posterior variance to the prior variance for MDR using COBLE and a single station (Figure 11b). Different lines are for different noise levels (average SNR); note that actual SNR decreases with increasing station distance. There is a unique distance for a single measurement and a given noise level that provides the minimum uncertainty in MDR, due to the competing effects of improving resolution with distance but decreasing signal. Note that in this figure, the noise level is fixed and the slip model is a dislocation (i.e., the fault is fully locked), so the measurement with the largest SNR is the one closest to the fault. However, it is actually measurements that are farther from the fault ( $x/D$  of order unity) and hence have a lower actual SNR that provide maximum resolution of the MDR in this case, consistent with Ziv *et al.* [2013].

In areas such as Southern California that have many parallel faults and the ocean in close proximity, it may be impossible to have geodetic measurements that resolve the MDR on a single fault. In this case, the bootstrap method, which requires stations in the far field for accuracy, cannot be used, but the COBE and COBLE methods can.

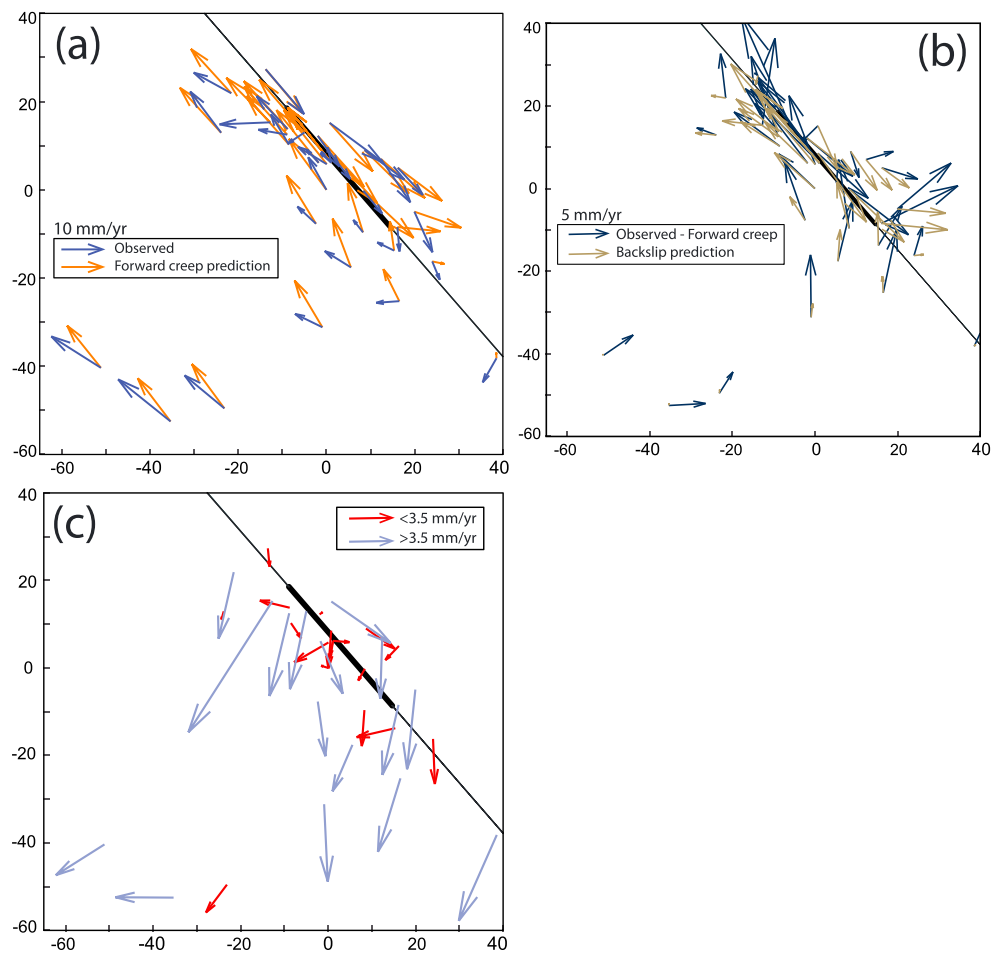
### 3. Application to Parkfield

We now employ the different methods discussed above to estimate the MDR of the Parkfield section of the San Andreas Fault (SAF), a segment well known for the repeated occurrence of magnitude 6 earthquakes for the past 150 years. It has been well instrumented for decades and several studies have looked at interseismic MDR and coseismic moment release. For the purpose of illustrating the methods described in this study, we use the data and fault model of Murray *et al.* [2001] to estimate the interseismic MDR for the time period spanning 1991–1998.

#### 3.1. Data and Model

The data consist of 35 GPS stations in a North America reference frame (data processing is fully described in Murray *et al.* [2001]). We subtract a reference velocity to put the velocities in a fault-relative reference frame, by interpolating velocities to a point on top of the model fault and subtracting the estimate from the rest of the velocity field. The result is shown in Figure 12a.

The fault model consists of 87 patches: two patches 14 km deep by 100 km long each, on either side of the Parkfield segment and representing the locked Carrizo plain section of the SAF to the southeast and the creeping section of the SAF (CSAF) to the northwest. One large patch 1000 km long and deep is located beneath the SAF to model the deep slip rate on the fault. The Parkfield segment is discretized into 84 patches, 7 deep and 12 long, each approximately  $2 \times 3$  km, for a total depth of 14 km and a width of 36 km, and total area 504 km<sup>2</sup>.

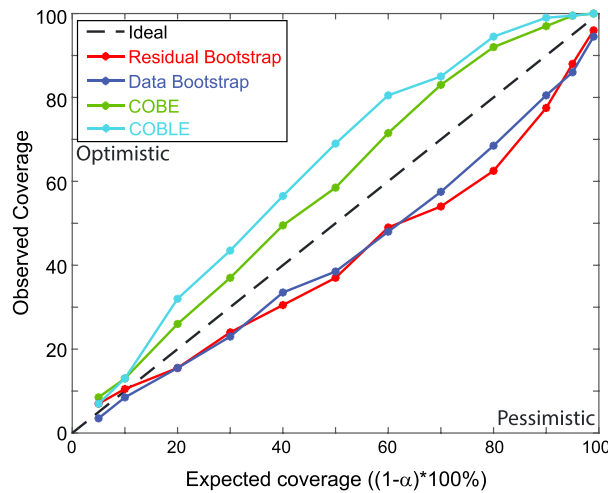


**Figure 12.** Fault and GNSS station geometry at Parkfield. Black line is the model fault trace; thicker portion is the Parkfield segment. Axes are in km. (a) Observed and predicted velocities for forward creep on all fault segments as described in the main text. (b) Remaining velocities after subtracting the forward prediction from the observed data in (a), and an example prediction from solving for back slip on the Parkfield segment, using an LTSR of 32 mm/yr. (c) Residuals after subtracting the back-slip prediction. 3.5 mm/yr ( $1\sigma$ ) arrows shown for scale.

To estimate MDR, we first subtract the predicted velocities due to forward creep on the entire fault model (Figure 12a). We assume creep at the long-term slip rate (LTSR) on the deep patch and the Parkfield segment, creep at 25 mm/yr on the CSAF, and zero creep (fully coupled interface) on the Carrizo plain segment (identical to *Murray and Segall* [2002]). This leaves only the contribution due to back slip on the Parkfield segment in the velocity field. These velocities are shown in Figure 12b, along with an example velocity prediction for back slip on just the Parkfield segment; the residuals from the back-slip model are shown in Figure 12c.

*Murray and Segall* [2002] estimated a range of acceptable LTSRs and transition depths using a simple 2-D model, then sampled from that distribution before solving for slip on the Parkfield segment. They used the bootstrap method to estimate MDR for each LTSR/depth pair and then integrated over samples to get the final distribution on MDR (Figure 2 in their paper). We use a fixed transition depth of 14 km and sample LTSRs from a Gaussian with mean 32 mm/yr and standard deviation 2 mm/yr, corresponding to the marginal distribution from their study for a 14 km locking depth.

We use the same data covariance matrix used by *Murray and Segall* [2002], but set a minimum variance threshold of 1 mm/yr and scale the covariance matrix by a factor of 3.5 such that the residuals of the best fitting model have approximately unit variance. There is a significant fault-perpendicular component in the residual velocities that cannot be explained by right-lateral slip on the Parkfield fault (Figure 12c); some of this is likely due to fault-normal shortening [e.g., *Harris and Segall*, 1987]. Additional uncertainty stems from assuming a spatially uniform creep rate on the CSAF; others have shown that there may be variations



**Figure 13.** Coverage plots for the Parkfield geometry and a single slip model using the COBE, COBLE/BEGS, and bootstrap methods.

in the creep rate at depth along strike of the creeping segment [e.g., *Ryder and Burgmann, 2008; Maurer and Johnson, 2014; Jolivet et al., 2015*].

We generated coverage plots with the COBE, COBLE/BEGS, residual bootstrap, and data bootstrap methods for the Parkfield fault and station geometry, using a single slip model similar to the best fitting from *Murray and Segall [2002]* (Figure 13). The bootstrapping methods both underestimate uncertainty in MDR, while the COBE and COBLE methods have conservative estimates of uncertainty. The percentage of distributions that include the true value at the 95% confidence level are 99.5% (COBE), 99% (COBLE), 88% (residual bootstrap), and 86% (data bootstrap).

### 3.2. Results

We use the COBE method to estimate the MDR for a suite of LTSRs and weight each PDF by the probability of the corresponding LTSR (assuming a Gaussian with mean 32 mm/yr and standard deviation 2 mm/yr) to get the joint distribution on MDR and LTSR (Figure 14). The results show that MDR is positively correlated with LTSR, as expected, and the uncertainty in MDR estimated using COBE is slightly greater for lower LTSRs. The maximum likelihood estimate for MDR varies by nearly a factor of 2 for opposite end-member LTSRs of 28 and 36 mm/yr.

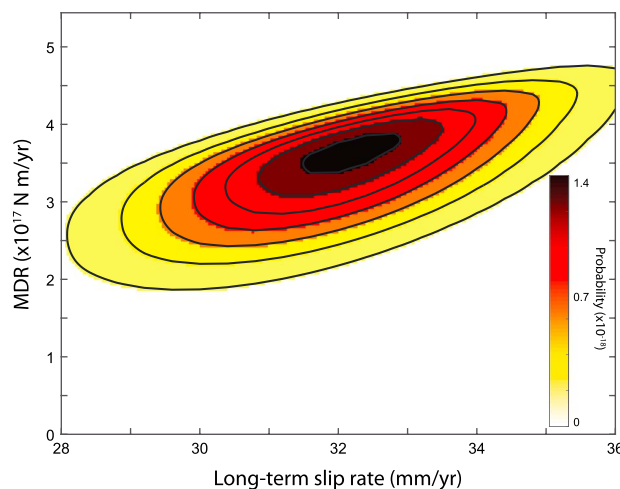
We also integrate over LTSR to get an overall estimate of MDR (Figure 15) using the following relation:

$$p(\dot{M}) = \int p(\dot{M}|\dot{s}_\infty)p(\dot{s}_\infty)d\dot{s}_\infty \quad (16)$$

where the integral is over the relevant range of possible slip rates.

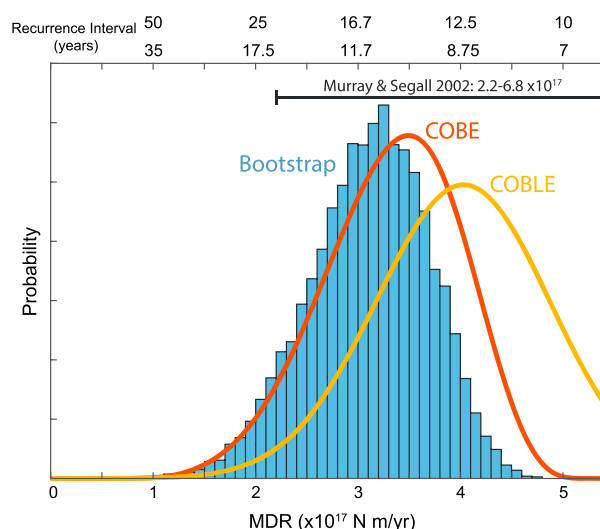
We show results for the residual bootstrap, COBE, and COBLE/BEGS, and compare to the original *Murray and Segall [2002]* results. For each sampled LTSR, the COBLE distribution assumes Gaussian priors on MDR and SDR, with a prior variance of  $\dot{M}_{\max}^2/12$ . For COBE and the residual bootstrap, slip is bounded on each fault patch to be between zero and the LTSR.

The results from COBE and COBLE are slightly more uncertain than the bootstrap, consistent with the synthetic coverage plot (Figure 13); the COBLE distribution in particular gives more probability to higher-MDR models.



**Figure 14.** Joint distribution for MDR and deep (long-term) slip rate, constructed by running COBE for a suite of long-term rates ( $32 \pm 4$  mm/yr) and multiplying the COBE PDF by the probability of each rate.

The maximum allowed MDR for our model at the highest LTSR (36 mm/yr) is  $5.44 \times 10^{17}$  N m/yr ( $M_w$  5.8), smaller than the 95% CI found by *Murray and Segall [2002]* ( $2.2\text{--}6.8 \times 10^{17}$  N m/yr [ $M_w$  5.5–5.9], shown by the horizontal black line in Figure 15), because they also sampled over deeper transition depths. Our 95% CI for annual MDR after integrating out LTSR is (in units of  $10^{17}$  N m/yr): (1.92–4.48,  $M_w$  5.5–5.8) (COBE), (1.9–4.1,  $M_w$  5.5–5.7) (bootstrap), and (2.33–5.44,  $M_w$  5.5–5.8)



**Figure 15.** Marginal distribution on the annual MDR obtained by integrating out long-term slip rate for COBE, COBLE/BEGS, and the residual bootstrap methods. Bottom axis is interseismic MDR in N m/yr; the top axis shows two sets of predicted recurrence times assuming a total moment during the 2004 earthquake of (lower numbers)  $3.5 \times 10^{18}$  N m ( $M_W$  6.3), and (top numbers)  $5 \times 10^{18}$  N m ( $M_W$  6.4). The horizontal black bar shows the 95% confidence bounds from Murray and Segall [2002], which goes off the plot to the right because that study included deeper transition depths. The observed recurrence time between the 1966 and the 2004 events is 38 years.

Langbein [2006] also modeled the 1966 event using a 14-km transition depth and found a slightly higher total moment ( $5.4 \times 10^{18}$  N m,  $M_W$  6.45) than the earlier work by Segall and Du [1993] using a 12-km transition depth and a slightly shorter fault model ( $4.4 \pm 0.4 \times 10^{18}$  N m,  $M_W$  6.4). Freed [2007] estimated the moment due to afterslip on the Parkfield segment for the time period 10 days to 2 years after the 2004 earthquake to correspond to a  $M_W$  6.3, assuming a transition depth of 21 km (but nearly all afterslip occurs above 15 km). His estimate gives a minimum value for the total moment (excluding the first 10 days after the earthquake) of  $\sim 4.3 \times 10^{18}$  N m ( $M_W$  6.4) [Freed, 2007].

We show the recurrence times in Figure 15 for a total coseismic plus postseismic moment of  $3.5 \times 10^{18}$  N m ( $M_W$  6.3, lower numbers), and  $5 \times 10^{18}$  N m ( $M_W$  6.4, upper numbers). For a recurrence time of 38 years (the observed interval between the 1966 and 2004 earthquakes) a total moment of  $3.5 \times 10^{18}$  N m and a simple time-predictable model gives an annual MDR of  $0.92 \times 10^{17}$  N m/yr ( $M_W$  5.3), whereas a total moment of  $5 \times 10^{18}$  N m gives an annual MDR of  $1.32 \times 10^{17}$  N m/yr ( $M_W$  5.4). COBE, COBLE, and bootstrap all exclude these values at greater than the 99% confidence level. To include the observed recurrence time at the 95% confidence level, assuming that the current interseismic MDR is steady throughout the cycle, would require a total moment during the 2004 Parkfield event of  $7.2 \times 10^{18}$  N m ( $M_W$  6.5), significantly more than inferred. There is a strong dependence of the MDR on the transition depth, which was fixed at 14 km for simplicity here; a shallower depth would reduce the estimated MDR but also reduce the estimate of the moment release on the fault. Furthermore, seismicity in this area occurs at depths up to about 15 km [e.g., Kim *et al.*, 2016]. This analysis confirms the Murray and Segall [2002] study that concluded the data are not consistent with a simple slip- or time-predictable model.

Johnson [2013a] suggested that the area of locked asperities accumulating stress can be much smaller than the area that slips at less than the LTSR (i.e., accumulate slip deficit). However, areas surrounding locked asperities are expected to release moment deficit aseismically, which should be accounted for in the afterslip estimates included above. A possibility is that the remaining slip deficit at Parkfield will not be made up until the next large, 1857-type earthquake, although the physics that would allow this are unclear.

(COBLE/BEGS, truncated at the maximum). The 99% CI are (1.51–4.72,  $M_W$  5.4–5.7) (COBE), (1.52–4.4,  $M_W$  5.4–5.7) (bootstrap), and (1.48–5.44,  $M_W$  5.4–5.8) (COBLE/BEGS, truncated at the maximum). The 99% range corresponds to approximately 28–100% of the maximum possible MDR given the range of LTSRs considered and a 14 km transition depth.

To estimate the predicted recurrence times for these MDR values, it is necessary to know the total moment due to an earthquake, including afterslip. A number of studies have estimated the total moment release during the 1966 and 2004 Parkfield earthquakes. Murray and Langbein [2006] estimate the moment for the 2004 earthquake plus the first 230 days of afterslip to be  $3.9 \times 10^{18}$  N m ( $M_W$  6.36) for a 14 km transition depth, with  $3.4 \times 10^{18}$  N m ( $M_W$  6.3) occurring on the part of the fault that was modeled by Murray and Segall [2002]. Murray and

#### 4. Conclusions

1. Geodetic measurements relatively far from the fault constrain total MDR; measurements close to the fault mainly resolve shallow slip.
2. The Bayesian Estimator with Gaussian priors and Scaled variance (BEGS), equivalently the Constrained Optimization Bounding Linear Estimator (COBLE), assumes Gaussian priors on SDR and MDR and scales the prior covariance on SDR. COBLE/BEGS give the optimal solution for the linear problem (without point-wise bounds on SDR), in that it is the maximum entropy estimate and has perfect coverage.
3. The Constrained Optimization Bounding Estimator (COBE) empirically meets the minimum criteria for a reliable estimator with uniform priors on SDR and MDR. COBE and COBLE are identical except for their respective priors on SDR and MDR, therefore when the data strongly constrain MDR the two PDFs are identical except at the bounds.
4. Assuming interseismic slip rate is everywhere bounded between locked and creeping, COBE reduces the uncertainty on MDR relative to the BEGS/COBLE solution.
5. Lacking strong data constraints, COBE yields conservative bounds on the MDR.
6. The bootstrap method performs well only when the data provide strong constraints on MDR. Synthetic tests are required to determine if the fault and network geometry are compatible with the bootstrap estimator.
7. COBE and COBLE can be extended to estimate covarying MDR on multiple fault segments to obtain estimates for both individual segments as well as correlations between segments.
8. Using COBE and comparing to COBLE/BEGS and the bootstrap methods, we estimate a moment deficit rate of about  $1.5-4.5 \times 10^{17}$  N m/yr on the Parkfield segment of the SAF, using the GPS data and fault model from Murray *et al.* [2001], to compare with MDR estimates from Murray and Segall [2002]. The estimate depends on the true long-term slip rate (LTSR) and can vary by almost a factor of 2 for LTSRs between 28 and 36 mm/yr. Coverage plots show that the station geometry at Parkfield is reasonable for constraining MDR. Data uncertainties were scaled due to large systematic fault-normal misfits; accounting for these requires estimation of fault-normal strain. Fault-parallel misfits are relatively small.
9. The inferred MDR at Parkfield exceeds the 2004 seismic moment (including afterslip), given the 38 year interseismic interval between 1966 and 2004.

In spite of the difficulties that remain, our study has shown that it is possible to eliminate problems in MDR estimates arising due to the choice of algorithm by adopting the COBE or COBLE method. Estimation of the MDR and associated uncertainty can and should become a routine part of fault hazard analysis and promises to play an important role along with paleoseismology and the historic earthquake record in constraining present-day seismic hazard.

#### Appendix A: Derivation of the Bayesian Linear Estimator

We seek a PDF on MDR conditioned on the data,  $p(\dot{\mathbf{M}}_d|\mathbf{d})$ , assuming Gaussian priors on SDR and MDR. We have  $\dot{\mathbf{s}} \sim \mathcal{N}(\mu_{\dot{\mathbf{s}}}, \mathbf{S})$ . Slip and data are related by  $\mathbf{d} = \mathbf{G}\dot{\mathbf{s}} + \varepsilon$ , where  $\varepsilon \sim \mathcal{N}(\mathbf{0}, \Sigma)$  is the Gaussian distribution on data uncertainty. From Bayes' Theorem, the PDF for SDR is given by

$$p(\dot{\mathbf{s}}|\mathbf{d}) = \frac{p(\mathbf{d}|\dot{\mathbf{s}})p(\dot{\mathbf{s}})}{p(\mathbf{d})} \quad (\text{A1})$$

In the following we use the rules for conditional and marginal distributions of normally distributed random vectors. First, the joint distribution on slip and data error is

$$\begin{bmatrix} \dot{\mathbf{s}} \\ \varepsilon \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_{\dot{\mathbf{s}}} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \Sigma \end{bmatrix}\right) \quad (\text{A2})$$

The distribution on the data is then

$$\mathbf{d} = \mathbf{G}\dot{\mathbf{s}} + \varepsilon = \begin{bmatrix} \mathbf{G} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{s}} \\ \varepsilon \end{bmatrix} \sim \mathcal{N}(\mathbf{G}\mu_{\dot{\mathbf{s}}}, \Sigma + \mathbf{G}\mathbf{S}\mathbf{G}^T) \quad (\text{A3})$$

Therefore, the joint distribution of slip and data is

$$\begin{bmatrix} \dot{\mathbf{s}} \\ \mathbf{d} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_{\dot{\mathbf{s}}} \\ \mathbf{G}\mu_{\dot{\mathbf{s}}} \end{bmatrix}, \begin{bmatrix} \mathbf{S} & \mathbf{S}\mathbf{G}^T \\ \mathbf{G}\mathbf{S} & \Sigma + \mathbf{G}\mathbf{S}\mathbf{G}^T \end{bmatrix}\right) \quad (\text{A4})$$

where for the off-diagonal entries of the covariance matrix,

$$\begin{aligned} \mathbb{E}[(\mathbf{d} - \mathbf{G}\mu_{\dot{\mathbf{s}}})(\dot{\mathbf{s}} - \mu_{\dot{\mathbf{s}}})^T] &= \mathbb{E}[(\boldsymbol{\varepsilon} + \mathbf{G}\dot{\mathbf{s}} - \mathbf{G}\mu_{\dot{\mathbf{s}}})(\dot{\mathbf{s}} - \mu_{\dot{\mathbf{s}}})^T] \\ &= \mathbf{G}\mathbf{S} \end{aligned}$$

Finally, using the rules for conditional probabilities ([e.g., Tarantola, 2005], see especially Tarantola's equations 3.37 and 3.38), we write the distribution on slip given data as

$$\dot{\mathbf{s}}|\mathbf{d} \sim \mathcal{N}(\mu_{\dot{\mathbf{s}}|\mathbf{d}}, \Sigma_{\dot{\mathbf{s}}|\mathbf{d}}) \quad (\text{A5})$$

where

$$\mu_{\dot{\mathbf{s}}|\mathbf{d}} = \mu_{\dot{\mathbf{s}}} + \mathbf{S}\mathbf{G}^T(\Sigma + \mathbf{G}\mathbf{S}\mathbf{G}^T)^{-1}(\mathbf{d} - \mathbf{G}\mu_{\dot{\mathbf{s}}})$$

and

$$\Sigma_{\dot{\mathbf{s}}|\mathbf{d}} = \mathbf{S} - \mathbf{S}\mathbf{G}^T(\Sigma + \mathbf{G}\mathbf{S}\mathbf{G}^T)^{-1}\mathbf{G}\mathbf{S}$$

The vector  $\mathbf{a}$  is the vector of patch areas times shear modulus, so the total area  $A = \mathbf{G}^{-1}\mathbf{a}^T\mathbf{e}$ , where  $\mathbf{e}$  is a vector of all ones. For convenience in what follows, we define

$$\begin{aligned} \mathbf{A} &\equiv \mathbf{S}^{-1} + \mathbf{G}^T\Sigma^{-1}\mathbf{G} \\ \mathbf{B} &\equiv \Sigma + \mathbf{G}\mathbf{S}\mathbf{G}^T \\ \mathbf{c} &\equiv \mathbf{G}^T\Sigma^{-1}\mathbf{d} + \mathbf{S}^{-1}\mu_{\dot{\mathbf{s}}} \end{aligned}$$

From the Sherman-Morrison-Woodbury formula,

$$\begin{aligned} \mathbf{A}^{-1} &= \mathbf{S} - \mathbf{S}\mathbf{G}^T\mathbf{B}^{-1}\mathbf{G}\mathbf{S} \\ \mathbf{B}^{-1} &= \Sigma^{-1} - \Sigma^{-1}\mathbf{G}\mathbf{A}^{-1}\mathbf{G}^T\Sigma^{-1} \end{aligned}$$

(Note that  $\mathbf{A}^{-1}$  corresponds to  $\tilde{\mathbf{C}}_M$  in equation 3.38 of Tarantola [2005]). Then  $\hat{M}_d = \mathbf{a}^T\dot{\mathbf{s}}$ , which is a linear transformation of  $\dot{\mathbf{s}}$ . Hence, the random variable  $\hat{M}_d|\mathbf{d} = \mathbf{a}^T(\dot{\mathbf{s}}|\mathbf{d})$  is also a Gaussian with mean and variance determined using the rules for affine transformation of a normal distribution:

$$\hat{M}_d \equiv \hat{M}_d|\mathbf{d} \sim \mathcal{N}(\mu_{\hat{M}_d|\mathbf{d}}, \Sigma_{\hat{M}_d|\mathbf{d}}) \quad (\text{A6})$$

where

$$\begin{aligned} \mu_{\hat{M}_d|\mathbf{d}} &= \mathbf{a}^T\mu_{\dot{\mathbf{s}}} + \mathbf{a}^T\mathbf{S}\mathbf{G}^T(\Sigma + \mathbf{G}\mathbf{S}\mathbf{G}^T)^{-1}(\mathbf{d} - \mathbf{G}\mu_{\dot{\mathbf{s}}}) \\ &= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{c} \end{aligned} \quad (\text{A7})$$

and

$$\begin{aligned} \Sigma_{\hat{M}_d|\mathbf{d}} &= \mathbf{a}^T(\mathbf{S} - \mathbf{S}\mathbf{G}^T\mathbf{B}^{-1}\mathbf{G}\mathbf{S})\mathbf{a} \\ &= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} \end{aligned} \quad (\text{A8})$$

Substituting these expressions into equation (10) gives the expression for the posterior on MDR.

## Appendix B: The COBLE PDF is Identical to the Bayes PDF

Now we show that  $\text{lik}(\hat{M}_d|\mathbf{d})$  in equation (14), the COBLE's likelihood function, is the same, up to normalization, as  $\hat{M}_d$  in equation (A6). The MLE problem again is

$$\begin{aligned} \max_{\dot{\mathbf{s}}} \quad & \text{lik}(\mathbf{d}|\dot{\mathbf{s}})p_{\mathcal{N}}(\dot{\mathbf{s}}; \mu_{\dot{\mathbf{s}}}, \mathbf{S}) \\ \text{s.t.} \quad & \mathbf{a}^T\dot{\mathbf{s}} = \hat{M}_d \end{aligned} \quad (\text{B1})$$

which can be rewritten in the following form:

$$\begin{aligned} \min_{\dot{\mathbf{s}}} \quad & R(\dot{\mathbf{s}}) \\ \text{s.t.} \quad & \mathbf{a}^T\dot{\mathbf{s}} = \hat{M}_d \end{aligned} \quad (\text{B2})$$

where

$$R(\dot{\mathbf{s}}) \equiv (\mathbf{G}\dot{\mathbf{s}} - \mathbf{d})^T\Sigma^{-1}(\mathbf{G}\dot{\mathbf{s}} - \mathbf{d}) + (\dot{\mathbf{s}} - \mu_{\dot{\mathbf{s}}})^T\mathbf{S}^{-1}(\dot{\mathbf{s}} - \mu_{\dot{\mathbf{s}}})$$

The Lagrangian corresponding to this problem (with a factor 1/2 for convenience) is

$$\mathcal{L}(\dot{\mathbf{s}}, y) \equiv \frac{1}{2}(\mathbf{G}\dot{\mathbf{s}} - \mathbf{d})^T\Sigma^{-1}(\mathbf{G}\dot{\mathbf{s}} - \mathbf{d}) + \frac{1}{2}(\dot{\mathbf{s}} - \mu_{\dot{\mathbf{s}}})^T\mathbf{S}^{-1}(\dot{\mathbf{s}} - \mu_{\dot{\mathbf{s}}}) + y(\mathbf{a}^T\dot{\mathbf{s}} - \hat{M}_d)$$

with Lagrange multiplier  $y$ . Taking derivatives with respect to  $\dot{\mathbf{s}}$  and  $y$  and setting the resulting equations equal to zero results in the Karush-Kuhn-Tucker (KKT) system for  $\dot{\mathbf{s}}$  and  $y$ :

$$\begin{bmatrix} \mathbf{A} & \mathbf{a} \\ \mathbf{a}^T & 0 \end{bmatrix} \begin{bmatrix} \dot{\mathbf{s}} \\ y \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \dot{M}_d \end{bmatrix} \quad (\text{B3})$$

where we have used  $\mathbf{c}$  and  $\mathbf{A}$  defined in Appendix A.

Using the relation for the inverse of a block matrix, the solution for  $\dot{\mathbf{s}}$  is

$$\dot{\mathbf{s}}^*(\dot{M}_d) = \mathbf{A}^{-1}(\mathbf{c} - \mathbf{a}(\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1}(\mathbf{a}^T \mathbf{A}^{-1} \mathbf{c} - \dot{M}_d)) \quad (\text{B4})$$

where we have made clear that  $\mathbf{s}^*$  is a function of  $\dot{M}_d$ . We then have

$$\text{lik}(\dot{M}_d | \mathbf{d}) \propto \exp\left(-\frac{1}{2}R(\dot{\mathbf{s}}^*(\dot{M}_d))\right). \quad (\text{B5})$$

Since  $R$  is a symmetric positive definite quadratic form in  $\dot{\mathbf{s}}$ ,  $\dot{\mathbf{s}}^*$  is an affine function of  $\dot{M}_d$ , and all random vectors are normal,  $\text{lik}(\dot{M}_d | \mathbf{d})$  is proportional to a normal distribution. Denote this normal distribution  $\hat{M}_d^{\text{COBLE}} \equiv \mathcal{N}(\mu_M, \sigma_M^2)$ , the PDF given by COBLE.

For the same reasons as above,  $R(\dot{\mathbf{s}}^*(M))$  can be written as scalar function  $f(M) = aM^2 + bM + k$ . If  $f(M)$  has exactly one zero, it can be written  $f(M) = a(M+b/(2a))^2 + k' = f(M) + k'$  for scalars  $a$  and  $b$  to be determined and constant  $k'$ . Since we have already established that the likelihood function is normal, the form of  $f(M)$  implies that in the normal distribution  $\hat{M}_d^{\text{COBLE}} \equiv \mathcal{N}(\mu_M, \sigma_M^2)$ , we can equate  $\mu_M = -b/(2a)$  and  $\sigma_M^2 = 1/a$ , for then  $R(\dot{\mathbf{s}}^*(M)) = a(M + b/(2a))^2 + k' = (M - \mu_M)^2 / \sigma_M^2 + k'$ .

To determine  $a$  and  $b$ , we first write

$$\mathbf{G}\dot{\mathbf{s}}^*(M) - \mathbf{d} = \mathbf{q} + rM$$

$$\dot{\mathbf{s}}^*(M) - \mu_{\dot{\mathbf{s}}} = \mathbf{v} + wM$$

for

$$\mathbf{q} \equiv \mathbf{G}(\mathbf{I} - \mathbf{A}^{-1}\mathbf{a}(\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1}\mathbf{a}^T)\mathbf{A}^{-1}\mathbf{c} - \mathbf{d}$$

$$r \equiv (\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1} \mathbf{G} \mathbf{A}^{-1} \mathbf{a}$$

$$\mathbf{v} \equiv (\mathbf{I} - \mathbf{A}^{-1}\mathbf{a}(\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1}\mathbf{a}^T)\mathbf{A}^{-1}\mathbf{c} - \mu_{\dot{\mathbf{s}}}$$

$$w \equiv \mathbf{A}^{-1}\mathbf{a}(\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1}$$

Substitute these into  $R$  and identify the result with  $f(M)$ :

$$\begin{aligned} R(\dot{\mathbf{s}}^*(M)) &= f(M) + k' \\ &= (\mathbf{q} + rM)^T \Sigma^{-1} (\mathbf{q} + rM) + (\mathbf{v} + wM)^T \mathbf{S}^{-1} (\mathbf{v} + wM) + k' \\ &= aM^2 + bM + k + k' \end{aligned}$$

From this equality we can determine  $a$ ,  $b$ :

$$\begin{aligned} a &= r^T \Sigma^{-1} r + w^T \mathbf{S}^{-1} w \\ &= (\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-2} \mathbf{a}^T \mathbf{A}^{-1} \mathbf{G}^T \Sigma^{-1} \mathbf{G} \mathbf{A}^{-1} \mathbf{a} + (\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-2} \mathbf{a}^T \mathbf{A}^{-1} \mathbf{S}^{-1} \mathbf{A}^{-1} \mathbf{a} \\ &= (\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-2} \mathbf{a}^T \mathbf{A}^{-1} (\mathbf{S}^{-1} + \mathbf{G}^T \Sigma^{-1} \mathbf{G}) \mathbf{A}^{-1} \mathbf{a} \\ &= (\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1} \end{aligned} \quad (\text{B6})$$

and

$$\begin{aligned} b/2 &= r^T \Sigma^{-1} \mathbf{q} + w^T \mathbf{S}^{-1} \mathbf{v} \\ &= (\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1} \mathbf{a}^T \mathbf{A}^{-1} [(\mathbf{S}^{-1} + \mathbf{G}^T \Sigma^{-1} \mathbf{G})(\mathbf{I} - \mathbf{A}^{-1}\mathbf{a}(\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1}\mathbf{a}^T)\mathbf{A}^{-1}\mathbf{c} - \mathbf{G}^T \Sigma^{-1} \mathbf{d} - \mathbf{S}^{-1} \mu_{\dot{\mathbf{s}}}] \\ &= (\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1} \mathbf{a}^T \mathbf{A}^{-1} [\mathbf{A}(\mathbf{I} - \mathbf{A}^{-1}\mathbf{a}(\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1}\mathbf{a}^T)\mathbf{A}^{-1} - \mathbf{I}]\mathbf{c} \\ &= -(\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a})^{-1} \mathbf{a}^T \mathbf{A}^{-1} \mathbf{c} \end{aligned} \quad (\text{B7})$$

Then  $\mu_M = -b/(2a)$  implies  $\mu_M = \mathbf{a}^T \mathbf{A}^{-1} \mathbf{c}$ , and  $\sigma_M^2 = 1/a$  implies  $\sigma_M^2 = \mathbf{a}^T \mathbf{A}^{-1} \mathbf{a}$ . Thus,  $\hat{M}_d^{\text{COBLE}} = \hat{M}_d$ .

## Appendix C: Coverage Test for the Linear Estimator

The synthetic coverage test for the linear estimator can be conducted as follows. First, sample from the distribution on slip  $\hat{\mathbf{s}} \sim \mathcal{N}(\boldsymbol{\mu}_s, \mathbf{S})$ . Sample the data errors  $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ , and form synthetic predicted data  $\mathbf{d} = \mathbf{G}\hat{\mathbf{s}} + \boldsymbol{\varepsilon}$ . Use equation (10) (equivalently, equation (14)) to compute  $\hat{M}_d$  and determine whether  $\hat{M}_d^{\text{true}} (= \mathbf{a}^T \hat{\mathbf{s}}^{\text{true}})$ , the true MDR, is within the centered CI  $\hat{\mu} \pm z\hat{\sigma}$  for the standard normal score  $z$  corresponding to a centered confidence level  $p$ . Do this for a suite of confidence levels, and repeat the entire process many times, accumulating these determinations for each  $p$ . Figure 7 shows that the COBLE/BEGS methods give perfect coverage, as we expect for a linear estimator with randomly sampled slip models.

### Acknowledgments

The authors thank Jessica Murray for providing the data and model for Parkfield and helpful advice, and Kaj Johnson for suggestions and insight on the methods we tested. We also thank Andreas Mavrommatis and Sarah Minson for helpful discussions during this project. Codes to replicate the synthetic tests presented here are freely available in the GitHub repository jlmaurer/mb-methods. Data related to the Parkfield study can be found in the relevant references. This work was supported by Proposal 14058 from the Southern California Earthquake Center. SCEC is funded by the NSF Cooperative Agreement EAR-0529922 and the USGS Cooperative Agreement 07HQAG0008. Sandia National Laboratories is a multi-timission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. The authors are not aware of any conflicts of interest from affiliations or funding sources.

### References

- Ampuero, J.-P., and F. A. Dahlen (2005), Ambiguity of the moment tensor, *Bull. Seismol. Soc. Am.*, *95*(2), 390–400, doi:10.1785/0120040103.
- Chuang, R. Y., and K. M. Johnson (2011), *Reconciling Geologic and Geodetic Model Fault Slip-rate Discrepancies in Southern California: Consideration of Nonsteady Mantle Flow and Lower Crustal Fault Creep*, vol. 39, pp. 627–630.
- Cover, T. M., and J. A. Thomas (2012), *Elements of Information Theory*, John Wiley, Hoboken, N. J.
- Diner, Ç., and A. Özgün Konca (2017), Moment tensor for seismic sources on a bimaterial interface: A hyperfunction approach, *Bull. Seismol. Soc. Am.*, *107*, 652–659, doi:10.1785/0120160044.
- Dmitrieva, K., P. Segall, and C. DeMets (2015), Network-based estimation of time-dependent noise in GPS position time series, *J. Geod.*, *89*(6), 591–606, doi:10.1007/s00190-015-0801-9.
- Duputel, Z., P. S. Agram, M. Simons, S. E. Minson, and J. L. Beck (2014), Accounting for prediction uncertainty when inferring subsurface fault slip, *Geophys. J. Int.*, *197*, 464–482, doi:10.1093/gji/ggt517.
- Efron, B., and R. J. Tibshirani (1994), *An Introduction to the Bootstrap*, CRC Press, New York.
- Evans, E. L., W. R. Thatcher, F. F. Pollitz, and J. R. Murray (2016), Persistent slip rate discrepancies in the eastern California (USA) shear zone, *Geology*, *44*, 691–694.
- Fialko, Y. (2006), Interseismic strain accumulation and the earthquake potential on the southern San Andreas Fault system, *Nature*, *441*(7096), 968–971, doi:10.1038/nature04797.
- Field, E. H., et al. (2014), Uniform California earthquake rupture forecast, version 3 (UCERF3): The time-independent model, *Bull. Seismol. Soc. Am.*, *104*(3), 1122–1180.
- Freed, A. M. (2007), Afterslip (and only afterslip) following the 2004 Parkfield, California, earthquake, *Geophys. Res. Lett.*, *34*, L06312, doi:10.1029/2006GL029155.
- Fukuda, J., and K. M. Johnson (2008), A fully Bayesian inversion for spatial distribution of fault slip with objective smoothing, *Bull. Seismol. Soc. Am.*, *98*(3), 1128–1146, doi:10.1785/0120070194.
- Harris, R. A., and P. Segall (1987), Detection of a locked zone at depth on the Parkfield, California, segment of the San Andreas Fault, *J. Geophys. Res.*, *92*(B8), 7945–7962, doi:10.1029/JB092iB08p07945.
- Hashimoto, C., A. Noda, T. Sagiya, and M. Matsu'ura (2009), Interplate seismicogenic zones along the Kuril-Japan trench inferred from GPS data inversion, *Nat. Geosci.*, *2*(2), 141–144.
- Hearn, E. H., F. Pollitz, W. Thatcher, and C. T. Onishi (2013), How do “ghost transients” from past earthquakes affect GPS slip rate estimates on Southern California faults?, *Geochem. Geophys. Geosyst.*, *14*, 828–838, doi:10.1002/ggge.20080.
- Johnson, H. O., D. C. Agnew, and K. Hudnut (1994), Extremal bounds on earthquake movement from geodetic data: Application to the Landers earthquake, *Bull. Seismol. Soc. Am.*, *84*(3), 660–667.
- Johnson, K. M. (2013a), Is stress accumulating on the creeping section of the San Andreas Fault?, *Geophys. Res. Lett.*, *40*, 6101–6105, doi:10.1002/2013GL058184.
- Johnson, K. M. (2013b), Slip rates and off-fault deformation in Southern California inferred from GPS data and models, *J. Geophys. Res. Solid Earth*, *118*, 5643–5664, doi:10.1002/jgrb.50365.
- Johnson, K. M., and J. Fukuda (2010), New methods for estimating the spatial distribution of locked asperities and stress-driven interseismic creep on faults with application to the San Francisco Bay area, California, *J. Geophys. Res.*, *115*, B12408, doi:10.1029/2010JB007703.
- Johnson, K. M., and P. Segall (2004), Viscoelastic earthquake cycle models with deep stress-driven creep along the San Andreas fault system, *J. Geophys. Res.*, *109*, B10403, doi:10.1029/2004JB003096.
- Jolivet, R., P. S. Agram, N. Y. Lin, M. Simons, M.-P. Doin, G. Peltzer, and Z. Li (2014), *Improving InSAR Geodesy Using Global Atmospheric Models*, vol. 119, pp. 2324–2341.
- Jolivet, R., M. Simons, P. S. Agram, Z. Duputel, and Z.-K. Shen (2015), Aseismic slip and seismicogenic coupling along the central San Andreas fault, *Geophys. Res. Lett.*, *42*, 297–306, doi:10.1002/2014GL062222.
- Johnson, K. M., Y.-J. Hsu, P. Segall, and S.-B. Yu (2001), Fault geometry and slip distribution of the 1999 Chi-Chi, Taiwan Earthquake imaged from inversion of GPS data, *Geophys. Res. Lett.*, *28*, 2285–2288, doi:10.1029/2000GL012761.
- Jónsson, S., H. Zebker, P. Segall, and F. Amelung (2002), Fault slip distribution of the 1999  $M_w$  7.1 Hector Mine, California, earthquake, estimated from satellite radar and GPS measurements, *Bull. Seismol. Soc. Am.*, *92*(4), 1377–1389, doi:10.1785/0120000922.
- Kim, W., T.-K. Hong, J. Lee, and T. Taira (2016), Seismicity and fault geometry of the San Andreas Fault around Parkfield, California and their implications, *Tectonophysics*, *677–678*, 34–44, doi:10.1016/j.tecto.2016.03.038.
- Kostrov, V. (1974), Seismic moment and energy of earthquakes, and seismic flow of rock, *Izv. Acad. Sci. USSR Phys. Solid Earth*, *1*, 23–44.
- Lindsey, E. O., V. J. Sahakian, Y. Fialko, Y. Bock, S. Barbot, and T. K. Rockwell (2014), Interseismic strain localization in the San Jacinto fault zone, *Pure Appl. Geophys.*, *171*, 2937–2954.
- Loveless, J. P., and B. J. Meade (2011), Spatial correlation of interseismic coupling and coseismic rupture extent of the 2011  $M_w = 9.0$  Tohoku-Oki earthquake, *Geophys. Res. Lett.*, *38*, L17306, doi:10.1029/2011GL048561.
- Maerten, F., P. Resor, D. Pollard, and L. Maerten (2005), Inverting for slip on three-dimensional fault surfaces using angular dislocations, *Bull. Seismol. Soc. Am.*, *95*, 1654–1665.
- Maurer, J., and K. Johnson (2014), Fault coupling and potential for earthquakes on the creeping section of the central San Andreas Fault, *J. Geophys. Res. Solid Earth*, *119*, 4414–4428, doi:10.1002/2013JB010741.
- McCaffrey, R. (2002), Crustal block rotations and plate coupling, in *Plate Boundary Zones*, *Geodyn. Ser.*, vol. 30, edited by S. Stein and J. Freymueller, pp. 101–122, AGU, Washington, D. C.



- McCaffrey, R., R. W. King, S. J. Payne, and M. Lancaster (2013), Active tectonics of northwestern U.S. inferred from GPS-derived surface velocities, *J. Geophys. Res. Solid Earth*, *118*, 709–723, doi:10.1029/2012JB009473.
- 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*, 1064–1075.
- Meade, B. J., and B. H. Hager (2005), Spatial localization of moment deficits in Southern California, *J. Geophys. Res.*, *110*, B04402, doi:10.1029/2004JB003331.
- Meade, B. J., and J. P. Loveless (2009), Block modeling with connected fault-network geometries and a linear elastic coupling estimator in spherical coordinates, *Bull. Seismol. Soc. Am.*, *99*(6), 3124–3139, doi:10.1785/0120090088.
- Minson, S. E., M. Simons, and J. L. Beck (2013), Bayesian inversion for finite fault earthquake source models I—Theory and algorithm, *Geophys. J. Int.*, *194*, 1701–1726, doi:10.1093/gji/ggt180.
- Murray, J., and J. Langbein (2006), Slip on the San Andreas Fault at Parkfield, California, over two earthquake cycles, and the implications for seismic hazard, *Bull. Seismol. Soc. Am.*, *96*(4B), S283–S303, doi:10.1785/0120050820.
- Murray, J., and P. Segall (2002), Testing time-predictable earthquake recurrence by direct measurement of strain accumulation and release, *Nature*, *419*(6904), 287–291, doi:10.1038/nature00984.
- Murray, J. R., P. Segall, P. Cervelli, W. Prescott, and J. Svarc (2001), Inversion of GPS data for spatially variable slip-rate on the San Andreas Fault near Parkfield, CA, *Geophys. Res. Lett.*, *28*(2), 359–362, doi:10.1029/2000GL011933.
- Murray, J. R., S. E. Minson, and J. L. Svarc (2014), Slip rates and spatially variable creep on faults of the northern San Andreas system inferred through Bayesian inversion of Global Positioning System data, *J. Geophys. Res. Solid Earth*, *119*, 6023–6047, doi:10.1002/2014JB010966.
- Nishimura, T., T. Hirasawa, S. Miyazaki, T. Sagiya, T. Tada, S. Miura, and K. Tanaka (2004), Temporal change of interplate coupling in northeastern Japan during 1995–2002 estimated from continuous GPS observations, *Geophys. J. Int.*, *157*(2), 901–916, doi:10.1111/j.1365-246X.2004.02159.x.
- Okada, Y. (1985), Surface deformation due to shear and tensile faults in a half-space, *Bull. Seismol. Soc. Am.*, *75*, 1135–1154.
- Page, M., S. Custodio, R. Archuleta, and J. Carlson (2009), Constraining earthquake source inversions with GPS data: 1. Resolution-based removal of artifacts, *J. Geophys. Res.*, *114*, B01314, doi:10.1029/2007JB005449.
- Prescott, W. H., J. C. Savage, J. L. Svarc, and D. Manaker (2001), Deformation across the Pacific-North America plate boundary near San Francisco, California, *J. Geophys. Res.*, *106*(B4), 6673–6682, doi:10.1029/2000JB900397.
- Ryder, I., and R. Burgmann (2008), Spatial variations in slip deficit on the central San Andreas Fault from InSAR, *Geophys. J. Int.*, *175*(3), 837–852, doi:10.1111/j.1365-246X.2008.03938.x.
- Savage, J., and R. Simpson (1997), Surface strain accumulation and the seismic moment tensor, *Bull. Seismol. Soc. Am.*, *87*(5), 1345–1353.
- Schmidt, D. A., R. Bürgmann, R. M. Nadeau, and M. d'Álessio (2005), Distribution of aseismic slip rate on the Hayward fault inferred from seismic and geodetic data, *J. Geophys. Res.*, *110*, B08406, doi:10.1029/2004JB003397.
- Segall, P., and Y. Du (1993), How similar were the 1934 and 1966 Parkfield earthquakes?, *J. Geophys. Res.*, *98*(B3), 4527–4538, doi:10.1029/92JB02408.
- Segall, P., and R. Harris (1986), Slip deficit on the San Andreas Fault at Parkfield, California, as revealed by inversion of geodetic data, *Science*, *233*(4771), 1409–1413, doi:10.1126/science.233.4771.1409.
- Segall, P., and R. Harris (1987), Earthquake deformation cycle on the San Andreas fault near Parkfield, California, *J. Geophys. Res.*, *92*(B10), 10,511–10,525, doi:10.1029/JB092iB10p10511.
- Suwa, Y., S. Miura, A. Hasegawa, T. Sato, and K. Tachibana (2006), Interplate coupling beneath NE Japan inferred from three-dimensional displacement field, *J. Geophys. Res.*, *111*, B04402, doi:10.1029/2004JB003203.
- Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Singapore.
- Tong, X., D. T. Sandwell, and B. Smith-Konter (2013), High-resolution interseismic velocity data along the San Andreas Fault from GPS and InSAR, *J. Geophys. Res. Solid Earth*, *118*, 369–389, doi:10.1029/2012JB009442.
- Ward, S. N. (1994), A multidisciplinary approach to seismic hazard in Southern California, *Bull. Seismol. Soc. Am.*, *84*(5), 1293–1309.
- Ward, S. N. (1998), On the consistency of earthquake moment rates, geological fault data, and space geodetic strain: The United States, *Geophys. J. Int.*, *134*(1), 172–186.
- Xue, L., S. Schwartz, Z. Liu, and L. Feng (2015), Interseismic megathrust coupling beneath the Nicoya Peninsula, Costa Rica, from the joint inversion of InSAR and GPS data, *J. Geophys. Res. Solid Earth*, *120*, 3707–3722, doi:10.1002/2014JB011844.
- Ziv, A., M.-P. Doin, and R. Grandin (2013), What can be learned from underdetermined geodetic slip inversions: The Parkfield GPS network example, *Geophys. J. Int.*, *134*, 1900–1908.