

A Bayesian approach to estimating tectonic stress from seismological data

Richard Arnold¹ and John Townend²

¹*School of Mathematics, Statistics, and Computer Science, Victoria University of Wellington, PO Box 600, Wellington, New Zealand.*

E-mail: richard.arnold@mcs.vuw.ac.nz

²*Institute of Geophysics, School of Geography, Environment, and Earth Sciences, Victoria University of Wellington, PO Box 600, Wellington, New Zealand*

Accepted 2007 April 30. Received 2007 April 29; in original form 2006 December 5

SUMMARY

Earthquakes are conspicuous manifestations of tectonic stress, but the non-linear relationships between the stresses acting on a fault plane, its frictional slip, and the ensuing seismic radiation are such that a single earthquake by itself provides little information about the ambient state of stress. Moreover, observational uncertainties and inherent ambiguities in the nodal planes of earthquake focal mechanisms preclude straightforward inferences about stress being drawn on the basis of individual focal mechanism observations. However, by assuming that each earthquake in a small volume of the crust represents a single, uniform state of stress, the combined constraints imposed on that stress by a suite of focal mechanism observations can be estimated. Here, we outline a probabilistic (Bayesian) technique for estimating tectonic stress directions from primary seismological observations. The Bayesian formulation combines a geologically motivated prior model of the state of stress with an observation model that implements the physical relationship between the stresses acting on a fault and the resultant seismological observation. We show our Bayesian formulation to be equivalent to a well-known analytical solution for a single, errorless focal mechanism observation. The new approach has the distinct advantage, however, of including (1) multiple earthquakes, (2) fault plane ambiguities, (3) observational errors and (4) any prior knowledge of the stress field. Our approach, while computationally demanding in some cases, is intended to yield reliable tectonic stress estimates that can be confidently compared with other tectonic parameters, such as seismic anisotropy and geodetic strain rate observations, and used to investigate spatial and temporal variations in stress associated with major faults and coseismic stress perturbations.

Key words: Bayesian data analysis, fault mechanics, focal mechanism, stress tensor, tectonic stress.

1 INTRODUCTION

1.1 Motivation

Focal mechanisms can serve as useful *in situ* gauges of the stresses causing seismogenic slip provided that the limited information about stress represented by each strike, dip, rake triple can be robustly interpreted. Existing algorithms produce estimates of the principal tectonic stress directions that are in good agreement with borehole measurements and computational models based on lithospheric buoyancy (Hickman & Zoback 2004; Townend & Zoback 2004; Townend 2006). Moreover, as the data obtained with dense regional seismic and geodetic networks accumulate, it is becoming increasingly feasible to compare the patterns of crustal deformation revealed by seismological and geodetic observations and in doing so identify distinct deformation processes such as fault zone locking and crustal collision (e.g. Becker *et al.* 2004; Boness & Zoback 2006; Townend & Zoback 2006).

Routine stress estimation using focal mechanism data has nevertheless been hampered to date by confusion over what information about ambient stresses each mechanism actually conveys and the directly related issue of how to calculate and represent the uncertainties associated with each estimated stress parameter (Townend 2006). The relationship between ambient tectonic stress and slip on a fault is generally specified in terms of the ‘Wallace–Bott (WB) hypothesis’ (see McKenzie 1969), which states that faults slip in the direction of the maximum resolved shear stress. If this assumption is valid, then in principle the single stress tensor most compatible with slip on a number of faults can be estimated.

In the canonical analysis of stress–fault slip relationships, McKenzie (1969) demonstrated that a single focal mechanism observation imposes minimal constraints on the orientation of the causative stress tensor, even when the slip vector ($\hat{\mathbf{u}}$) and fault normal vector ($\hat{\mathbf{n}}$) can be distinguished. Specifically, in the absence of any prior information about the relative magnitudes of the principal stresses, all that can be said based on a single focal mechanism is that the axes of maximum and minimum compressive stress lie within the dilatational quadrant (bisected by the P -axis, which is parallel to $(\hat{\mathbf{u}} - \hat{\mathbf{n}})/2$) and the compressional quadrant (bisected by the T -axis, parallel to $(\hat{\mathbf{u}} + \hat{\mathbf{n}})/2$), respectively.

The caveats highlighted by McKenzie (1969) for interpreting focal mechanism data in terms of stress have been somewhat overlooked or misinterpreted in recent years, leading a number of authors to assert, incorrectly, that the P - and T -axes are equivalent to the axes of maximum (S_{\max}) and minimum compressive stress (S_{\min}), respectively (*cf.* Scholz 2002; Townend 2006, for discussion). On the contrary, the P - and T -axes are geometric properties of an individual earthquake's moment tensor, and directly related to the fault plane normal and slip vectors, whereas the S_{\max} and S_{\min} axes pertain to a small volume of the crust within which faults of any number of orientations may be activated.

This limitation on stress inversion procedures becomes more acute the closer one focuses on a major plate-bounding structure such as the San Andreas fault, where on-fault earthquakes tend to exhibit very similar focal mechanisms and the information about stress directions contained in a suite of focal mechanisms may differ little from that implied by a single mechanism (Hardebeck 2006). In such circumstances, two commonly used algorithms (Gephart & Forsyth 1984; Michael 1984, 1987) tend to return the average P -axis of the focal mechanisms as the axis of maximum compressive stress with very small confidence intervals (Townend 2003). It will be shown below that this result represents the most probable S_{\max} geometry in underconstrained cases. Consequently, when focal mechanism diversity is very low, the estimate of the stress tensor's orientation obtained with either of these algorithms may appear to be reliable but in fact be completely misleading. This issue is particularly problematic when fault strength and, specifically, the orientation of the crustal stress tensor with respect to a major fault are under consideration (e.g. Hardebeck & Michael 2004; Townend & Zoback 2004; Townend 2006, and references therein).

Previous methods of estimating the posterior uncertainties in stress parameters have proven unsatisfactory in various ways. As noted above, two commonly used algorithms (Gephart & Forsyth 1984; Michael 1984, 1987) yield smaller uncertainties in the principal stress axes' orientations than the limiting analytical solution reveals are appropriate when the focal mechanisms being analysed have similar geometries (Townend 2003). Similarly, both these algorithms and that of Xu (2004) are unable to accommodate observational errors in the focal mechanism data themselves. Michael (1984, 1987) advocated a bootstrap procedure for estimating posterior uncertainties, but this approach does not enable any specific focal mechanism uncertainty estimates to be fully utilized, and is complicated by the need to identify a fault plane. Abers & Gephart (2001) developed a method of solving for the set of stress parameters most compatible with a suite of focal mechanisms, and each of the focal mechanisms themselves, based on first motion polarity data. This approach is similar to that of Rivera & Cisternas (1990), and while feasible in the case of small numbers of inversions (e.g. Balfour *et al.* 2005), becomes prohibitively slow when applied to dozens or hundreds of focal mechanism suites (Hardebeck & Michael 2006). Moreover, the joint focal mechanism–stress estimation problem, as currently implemented, does not allow for sources of uncertainty in the focal mechanism parameters other than those stemming from first motion polarity uncertainties.

A number of similar algorithms have been developed for use with fault-slip data collected from exhumed faults (e.g. Angelier 1979, 1984). The key distinction between fault-slip algorithms and those designed for use with focal mechanism data is that fault-slip data do not suffer from nodal plane ambiguity: that is, the fault plane normal and slip vectors are distinguishable. While this aspect of the stress estimation problem is, therefore, simpler than in the case of focal mechanism data, fault-slip data may well represent long periods of intermittent faulting that span more than one state of stress. Unless stratigraphic or structural relationships constrain the faulting episodes to a single, reasonably brief period of geological time, the reliability of such inversions is likely to be severely compromised.

Several recent studies of tectonic stress motivate this work, and in particular our attempt to describe the generalized focal mechanism stress estimation problem within a Bayesian framework and to obtain robust posterior uncertainties. Regional seismic and geodetic networks now enable tectonic stress directions to be related to other geophysical observables, notably crustal strain rate (Townend & Zoback 2006) and seismic anisotropy directions (Boness & Zoback 2006), or to be incorporated in dislocation models of deformation (e.g. Becker *et al.* 2004). Hardebeck & Michael (2006) have recently demonstrated the benefits of treating regional stress field estimation in generalized inverse problem terms, and particularly the use of damped inversion algorithms that enable spatially adjacent observations to influence one another in smoothed inversions. Furthermore, as Hardebeck (2006) has shown, seismically active fault populations appear to be rather homogeneous at scales of less than ~ 10 km, meaning that low focal mechanism diversity may be a consistent feature of small-scale focal mechanism studies.

We are also motivated by the prospect of measuring temporal changes in principal stress directions in response to earthquakes or volcanic eruptions (e.g. Hardebeck & Hauksson 1999; Provost & Houston 2003; Ratchkovski 2003; Sipkin & Silver 2003; Ratchkovski *et al.* 2004; Roman *et al.* 2004; Bohnhoff *et al.* 2006). These and other studies have reported on apparent temporal changes in stress, but it is unclear what significance these results have in view of the techniques used and the lack of a reliable model of the uncertainties associated with each stress estimate (Townend & Zoback 2001, 2006; Townend 2006; Lund & Townend 2007). A similar comment may be made about interpreting spatial variations in stress.

In this paper, we present a fully Bayesian analysis (*cf.* Sivia 1996; Tarantola 2005; Wéber 2006) of the stress estimation problem, based on a preliminary model described by Arnold *et al.* (2005). Our objective is to describe in probabilistic terms the distribution of resolvable

stress parameters consistent with a suite of focal mechanism observations. We explicitly account for the fact that in general the two nodal planes of each focal mechanism are indistinguishable, and presume all pertinent observations to be made with some uncertainty.

Our approach is parsimonious in the sense that it is based on combining small but reliable increments of constraint rather than the larger but more subjective constraints provided by introducing an *ad hoc* objective function or by linearizing the stress–fault slip relationship.

Hardebeck (2006) has recently demonstrated that in Northern California only a restricted set of fault planes are represented in focal mechanisms. Moreover, it is commonly the case that some prior knowledge exists about the stress tensor geometry (e.g. a regional axis of maximum horizontal compressive stress, or one principal stress axis being subvertical). However, as a first step in addressing the full complexity of the stress estimation problem, the analysis presented below focuses on the case in which there is no prior information about the orientation of fault planes or about any stress tensor parameters.

1.2 Outline

We first define coordinate transformations between the three coordinate systems of interest—geographic, principal stress and fault plane (Section 2.1)—and obtain algebraic relationships between observational parameters (namely the strike, dip and rake of a number of focal mechanisms) and model parameters (the stress tensor’s principal axes and a single stress magnitude parameter; Sections 2.2–2.4). The relationships constitute our model in that they encapsulate the constraint on parameters of interest (stress) imposed by specific observations (focal mechanisms). Next, we present statistical representations of the stress tensor parameters in the absence of any focal mechanism observations, and the range of possible focal mechanisms that can be observed (Section 2.5). In Section 3, we combine the data generation model and priors to obtain a posterior model of stress given a suite of focal mechanism observations; we first consider the case of a single, errorless focal mechanism observation (Section 3.1), and then single and multiple focal mechanism observations containing measurement errors (Sections 3.2 and 3.3). Section 4 illustrates the application of the last of these models to a focal mechanism data set from Japan (Townend & Zoback 2006) and Section 5 outlines avenues of future research in which we will consider more informative—that is, prescriptive—prior models.

Table 1 provides a list of symbols used in this paper.

Table 1. List of symbols.

Symbol	Explanation
$\mathbb{G}\{\hat{\mathbf{g}}_1, \hat{\mathbf{g}}_2, \hat{\mathbf{g}}_3\}$	Geographical coordinate system: axes oriented north, east, and down
$\mathbb{S}\{\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \hat{\mathbf{s}}_3\}$	Principal stress coordinate system: axes parallel to the principal stress directions (1 = min, 3 = max)
$\mathbb{F}\{\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \hat{\mathbf{f}}_3\}$	Fault plane coordinate system: axes parallel to the slip vector $\hat{\mathbf{u}}$, null vector $\hat{\mathbf{n}} \times \hat{\mathbf{u}}$ and fault normal $\hat{\mathbf{n}}$
$R_{\mathbb{B}\mathbb{A}}$	Transformation matrix from coordinate system \mathbb{A} to coordinate system \mathbb{B}
$W = R_{\mathbb{S}\mathbb{F}}$	Wallace-Bott (WB) matrix: transforms between \mathbb{F} and \mathbb{S} coordinates
$\Theta = (\xi, \delta, \lambda)$	Strike, dip, and rake angles of a focal mechanism
$\Phi = (\phi, \theta, \psi)$	Euler angles
Φ_G, Φ_F	Euler angles of a focal mechanism in \mathbb{G} and \mathbb{F} coordinates
Φ_S	Euler angles of the stress tensor in \mathbb{G} coordinates
$E(\Phi) = B_3(\phi)B_1(\theta)B_3(\psi)$	Euler rotation matrix
\mathbf{S}	Compressive stress tensor
$S_{\min} \leq S_{\text{mid}} \leq S_{\max}$	Principal components of compressive stress \mathbf{S}
\mathbf{M}	Sesimic moment tensor
P_f	Ambient fluid pressure
\mathbf{S}'	Effective compressive stress tensor equal to $\mathbf{S} - P_f \mathbf{I}$
$S'_{\min} \leq S'_{\text{mid}} \leq S'_{\max}$	Principal components of effective compressive stress \mathbf{S}'
D	Differential stress $S_{\max} - S_{\min}$
$(\hat{\mathbf{u}}, \hat{\mathbf{a}}, \hat{\mathbf{n}})$	Fault slip, null and normal unit vectors
\mathbf{c}'	Effective traction vector
$\mathbf{c}'_n, \mathbf{c}'_s$	Effective normal and shear traction vectors
c'_n, c'_s	Magnitudes of the effective normal and shear tractions
v	Stress magnitude ratio given by $(S_{\max} - S_{\text{mid}})/(S_{\max} - S_{\min})$
κ, λ	Fault geometry parameters
$f(\cdot), p(\cdot)$	Probability density functions
$\delta(\cdot), \delta^3(\cdot)$	Delta functions in one and three dimensions
τ	Precision of a Matrix–Fisher distribution
$c(\tau)$	Normalizing constant of a Matrix–Fisher distribution
$I(A)$	Indicator function of set A (=1 if A is true, 0 otherwise)
$\text{tr}(B)$	Trace of the matrix B
C_2	Matrix which swaps the x - and z -axes, and reverses the y -axis

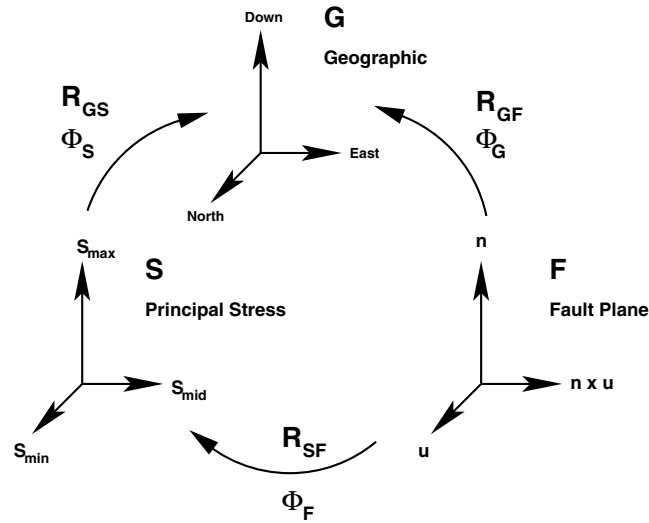


Figure 1. Coordinate systems and transformation matrices used in the analysis.

2 MODEL SPECIFICATION

2.1 Coordinate systems

Three right-handed 3-D Cartesian coordinate systems are relevant (Fig. 1).

- (i) *Geographical coordinates*: \mathbb{G} with unit vectors $\{\hat{\mathbf{g}}_1, \hat{\mathbf{g}}_2, \hat{\mathbf{g}}_3\}$ aligned, respectively, towards the north, the east and downwards;
- (ii) *Principal stress coordinates*: \mathbb{S} with unit vectors $\{\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \hat{\mathbf{s}}_3\}$ aligned along the principal axes of the stress tensor \mathbf{S} , in order of increasing compressive stress;
- (iii) *Fault plane coordinates*: \mathbb{F} with unit vectors $\{\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \hat{\mathbf{f}}_3\}$, respectively, aligned along the slip vector $\hat{\mathbf{u}}$, the ‘null’ vector $\hat{\mathbf{n}} \times \hat{\mathbf{u}}$, and the fault normal $\hat{\mathbf{n}}$.

We define three orthogonal rotation matrices $R_{\mathbb{G}\mathbb{F}}$, $R_{\mathbb{G}\mathbb{S}}$ and $R_{\mathbb{S}\mathbb{F}}$ which can be used to transform the representation of vectors from one coordinate system to another. For example a vector \mathbf{x} has representation $\mathbf{x}_{\mathbb{F}}$ in fault plane coordinates, but representation $\mathbf{x}_{\mathbb{G}} = R_{\mathbb{G}\mathbb{F}}\mathbf{x}_{\mathbb{F}}$ in geographical coordinates. These three matrices satisfy the relationship

$$R_{\mathbb{G}\mathbb{F}} = R_{\mathbb{G}\mathbb{S}}R_{\mathbb{S}\mathbb{F}}. \quad (1)$$

Any rotation matrix R can be represented by an equivalent set of three Euler angles $\Phi = (\phi, \theta, \psi)$ such that

$$R = E(\Phi) = B_3(\phi)B_2(\theta)B_3(\psi) \quad (2)$$

(see Section A1 of Appendix A). We define three sets of Euler angles Φ_G , Φ_S and Φ_F such that $R_{\mathbb{G}\mathbb{F}} = E(\Phi_G)$, $R_{\mathbb{G}\mathbb{S}} = E(\Phi_S)$ and $R_{\mathbb{S}\mathbb{F}} = E(\Phi_F)$ and hence eq. (1) is equivalent to the expression

$$E(\Phi_G) = E(\Phi_S)E(\Phi_F). \quad (3)$$

2.2 Stress tensor properties

The stress tensor \mathbf{S} is defined by six parameters: the three Euler angles Φ_S which define its orientation with respect to geographical coordinates \mathbb{G} , and its three principal stresses ($S_{\min} \leq S_{\text{mid}} \leq S_{\max}$). We adopt the convention that compressive stress is positive, and that the minimum and maximum stresses are, respectively, aligned along the x - and z -axes in principal stress coordinates \mathbb{S} . In those coordinates the stress tensor \mathbf{S} is by definition diagonal:

$$\mathbf{S}_{\mathbb{S}} = \begin{bmatrix} S_{\min} & 0 & 0 \\ 0 & S_{\text{mid}} & 0 \\ 0 & 0 & S_{\max} \end{bmatrix}_{\mathbb{S}}. \quad (4)$$

We define the differential stress, D , and the stress ratio, ν , as

$$D = S_{\max} - S_{\min} \quad \nu = \frac{S_{\max} - S_{\text{mid}}}{S_{\max} - S_{\min}}, \quad (5)$$

where $D \geq 0$ and $0 \leq \nu \leq 1$. It follows that

$$\mathbf{S}_s = S_{\max} \mathbf{I}_s - D \mathbf{J}_s \quad \text{where} \quad \mathbf{J}_s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \nu & 0 \\ 0 & 0 & 0 \end{bmatrix}_s \quad (6)$$

and \mathbf{I}_s is the identity matrix. The effective stress tensor \mathbf{S}' is obtained by subtracting the ambient fluid pressure P_f from the stress tensor, that is,

$$\mathbf{S}'_s = \mathbf{S}_s - P_f \mathbf{I}_s = S'_{\max} \mathbf{I}_s - D \mathbf{J}_s, \quad (7)$$

where $S'_{\max} = S_{\max} - P_f$.

2.3 Fault plane compressive stress and the Wallace–Bott hypothesis

If we define a fault normal in principal stress coordinates \mathbb{S} as

$$\hat{\mathbf{n}}_s = [n_1 \quad n_2 \quad n_3]_s^T \quad \text{where} \quad n_1^2 + n_2^2 + n_3^2 = 1, \quad (8)$$

then the effective compressive stress \mathbf{c}' on the fault caused by the effective stress tensor \mathbf{S}' can be resolved into normal and shear components

$$\mathbf{c}' = \mathbf{S}' \hat{\mathbf{n}} = c'_n \hat{\mathbf{n}} + \mathbf{c}_s. \quad (9)$$

Since we have adopted the convention that compressive stress is positive, $\mathbf{c}' = -\mathbf{t}'$ where \mathbf{t}' is the traction \mathbf{t} acting on the fault. Note that we have omitted the prime on \mathbf{c}_s because it has the same value whether we work with actual or effective stresses.

The magnitude of the normal stress is

$$c'_n = \hat{\mathbf{n}} \cdot \mathbf{c}' = \hat{\mathbf{n}} \cdot (S'_{\max} \mathbf{I}_s - D \mathbf{J}_s) \hat{\mathbf{n}} = S'_{\max} - \kappa D, \quad (10)$$

where $\kappa = \hat{\mathbf{n}}_s^T \mathbf{J}_s \hat{\mathbf{n}}_s = n_1^2 + \nu n_2^2$, ($0 \leq \kappa \leq 1$). The shear stress is given by

$$\mathbf{c}_s = c_s \hat{\mathbf{c}}_s = \hat{\mathbf{n}} \times (\mathbf{c}' \times \hat{\mathbf{n}}) = D \begin{bmatrix} -(1 - \kappa)n_1 \\ (\kappa - \nu)n_2 \\ \kappa n_3 \end{bmatrix}_s \quad (11)$$

which has magnitude

$$c_s = D \sqrt{\kappa(1 - \kappa) - \nu(1 - \nu)n_2^2}. \quad (12)$$

The WB hypothesis (after Wallace 1951; Bott 1959) asserts that in the presence of a uniform ambient stress field rigid blocks slide linearly with respect to each other along a planar fault in the direction of the maximum resolved shear stress. The slip vector $\hat{\mathbf{u}}$ of the hanging wall of the fault is antiparallel to the shear stress vector \mathbf{c}_s .

The fault normal $\hat{\mathbf{n}}$ together with the slip vector $\hat{\mathbf{u}}$ define the fault plane coordinate system \mathbb{F} . The slip vector and normal vectors lie along the x - and z -axes, and the null vector $\hat{\mathbf{a}} = \hat{\mathbf{n}} \times \hat{\mathbf{u}}$ along the y -axis. The transformation matrix $R_{\mathbb{S}\mathbb{F}}$ between \mathbb{F} and principal stress coordinates \mathbb{S} is an orthogonal matrix with columns that are the representations of $\hat{\mathbf{u}}$, $\hat{\mathbf{a}}$ and $\hat{\mathbf{n}}$ in \mathbb{S} :

$$W = R_{\mathbb{S}\mathbb{F}} = [\hat{\mathbf{u}} \quad \hat{\mathbf{a}} \quad \hat{\mathbf{n}}]_s = \begin{bmatrix} \lambda(1 - \kappa)n_1 & -\lambda\nu n_2 n_3 & n_1 \\ -\lambda(\kappa - \nu)n_2 & \lambda n_1 n_3 & n_2 \\ -\lambda\kappa n_3 & -\lambda(1 - \nu)n_1 n_2 & n_3 \end{bmatrix}_s, \quad (13)$$

where λ is a positive normalization constant defined as $\lambda^2 \equiv 1/[\kappa(1 - \kappa) - \nu(1 - \nu)n_2^2]$ and $\hat{\mathbf{u}} = -(\lambda/D)\mathbf{c}_s$. We refer to the matrix $W = R_{\mathbb{S}\mathbb{F}}$ as the WB matrix since it embodies the consequences of the WB hypothesis.

Note that the WB matrix depends on the principal values of the stress tensor only through the stress ratio ν : that is, it is independent of both the differential stress (D) and the maximum stress (S'_{\max}). This result expresses the fact that the WB hypothesis only allows inference about ν and Φ_s : in other words the values of D and S'_{\max} are not constrained by focal mechanism observations.

The WB matrix $W = R_{\mathbb{S}\mathbb{F}} = E(\Phi_F)$ can also be written in terms of three Euler angles $\Phi_F = (\phi_F, \theta_F, \psi_F)$, and it follows firstly from eq. (A4) (Section A1) that

$$\hat{\mathbf{n}}_s = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}_s = \begin{bmatrix} \cos \phi_F \sin \theta_F \\ \sin \phi_F \sin \theta_F \\ \cos \theta_F \end{bmatrix}_s \quad (14)$$

$$\kappa = n_1^2 + \nu n_2^2 = \sin^2 \theta_F (\cos^2 \phi_F + \nu \sin^2 \phi_F) \quad (15)$$

and secondly, using eq. (A5),

$$\tan \psi_F = \frac{W_{32}}{-W_{31}} = \frac{-\lambda(1-\nu)n_1n_2}{\lambda\kappa n_3} = \frac{-(1-\nu)\sin\phi_F \cos\phi_F}{(\cos^2\phi_F + \nu \sin^2\phi_F)\cos\theta_F} \quad (16)$$

which can be rearranged to give

$$\nu = -\frac{W_{12}W_{13}}{W_{22}W_{23}} = \frac{\cos\theta_F \tan\psi_F + \tan\phi_F}{(1 - \tan\phi_F \cos\theta_F \tan\psi_F)\tan\phi_F}. \quad (17)$$

Eqs (16) and (17) are equivalent to eq. (19) of McKenzie (1969). The WB hypothesis states that once a fault normal $\hat{\mathbf{n}}$ is specified, which is equivalent to choosing the two Euler angles (ϕ_F, θ_F) in \mathbb{S} , then the slip vector $\hat{\mathbf{u}}$ can be calculated, and hence the third Euler angle ψ_F . The WB matrix is thus a matrix representation of a focal mechanism in principal stress coordinates \mathbb{S} .

The condition that the slip vector $\hat{\mathbf{u}} = -(D/\lambda)\mathbf{c}_s$ be antiparallel to the shear stress \mathbf{c}_s implies that $D/\lambda \geq 0$. The differential stress D is non-negative by definition (eq. 5) and so from eq. (13) we have the following constraint on the WB matrix:

$$\lambda = \frac{\lambda n_1 n_3}{n_1 n_3} = \frac{W_{22}}{W_{13} W_{33}} \geq 0. \quad (18)$$

Any orthogonal matrix W with $\det W = +1$ which satisfies eq. (18) represents a feasible focal mechanism in \mathbb{S} .

Using eqs (A4) and (16) it follows from eq. (18) that W must satisfy

$$\frac{W_{22}}{W_{13} W_{33}} = \frac{\cos\psi_F}{\sin\theta_F \cos\theta_F (\cos^2\phi_F + \nu \sin^2\phi_F)} \geq 0 \quad (19)$$

and given that $\sin\theta_F \geq 0$, condition (18) is equivalent to

$$\frac{\cos\psi_F}{\cos\theta_F} \geq 0 \Leftrightarrow \begin{cases} 0 \leq \theta_F \leq \frac{\pi}{2} & -\frac{\pi}{2} \leq \psi_F \leq \frac{\pi}{2}, \text{ or} \\ \frac{\pi}{2} < \theta_F \leq \pi & +\frac{\pi}{2} < \psi_F \leq \frac{3\pi}{2}. \end{cases} \quad (20)$$

This constraint limits the slip azimuth ψ_F , conditional on the colatitude θ_F of the fault normal $\hat{\mathbf{n}}$, to half of its possible range. The consequence of this for the inverse problem is that a single focal mechanism rules out half of the sphere as possible orientations for the principal axis of maximum stress, as was shown by McKenzie (1969).

2.4 Focal mechanism observations

An observation of a focal mechanism can be represented as a set of three angles $\Theta = (\xi, \delta, \lambda)$ —namely the strike, dip and rake of the focal mechanism (see Section A3). These angles describe the orientation of fault plane coordinates \mathbb{F} with respect to the Geographic coordinates \mathbb{G} . The strike, dip and rake can be transformed into standard Euler angles $\Phi_G = (\phi_G, \theta_G, \psi_G)$ by the formulae:

$$\phi_G = \xi + \frac{\pi}{2}; \quad \theta_G = \pi - \delta; \quad \psi_G = \lambda - \frac{\pi}{2}. \quad (21)$$

The rotation matrix $R_{\mathbb{G}\mathbb{F}} = E(\Phi_G)$ corresponding to $\Theta \Leftrightarrow \Phi_G$ has columns which are the representations of the slip, null and fault normal vectors expressed in geographic coordinates:

$$R_{\mathbb{G}\mathbb{F}} = E(\Phi_G) = [\hat{\mathbf{u}} \ \hat{\mathbf{a}} \ \hat{\mathbf{n}}]_{\mathbb{G}} \quad (22)$$

There are two ambiguities in focal mechanism solutions. One ambiguity is geometrical: the fault plane of focal mechanism Φ_{G1} can be equally well represented with the fault normal in either direction: $\pm\hat{\mathbf{n}}$. This swaps the identity of the hanging and footwalls of the fault, and reverses the slip vector. Consequently,

$$E(\Phi_{G2}) = [-\hat{\mathbf{u}} \ +\hat{\mathbf{a}} \ -\hat{\mathbf{n}}]_{\mathbb{G}} \quad \text{where } A_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = E(\Phi_{G1})A_2 \quad (23)$$

is an equivalent representation of the focal mechanism. The angles Φ_{G2} in this representation can be calculated from Φ_{G1} as follows:

$$\phi_{G2} = \phi_{G1} + \pi; \quad \theta_{G2} = \pi - \theta_{G1}; \quad \psi_{G2} = \pi - \psi_{G1}. \quad (24)$$

The two sets of Euler angles Φ_{G1} and Φ_{G2} represent the same focal mechanism, and thus lead to identical inferences about the stress tensor. We need only work with one or other of these two representations, and we can exploit this geometrical ambiguity to adopt the usual convention that the dip δ always lies in the range $0 \leq \delta \leq \frac{\pi}{2}$ (and consequently that θ_G always lies in the range $\frac{\pi}{2} \leq \theta_G \leq \pi$).

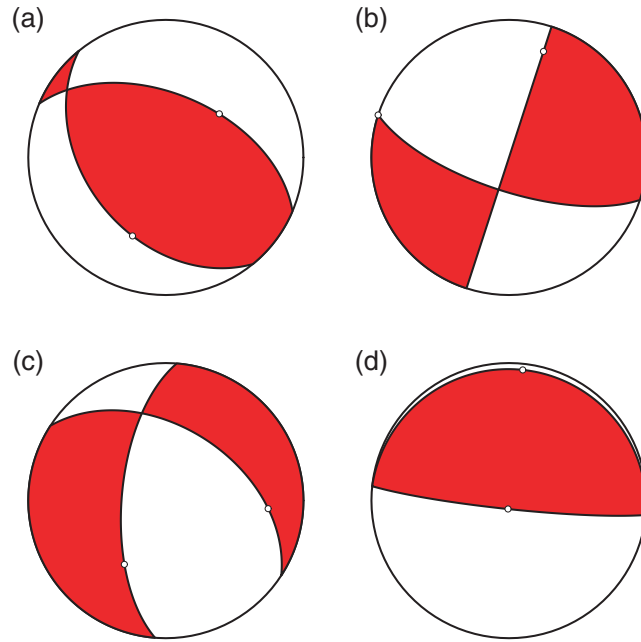


Figure 2. Four earthquake focal mechanisms from the 1971 San Fernando sequence (data reproduced from Gephart & Forsyth 1984) represented as beach ball diagrams. The focal mechanisms have (ξ, δ, λ) values of $(293^\circ, 52^\circ, 72^\circ)$, $(198^\circ, 90^\circ, 20^\circ)$, $(303^\circ, 46^\circ, -142^\circ)$ and $(96^\circ, 85^\circ, 90^\circ)$. The fault normal and slip vectors are shown with small open circles, and cannot be distinguished.

The second ambiguity in a focal mechanism is its physical interpretation: the fault normal and slip vector cannot be distinguished in general. If Φ_{G1} is a possible interpretation of the focal mechanism data, then so is Φ_{Ga1} defined by

$$E(\Phi_{Ga1}) = [+ \hat{\mathbf{n}} \quad - \hat{\mathbf{a}} \quad + \hat{\mathbf{u}}]_{\mathbb{G}} \quad \text{where} \quad C_2 = \begin{bmatrix} 0 & 0 & +1 \\ 0 & -1 & 0 \\ +1 & 0 & 0 \end{bmatrix}. \quad (25)$$

$= E(\Phi_{G1})C_2,$

A fourth interpretation Φ_{Ga2} of the observation can be constructed by reversing the axes associated with Φ_{Ga1} giving $E(\Phi_{Ga2}) = E(\Phi_{Ga1})A_2$. Again Φ_{Ga1} and Φ_{Ga2} lead to identical inferences about the stress tensor.

The seismic moment tensor \mathbf{M} , where $M_{ij} = M_0(u_i n_j + u_j n_i)$, displays these same symmetries: we can exchange $\hat{\mathbf{n}}$ and $\hat{\mathbf{u}}$, or simultaneously change their signs, and the value of the tensor is not affected.

In the statistical model we describe below, we need only work with two of the four representations of any given focal mechanism namely one of each pair (Φ_{G1}, Φ_{G2}) and (Φ_{Ga1}, Φ_{Ga2}) . One exception is where the slip and normal vectors *can* be distinguished, such as slickenfibres observations on outcropping planes. In such cases only one observation from the appropriate pair should be used.

Fig. 2 shows four earthquake focal mechanisms from the San Fernando sequence (Gephart & Forsyth 1984) displayed as standard beach ball diagrams (see, e.g. Stein & Wysession 2003). The fault normal and slip vectors are shown with small open circles, although in each case it is not known which is which. We use these four focal mechanisms in subsequent figures to demonstrate progressive steps in the stress calculations.

2.5 Statistical model

Given the definitions in the previous sections, we can now construct a statistical model which generates earthquake focal mechanism observations Φ_G in a volume of constant but unknown stress \mathbf{S} containing faults of all possible orientations. For this we need prior distributions on the parameters of the stress tensor, and on the orientations of faults (ϕ_F, θ_F) .

2.5.1 Stress tensor parameters

Under the WB hypothesis the only four parameters of the stress tensor that are relevant are the three orientation angles $\Phi_S = (\phi_S, \theta_S, \psi_S)$ and the stress ratio ν . It can be shown that if the three effective principal stresses $(S'_{\min}, S'_{\text{mid}}, S'_{\max})$ are the order statistics of three independent draws from a uniform distribution over the physical range of effective stresses $(0, S'_T)$, then the stress ratio ν follows a uniform distribution:

$$\nu \sim \text{Uniform}(0, 1). \quad (26)$$

In the absence of any prior information about the orientation of the stress tensor, we adopt a uniform distribution for its orientation. For Euler angles $\Phi_S = (\phi_S, \theta_S, \psi_S)$, a uniform distribution is achieved by taking

$$\begin{aligned} \phi_S &\sim \text{Uniform}(0, 2\pi) & 0 \leq \phi_S \leq 2\pi \\ \cos \theta_S &\sim \text{Uniform}(0, 1) & 0 \leq \theta_S \leq \frac{\pi}{2} \\ \psi_S &\sim \text{Uniform}(0, \pi) & 0 \leq \psi_S \leq \pi. \end{aligned} \quad (27)$$

In specifying this prior we have exploited the reflectional symmetry of the stress tensor to restrict the ranges of the angles $(\theta_S, \phi_S, \psi_S)$. In particular, we have required that the positive direction of the maximum principal axis S_{\max} point into the half-space with positive z (so that $0 \leq \theta_S \leq \pi/2$). Furthermore, we have required that the second principal axis is also directed into the positive z half-space (so that ψ_S is restricted to $[0, \pi]$). This has the consequence that the sines of θ_S and ψ_S are both non-negative.

The three angles $(\theta_S, \phi_S, \psi_S)$ define the transformation matrix $R_{\text{GS}} = E(\Phi_S)$.

2.5.2 Orientation of the fault plane

We assume that fault planes are randomly oriented, and hence that in principal stress coordinates \mathbb{S} we can select the fault plane normal $\hat{\mathbf{n}}$ by drawing (ϕ_F, θ_F) from the distributions:

$$\begin{aligned} \phi_F &\sim \text{Uniform}(0, 2\pi) & 0 \leq \phi_F \leq 2\pi \\ \cos \theta_F &\sim \text{Uniform}(0, 1) & 0 \leq \theta_F \leq \frac{\pi}{2}. \end{aligned} \quad (28)$$

Here, once again, the reflectional symmetry of the situation means that we have been able to restrict the fault normal to the half of the unit sphere with positive z in principal stress coordinates (i.e. $0 \leq \theta_F \leq \pi/2$). Given the two angles (ϕ_F, θ_F) we can calculate the remaining rotation angle ψ_F using eq. (16) in order to define fully the transformation matrix $R_{\text{GF}} = E(\Phi_F)$.

2.5.3 Focal mechanism observation model

Given the values of stress orientation Φ_S and fault plane coordinate orientation Φ_F , the values of the observation Φ_G follow from eq. (3). That equation can be rearranged to give the vector of angles Φ_G as a function of (Φ_S, Φ_F) :

$$\Phi_G = \mathbf{F}^{(1)}(\Phi_S, \Phi_F(v)). \quad (29)$$

If Φ_G is measured without error, then the likelihood of the observation Φ_G conditional on (Φ_S, Φ_F) is a delta function

$$p(\Phi_G | \Phi_S, \Phi_F) = \delta^3[\Phi_G - \mathbf{F}^{(1)}(\Phi_S, \Phi_F)]. \quad (30)$$

In the usual case in which the slip vector and the fault normal cannot be distinguished, the likelihood is the sum of two delta functions, one for each interpretation:

$$p(\Phi_G | \Phi_S, v, \phi_F, \theta_F) = \frac{1}{2} \sum_{\ell=1}^2 \delta^3[\Phi_G - \mathbf{F}^{(\ell)}(\Phi_S, \Phi_F)], \quad (31)$$

where $\mathbf{F}^{(1)}(\cdot)$ is as given in eq. (29) and $\mathbf{F}^{(2)}(\cdot)$ is the equivalent rearrangement of

$$R_{\text{GF}} = E(\Phi_G)C_2 = E(\Phi_S)E(\Phi_F)C_2 \quad (32)$$

in favour of Φ_G with C_2 as defined in eq. (25).

Where a focal mechanism is observed with error, we assume that the error follows a Matrix–Fisher distribution with a specified precision parameter τ (eq. A9; see also Mardia & Jupp 2000). That is, the observed angle Φ_G^o conditional on the true value Φ_G follows the distribution

$$p(\Phi_G^o | \Phi_G, \tau) = c(\tau) \sin \theta_G^o \exp\{\tau \text{tr}[E(\Phi_G)^T E(\Phi_G^o)]\} \quad (33)$$

The concentration parameter τ specifies the precision of the observation. In the limit as $\tau \rightarrow \infty$ the distribution becomes a delta function centred at Φ_G , and if $\tau = 0$ the distribution is uniform. Averaging over all possible values of Φ_G we obtain the likelihood equivalent to eq. (31) in the case of an observation Φ_G^o made with error:

$$\begin{aligned} p(\Phi_G^o | \Phi_S, v, \phi_F, \theta_F, \tau) &= \int d\Phi_G p(\Phi_G^o | \Phi_G, \tau) p(\Phi_G | \Phi_S, v, \phi_F, \theta_F) \\ &= \frac{1}{2} \sum_{\ell=1}^2 p(\Phi_G^o | \Phi_G = \mathbf{F}^{(\ell)}(\Phi_S, \Phi_F(v)), \tau) \\ &= \frac{1}{2} \sum_{\ell=1}^2 c(\tau) \sin \theta_G^o \exp\left(\tau \text{tr}\left\{E(\mathbf{F}^{(\ell)}(\Phi_S, \Phi_F(v)))^T E(\Phi_G^o)\right\}\right). \end{aligned} \quad (34)$$

3 BAYESIAN INFERENCE

Our goal is to make inferences about the stress tensor parameters (Φ_S, ν) in a particular region of the earth's crust, using a set of focal mechanism observations $\{\Phi_{Gi}^o\}$ and their precisions $\{\tau_i\}$ (for $i = 1, \dots, n$). That is, we wish to evaluate the posterior distribution $p[\Phi_S, \nu | \{\Phi_{Gi}^o, \tau_i\}]$.

From Bayes' Theorem we can derive this posterior in unnormalized form given appropriate priors on the parameters (Φ_S, ν) and assuming independent observations (conditional on the stress model):

$$p[\Phi_S, \nu | \{\Phi_{Gi}^o, \tau_i\}] \propto p(\Phi_S, \nu) p[\{\Phi_{Gi}^o\} | \Phi_S, \nu, (\tau_i)] \propto p(\Phi_S, \nu) \prod_{i=1}^n p(\Phi_{Gi}^o | \Phi_S, \nu, \tau_i). \quad (35)$$

We assume that the stress tensor is constant through the region, and we put uniform priors on ν , Φ_S and $(\phi_F, \cos \theta_F)$ as defined in eqs (26)–(28). Each observed focal mechanism Φ_{Gi}^o is generated from the true focal mechanism Φ_{Gi} through the Matrix–Fisher distribution in eq. (33) with precision parameter τ_i .

In this section, we solve this inference problem in stages. We firstly consider a single observation measured without error (Model 1), then a single observation with error (Model 2), and finally multiple observations with errors (Model 3).

3.1 Model 1: single errorless observation

For this simplest case, we start with the posterior distribution in eq. (35) for a single observation observed without error and for which the fault normal and slip vector can be distinguished. We introduce the Euler angles of the fault plane $\Phi_F = (\phi_F, \theta_F, \psi_F)$ as nuisance parameters that we immediately integrate out:

$$p(\Phi_S, \nu | \Phi_G) \propto p(\Phi_S, \nu) \int d\Phi_F p(\Phi_G | \Phi_F, \Phi_S) p(\psi_F | \phi_F, \theta_F, \Phi_S, \nu) p(\phi_F, \theta_F). \quad (36)$$

The uniform priors imply that $p(\Phi_S, \nu) \propto \sin \theta_S$ and $p(\phi_F, \theta_F) \propto \sin \theta_F$, leading to

$$p(\Phi_S, \nu | \Phi_G) \propto \sin \theta_S \int \sin \theta_F d\Phi_F \delta^3 \{ \Phi_G - \mathbf{F}^{(1)}[\Phi_S, \Phi_F(\nu)] \} \delta [\psi_F - G(\nu, \phi_F, \theta_F)] I(-\hat{\mathbf{u}} \cdot \mathbf{c}_s \geq 0). \quad (37)$$

The function $G(\nu, \phi_F, \theta_F)$ comes from eq. (16) rearranged as

$$\psi_F = G(\nu, \phi_F, \theta_F) = \tan^{-1} \left[\frac{-(1-\nu) \sin \phi_F \cos \phi_F}{(\cos^2 \phi_F + \nu \sin^2 \phi_F) \cos \theta_F} \right]. \quad (38)$$

The condition that the slip vector $\hat{\mathbf{u}}$ be antiparallel to the direction of maximum shear compressive stress \mathbf{c}_s in the fault plane must be explicitly included in the form of the indicator function $I(-\hat{\mathbf{u}} \cdot \mathbf{c}_s \geq 0)$ with $I(\cdot)$ being a function with the value one if its condition is fulfilled and zero otherwise. We have already shown (eqs 19 and 20) that

$$I(-\hat{\mathbf{u}} \cdot \mathbf{c}_s \geq 0) = I\left(\frac{W_{22}}{W_{13}W_{33}} \geq 0\right) = I\left(\frac{\cos \psi_F}{\cos \theta_F} \geq 0\right). \quad (39)$$

We now change the variable of integration from Φ_F to Φ_G , and transform the delta function in ψ_F to one in ν . The Jacobian of the transformation between Euler angles Φ_G and Φ_S which are related by $E(\Phi_G) = E(\Phi_S)R$, for any rotation matrix R , is

$$\left| \frac{\partial(\Phi_F)}{\partial(\Phi_G)} \right| = \left| \frac{\partial(\phi_F, \theta_F, \psi_F)}{\partial(\phi_G, \theta_G, \psi_G)} \right| = \left| \frac{\sin \theta_G}{\sin \theta_F} \right|. \quad (40)$$

Now since W is the matrix $R_{\text{SF}} = E(\Phi_F)$ the angles Φ_F can be extracted from its components by eq. (A5). From eq. (13), the fault normal $\hat{\mathbf{n}}$ is the third column of W and also

$$\nu = -\frac{W_{12}W_{13}}{W_{22}W_{23}} \quad \text{and} \quad \frac{\partial \nu}{\partial \psi_F} = \frac{W_{13}W_{33}}{W_{22}^2W_{23}}. \quad (41)$$

Combining the results in eqs (18), (40) and (41), and carrying out the integration, we obtain

$$p(\Phi_S, \nu | \Phi_G) \propto \sin \theta_S \int d\Phi_G \left| \frac{\partial(\Phi_F)}{\partial(\Phi_G)} \right| \sin \theta_F \delta^3(\Phi_G - \mathbf{F}^{(1)}(\Phi_S, \Phi_F)) \delta(\nu - g(\Phi_F)) \left| \frac{\partial \nu}{\partial \psi_F} \right| I(-\hat{\mathbf{u}} \cdot \mathbf{c}_s \geq 0) \\ = b \sin \theta_S \delta \left(\nu + \frac{W_{12}W_{13}}{W_{22}W_{23}} \right) \left| \frac{W_{13}W_{33}}{W_{22}^2W_{23}} \right| I \left(\frac{W_{22}}{W_{13}W_{33}} \geq 0 \right), \quad (42)$$

where $W = E(\Phi_S)^T E(\Phi_G)$ and $g(\Phi_F)$ is the right hand side of eq. (17). The normalizing factor b may depend on Φ_G but not on Φ_S or ν . However, by observing that $p(\Phi_S, \nu)$ is of the form $b(\Phi_G) \sin \theta_S f(\nu, W)$, it can be shown that b is also independent of Φ_G . To see this we make a change of variables from (Φ_S, Φ_G) to (Φ'_S, Φ'_G) using an arbitrary rotation matrix R :

$$E(\Phi'_S) = RE(\Phi_S) \quad \text{and} \quad E(\Phi'_G) = RE(\Phi_G). \quad (43)$$

It follows that the value of the WB matrix is unchanged

$$W = E(\Phi_S)^T E(\Phi_G) = E(\Phi'_S)^T E(\Phi'_G) = W' \quad (44)$$

and hence

$$\frac{p(\Phi'_S, \nu | \Phi'_G)}{p(\Phi_S, \nu | \Phi_G)} \left| \frac{\partial(\Phi'_S)}{\partial(\Phi_S)} \right| = 1 = \frac{b(\Phi'_G)}{b(\Phi_G)} \quad (45)$$

for any arbitrary pair of Euler angle sets (Φ'_G, Φ_G) , implying that b is independent of Φ_G . Moreover, if we take $\Phi'_G = (0, 0, 0)$, which corresponds to a focal mechanism precisely aligned with geographic coordinates \mathbb{G} , it follows that we can write

$$p(\Phi_S, \nu | \Phi_G) = \frac{\sin \theta_S}{\sin \theta'_S} p(\Phi'_S, \nu | \mathbf{0}) \text{ where } E(\Phi'_S) = E(\Phi_G)^T E(\Phi_S). \quad (46)$$

We will exploit this useful property to evaluate the posterior distribution for arbitrary focal mechanisms Φ_G , and so for the moment we restrict attention to the case where $\Phi_G = (0, 0, 0)$ so that $W = E(\Phi_S)^T$. Referring to the definition of an Euler matrix (eq. (A4)), the posterior distribution then becomes

$$p(\Phi_S, \nu | \mathbf{0}) = \frac{b \sin \theta_S}{|\sin \phi_S \cos \phi_S|} \frac{x(x^2 + \gamma^2)}{(x + \gamma^2)^2} \delta[\nu - g(x, \gamma)] I(\cos \phi_S \cos \theta_S \leq 0) I(0 \leq x \leq 1). \quad (47)$$

Here we have defined

$$\gamma = \gamma(\phi_S, \theta_S) = -\tan \phi_S \cos \theta_S \quad x(\phi_S, \theta_S, \psi_S) = \frac{\gamma(\phi_S, \theta_S)}{\tan \psi_S} \quad (48)$$

so that all of the dependence on ψ_S is contained in x . It follows that

$$g(x, \gamma) = -\frac{W_{12}W_{13}}{W_{22}W_{23}} = \frac{x(1-x)}{x + \gamma^2}. \quad (49)$$

There are two conditions in eq. (47). The first, $\cos \theta_S \cos \phi_S \leq 0$, is derived from the constraint $W_{22}/W_{13}W_{33} \geq 0$. It implies that for $0 \leq \theta_S \leq \frac{\pi}{2}$ then $\frac{\pi}{2} \leq \phi_S \leq \frac{3\pi}{2}$, and if $\frac{\pi}{2} \leq \theta_S \leq \pi$ then $-\frac{\pi}{2} \leq \phi_S \leq \frac{\pi}{2}$. This restricts feasible values of the angles (ϕ_S, θ_S) defining the direction of principal axis of maximum stress to two quadrants of the sphere.

The second condition, $0 \leq x \leq 1$ follows from eq. (49): only in this range of x values is the stress ratio ν in the required range 0 to 1. In the interval $x \in [0, 1]$, ν is zero at $x = 0$ and $x = 1$, and ν attains its maximum value of

$$\begin{aligned} \nu_{\max} &= 1 + 2\gamma^2 - 2\sqrt{\gamma^2(1 + \gamma^2)} \\ &= 1 + 2 \tan^2 \phi_S \cos^2 \theta_S - 2\sqrt{\tan^2 \phi_S \cos^2 \theta_S (1 + \tan^2 \phi_S \cos^2 \theta_S)} \end{aligned} \quad (50)$$

when $x = x_{\max}$, which is given by

$$\begin{aligned} x_{\max} &= -\gamma^2 + \sqrt{\gamma^2(1 + \gamma^2)} \\ &= -\tan^2 \phi_S \cos^2 \theta_S + \sqrt{\tan^2 \phi_S \cos^2 \theta_S (1 + \tan^2 \phi_S \cos^2 \theta_S)}. \end{aligned} \quad (51)$$

The stereonet in Fig. 3 is a contour map of the constraint in eq. (48) and reproduces Fig. 3 of McKenzie (1969) who first derived this result. Note that for a given orientation of Φ_S , the allowed range of the stress ratio ν is strongly restricted by a single focal mechanism.

Integrating eq. (47) over ν removes the delta function, and then integrating again over ψ_S (taking into account the restriction $0 \leq x \leq 1$) yields the posterior probability density for the orientation (ϕ_S, θ_S) of the principal axis of greatest stress:

$$p(\phi_S, \theta_S | \mathbf{0}) = \frac{\sin \theta_S |\cos \theta_S|}{\pi \cos^2 \phi_S} I(\cos \phi_S \cos \theta_S \leq 0) \left[\log \frac{1 + \tan^2 \phi_S \cos^2 \theta_S}{\tan^2 \phi_S \cos^2 \theta_S} - \frac{1}{1 + \tan^2 \phi_S \cos^2 \theta_S} \right]. \quad (52)$$

This density is singular along $\phi_S = 0$ and $\phi_S = \pi$, but the singularity is integrable. We have used the condition that the density integrate to 1 to evaluate the normalizing constant: $b = 1/\pi$.

Fig. 4 is a contour map of the posterior distribution of the angles (ϕ_S, θ_S) for a single error-free observation with $\Phi_G = (\phi_G, \theta_G, \psi_G) = (0, 0, 0)$ as given in eq. (52). For this observation the xy -plane is the fault plane, and the slip is along the positive x -axis.

As we noted earlier, the density diverges along $(\phi_S = \pi, \theta_S < \frac{\pi}{2})$, and we have, therefore, arbitrarily truncated this divergence in order to produce Fig. 4. Most of the probability is concentrated between the reverse of the slip vector (i.e. along the negative x -axis) and the fault normal (along the z -axis: at the centre of the diagram, the most likely orientation being halfway between the negative x - and positive z -axes (the P -axis).

There is a second maximum of the probability distribution close to the y -axis. This arises because when the axis of greatest stress lies in the fault plane, slip along the x -axis will be caused by the shear stress associated with the intermediate stress axis. In that case the most likely orientation of that intermediate axis is in the xz -plane, and hence the axis of greatest stress must lie on or near the y -axis.

Eqs (47) and (52) apply when the slip vector and fault normal can be distinguished. If, as is usual in seismological contexts, these vectors cannot be distinguished, we add this posterior to one where the x - and z -axes in fault plane coordinates are swapped and the y -axis is reversed. We have assumed the focal mechanism has $\Phi_G = (0, 0, 0)$ and $E(\Phi_G) = I$. The alternative interpretation has $\Phi_{Ga} = (0, \frac{\pi}{2}, \pi)$ and $E(\Phi_G) = C_2$ (with C_2 as defined in eq. 25). To evaluate the posterior at (Φ_S, ν) given this alternative observation, we use eq. (47) and evaluate $p(\Phi_{Sa}, \nu | \mathbf{0})$ with $E(\Phi_{Sa}) = C_2 E(\Phi_S)$. The full posterior density is then

$$p^{\text{full}}(\Phi_S, \nu | \mathbf{0}) = \frac{1}{2} p(\Phi_S, \nu | \mathbf{0}) + \frac{1}{2} p(\Phi_{Sa}, \nu | \mathbf{0}). \quad (53)$$

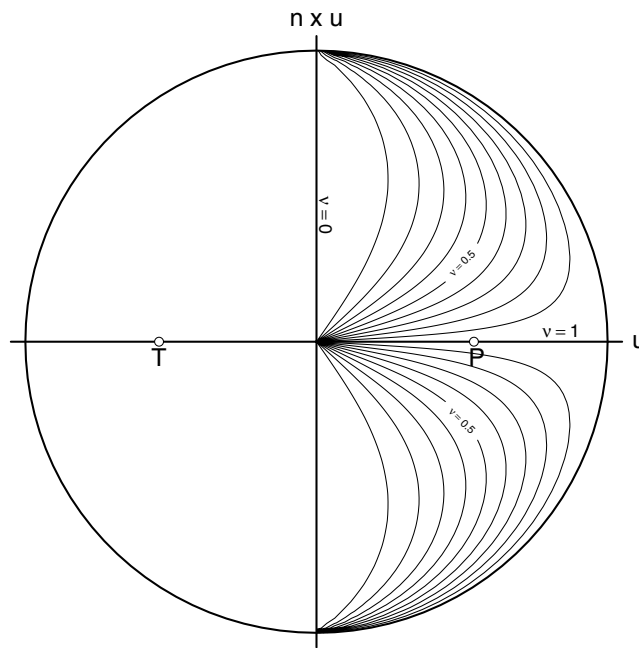


Figure 3. For an earthquake focal mechanism where the fault and slip vector are unambiguously distinguished there is a constraint on the maximum value of the stress ratio ν , depending on the orientation of the principal axis (ϕ_S, θ_S) . This stereonet is a contour map of the function in eq. (48) and reproduces the result derived by McKenzie (1969) (Our diagram is a modified version of Fig. 3 of that paper). The slip vector $\hat{\mathbf{u}}$ is to the right, along the positive x -axis, the fault normal $\hat{\mathbf{n}}$ is vertical (out of the paper) and the null vector $\hat{\mathbf{n}} \times \hat{\mathbf{u}}$ lies in the plane of the paper, along the positive y -axis.

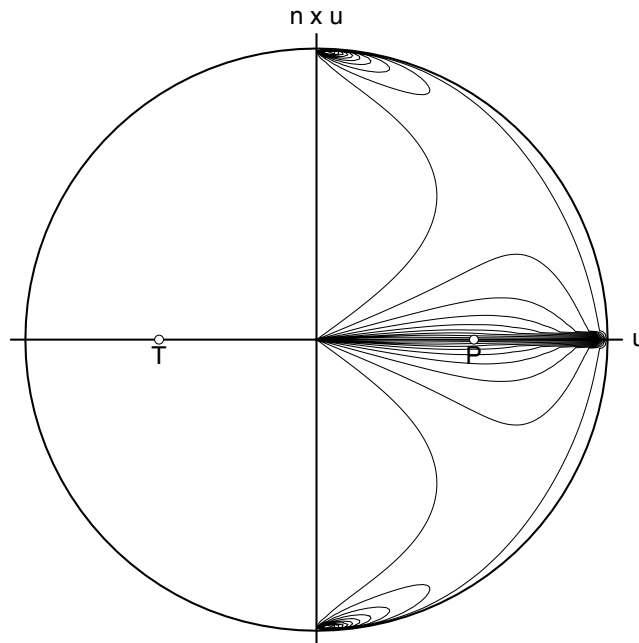


Figure 4. Posterior distributions for the orientation of the principal axis of greatest stress (ϕ_S, θ_S) when the slip and normal vectors are distinguishable. The slip vector is to the right. See text for details.

Fig. 5 shows the full posterior for the case of an observation $\Phi_G = (0, 0, 0)$, when the slip and normal vectors cannot be distinguished. It is very similar to Fig. 4, but the probability density is less concentrated towards the xy -plane, and instead lies symmetrically between the xy - and the yz -planes.

Fig. 6 shows the posterior distributions $p_{\text{full}}(\phi_S, \theta_S | \Phi_G)$ for the four earthquakes shown in Fig. 2. The four focal mechanisms are quite different, and apparently impose quite strong constraints on the direction on the direction of greatest stress. However, it should be recognized that these are the 2-D marginal distributions $p(\phi_S, \theta_S | \Phi_G)$ of the 4-D posterior distribution $p(\phi_S, \theta_S, \psi_S, \nu | \Phi_G)$, and their degree of consistency or otherwise also depends on the distributions of ψ_S and ν .

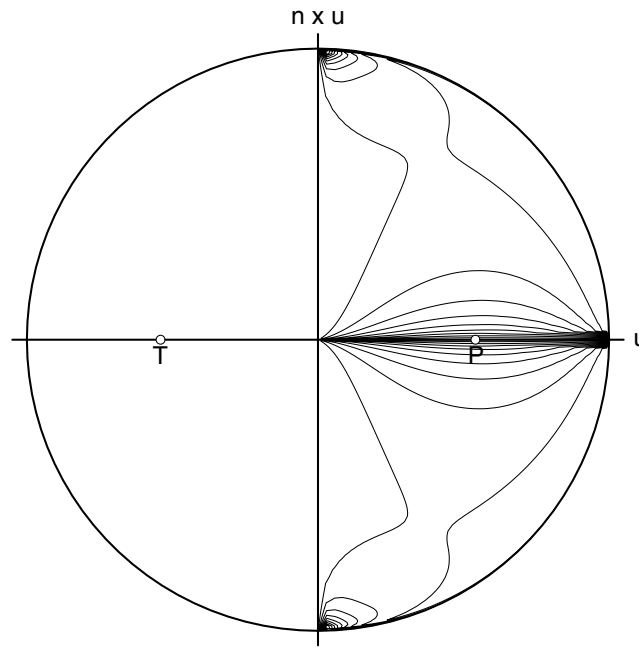


Figure 5. Posterior distributions for the orientation of the principal axis of greatest stress (ϕ_S, θ_S) when the slip and normal vectors cannot be distinguished. See text for details.

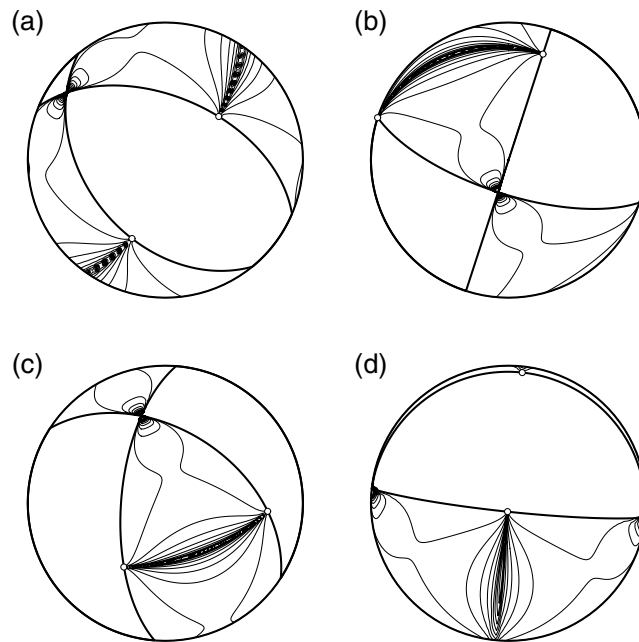


Figure 6. Posterior probability densities of the axes of principal stress for the four earthquakes in Fig. 2 assuming the observations have been made without error.

The main result of this section is the following expression for the normalized posterior probability density of the stress tensor parameters (Φ_S, ν) for a single errorless observation $\Theta \Leftrightarrow \Phi_G$, allowing for fault plane ambiguity:

$$p(\Phi_S, \nu | \Phi_G) = \frac{\sin \theta_S}{2\pi} \sum_{\ell=1}^2 \delta \left[\nu + \frac{W_{12}^{(\ell)} W_{13}^{(\ell)}}{W_{22}^{(\ell)} W_{23}^{(\ell)}} \right] \left| \frac{W_{13}^{(\ell)} W_{33}^{(\ell)}}{(W_{22}^{(\ell)})^2 W_{23}^{(\ell)}} \right| I \left[\frac{W_{13}^{(\ell)} W_{33}^{(\ell)}}{W_{22}^{(\ell)}} \geq 0 \right]. \quad (54)$$

Here the matrices $W^{(\ell)}$ are:

$$W^{(1)} = E(\Phi_S)^T E(\Phi_G) \quad \text{and} \quad W^{(2)} = E(\Phi_S)^T E(\Phi_G) C_2. \quad (55)$$

We refer to this case as Model 1. The presence of the delta function concentrates the likelihood on a 3-D subspace of the 4-D parameter space (Φ_S, ν) . In the absence of error, this is a strong restriction on the possible values. In the next section we consider the consequences of allowing for observational error.

3.2 Model 2: single observation with error

We now generalize Model 1 to include errors in the observation of the true focal mechanism parameters Φ_G . We assume that the observed values Φ_G^o are drawn from the Matrix–Fisher distribution with density $p(\Phi_G^o | \Phi_G, \tau)$ where the parameter τ is known (eq. 33). Given the likelihood of a single observation (eq. 34) derived earlier, we can now write down the posterior distribution of the stress tensor properties conditional on a single observation made with error parameter τ :

$$\begin{aligned} p(\Phi_S, \nu | \Phi_G^o, \tau) &\propto p(\Phi_S, \nu) p(\Phi_G^o | \Phi_S, \nu, \tau) \\ &\propto p(\Phi_S, \nu) \int d\Phi_G \int d\Phi_F p(\Phi_G^o | \Phi_G, \tau) p(\Phi_G | \Phi_F, \Phi_S) p(\phi_F, \theta_F) p(\psi_F | \phi_F, \theta_F, \nu) \\ &\propto p(\Phi_S, \nu) \sum_{\ell=1}^2 \int d\Phi_G \int d\Phi_F p(\Phi_G^o | \Phi_G, \tau) \delta^3[\Phi_G - \mathbf{F}^{(\ell)}(\Phi_F, \Phi_S)] p(\phi_F, \theta_F) \delta[\psi_F - G(\phi_F, \theta_F, \nu)]. \end{aligned} \quad (56)$$

The triple integration over Φ_G is a simple integration over three delta functions which can be carried out immediately, and assuming uniform priors the result takes the form

$$p(\Phi_S, \nu | \Phi_G^o, \tau) \propto \sin \theta_S \sum_{\ell=1}^2 \int \int d\phi_F d\theta_F \sin \theta_F p\{\Phi_G^o | \Phi_G = \mathbf{F}^{(\ell)}[\Phi_F(\nu), \Phi_S], \tau\}. \quad (57)$$

This equation represents Model 2: a single observation with error. With the Matrix–Fisher error mechanism this leads to

$$p(\Phi_S, \nu | \Phi_G^o, \tau) \propto \sin \theta_S \sum_{\ell=1}^2 \int \int d\phi_F d\theta_F \sin \theta_F \exp[\tau \operatorname{tr}(E(\Phi_F(\nu))^T E(\Phi_S)^T E(\Phi_G^o) C_2^{\ell-1})]. \quad (58)$$

The double integral over (ϕ_F, θ_F) must be done numerically (e.g. using an adaptive integrator). If the precision parameter τ is sufficiently large that the function $p(\Phi_G^o | \Phi_G, \tau)$ may be approximated by a delta function, then we recover Model 1.

As in Model 1, aside from the factor $\sin \theta_S$ the stress orientations Φ_S and the observation Φ_G^o appear in eq. (55) *only* in the combination $E(\Phi_S)^T E(\Phi_G^o)$. As before, if we define angles Φ_S' such that $E(\Phi_S') = E(\Phi_G^o)^T E(\Phi_S)$ and let $\Phi_G^o = (0, 0, 0)$ so that $E(\Phi_G^o) = I$, then we have

$$p(\Phi_S, \nu | \Phi_G^o, \tau) = p(\Phi_S', \nu | \mathbf{0}, \tau) \frac{\sin \theta_S}{\sin \theta_S'} \quad (59)$$

For fixed τ it is sufficient to tabulate the function

$$q(\Phi_S, \nu | \tau) = \int \int d\phi_F d\theta_F \sin \theta_F \exp[\tau \operatorname{tr}(E(\Phi_F(\nu))^T E(\Phi_S)^T)] \quad (60)$$

on a grid of (Φ_S, ν) values, and use this tabulation as a look-up table to evaluate by interpolation

$$p(\Phi_S, \nu | \Phi_G^o, \tau) \propto \sin \theta_S \sum_{\ell=1}^2 q(\Phi_S^{(\ell)}, \nu | \tau), \quad (61)$$

where $E(\Phi_S^{(1)}) = C_2 E(\Phi_S^{(2)}) = E(\Phi_G^o)^T E(\Phi_S)$.

Figs 7 and 8 show the constraint on the orientation of the axis of greatest stress for the same four focal mechanisms from the San Fernando sequence that were plotted in Fig. 6. The most notable difference is the spreading out of the contours by the observational errors. The plots show the effect of changing the precision parameter from $\tau = 20$ (observational errors of approximately 10°) to $\tau = 5$ (20° error).

3.3 Model 3: multiple observations with error

The extension of Model 2 to multiple observations with error is straightforward. Given n observations $\{\Phi_{Gi}^o\}$ ($i = 1, \dots, n$) with their Matrix–Fisher precisions $\{\tau_i\}$ the posterior distribution we seek is

$$\begin{aligned} p[\Phi_S, \nu | \{\Phi_{Gi}^o, \tau_i\}] &\propto p(\Phi_S, \nu) p[\{\Phi_{Gi}^o\} | \Phi_S, \nu, \{\tau_i\}] \\ &\propto p(\Phi_S, \nu) \prod_{i=1}^n p(\Phi_{Gi}^o | \Phi_S, \nu, \tau_i) \\ &\propto \sin \theta_S \prod_{i=1}^n \sum_{\ell=1}^2 \int \int d\phi_F d\theta_F \sin \theta_F p\{\Phi_{Gi}^o | \Phi_G = \mathbf{F}^{(\ell)}[\Phi_F(\nu), \Phi_S], \tau_i\} \end{aligned}$$

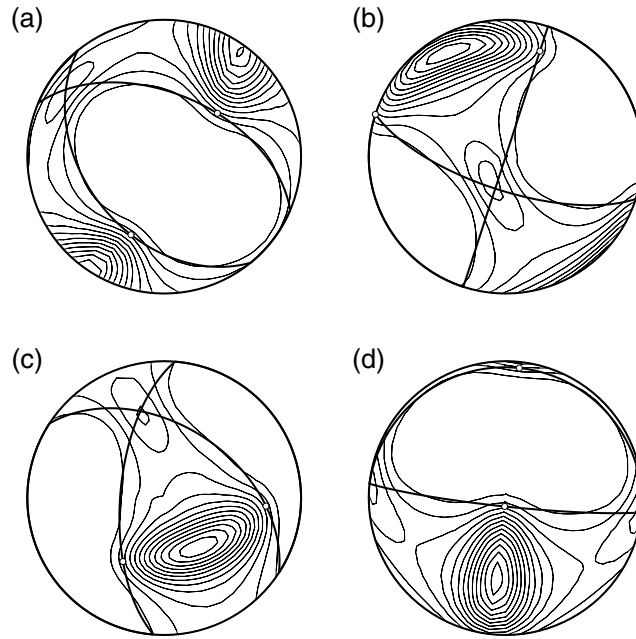


Figure 7. Posterior probability densities of the principal axis of maximum stress for the four earthquakes in Fig. 2 assuming the observations have been made with 10° error ($\tau = 20$).

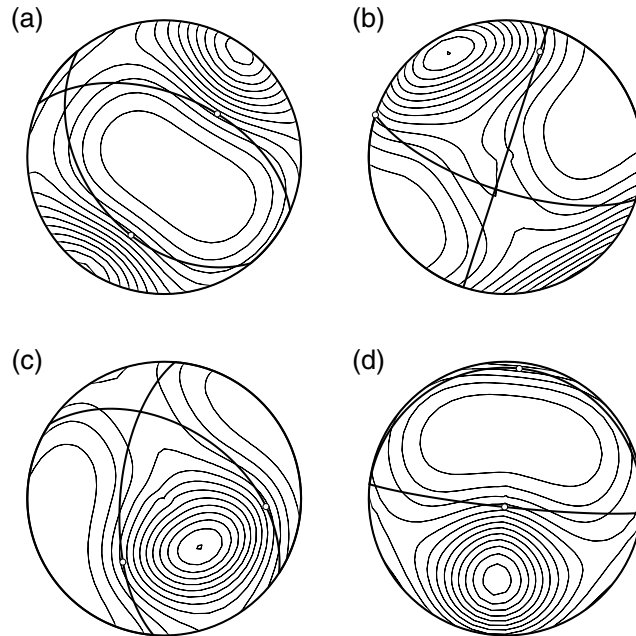


Figure 8. Posterior probability densities of the principal axis of maximum stress for the four earthquakes in Fig. 2 assuming the observations have been made with 20° error ($\tau = 5$).

$$\begin{aligned} &\propto \sin \theta_S \prod_{i=1}^n \sum_{\ell=1}^2 \int \int d\phi_F d\theta_F \sin \theta_F \exp \left[\tau_i \operatorname{tr} \left(E(\Phi_F(\nu))^T E(\Phi_S)^T E(\Phi_{Gi}^o) C_2^{\ell-1} \right) \right] \\ &\propto \sin \theta_S \prod_{i=1}^n \sum_{\ell=1}^2 q(\Phi_{Si}^{(\ell)}, \nu | \tau_i) \end{aligned}$$

where $E(\Phi_{Si}^{(1)}) = E(\Phi_{Gi}^o)^T E(\Phi_S)$, and $C_2 E(\Phi_{Si}^{(2)}) = E(\Phi_{Gi}^o)^T E(\Phi_S)$.

The prior on the stress parameters $p(\Phi_S, \nu)$ appears as an overall multiplier in front of these expressions. In large datasets only a highly informative (strongly peaked) prior distribution will affect the posterior.

Aside from the factor of $\sin \theta_S$, the posterior distribution is a product of the individual posterior distributions of each of the focal mechanisms. For example, the individual posterior distributions for the four earthquake focal mechanisms from the San Fernando sequence

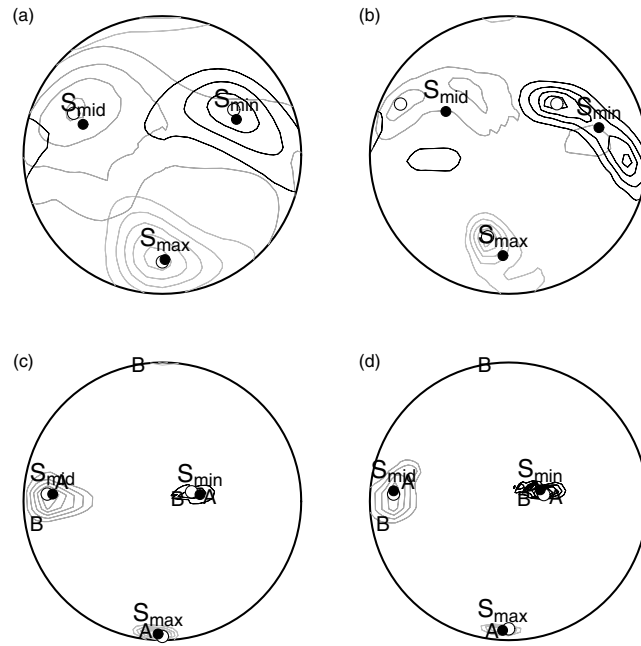


Figure 9. Posterior probability densities of the three axes of principal stress for all earthquakes in the San Fernando sequence. Panels (a) and (b) use data from four earthquakes only, and assume the observations have been made with 20° error ($\tau = 5$) and 10° ($\tau = 20$) error, respectively. Panels (c) and (d) show the equivalent posterior densities using all 76 earthquakes. The open circles show the positions of the marginal posterior modes of the three axes separately, and the filled circles show the posterior mean of all three axes simultaneously (see text for details). The two best-fitting solutions of Gephart & Forsyth (1984) are marked A and B in panels (c) and (d).

from Gephart & Forsyth (1984) shown in Figs 7 and 8 can be simply combined, and Figs 9(a) and (b) show the corresponding posterior distributions for the directions of the principal stress axes. In these diagrams the contours are the three separate marginal distributions of the directions of the three axes of principal stress. As is to be expected, in the case when the precision is low ($\tau = 5$, Fig. 9a), the contours spread more widely than where the precision is high ($\tau = 20$, Fig. 9b).

An open circle marks the posterior mode in each case. For example, the marginal distribution of the the direction of S_{\max} is the distribution of the pair of angles (ϕ_S, θ_S) which is calculated by integrating the posterior distribution over ψ_S and ν :

$$p(\phi_S, \theta_S | \{\Phi_{G_i}^o, \tau_i\}) = \int \int p(\Phi_S, \nu | \{\Phi_{G_i}^o, \tau_i\}) d\psi_S d\nu. \quad (63)$$

The three posterior modes are not guaranteed to be orthogonal to one another. The filled circles in Fig. 9 show the posterior mean of the Euler angles Φ_S calculated using the axial data method described in Section A5. These directions are orthogonal by definition. There is close agreement between the marginal posterior modes and the posterior mean in the low precision solution (Fig. 9a), but some disagreement in the high precision solution, especially for the intermediate stress axis S_{mid} , suggesting that the high precision solution is not justified by the data. The misfit angles found by Gephart & Forsyth (1984), which are typically in the range $10\text{--}30^\circ$ also suggest that the low precision solution ($\tau = 5$, 20° error) is more appropriate than the high precision one ($\tau = 20$, 10° error).

When all 76 focal mechanisms from the sequence are combined, as shown in Figs 9(c) and (d), the posterior distributions narrow significantly. We have plotted (A and B) the two best-fitting solutions found by Gephart & Forsyth (1984) and our results agree well with their solution A.

3.4 Posterior summaries

The full posterior density distribution $p[\Phi_S, \nu | (\Phi_{G_i}^o, \tau_i)]$ for the 76 earthquake San Fernando sequence is shown as six 2-D marginal densities in Fig. 10. The diagram clearly shows how the symmetries in the orientation angles lead to multiple modes in the posterior, although these modes are all equivalent.

A posterior distribution such as this can be summarized in various convenient ways. The maximum *a posteriori* (MAP) estimate, or posterior mode, can be located on a grid of values or by using an optimizing algorithm. The posterior mean of ν is also straightforward:

$$\hat{\nu} = \int_0^1 d\nu \int \int \int d\Phi_S \nu p[\Phi_S, \nu | (\Phi_{G_i}^o, \tau_i)]. \quad (64)$$

However, care must be taken when computing the posterior mean of the angles of the orientation of the stress tensor Φ_S . The multiple modes mean that simply averaging the values of Φ_S using the posterior density will lead to the wrong result. Averaging angles is in any case problematic because of their cyclic nature. We summarize methods for averaging angular data in Section A5. For the case of the stress

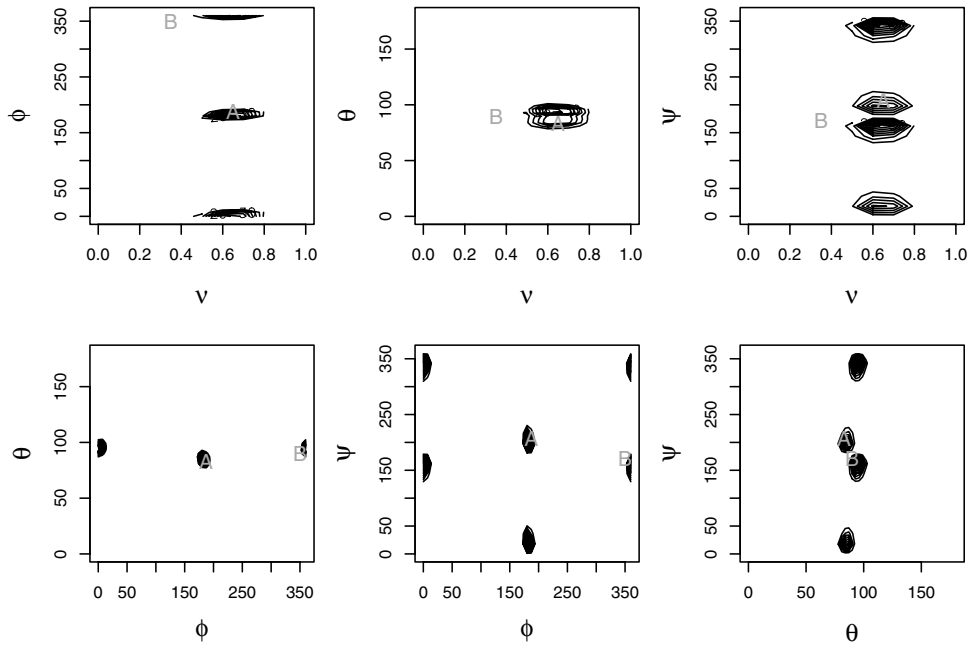


Figure 10. 2-D marginal posterior probability densities for all 76 earthquakes in the San Fernando sequence assuming the observations have been made with 20° error ($\tau = 5$). The two best-fitting solutions of Gephart & Forsyth (1984) are marked A and B.

orientation angles it is appropriate to compute the posterior mean as follows (*cf.* eq. A19). Form three matrices A_k ($k = 1, 2, 3$)

$$A_k = \int_0^1 d\nu \int \int \int d\Phi_S E(\Phi_S)_{\bullet k} [E(\Phi_S)_{\bullet k}]^T p[\Phi_S, \nu | (\Phi_{Gi}^o, \tau_i)] \quad (65)$$

where $E(\Phi_S)_{\bullet k}$ is the k^{th} column of $E(\Phi_S)$. Let $\hat{\nu}_k$ be the (normalized) eigenvector corresponding to the largest eigenvalue of A_k and let X be the matrix formed by combining these three eigenvectors:

$$X = [\hat{\nu}_1 \ \hat{\nu}_2 \ \hat{\nu}_3] \quad (66)$$

Then the posterior mean $\hat{\Phi}_S$ corresponds to the polar (orthogonal) part of the matrix X , that is,

$$E(\hat{\Phi}_S) = X[(X^T X)^{1/2}]^{-1} \quad (67)$$

and $(X^T X)^{1/2}$ is the positive definite symmetric square root (see Section A5 and chapter 13 of Mardia & Jupp 2000).

The marginal posterior density of an arbitrary function of the stress parameters $g(\Phi_S, \nu)$ can be calculated from a tabulation of the posterior on a grid of (Φ_S, ν) values. First compute g everywhere on the grid, and then form a weighted kernel density estimate of the distribution of g using the posterior probability of each grid point as the weight of that point. Posterior densities such as these can be used to form interval estimates ('Bayesian credible intervals') of the uncertainties of scalar parameters.

Marginal posterior densities of the stress ratio ν and the azimuths of the horizontal components of three principal stress axes are given for the San Fernando sequence in Fig. 11. Also plotted is the distribution of the azimuth of maximum horizontal compressive stress $\phi_{SH_{\max}}$ (as defined by eq. 70 in Section 4.1 below).

4 APPLICATION TO STRESS MAPPING

In this section we apply the Bayesian inversion technique described in Section 2 to focal mechanism observations from Japan, and map the azimuth $\phi_{SH_{\max}}$ of maximum horizontal compressive stress $S_{H_{\max}}$. Our purpose here is to show that under the conditions of uniform priors that we have assumed, our method (Model 3) obtains results similar to those from the widely used alternative of Michael (1984, 1987).

4.1 Computing the axis of maximum horizontal compressive stress

Given a stress tensor \mathcal{S} oriented at Φ_S with respect to geographic coordinates \mathbb{G} we can compute the direction of maximum compressive horizontal compressive stress, $S_{H_{\max}}$, even when none of the three principal stress directions is strictly horizontal and the only information about stress magnitudes is that represented by ν (Lund 2000; Lund & Townend 2007).

Consider a fault normal which is horizontal in geographic coordinates, at azimuth ϕ_G (i.e. $\theta_G = \frac{\pi}{2}$). Then

$$\hat{\mathbf{n}}_{\mathbb{G}} = [\cos \phi_G \ \sin \phi_G \ 0]_{\mathbb{G}}^T. \quad (68)$$

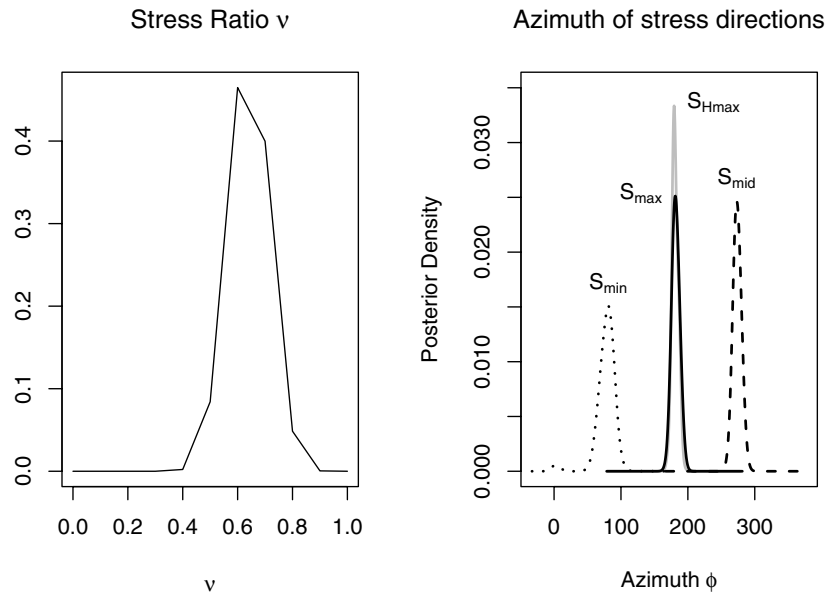


Figure 11. Posterior distributions of the stress ratio ν , the azimuths of the of principal stress directions, and the direction of maximum horizontal compressive stress ϕ_{SHmax} for all 76 earthquakes in the San Fernando sequence.

Estimating the direction of maximum horizontal compression is equivalent to identifying the plane whose normal stress (acting in the direction of $\hat{\mathbf{n}}$) is greatest. As described in more detail by Lund & Townend (2007), this can be done by representing $\hat{\mathbf{n}}$ with respect to principal stress coordinates and obtaining an expression for the (effective) normal stress c'_n . Eq. (10) implies that maximizing c'_n is equivalent to minimizing $\kappa = n_{1S}^2 + \nu n_{2S}^2$, that is,

$$\kappa = (R_{11} \cos \phi_G + R_{21} \sin \phi_G)^2 + \nu (R_{12} \cos \phi_G + R_{22} \sin \phi_G)^2, \quad (69)$$

where R_{ij} are the components of $R_{GS} = E(\Phi_S)$. On differentiation, this yields

$$\tan(2\phi_G^m) = \frac{2[R_{11}R_{21} + \nu R_{12}R_{22}]}{(R_{11}^2 - R_{21}^2) + \nu(R_{12}^2 - R_{22}^2)} \quad (70)$$

and we check whether this yields the desired minimum, and hence the azimuth of maximum horizontal compressive stress, $\phi_{SHmax} = \phi_G^m$, using a second derivative test. If ϕ_G^m yields a maximum, then $\phi_{SHmax} = \phi_G^m + \pi/2$.

4.2 Application to a focal mechanism data set from Japan

The data used here are presented and discussed in some detail in Townend & Zoback (2006). Briefly, the data are 316 clusters of 15 657 focal mechanisms shallower than 35 km in central Japan. Each cluster contains at least 20 focal mechanisms, and no filtering has been undertaken on the basis of focal mechanism uncertainties. We assume that all earthquakes in a given cluster were caused by a single stress tensor.

Fig. 12(a) shows the orientation of S_{Hmax} estimated for the 316 clusters of focal mechanisms. For each cluster a wedge-symbol representing an 80 per cent posterior credible interval is shown. There is strong spatial correlation between the orientations of neighbouring clusters, and also a pronounced rotation across the map. Fig. 12(b) displays the orientation of S_{Hmax} for the same data, but estimated with the Michael (1984, 1987) algorithm and using the bootstrap to compute 80 per cent intervals. Note that neither the Bayesian nor the Michael bootstrap intervals are necessarily symmetric about the estimate.

The differences between the estimates obtained with the two methods are shown in Fig. 12(c), and we see that there is good agreement between the two approaches. The ϕ_{SHmax} estimates themselves are plotted as white (open) circles in Fig. 12(d), and are distributed tightly around the 45° line of equality. Also shown in Fig. 12(d) using black (closed) circles are the widths of the 80 per cent interval estimates. There is also close agreement here, although the Bayesian uncertainties tend to be larger than those from the Michael method where the uncertainty is less than about 30°. Otherwise the Bayesian intervals are narrower than the Michael intervals.

5 CONCLUSIONS

There have been a number of approaches to the problem of using data to put realistic constraints on the properties of the stress tensor. The Bayesian probabilistic approach we have taken here provides a coherent means of combining focal mechanism data, clearly laying out the assumptions that are made, and modelling the entire data generation process. In particular, the most important assumptions are the WB

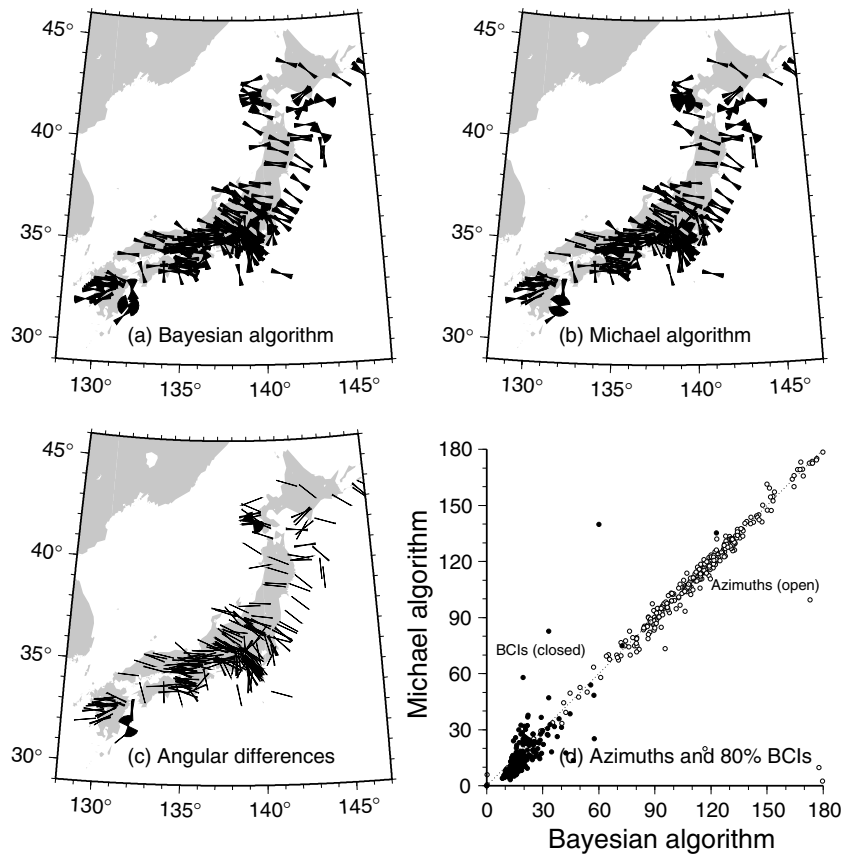


Figure 12. A summary of the results obtained using the Bayesian inversion method of Section 2 (‘Bayesian’) and the Michael (1984, 1987) algorithm (‘Michael’) when applied to 316 clusters of focal mechanisms in Japan. (a) Map of the four main islands of Japan showing the axis of maximum horizontal compression (S_{Hmax}) calculated using eq. (67). The wedge-symbols span the 80 per cent confidence interval for each S_{Hmax} estimate. (b) S_{Hmax} results calculated using the Michael (1984, 1987) algorithm with duplicated fault planes (cf. Fig. 4 of Townend & Zoback 2006). (c) Map showing the azimuthal difference between the two sets of results at each cluster location (i.e. the acute angle between the two estimated directions). (d) Cross-plot illustrating the close agreement between the S_{Hmax} azimuths (open circles) and 80 per cent Bayesian credible interval widths (‘BCIs’) estimated using the two algorithms. The dashed line indicates equal values.

hypothesis (Section 2.3), randomly oriented fault planes (Section 2.5) and the assumption that sets of neighbouring earthquakes are driven by the same stress tensor. Our method also enables us to combine observations of different quality, by including a precision parameter τ_i for each observation (Section 2.4).

The output of the method is the 4-D posterior density function $p(\Phi_S, \nu | \text{data})$ which can be summarized in any convenient manner. We have used the direction of maximum horizontal compressive stress ϕ_{SHmax} as an example, since it is an informative scalar indicator of the stress field that is easily mapped.

There are several obvious extensions to the model we have presented here. Although the model framework we have constructed is fully general, we have demonstrated its properties with non-informative priors, and a Matrix–Fisher error distribution with a scalar parameter for the error distribution. We have chosen the Matrix–Fisher distribution for its simplicity and tractability, and the fact that it is a distribution over orthogonal matrices and is therefore free of the known singularities of distributions specified by, say, Euler angles. Moreover, since focal mechanism data are typically reported with only a scalar measure of accuracy, it is appropriate to use a distribution characterized by a single scalar τ . More complex error distributions, in particular those which allow the errors in strike, dip and rake to be unequal, can be straightforwardly incorporated, although at increased computational cost since the convenient invariance property (eq. 59) can no longer be exploited. Little research has been done into the nature of focal mechanism error distributions, but future research in this area should lead to methods for estimating τ , or the parameters of more complex error distributions if these can be identified.

We noted above (Section 3.3) that the prior information $p(\Phi_S, \nu)$ on the stress tensor appears in the posterior distribution as an overall multiplier. The posterior is, therefore, less sensitive to any but the most concentrated priors on the stress parameters. Moreover, modifying the stress prior is simpler than modifying the prior on the faults because the fault prior is embedded inside the integrals that must be evaluated. A more restrictive prior distribution on the orientation of faults $p(\phi_F, \theta_F | \Phi_S)$ would be appropriate in cases where the fault diversity is very low (e.g. Hardebeck 2006). Furthermore, the idealized distribution of focal mechanism uncertainties represented by eq. (33) will be relaxed in ongoing work to better represent real observational errors. It is also possible to allow the likelihood of an earthquake to depend on the frictional properties of the fault. In that case, all three parameters of the effective stress tensor (S'_{max}, D, ν) appear in the model.

In future work it will be desirable to incorporate the spatial correlation in stress between neighbouring locations visible in maps such as Fig. 12(a). This topic has been addressed recently by Hardebeck & Michael (2006), and enables prior knowledge of, for example, the regional azimuth of maximum horizontal compressive stress determined from borehole data or by other means to be incorporated in inversions.

In summary, we have presented a general framework for the focal mechanism-stress inversion problem which models the entire data generation process. It allows for the explicit inclusion of prior information, and empirically determined focal mechanism error distributions neither of which is generally possible with existing methods.

ACKNOWLEDGMENTS

We thank Björn Lund for feedback on early drafts of this manuscript and Thorsten Becker, Jeanne Hardebeck and an anonymous reviewer for their helpful comments.

REFERENCES

- Abers, G.A. & Gephart, J.W., 2001. Direct inversion of earthquake first motions for both the stress tensor and focal mechanisms and application to southern California, *J. Geophys. Res.*, **106**, 26 523–26 540.
- Aki, K. & Richards, P.G., 1980. *Quantitative Seismology*, 1st edn, W.H. Freeman and Co., San Francisco.
- Angelier, J., 1979. Determination of the mean principal directions of stresses for a given fault population, *Tectonophysics*, **56**, T17–T26.
- Angelier, J., 1984. Tectonic analysis of fault slip data sets, *J. Geophys. Res.*, **89**, 5835–5848.
- Arnold, A., Townend, J. & Vignaux, T., 2005. Mapping tectonic stress using earthquakes, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, San José, CA, Vol. CP803, eds Knuth, K.H., Abbas, A.E., Morris, R.D. & Castle, J.P., American Institute of Physics, Melville, NY.
- Balfour, N.J., Savage, M.K. & Townend, J., 2005. Stress and crustal anisotropy in Marlborough, New Zealand: evidence for low fault strength and structure-controlled anisotropy, *Geophys. J. Int.*, **163**, doi:10.1111/j.1365-246X.2005.02783.x.
- Becker, T., Hardebeck, J. & Anderson, G., 2004. Constraints on fault slip rates of the southern California plate boundary from GPS velocity and stress inversions, *Geophys. J. Int.*, **160**, 634–650, doi:10.1111/j.1365-246X.2004.02528.x.
- Bohnhoff, M., Gresser, H. & Dresen, G., 2006. Strain partitioning and stress rotation at the North Anatolian fault zone from aftershock focal mechanisms of the 1999 Izmit $M_w = 7.4$ earthquake, *Geophys. J. Int.*, **166**, 373–385, doi:10.1111/j.1365-246X.2006.03027.x.
- Boness, N.L. & Zoback, M.D., 2006. Mapping stress and structurally controlled crustal shear velocity anisotropy in California, *Geology*, **34**, 825–828, doi:10.1130/G22309.1.
- Bott, M.H.P., 1959. The mechanisms of oblique slip faulting, *Geol. Mag.*, **96**, 109–117.
- Gephart, J.W. & Forsyth, D.W., 1984. An improved method for determining the regional stress tensor using earthquake focal mechanism data: application to the San Fernando earthquake sequence, *J. Geophys. Res.*, **89**, 9305–9320.
- Hardebeck, J., 2006. Homogeneity of small-scale earthquake faulting, stress, and fault strength, *Bull. Seism. Soc. Am.*, **96**, 1675–1688, doi:10.1785/0120050257.
- Hardebeck, J. & Michael, A., 2006. Damped regional-scale stress inversions: methodology and examples for southern California and the Coalinga aftershock sequence, *J. Geophys. Res.*, **111**, doi:10.1029/2005JB004144.
- Hardebeck, J.L. & Hauksson, E., 1999. Role of fluids in faulting inferred from stress field signatures, *Science*, **285**, 236–239.
- Hardebeck, J.L. & Michael, A.J., 2004. Stress orientations at intermediate angles to the San Andreas fault, California, *J. Geophys. Res.*, **109**, doi:10.1029/2004JB003239.
- Hickman, S. & Zoback, M.D., 2004. Stress orientations and magnitudes in the SAFOD pilot hole, *Geophys. Res. Lett.*, **31**, doi:10.1029/2004GL020043.
- Lund, B., 2000. Crustal stress studies using microearthquakes and boreholes, *PhD thesis*, Uppsala University.
- Lund, B. & Townend, J., 2007. Calculating horizontal stress orientations with full or partial knowledge of the tectonic stress tensor, *Geophys. J. Int.*, submitted, doi:10.1111/j.1365-246X.2007.03468.x.
- Mardia, K.V. & Jupp, P.E., 2000. *Directional Statistics*, Wiley, Chichester.
- McKenzie, D., 1969. The relationship between fault plane solutions for earthquakes and the directions of the principal stresses, *Bull. Seism. Soc. Am.*, **59**, 591–601.
- Michael, A.J., 1984. Determination of stress from slip data: faults and folds, *J. Geophys. Res.*, **89**, 11 517–11 526.
- Michael, A.J., 1987. Use of focal mechanisms to determine stress: a control study, *J. Geophys. Res.*, **92**, 357–368.
- Provost, A.-S. & Houston, H., 2003. Investigation of temporal variations in stress orientations before and after four major earthquakes in California, *Phys. Earth planet. Inter.*, **139**, 255–267.
- Ratchkovski, N.A., 2003. Change in stress directions along the central Denali fault, Alaska, after the 2002 earthquake sequence, *Geophys. Res. Lett.*, **30**, doi:10.1029/2003GL017905.
- Ratchkovski, N., Wiemer, S. & Hansen, R., 2004. Seismotectonics of the central Denali fault, Alaska, and the 2002 Denali fault earthquake sequence, *Bull. Seism. Soc. Am.*, **94**, S156–S174.
- Rivera, L. & Cisternas, A., 1990. Stress tensor and fault plane solutions for a population of earthquakes, *Bull. Seism. Soc. Am.*, **80**, 600–614.
- Roman, D.C., Moran, S.C., Power, J.A. & Cashman, K.V., 2004. Temporal and spatial variation of local stress fields before and after the 1992 eruptions of Crater Peak vent, Mount Spurr volcano, Alaska, *Bull. Seism. Soc. Am.*, **94**, 2366–2379.
- Scholz, C.H., 2002. *The mechanics of earthquakes and faulting*, 2nd edn, Cambridge University Press, Cambridge.
- Sipkin, S.A. & Silver, P.G., 2003. Characterization of the time-dependent strain field at seismogenic depths using first-motion focal mechanisms: observations of large-scale decadal variations in stress along the San Andreas fault system, *J. Geophys. Res.*, **108**, doi:10.1029/2002JB002064.
- Sivia, D., 1996. *Data Analysis: A Bayesian Tutorial*, Oxford University Press, Oxford.
- Stein, S. & Wysession, M., 2003. *An Introduction to Seismology, Earthquakes, and Earth Structure*, 1st edn, Blackwell Publishing, Malden, MA.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia.
- Townend, J., 2003. Mechanical constraints on the strength of the lithosphere and plate-bounding faults, *PhD thesis*, Stanford University.
- Townend, J., 2006. What do faults feel? Observational constraints on the stress acting on seismogenic faults, in *Earthquakes: Radiated Energy and Physics of Faulting*, Vol. 170, Geophysical Monograph Series, pp. 313–327, eds Abercrombie, R., McGarr, A., Kanamori, H. & Di Toro, G., AGU, Washington, doi:10.1029/170GM31.
- Townend, J. & Zoback, M.D., 2001. Implications of earthquake focal mechanisms for the frictional strength of the San Andreas fault system, in *The Nature and Significance of Fault Zone Weakening*, Vol. 186, pp. 13–21, eds Holdsworth, R.E., Strachan, R.A., Macloughlin, J. & Knipe, R.J., Geological Society of London Special Publication.
- Townend, J. & Zoback, M.D., 2004. Regional tectonic stress near the San Andreas fault in central and southern California, *Geophys. Res. Lett.*, **31**, doi:10.1029/2003GL018918.

- Townend, J. & Zoback, M.D., 2006. Stress, strain, and mountain building in central Japan, *J. Geophys. Res.*, **111**, doi:10.1029/2005JB003759.
- Wallace, R.E., 1951. Geometry of shearing stress and relation to faulting, *J. Geol.*, **59**, 118–130.
- Wéber, Z., 2006. Probabilistic local waveform inversion for moment tensor and hypocentral location, *Geophys. J. Int.*, **165**, 607–621, doi:10.1111/j.1365-246X.2006.02934.x.
- Xu, P., 2004. Determination of regional stress tensors from fault-slip data, *Geophys. J. Int.*, **157**, 1316–1330, doi:10.1111/j.1365-246X.2004.02271.x.

APPENDIX A: PROPERTIES OF ROTATION MATRICES

A1 Euler angles

A general rotation in three dimensions can be described by a set of three Euler angles $\Phi = (\phi, \theta, \psi)$ with

$$0 \leq \phi \leq 2\pi, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \psi \leq 2\pi. \quad (\text{A1})$$

A rotation matrix R is formed from these angles by

$$R = E(\Phi) = E(\phi, \theta, \psi) = B_3(\phi)B_2(\theta)B_3(\psi) \quad (\text{A2})$$

where the standard matrices $B_i(\alpha)$ describe a rotation of angle α anticlockwise about axis i :

$$B_2(\alpha) = \begin{bmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{bmatrix}; \quad B_3(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (\text{A3})$$

It follows that

$$E(\Phi) = \begin{bmatrix} \cos \phi \cos \theta \cos \psi - \sin \phi \sin \psi \\ \sin \phi \cos \theta \cos \psi + \cos \phi \sin \psi \\ -\sin \theta \cos \psi \\ -\cos \phi \cos \theta \sin \psi - \sin \phi \cos \psi \cos \phi \sin \theta \\ -\sin \phi \cos \theta \sin \psi + \cos \phi \cos \psi \sin \phi \sin \theta \\ \sin \theta \sin \psi \quad \cos \theta \end{bmatrix}. \quad (\text{A4})$$

A rotation matrix R is always orthogonal (i.e. $R^T R = I$: the columns and rows of R are orthonormal), and has $\det R = +1$ (i.e. we disallow inversions: transformations where space is reflected in any plane).

Any rotation matrix R can be decomposed according to eq. (A2), and hence there always exist angles Φ equivalent to rotation matrix R :

$$\phi = \tan^{-1} \frac{R_{23}}{R_{13}} \quad \theta = \cos^{-1} R_{33} \quad \psi = \tan^{-1} \frac{R_{32}}{-R_{31}}. \quad (\text{A5})$$

A2 Rotation and misfit angles

A rotation matrix R rotates a coordinate system around an axis $\hat{\mathbf{x}}$ by an angle α defined by

$$\alpha = \pm \cos^{-1} \left(\frac{\text{tr}\{R\} - 1}{2} \right). \quad (\text{A6})$$

The three eigenvalues of R are: $\{1, e^{\pm i\alpha}\}$ and the rotation axis $\hat{\mathbf{x}}$ is the eigenvector associated with the eigenvalue 1.

Since the product of two orthogonal matrices is also orthogonal, we can define a scalar distance measure between two orthogonal matrices R_1 and R_2 as the rotation angle or ‘misfit angle’ α corresponding to the orthogonal matrix $R = R_1^T R_2$. In general, α is the angle by which the coordinate system defined by R_1 must be rotated to coincide with that of R_2 .

The entries of R are all the possible scalar products of the columns of R_1 and R_2 . If R_1 and R_2 are identical then R is the identity matrix and the misfit angle is zero.

A3 Geographic coordinates and focal mechanisms

Earthquake focal mechanisms are by convention represented by the set of angles $\Theta = (\xi, \delta, \lambda)$ called the strike, dip, and rake (e.g. Aki & Richards 1980). These three angles define fault plane coordinates \mathbb{F} with unit vectors along the slip, null and normal vectors of the focal mechanism. The angles Θ are equivalent to a set of Euler angles $\Phi = (\phi, \theta, \psi)$:

$$\begin{aligned} \phi &= \xi + \frac{\pi}{2} & \xi &= \phi - \frac{\pi}{2} \\ \theta &= \pi - \delta & \delta &= \pi - \theta \\ \psi &= \lambda - \frac{\pi}{2} & \lambda &= \psi + \frac{\pi}{2} \end{aligned} \quad (\text{A7})$$

That is, the rotation matrix R associated with the angles Θ is, therefore, $E[\Phi(\Theta)] = B_3(\xi + \frac{\pi}{2})B_2(\pi - \delta)B_3(\lambda - \frac{\pi}{2})$.

A4 Random rotation matrices

The Matrix–Fisher distribution $R \sim \text{Matrix–Fisher}(R_0, \tau)$ with orthogonal matrix mode R_0 and scalar concentration parameter τ describes a probability distribution of the orthogonal matrix R , or equivalently the distribution of the unit vectors of a coordinate system. It has probability density proportional to

$$f(R | R_0, \tau) \propto e^{\tau \text{tr}\{R_0^T R\}}. \quad (\text{A8})$$

This form is a special case of the distribution defined in Mardia & Jupp (2000, §13.2.3).

If we represent R and R_0 by their corresponding sets of Euler angles Φ and Φ_0 then the probability density in Φ can be written

$$\begin{aligned} f(\Phi | \Phi_0, \tau) &= f(\phi, \theta, \psi | \phi_0, \theta_0, \psi_0, \tau) \\ &= c(\tau) \sin \theta e^{\tau \text{tr}\{R_0^T R\}} \\ &\quad \text{for } (0 \leq \phi \leq 2\pi, 0 \leq \theta \leq \pi, 0 \leq \psi \leq 2\pi), \end{aligned} \quad (\text{A9})$$

where $c(\tau)$ is a normalizing constant which must be determined numerically. The columns of R_0 are the modes for the columns of R .

The case where the mode R_0 is the identity matrix has $(\phi_0, \theta_0, \psi_0) = (0, 0, 0)$ and density

$$\begin{aligned} f_0(\phi, \theta, \psi | \tau) &= c(\tau) \sin \theta e^{\tau \text{tr}\{R\}} \\ &= c(\tau) \sin \theta \exp\{\tau \cos \theta [1 + \cos(\phi + \psi)] \\ &\quad + \tau \cos(\phi + \psi)\} \end{aligned} \quad (\text{A10})$$

A uniform distribution in R has $\tau = 0$ and

$$f_0(\phi, \theta, \psi | 0) = \frac{\sin \theta}{8\pi^2} \quad (\text{A11})$$

which is equivalent to $\phi \sim \text{Uniform}(0, 2\pi)$, $\cos \theta \sim \text{Uniform}(-1, 1)$, and $\psi \sim \text{Uniform}(0, 2\pi)$.

We have determined the following rough relationship between τ and the standard deviation σ_Φ of the angles Φ (expressed in degrees):

$$\sigma_\Phi = e^{3.9155 - 0.5659 \log \tau}. \quad (\text{A12})$$

In particular, standard deviations of $\sigma_\Phi = 1^\circ, 5^\circ, 10^\circ, 15^\circ, 20^\circ$ and 30° correspond, respectively, to τ values of 1000, 60, 17, 8, 5 and 2.5. Note that these standard deviations apply for values of θ close to $\frac{\pi}{2}$, but are less accurate for all angles close to a pole in the coordinate system, that is, when θ is near 0 or π .

A5 Averaging angles

When computing the average of directional data represented by angles in the range $[0, 2\pi]$, care must be taken to ensure that the cyclic nature of angles does not lead to the wrong result. For example, if a sample of angles scattered evenly either side of zero is simply averaged, the result will be π since half of the sample will be just greater than zero, and half just less than 2π . Moreover, if axial data are averaged (i.e. data where the directions ϕ and $\phi + \pi$ are equivalent), then simple approaches to averaging must also be modified.

In this section, we give solutions to the averaging problem for directional data of various forms (see also Mardia & Jupp 2000).

In each case we distinguish directional data from axial data, where for the latter any vector \mathbf{x} is equivalent to $-\mathbf{x}$. The symbol $\langle \dots \rangle$ represents an average over the relevant quantity. The average may be weighted or unweighted.

A5.1 Azimuths

$\{\phi_i\}, i = 1, \dots, n.$

(i) Directional data.

$$\bar{\phi} = \tan^{-1} \left(\frac{\langle \sin \phi \rangle}{\langle \cos \phi \rangle} \right) \quad (\text{A13})$$

(ii) Axial data.

$$\bar{\phi} = \frac{1}{2} \tan^{-1} \left(\frac{\langle \sin 2\phi \rangle}{\langle \cos 2\phi \rangle} \right) \quad (\text{A14})$$

A5.2 Spherical polar coordinates

$\{\phi_i, \theta_i\}, i = 1, \dots, n$, with unit vectors

$$\hat{\mathbf{r}}_i = (\sin \theta_i \cos \phi_i, \sin \theta_i \sin \phi_i, \cos \theta_i)^T \quad (\text{A15})$$

(i) Directional data.

$$\bar{\phi} = \tan^{-1} \left(\frac{\langle \sin \phi \rangle}{\langle \cos \phi \rangle} \right)$$

$$\bar{\theta} = \cos^{-1} \left(\frac{\langle \cos \theta \rangle}{\sqrt{\langle \sin \theta \cos \phi \rangle^2 + \langle \sin \theta \sin \phi \rangle^2 + \langle \cos \theta \rangle^2}} \right) \quad (\text{A16})$$

(ii) Axial data. The unit vector in the direction of the mean is

$$\hat{\mathbf{r}} = (\sin \bar{\theta} \cos \bar{\phi}, \sin \bar{\theta} \sin \bar{\phi}, \cos \bar{\theta})^T, \quad (\text{A17})$$

where $\hat{\mathbf{r}}$ is the normalized eigenvector corresponding to the largest eigenvalue of the matrix

$$A = \langle \hat{\mathbf{r}}_i \hat{\mathbf{r}}_i^T \rangle. \quad (\text{A18})$$

A5.3 Euler angles

$\{\Phi_i = (\phi_i, \theta_i, \psi_i)\}, i = 1, \dots, n$, with Euler rotation matrices $E(\Phi_i)$ as defined in eq. (A4).

(i) Directional data. The Euler matrix corresponding to the mean $\bar{\Phi}$ is the polar (orthogonal) part of the matrix $X = \langle E(\Phi_i) \rangle$, that is,

$$E(\bar{\Phi}) = X[(X^T X)^{1/2}]^{-1}. \quad (\text{A19})$$

Here $(X^T X)^{1/2}$ is the positive definite symmetric square root MDM^T where $(\lambda_1, \lambda_2, \lambda_3)$ are the eigenvalues of $X^T X$, M is the matrix of (normalized) eigenvectors, and $D = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3})$ (Mardia & Jupp 2000, §13.2.3).

(ii) Axial data. The Euler matrix corresponding to the mean $\bar{\Phi}$ is the polar (orthogonal) part of the matrix

$$X = [\hat{\mathbf{v}}_1 \ \hat{\mathbf{v}}_2 \ \hat{\mathbf{v}}_3]. \quad (\text{A20})$$

Here $\hat{\mathbf{v}}_k (k = 1, 2, 3)$ is the eigenvector corresponding to the largest eigenvalue of the matrix

$$A_k = \langle E(\Phi_i)_{\bullet k} [E(\Phi_i)_{\bullet k}]^T \rangle \quad (\text{A21})$$

and $E(\Phi_i)_{\bullet k}$ is the k^{th} column of $E(\Phi_i)$.