

# What can be learned from underdetermined geodetic slip inversions: the Parkfield GPS network example

A. Ziv,<sup>1</sup> M.-P. Doin<sup>2</sup> and R. Grandin<sup>2</sup>

<sup>1</sup>Geophysics and Planetary Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel.

E-mail: zivalon@tau.ac.il

<sup>2</sup>Laboratoire de Géologie, Ecole Normale Supérieure, 24 Rue Lhomond, 75231 Paris Cedex 5, France

Accepted 2013 May 20. Received 2013 March 13; in original form 2012 May 2

## SUMMARY

Often geodetic data are inverted for fault slip using less independent constraints than model parameters, and the solution is non-unique. That underdetermined geodetic slip inversions cannot provide unique slip distributions does not mean that they cannot provide other unique information regarding the slip distribution. In order to see which of the slip distribution attributes are obtainable by underdetermined inversions, we considered a synthetic GPS data set and inverted it for slip. We set the fault and network geometries to be identical to those of the Parkfield segment and the 14 SCIGN GPS sites next to it. We show that while slip inversions of such data yield robust estimate of the geodetic potency and the moment centroid, neither the spreadness nor the skewness may be resolved given the SCIGN network configuration. Furthermore, we show that randomly constructed networks are better configured than the Parkfield network, in the sense that they better recover the macroscopic attributes of the slip distribution. Finally, we show that the moment magnitude may be recovered using individual GPS stations, provided that these stations are not located in close proximity to the fault zone.

**Key words:** Inverse theory; Earthquake source observations; North America.

## 1 INTRODUCTION

Detailed mapping of seismic and aseismic slip distributions is important for addressing several fundamental questions in earthquake physics. For example, it may shed light on the mechanics of rupture initiation and arrest, the mechanics of aftershock triggering and stress transfer, and it may help to constrain the fault constitutive properties. Resolving fault slip on spatial and temporal scales that are of interest to earthquake physicists is a difficult task. Thus, earthquake slip inversion is the subject of intense research.

Often geodetic data are inverted for fault slip using less independent data than model parameters. In such cases, the solution is non-unique and the problem is ill conditioned. Ways to address this problem include the incorporation of non-geodetic data sets into the inversion (Custódio *et al.* 2009; Simons *et al.* 2011; Ziv 2012), addition of a non-negativity constraint (Page *et al.* 2009) and reduction of the model space (Page *et al.* 2009). Indeed, these approaches reduce the null-space (i.e. the number of unresolvable parameters) and narrow the range of possible solutions.

Geodetic data contain noise whose relative importance depends on various factors, including the size of the slip event and the distance between its centroid and the observation point. A well-known consequence of noisy data inversion is that the noise is mapped to the model space, the effect of which is to further enhance the non-uniqueness of the problem. A common approach to reduce the effect of noise is to constrain the solution to be smooth. Yet, the

use of smoothing is problematic as the roughness of the true slip distribution is not *a priori* known.

That underdetermined slip inversions of (noisy) geodetic data cannot provide unique slip distributions, does not mean that they cannot provide other unique information regarding the slip distribution. We would like to know which attributes of the slip distribution are resolvable, and what properties are in common to very different solutions that satisfy the data similarly well. To address these questions we constructed a synthetic GPS-like data set, inverted it for slip using different numerical approaches, and obtained a suite of solutions. We compared the solutions, and identified a set of uniquely determined parameters that characterize the slip distribution.

## 2 FORMULATION OF THE LINEAR INVERSE PROBLEM

We now describe the simplest and most common approach for static slip inversion using geodetic data. The scheme described here is applicable to cases where the fault geometry is known, and therefore the fault parameters may be pre-determined. In such cases the ground displacements are linearly related to the fault slip, and the inversion may be formulated as a linear inverse problem (Menke 1989).

The fault plane is discretized uniformly into a set of  $M$  rectangular dislocations (i.e. elements of uniform slip) embedded within

a uniform elastic half-space. Slip on the fault,  $u$ , causes ground displacement,  $d$ , at  $N$  sites according to

$$G_{ij}^{kl} u_j^l = d_i^k, \quad (1)$$

where  $G$  is an elastic kernel for dislocations (Okada 1985) relating a unit slip on dislocation  $j$  with ground displacement at site  $i$ . The index  $k$  stands for the three ground displacement components, that is, east, north and up, and the index  $l$  signifies the three components of fault slip, that is, strike-parallel, dip-parallel and fault-perpendicular. The problem is said to be underdetermined if  $N < M$ . Often the fault is known to slip at a certain direction. In such cases it is sensible to reduce  $l$  from three slip components to a single component, and the problem is underdetermined if  $3N < M$ .

If the data were noise-free or if the distribution of noise were to be the same at all sites and along all directions, the inversion would simply amount to solving the above equation for  $u$ . To account for the noise being site and direction dependent, it is useful to assign noise-dependent weights to each equation in (1) as follows:

$$W_i^k G_{ij}^{kl} u_j^l = W_i^k d_i^k, \quad (2)$$

with  $W$  being a weighting matrix whose diagonal entries are inversely proportional to the data uncertainty squared.

The solution vector contains spurious structures due to the mapping of noise from the data-space to the model-space. In order to suppress this effect, it is useful to impose smoothing. In this study, we use a Laplacian smoothing

$$\kappa \nabla^2 u_j^l = 0, \quad (3)$$

where  $\kappa$  is the smoothing coefficient. Apart from suppressing spurious structures due to noise, the addition of the smoothing constraint adds  $M$  equations, that is, it turns the problem from an underdetermined to an overdetermined one.

Solution to the overdetermined problem may be obtained using a variety of minimization schemes, whereas solution to the underdetermined problem is less straight forward and is done here in two ways. The first is by use of the non-negative least squares (NNLS) algorithm of Lawson & Hanson (1995), to get a solution to (2) or (2)–(3) subject to  $u \geq 0$ . The second is by multiplying either side of the set of equations by the generalized inverse (also known as the Moore–Penrose pseudo-inverse). The generalized inverse is calculated here by singular value decomposition (SVD) of  $G$  to  $U\Lambda V^T$ , where  $\Lambda$  is a diagonal matrix of the singular values, and  $U$  and  $V$  are orthogonal matrices. Following SVD and after setting to zero (or dumping) the reciprocal of unstable singular values, the generalized inverse is equal to:  $V\Lambda^{-1}U^T$ . Note that the number of non-zero singular values cannot exceed the number of equations, that is,  $3N$ . Use of this approach minimizes both the sum of the data residuals squared, and the length of the solution vector, which may include both positive and negative components.

In this study we solve for  $u$  in (2) subject to the smoothing constraint (3), using either the generalized inverse matrix or the NNLS algorithm. For each solution, the percentage of model-space and data-space residuals are calculated according to

$$\text{misfit percentage} = 100 \frac{\|D^{\text{syn}} - D^{\text{est}}\|}{\|D^{\text{syn}}\|}, \quad (4)$$

where  $D$  stands for the ground displacement vector when calculating the percentage of data-space residual, and it stands for the fault slip vector when calculating the percentage of model-space residual, the superscripts ‘syn’ and ‘est’ denote ‘synthetic’ and ‘estimated’, respectively, and the double vertical bars signify the  $L_2$  norm. Because the synthetic data that we use is constructed from a pure strike-slip

model (see further details in Section 3.2), we constrain the solution to be strike-parallel.

### 3 WHICH ATTRIBUTES OF THE SLIP DISTRIBUTION ARE RESOLVABLE?

#### 3.1 The set of slip distribution functions

The geodetic potency of a single slip event is equal to the integral of slip over the fault plane:

$$p = A \sum_j u_j, \quad (5)$$

where  $A$  is the elementary dislocation area, and the summation over the slip is vectorial. Provided that the slip is entirely seismic and is everywhere and at all times not flipping its direction, the seismic and geodetic potencies are exactly the same. In that case multiplying the geodetic potency by the rigidity,  $\mu$ , gives the seismic moment, which may then be used to compute the moment magnitude of the slip event according to (Hanks & Kanamori 1979):  $M_w = \frac{2}{3}[\log(\mu p) - 9.1] = \frac{2}{3}[\log(M_0) - 9.1]$ , where  $M_0$  is the seismic moment in Joules.

Another important attribute of the slip distribution is the moment centroid, whose components are

$$\begin{aligned} C_x &= \sum_j u_j x_j / \sum_j u_j \\ C_y &= \sum_j u_j y_j / \sum_j u_j, \end{aligned} \quad (6)$$

where  $x$  and  $y$  are the along-strike and downdip dislocation coordinates, respectively.

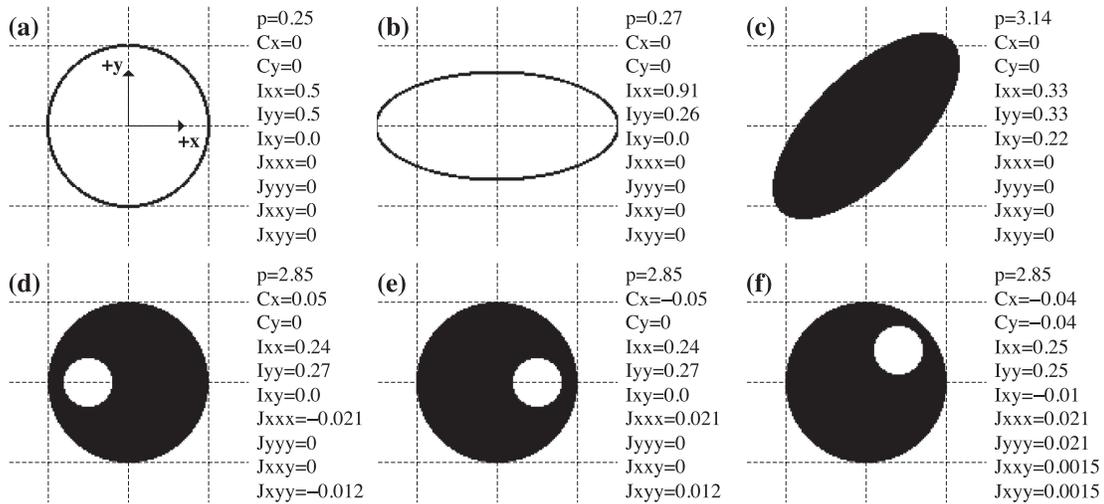
The extent to which the slip distribution is spread with respect to the centroid is quantified using the spreadness tensor, whose components are

$$\begin{aligned} I_{xx} &= \sum_j u_j (x_j - C_x)^2 / \sum_j u_j \\ I_{yy} &= \sum_j u_j (y_j - C_y)^2 / \sum_j u_j \\ I_{xy} &= \sum_j u_j (x_j - C_x)(y_j - C_y) / \sum_j u_j. \end{aligned} \quad (7)$$

Finally, the skewness of the slip with respect to the centroid is quantified using the followings:

$$\begin{aligned} J_{xxx} &= \sum_j u_j (x_j - C_x)^3 / \sum_j u_j \\ J_{yyy} &= \sum_j u_j (y_j - C_y)^3 / \sum_j u_j \\ J_{xxy} &= \sum_j u_j (x_j - C_x)^2 (y_j - C_y) / \sum_j u_j \\ J_{yyx} &= \sum_j u_j (y_j - C_y)^2 (x_j - C_x) / \sum_j u_j. \end{aligned} \quad (8)$$

Note that the spreadness and the skewness tensors may be thought of as analogous to the second and third moment tensors of inertia, respectively, whereas the potency (divided by fault area) and the centroid are equivalent to the total mass and the centre of inertia, respectively. While the potency and the moment centroids are commonly used and are easily grasped attributes of the



**Figure 1.** Sketches of several unrealistically simple slip distributions and their respective attributes. (a) Circular ring. (b) Elliptical ring. (c) Inclined ellipse. (d) Left skewed ring. (e) Right skewed ring. (f) Top-right skewed ring. The coordinate system is shown in panel (a).

slip distribution, the spreadness and the skewness are not. Thus in Fig. 1, in order to guide the intuition, we present sketches of several unrealistically simple slip distributions and their respective attributes.

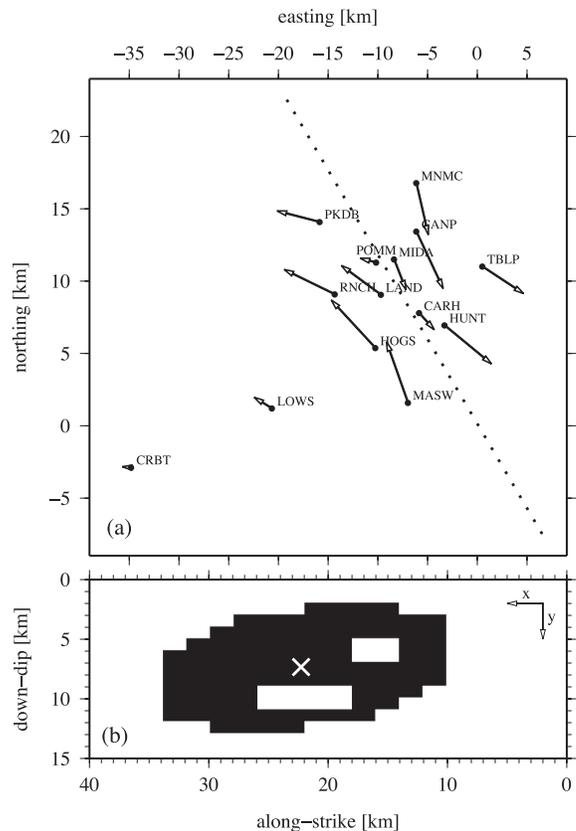
### 3.2 The synthetic data

In order to see which attributes of the slip distribution are resolvable and are shared by different solutions that produce very similar ground displacements, we construct a synthetic slip distribution and a synthetic ground displacement associated with it. We first set the fault and network geometries to be identical to those of the Parkfield segment and the 14 SCIGN GPS sites next to it (black circles in Fig. 2a). The obvious added value of generating a synthetic data set that is similar to that of the 2004 Parkfield earthquake is that the result of this study may be used to draw conclusions regarding one of the best studied earthquakes to date, and the very likely site of a future magnitude 6 earthquake (Bakun *et al.* 2005; Barbot *et al.* 2012). Next, after having assessed the degree to which the slip distribution functions listed in Section 3.1 may be recovered using the Parkfield network, we compare the Parkfield network with a set of randomly distributed networks, each of which consists of 14 GPS antennas.

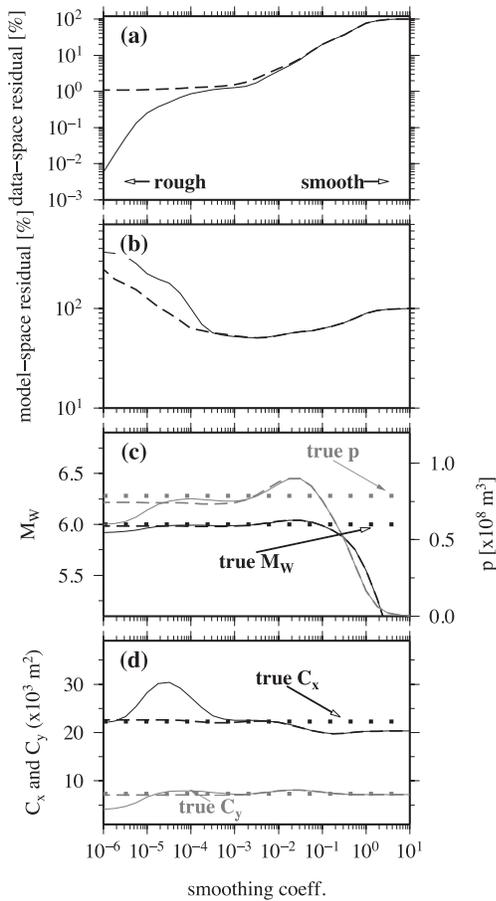
The synthetic slip distribution is shown in Fig. 2(b). Similar to the 2004 Parkfield earthquake, the moment magnitude of the synthetic earthquake is equal to 6 and the slip is right-lateral. The synthetic slip distribution consists of an inclined elliptical patch of uniform slip of about 0.4 m that is disrupted by two smaller patches of zero slip (i.e. barriers). While this slip distribution may differ from any of the past coseismic slip distributions along the Parkfield segment, it has the advantage of including both long- and short-wavelength features that are generally targeted by coseismic slip inversion resolution tests. We use eq. (1) to compute the ground displacements corresponding to the synthetic slip distribution. To the synthetic ground displacement we add uncorrelated Gaussian random noise. In generating the noise we use different standard deviation for each ground displacement component, and these are set to be equal to the corresponding average standard deviations reported on the Parkfield's webpage posted by the Scripps Orbit and Permanent Array Center (SOPAC).

### 3.3 Parkfield network configuration

We now examine consequences of the fault and network geometries that are identical to those of the Parkfield segment and the 14 SCIGN GPS sites next to it (Fig. 2a). Previous researchers have



**Figure 2.** The GPS network configuration and the synthetic slip distribution. (a) A map showing the synthetic network configuration, with the dark circles denoting the location of Parkfield's 14 GPS sites, and the dotted line indicating the surface trace of the model fault. (b) The synthetic slip distribution, with shaded area indicating an area of uniform slip amounting to 0.42 m and white cross indicating the centroid location. The model fault is 40 km long and 15 km wide.

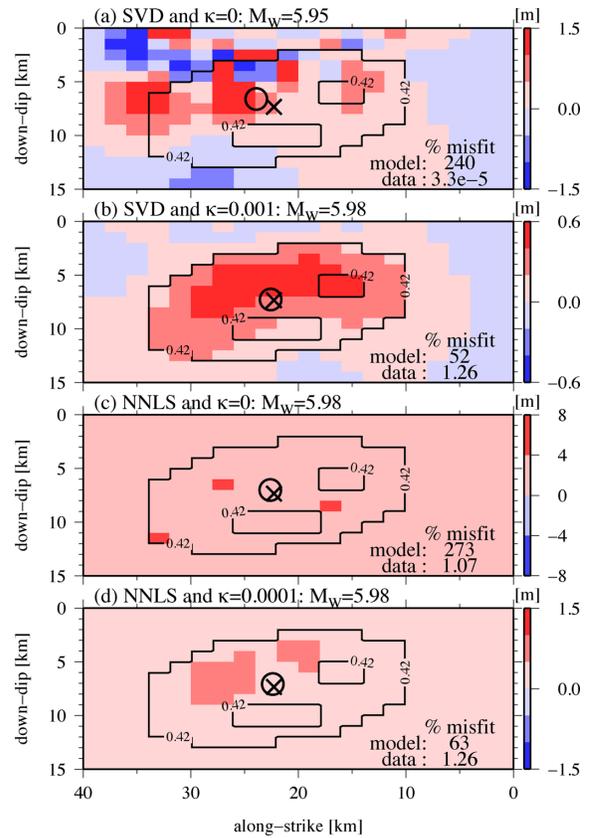


**Figure 3.** Parkfield's inversion results as a function of the smoothing coefficient using the NNLS algorithm (dashed curves) and the generalized inverse approach (solid curves). (a) Percentage of data-space residual. (b) Percentage of model-space residual. (c) Moment magnitude (black) and geodetic potency (grey). (d) Centroids. Horizontal dotted lines in panels (c) and (d) indicate the true synthetic values. Model-space and data-space residuals are calculated using eq. (4).

shown that this network cannot resolve the actual slip distribution below 3 km and near the segment ends (Barbot *et al.* 2009; Page *et al.* 2009; Ziv 2012). In Fig. 3, we show the percentage of data-space and model-space residuals as a function of the smoothing coefficient. The data misfit remains less than about 1 per cent over a range of smoothing for which the model-space residual varies between 55 and 400 per cent. That radically different slip distributions provide nearly equally good fit to the data is a well-known consequence of the problem being underdetermined and the data containing noise (Savage 1990; Du *et al.* 1992; Bos & Spakman 2003).

### 3.3.1 The geodetic potency and the centroid

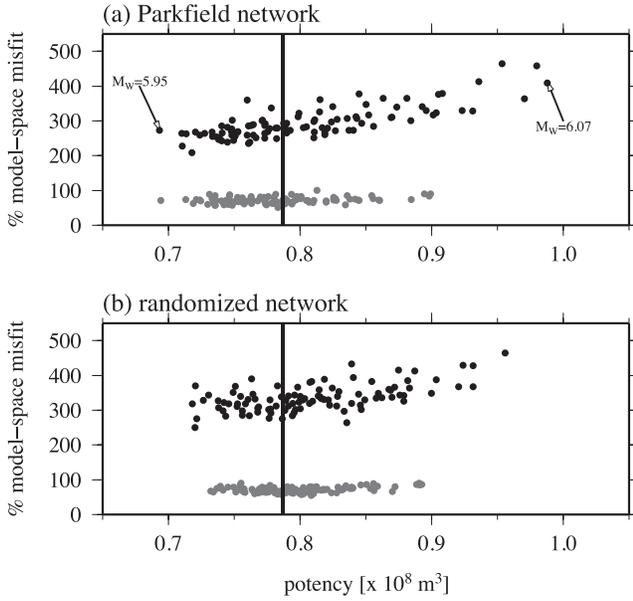
We find that the geodetic potency (and the moment magnitude) and the centroids remain nearly constant and nearly equal to the synthetic values over a wide range of smoothing (Figs 3c and d). Surprisingly, even solutions providing poor fit to the geodetic data, say 70 per cent, successfully recover these slip attributes. To visualize the extent to which radically different slip distributions provide not only practically equal good fit to the ground displacement, but also to the seismic moment and the centroid, we present in Fig. 4 a set of four different solutions for the slip distribution, all satis-



**Figure 4.** Example slip distributions satisfying at least 99 per cent of the data. (a) Solution obtained using the generalized inverse and  $\kappa = 0$  (i.e. no smoothing). (b) Solution obtained using the generalized inverse and  $\kappa = 10^{-3}$ . (c) Solution obtained using the NNLS algorithm and  $\kappa = 0$ . (d) Solution obtained using the NNLS algorithm and  $\kappa = 10^{-4}$ . Geodetic moment magnitude as well as the percentage of data-space and model-space residuals are indicated on each panel. The crosses and circles indicate the true and estimated moment centroids, respectively. The model fault is represented by an array of 300 dislocations, whose along-strike and down-dip dimensions are 2 and 1 km, respectively. The solid contour indicates the target (i.e. synthetic) slip distribution. Note the difference in colour code of different panels.

fying  $\geq 99$  per cent of the geodetic data. A set of two solutions is obtained using the generalized inverse matrix (Figs 4a and b), and another set of two solutions is obtained using the NNLS algorithm (Figs 4c and d). One solution in each set is subject to the smoothing constraint of (3), while the other is not. Note that slip distributions obtained using the generalized inverse are considerably smoother than those obtained using the NNLS algorithm. For example, the solution shown in Fig. 4(b) using the generalized inverse is notably smoother than that shown in Fig. 4(d) using the NNLS algorithm, despite the latter being subject to 10-fold stronger smoothing than the former. Although the four solutions presented here satisfy more than 99 per cent of the data, the misfit in the model-space varies widely between 54 and 492 per cent. Despite this great variability, the moment magnitudes of the 4 slip distributions (reported on each panel) are between 5.95 and 5.98, and a close match is found between synthetic and estimated centroids. Note that hereafter results will be obtained using the NNLS algorithm.

In order to see how the data uncertainties are mapped to uncertainties in the slip distribution attributes, we produce a set of Monte Carlo realizations of the data errors. As before, the synthetic errors are spatially uncorrelated and their distribution is Gaussian,

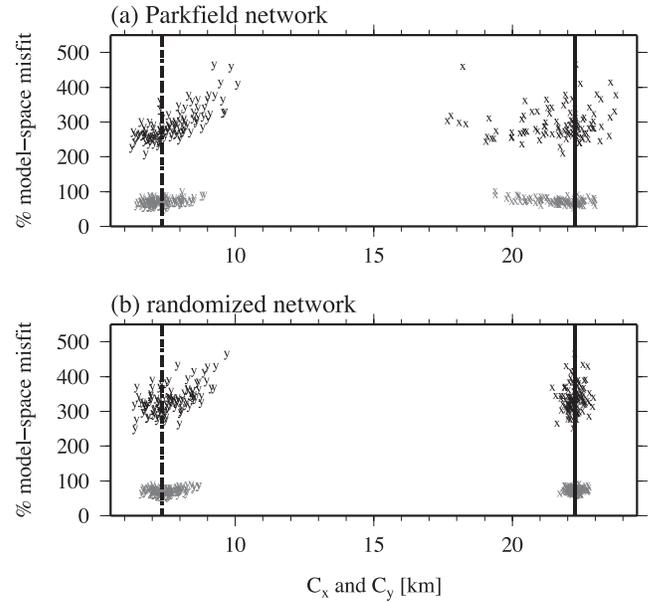


**Figure 5.** Percentage of model-space misfit versus geodetic potency resulting from 100 Monte Carlo tests employing smoothing strength  $\kappa = 0$  (dark symbols) and 100 tests employing  $\kappa = 10^{-4}$  (grey symbols). (a) Parkfield network configuration, with the largest and the smallest estimated moment magnitudes indicated. (b) Random network configuration. The vertical dashed lines indicate the true synthetic value.

with different standard deviations for different displacement components (set equal to the corresponding average standard deviations reported on SOPAC's website). Each Monte Carlo realization of the data error is then added to the noise-free synthetic data, and inverted for slip using the NNLS algorithm. In Fig. 5(a), we show the geodetic potency as a function of the percentage of model-space misfit of 100 such inversions. The percentage of model misfit of results obtained without smoothing (black circles) are in the range of 200 and 450 per cent, and that of results obtained using  $\kappa = 10^{-4}$  (grey circles) are between 60 and 100 per cent. Note that here, because the synthetic slip distribution is smooth, the effect of smoothing is to reduce both the model misfit and the spread in the moment magnitude estimates. The result of the Monte Carlo tests shows that uncertainties in the geodetic potency determination due to data uncertainties are 20–30 per cent. When translated to moment magnitude using the moment-magnitude relation of Hanks & Kanamori (1979) and a shear modulus of 16 GPa, the largest and smallest seismic moment of the set are equal to 6.07 and 5.95, respectively, indicating that geodetic magnitude determinations are practically insensitive to data uncertainties. Interestingly, despite the GPS sites being strongly clustered near the fault centre, the Monte Carlo tests yield similar scattering for the  $C_x$  and  $C_y$ , indicating that the two components are similarly sensitive to data uncertainties (Fig. 6).

### 3.3.2 The spreadness and the skewness tensors

The estimated components of the spreadness and the skewness tensors as a function of the model misfit for 100 Monte Carlo realizations are shown in Figs 7(a) and 8(a), respectively. Because the Parkfield GPS network is densely distributed around the fault centre with no stations near the fault ends, the control on the rupture lateral extent is rather poor. Consequently, the components that are a function of the along-strike coordinate (i.e.  $I_{xx}$ ,  $I_{yy}$ ,  $J_{xxx}$  and  $J_{yyy}$ )



**Figure 6.** Percentage of model-space misfit versus centroids resulting from 100 Monte Carlo tests employing  $\kappa = 0$  (dark symbols) and 100 tests employing  $\kappa = 10^{-4}$  (grey symbols). (a) Parkfield network configuration. (b) Random network configuration. The vertical dashed and dotted lines indicate the true synthetic values of the along-strike and the downdip components of the centroids, respectively.

are less well constrained than those that are not (i.e.  $I_{yy}$  and  $J_{yyy}$ ). We conclude that neither the spreadness nor the skewness may be resolved given the Parkfield network configuration.

### 3.4 Comparison with randomly generated networks

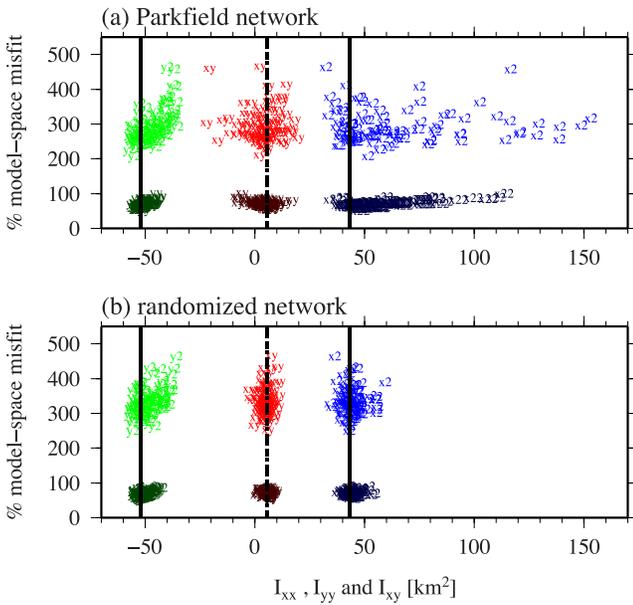
Because our ability to resolve the slip distribution functions (eqs 5–8) is strongly influenced by the spatial distribution of the observation points, it is instructive to consider alternative network configurations. In the previous section we have seen that important attributes of the synthetic slip distribution cannot be resolved using the Parkfield's GPS network, and we will now see how use of randomly generated GPS networks affects the resolution of these attributes.

Using a uniformly distributed random number generator, we constructed a set of 10 000 random networks, each of which consists of 14 sites, whose position is constrained to a  $35 \times 35$  km<sup>2</sup> squared region enclosing the model fault. For each of the random network configurations, we then computed a set of 100 different synthetic data, containing uncorrelated random noise as is explained in Section 3.3. Next we used the NNLS algorithm to invert each data set to get the slip distribution, and used that slip distribution to compute the set of slip distribution functions (eqs 5–8).

We define a network discrepancy parameter that quantifies, for each network configuration, the percentage of the average discrepancy between the estimated and the true slip distribution functions as follows:

network discrepancy parameter =

$$100 \frac{1}{m} \sum_{l=1}^m \sqrt{\frac{1}{10} \sum_{n=1}^{10} \left( \frac{f_n^{\text{syn}} - f_n^{\text{est}}}{f_n^{\text{syn}}} \right)^2}, \quad (9)$$

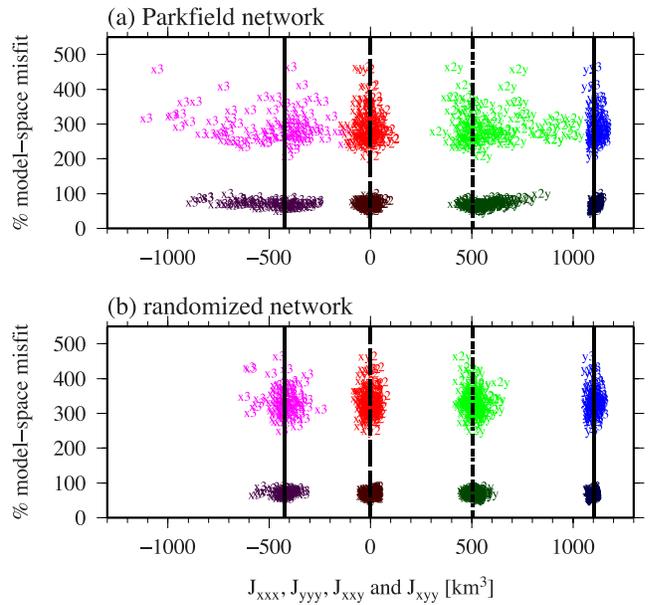


**Figure 7.** Percentage of model-space misfit versus the spreadness resulting from 100 Monte Carlo tests employing  $\kappa = 0$  (light coloured symbols) and 100 tests employing  $\kappa = 10^{-4}$  (dark coloured symbols). (a) Parkfield network configuration. (b) Random network configuration. The vertical lines indicate the true synthetic values. In order to avoid overlap between symbols of different components, values of  $I_{yy}$  were shifted to the left by a constant amount.

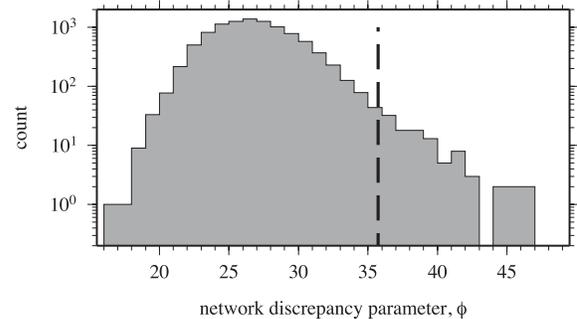
where  $m$  is the number of data realizations for each random network configuration, the superscripts ‘syn’ and ‘est’ denote ‘synthetic’ and ‘estimated’, respectively, and  $f_1$  through  $f_{10}$  are the 10 slip distribution functions defined in eqs (5)–(8). In Fig. 9, we compare the distribution of the network discrepancy parameters corresponding to 10 000 random network configurations to the network discrepancy parameter corresponding to the Parkfield network (indicated by the vertical dashed line). Remarkably, we find that more than 99 per cent of the random configurations result in a network discrepancy parameter that is smaller than that of the Parkfield network. On the basis of this analysis we conclude that randomly constructed networks are better configured than the Parkfield network, in the sense that they better recover the macroscopic attributes of the slip distribution.

The four configurations that yield the smallest network discrepancy parameter are shown in Fig. 10. Despite these configurations sharing very similar value of discrepancy parameter, the differences between them are notable. A count of station-to-centroid distances reveals a clear tendency for well-configured random networks to peak between 15 and 20 km (Fig. 11). In contrast, the Parkfield network is much more clustered near the fault centre, with two sites practically located inside the fault zone.

Estimated slip parameters obtained through inversions of 100 Monte Carlo data realizations computed for the Parkfield network and the random network configuration that yields the smallest discrepancy parameter are compared in Figs 5–8. Note that the distribution of the estimated slip parameters corresponding to this random network is far less scattered and is in much better agreement with the synthetic slip parameters than those corresponding to the Parkfield network. While use of the random network configuration clearly improves the fit between the estimated and the actual slip distribution functions (eqs 5–8), the fit to the slip distribution remains poor.



**Figure 8.** Percentage of model-space misfit versus the skewness resulting from 100 Monte Carlo tests employing  $\kappa = 0$  (light coloured symbols) and 100 tests employing  $\kappa = 10^{-4}$  (dark coloured symbols). (a) Parkfield network configuration. (b) Random network configuration. The vertical lines indicate the true synthetic values. In order to reduce overlap between symbols of different components, values of different components were shifted horizontally by different amount.

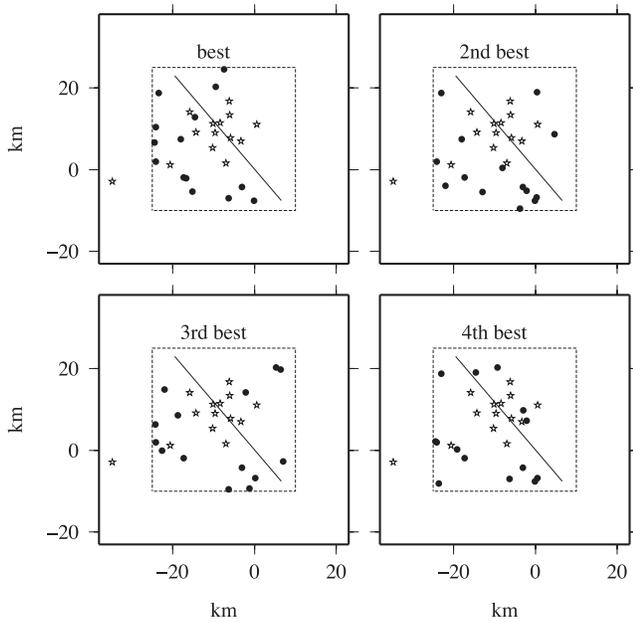


**Figure 9.** A histogram showing the distribution of network discrepancies corresponding to 10 000 random network configurations. The vertical dashed line indicates the average discrepancy corresponding to the Parkfield network.

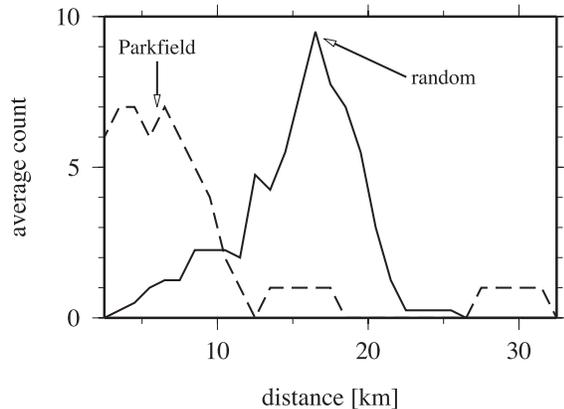
#### 4 SINGLE-SITE MAGNITUDE DETERMINATION

We have shown that while the extent to which the spreadness and the skewness may be resolved depends strongly on the spatial distribution of the observation sites, the geodetic potency and the centroid are better resolved and are less sensitive to the network configuration. To further illustrate the robustness of the magnitude and the centroid determinations, we solve the set of equations for the slip distribution using the 3-D ground displacement at each of the 14 GPS sites separately.

In Fig. 12(a), we show the 14 geodetic potency versus smoothing curves that we obtained using the NNLS algorithm. Note that for most GPS sites, the geodetic potency is nearly independent of smoothing for smoothing coefficient smaller than  $10^{-2}$ . Beyond that point, the potency decreases rapidly with increasing smoothing. In Fig. 12(b) we show the 14 moment magnitudes that we obtained for each station separately by solving eq. (2) only, that is, without the

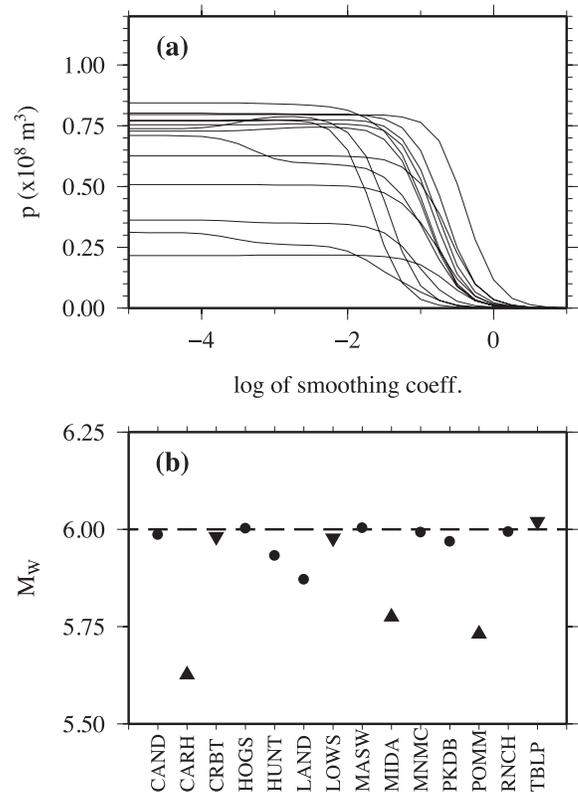


**Figure 10.** Maps of four random networks yielding the smallest network discrepancy parameters, with solid circles indicating the position of the randomly positioned GPS sites and open stars indicating the position of the Parkfield GPS stations.



**Figure 11.** A comparison between the station-to-centroid distance count of the Parkfield network (dashed curve) and the four random networks yielding the smallest network discrepancy parameters (solid curve).

smoothing constraint. Note that 10 of the 14 single-site magnitude estimates are within  $\pm 0.05$  magnitude unit of the synthetic magnitude,  $\pm 20$  percent of the synthetic potency (or seismic moment). In addition, while the single-site magnitude estimates of GPS sites located closest to the model fault plane are rather poor (i.e. CARH, POMM and MIDA, indicated by upward-pointing triangles), those of sites located further away from the fault are extremely close to the ‘true’ magnitude (i.e. CRBT, LOWS and TBLP, indicated by downward-pointing triangles). In Fig. 13, we show the single-site  $C_x$  and  $C_y$  determinations. Similar to the single-site magnitude determination, the centroid determinations too are poor for GPS stations located in the vicinity of the fault plane, but are in good/fair agreement with the synthetic model for stations located further away from it. In summary, while having as many constraints as possible helps to account for uncertainties in fault geometry and measurement errors, our analysis shows that use of individual sites for earthquake magnitude determination is robust, provided that these sites are not



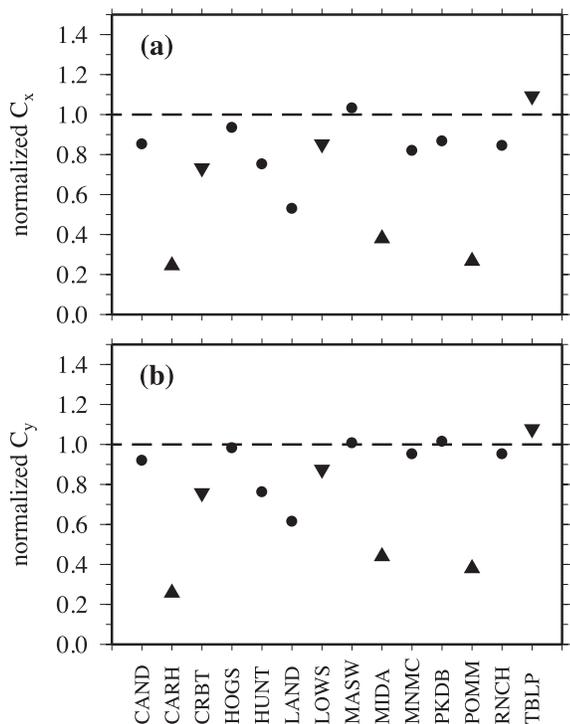
**Figure 12.** The results of single-site size determination. (a) Curves of geodetic potency versus smoothing coefficient. (b) The 14 moment magnitude estimates as a function of GPS site, with upward and downward pointing triangles indicating the two closest and the two most distant sites, respectively. The dashed line indicates the true synthetic magnitude.

located within or in close proximity to the fault zone. Next we discuss the implications of this result for earthquake early warning using real-time GPS.

## 5 IMPLICATIONS FOR EARTHQUAKE EARLY WARNING

Earthquake Early Warning Systems (EWS) are rapid co- or post-rupture magnitude and location determination systems that issue a hazard assessment ahead of the destructive ground shaking, during which various life and infrastructure saving measures may be taken. Such systems are currently being implemented in Japan (Nakamura 1988), Mexico (Aranda *et al.* 1995) and Taiwan (Wu & Kanamori 2005). From a seismological perspective, the most challenging task in the design of an EWS is the development of a robust algorithm for rapid magnitude assessment. To date, the data that are utilized by such algorithms are seismograms and/or accelerograms. Recent advances in continuous GPS analysis techniques make it now possible to obtain a close to real-time estimate of the 3-D ground displacement (Bock *et al.* 2000). Can real-time estimates of GPS ground displacement be harnessed for rapid magnitude determination?

We have shown that underdetermined slip inversions, even those utilizing just a single GPS site (provided that this site is not located within or in close proximity to the fault zone), yield robust estimate of the moment magnitude. This result implies that real-time GPS ground displacement can be used for rapid magnitude determination. Guided by this thinking, Allen & Ziv (2011) developed a simple algorithm that can be applied in real-time, which extracts



**Figure 13.** The results of single-site centroid determination. (a) The 14 along-strike centroid coordinates as a function of GPS site. (b) The 14 down-dip centroid coordinates as a function of GPS site. Centroids are normalized by the true synthetic values. Upward and downward pointing triangles indicate the two closest and the two most distant sites, respectively.

the static offsets from the GPS total displacement waveforms by averaging the dynamic oscillations. The static offsets are then used as input for linear slip inversions similar to the one used here, using a model fault plane that is chosen from a list of pre-recognized faults on the basis of proximity to the real-time estimated hypocentre. The estimated magnitude is then calculated directly from the slip distribution. Both the static offsets and the magnitude estimates are being repeatedly updated every second. The method has been tested and successfully implemented on real-time displacement waveforms of the 2010 April 4,  $M_w$  7.2 El Mayor-Cucapah earthquake, showing that real-time high-rate GPS can provide a useful and independent assessment of earthquake magnitude for the purpose of earthquake early warning and real-time earthquake information systems (Crowell *et al.* 2009).

## 6 SUMMARY AND CONCLUSIONS

In order to see which of the slip distribution attributes are obtainable, and are shared by different underdetermined models providing satisfactory fit to the data, we constructed a synthetic GPS data set and inverted it for slip. The geometry of that case study corresponds to that of the Parkfield segment and the 14 SCIGN GPS sites next to it. We show that slip inversions of that data yield robust estimate of both the geodetic potency and the moment centroid. In contrast, the slip inversions of the same data failed to resolve the spreadness and the skewness of the slip distribution with respect to the centroid, especially those components that are a function of the along-strike position. This latter conclusion is not surprising, since the Parkfield network is densely distributed around the fault centre with no stations near the fault ends.

We examined whether use of randomly generated GPS networks improves or worsen the resolution of the slip distribution functions (eqs 5–8). We find that randomly constructed networks are better configured than the Parkfield network, in the sense that they better recover the true slip distribution functions. Nevertheless, use of the random network configuration does not improve the fit between the estimated and the synthetic input slip distribution.

We show that the moment magnitude may be recovered using only one GPS station, provided that this station is not located too close to the fault plane. Some researchers have added a seismic moment constraint to the set of equations (Johnson *et al.* 1994; Lundgren & Stramondo 2002), whereas the results of this study suggest that such a constraint is unneeded.

Near-fault GPS stations are critical for constraining shallow slip during inter- and post-seismic intervals. For constraining the slip attributes of moderate-to-large earthquakes, however, it is important to have stations further away from the fault plane. Since the 2004 Parkfield earthquake, the number of GPS antennas within a distance range of 35 km from the Parkfield epicentre has tripled. These stations are now part of the Plate Boundary Observatory, which is a component of the EarthScope project. Additional GPS stations that were deployed beyond that distance range may, however, not be useful for monitoring the next Parkfield earthquake, as the expected ground displacement at these stations due to  $M_w$  6 along the Parkfield segment and the data accuracy are of similar magnitude. Despite the dramatic increase in the number of stations, detailed mapping of the seismic slip distribution of the next Parkfield earthquake, on spatial and temporal scales that are of interest to earthquake physicists, will not be feasible. Nevertheless, thanks to the addition of stations at the intermediate range (i.e. between 10 and 20 km of the epicentre), it is expected that the macroscopic attributes of that earthquake, including the spreadness and the skewness, will be well resolved.

## ACKNOWLEDGEMENTS

We thank the Editor, Bert Vermeersen, and the reviewers for their constructive remarks.

## REFERENCES

- Allen, R.M. & Ziv, A., 2011. Application of real-time GPS to earthquake early warning, *Geophys. Res. Lett.*, **38**, L16310, doi:10.1029/2011GL047947.
- Aranda, J.M.E., Jiménez, A., Ibarrola, G., Alcántar, F., Aguilar, A., Inostroza, M. & Maldonado, S., 1995. Mexico city seismic alert system, *Seism. Res. Lett.*, **66**, 42–53.
- Bakun, W.H. *et al.*, 2005. Implications for prediction and hazard assessment from the 2004 Parkfield earthquake, *Nature*, **437**, 969–974.
- Barbot, S., Fialko, Y. & Bock, Y., 2009. Postseismic deformation due to the Mw6.0 2004 Parkfield earthquake: stress-driven creep on a fault with spatially variable rate-and-state friction parameters, *J. geophys. Res.*, **114**, B07405, doi:10.1029/2008JB005748.
- Barbot, S., Lapusta, N. & Avouac, J.-P., 2012. Under the hood of the earthquake machine: toward predictive modeling of the seismic cycle, *Science*, **336**, 707–710.
- Bock, Y., Nikolaidis, R., Jonge, P.J. & Bevis, M., 2000. Instantaneous geodetic positioning at medium distances with the global positioning system, *J. geophys. Res.*, **105**, 28 223–28 253.
- Bos, A.G. & Spakman, W., 2003. The resolving power of coseismic surface displacement data for fault slip distribution at depth, *Geophys. Res. Lett.*, **30**, 2110, doi:10.1029/2003GL017946.

- Crowell, B.W., Bock, Y. & Squibb, M.B., 2009. Earthquake early warning using total displacement waveforms from real-time GPS networks, *Seism. Res. Lett.*, **80**, 772–782.
- Custódio, S., Page, M. & Archuleta, R., 2009. Constraining earthquake source inversions with GPS data 2: a two-step approach to combine seismic and geodetic datasets, *J. geophys. Res.*, **114**, B01315, doi:10.1029/2008JB005746.
- Du, Y., Aydin, A. & Segall, P., 1992. Comparison of various inversion techniques as applied to the determination of a geophysical deformation model for the 1983 borah peak earthquake, *Bull. seism. Soc. Am.*, **82**, 1840–1866.
- Hanks, T.C. & Kanamori, H., 1979. A moment magnitude scale, *J. geophys. Res.*, **84**, 2348–2350.
- Johnson, H.O., Agnew, D.C. & Hudnut, K., 1994. Extremal bounds on earthquake movement from geodetic data: application to the Landers earthquake, *Bull. seism. Soc. Am.*, **84**, 660–667.
- Lawson, C.L. & Hanson, R.J., 1995. *Solving Least Squares Problems*, Society for Industrial and Applied Mathematics, Philadelphia, PA.
- Lundgren, P. & Stramondo, S., 2002. Slip distribution of the 1997 Umbria-Marche earthquake sequence: joint inversion of GPS and synthetic aperture radar interferometry data, *J. geophys. Res.*, **107**, 2316, doi:10.1029/2000JB000103.
- Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory*, Vol. 45 of International Geophysics Series, Academic Press, San Diego, CA.
- Nakamura, Y., 1988. On the urgent earthquake detection and alarm system (UrEDAS), in *Proceedings of Ninth World Conference on Earthquake Engineering*, Tokyo-Kyoto, Japan, Vol. VII, pp. 673–678, CEA.
- Okada, Y., 1985. Surface deformation due to shear and tensile faults in a half-space, *Bull. seism. Soc. Am.*, **75**, 1135–1154.
- Page, M.T., Custódio, S., Archuleta, R.J. & Carlson, J.M., 2009. Constraining earthquake source inversions with GPS data 1: resolution based removal of artifacts, *J. geophys. Res.*, **114**, B01314, doi:10.1029/2007JB005449.
- Savage, J.C., 1990. Equivalent strike-slip earthquake cycles in half-space and lithosphere-asthenosphere Earth models, *J. geophys. Res.*, **95**, 4873–4879.
- Simons, M. *et al.*, 2011. The 2011 magnitude 9.0 Tohoku-Oki earthquake: mosaicking the megathrust from seconds to centuries, *Science*, **332**, 1421–1425.
- Wu, Y.M. & Kanamori, H., 2005. Experiment on an onsite early warning method for the Taiwan early warning system, *Bull. seism. Soc. Am.*, **95**, 347–353.
- Ziv, A., 2012. Inference of co-seismic slip via joint inversion of GPS and aftershock data: the 2004 Parkfield example, *J. geophys. Res.*, **117**, B03307, doi:10.1029/2011JB008400.